# SUPPORTING INFORMATION

## Gene-guided discovery and engineering of branched cyclic peptides in plants

Roland D. Kersten[a,1], Jing-Ke Weng[a,b,1]

[a] Whitehead Institute for Biomedical Research, 455 Main Street, Cambridge, MA, United States
[b] Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, United States

[1] Corresponding authors: Roland D. Kersten, E-mail: rkersten@wi.mit.edu, Jing-Ke Weng, E-mail: wengj@wi.mit.edu

www.pnas.org/cgi/doi/10.1073/pnas.1813993115

## Table of contents

## SUPPORTING TEXT

### Chemotyping of lyciumin peptides from plant material

For comparative chemotyping of lyciumin concentrations in different plant tissues, peptides were extracted from plant tissues as described above from three different plants of the same age. Analyzed tissues of *Amaranthus hypochondriacus* and *Chenopodium quinoa* (three month old) were flower, leaf, root, seed and stem. Analyzed tissues of *Beta vulgaris* (three month old) were leaf, root, seed and stem. Analyzed tissues of *Glycine max* (three month old) were bean, leaf, pod, root and stem. Analyzed tissues for *Solanum tuberosum* (three week old) were sprout and tuber. Peptide extracts were subjected to low resolution MS analysis by selected-ion monitoring (SIM) of lyciumin masses specific to each plant with the following LC-MS parameters: LC – Phenomenex Kinetex® 2.6 µm C18 reverse phase 100 Å 150 x 3 mm LC column, LC gradient: solvent A – 0.1% formic acid, solvent B – acetonitrile (0.1% formic acid), 0-1 min: 5% B, 1-8 min: 5-95% B, 8-10 min: 95% B, 10-15 min: 5% B, MS – positive ion mode, SIM (*Amaranthus hypochondriacus*: 872.8-873.8 m/z and 963.8-964.8 m/z, *Chenopodium quinoa*: 869.8-870.8 m/z and 972.8-973.8 m/z, *Beta vulgaris*: 894.8-895.8 m/z, *Glycine max*: 910.8-911.8 m/z and 993.8-994.8 m/z, *Solanum tuberosum*: 880.8-881.8 m/z, 896.8-897.8 m/z, 913.8-914.8 m/z, 922.8-923.8 m/z, 947.8-948.8 m/z, 972.8-973.8 m/z and 1048.8-1049.8 m/z). Lyciumin ion abundance values were determined by peak area integration from each lyciumin SIM chromatogram in QualBrowser in the Thermo Xcalibur software package (version 3.0.63, ThermoScientific).

### Phylogenetic analysis of lyciumin precursor peptides

Protein sequences of characterized and predicted lyciumin precursors from genomes (Dataset S1, except 3′-partial sequences) and four founding members of the BURP domain family (NP_001303011.1 - BURP domain-containing protein BNM2A precursor [*Brassica napus*],NP_001234835.1 - Polygalacturonase-1 non-catalytic subunit beta precursor [*Solanum lycopersicum*], CAA31603.1/CAA31602.1 - Embryonic abundant protein USP87 / Embryonic abundant protein USP92 [*Vicia faba*], NP_197943.1 - BURP domain protein RD22 [*Arabidopsis thaliana*]) (*1-4*) were reduced to their BURP domain (Pfam PF03181) and aligned using Muscle algorithm (*5*) in MEGA (ver. 7.0.9) (*6*). A neighbor-joining phylogenetic tree was generated with 2000 bootstrap generations using the p-distance method (*7*) in MEGA.

### Lyciumin engineering in *Nicotiana benthamiana*

IDT gblocks® of *Sali3-2* for lyciumin engineering (from Figure 4B, core peptide sequences are in [brackets]):

```
>Sali3-2
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAACCTTATGGTGTATATACATGGTTAACGGATATTAAA
GACACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCC
ATTTCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGT
CATGGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCT
GTTTGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTA
TGGGTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QPWGVYTW]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAACCTTGGGGTGTATATACATGGTTAACGGATATTAAA
GACACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCC
ATTTCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGT
CATGGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCT
GTTTGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTA
TGGGTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QPYGVYFW]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAACCTTATGGTGTATATTTCTGGTTAACGGATATTAAA
GACACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCC
ATTTCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGT
```

CATGGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCT
GTTTGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTA
TGGGTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QPWGVGAW]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAACCTTGGGGTGTAGGTGCATGGTTAACGGATATTAAA
GACACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCC
ATTTCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGT
CATGGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCT
GTTTGCCACTCAGATACTTCTGGAATGAATCACACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTA
TGGGTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QPFGVYTW]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAACCTTTTGGTGTATATACATGGTTAACGGATATTAAA
GACACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCC
ATTTCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGT
CATGGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCT
GTTTGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTA
TGGGTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QPWGVGTW]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAACCTTGGGGTGTAGGTACATGGTTAACGGATATTAAA
GACACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCC
ATTTCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGT
CATGGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCT
GTTTGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTA
TGGGTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QAYGVYTW]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAAGCTTATGGTGTATATACATGGTTAACGGATATTAAA
GACACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCC
ATTTCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGT
CATGGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCT
GTTTGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTA
TGGGTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QPAGVYTW]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAACCTGCTGGTGTATATACATGGTTAACGGATATTAAA
GACACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCC
ATTTCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGT
CATGGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCT
GTTTGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTA
TGGGTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QPFGFFSW]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAACCTTTTGGTTTCTTCTCATGGTTAACGGATATTAAA
GACACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCC
ATTTCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGT
CATGGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCT
GTTTGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTA
TGGGTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QPYGVYW]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAACCTTATGGTGTATATTGGTTAACGGATATTAAAGAC
ACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCCATT
TCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGTCAT
GGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCTGTT
TGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTATGG
GTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QPYGVYTAW]

ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAACCTTATGGTGTATATACAGCATGGTTAACGGATATT
AAAGACACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTT
GCCATTTCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATG
TGTCATGGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTT
GCTGTTTGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATT
TTATGGGTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QPWGVYSW]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAACCTTGGGGTGTATATTCATGGTTAACGGATATTAAA
GACACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCC
ATTTCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGT
CATGGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCT
GTTTGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTA
TGGGTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QPYAVYTW]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAACCTTATACTGTATATACATGGTTAACGGATATTAAA
GACACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCC
ATTTCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGT
CATGGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCT
GTTTGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTA
TGGGTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QPYTVYTW]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAACCTTATACTGTATATACATGGTTAACGGATATTAAA
GACACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCC
ATTTCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGT
CATGGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCT
GTTTGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTA
TGGGTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QPYGAYTW]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAACCTTATGGTGCATATACATGGTTAACGGATATTAAA
GACACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCC
ATTTCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGT
CATGGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCT
GTTTGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTA
TGGGTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QPYGVATW]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAACCTTATGGTGTAGCTACATGGTTAACGGATATTAAA
GACACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCC
ATTTCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGT
CATGGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCT
GTTTGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTA
TGGGTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QPYGVYAW]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAACCTTATGGTGTATATGCATGGTTAACGGATATTAAA
GACACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCC
ATTTCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGT
CATGGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCT
GTTTGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTA
TGGGTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QPYGVYTA]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAACCTTATGGTGTATATACAGCGTTAACGGATATTAAA
GACACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCC
ATTTCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGT

CATGGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCT
GTTTGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTA
TGGGTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QYGVYTW]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAATATGGTGTATATACATGGTTAACGGATATTAAAGAC
ACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCCATT
TCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGTCAT
GGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCTGTT
TGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTATGG
GTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QGVYTW]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAAGGTGTATATACATGGTTAACGGATATTAAAGACACC
TCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTTGCCATTTCA
AAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATGTGTCATGGG
CTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTTGCTGTTTGC
CACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATTTTATGGGTA
CCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

>Sali3-2-[QAPYGVYTW]
ATGGAATTTCGATGCTCAGTCATCTCTTTTACCATTCTCTTCTCTCTTGCTCTTGCAGGAGAGAGCCATGTCCATGCATCGCTACCTGAGGAAGATTATTGGGAAGCTGTT
TGGCCAAACACTCCCATTCCCACTGCACTGCGAGAGCTTCTAAAGCCTCTCCCTGCAGGTGTTGAAATCGATGAACTCCCTAAGCAAATTGATGATACACAGTACCCAAAA
ACATTCTTCTATAAAGAAGACCTTCATCCAGGCAAAACAATGAAAGTACAATTCACCAAGCGTCCCTATGCACAAGCACCTTATGGTGTATATACATGGTTAACGGATATT
AAAGACACCTCTAAAGAAGGATATAGTTTTGAAGAGATATGCATCAAGAAAGAAGCGTTTGAGGGAGAAGAGAAGTTTTGTGCAAAATCCTTGGGAACAGTAATTGGTTTT
GCCATTTCAAAGCTGGGAAAGAACATTCAAGTACTTTCAAGTTCCTTTGTCAATAAGCAAGAGCAATACACTGTGGAAGGAGTGCAGAATCTTGGAGACAAAGCAGTGATG
TGTCATGGGCTAAATTTCAGAACTGCAGTATTTTACTGCCATAAAGTCCGTGAAACAACAGCTTTCATGGTTCCATTGGTGGCTGGTGATGGAACCAAAACTCAGGCACTT
GCTGTTTGCCACTCAGATACTTCTGGAATGAATCATCACATGCTTCATGAACTCATGGGAGTTGATCCTGGAACTAACCCTGTTTGCCATTTCCTTGGAAGCAAGGCCATT
TTATGGGTACCCAATTTATCTATGGACACTGCCTATCAGACTAACGTTGTTGTTTAA

**Purification and structure elucidation of lyciumins**

For lyciumin A, B and D isolation, *Lycium barbarum* roots (100 g wet weight) were ground with a tissue homogenizer and extracted for 16 h with methanol shaking at 225 rpm and 37 ˚C. For lyciumin C isolation, amaranth grain (4.5 kg) was ground in a tissue homogenizer and extracted for 16 h with methanol shaking at 225 rpm and 37 ˚C. Methanol extracts were filtered and dried *in vacuo*. Dried methanol extracts were resuspended in water and partitioned twice with hexane and twice with ethyl acetate and then extracted twice with n-butanol. n-butanol extracts were dried *in vacuo*. Dried n-butanol extracts were resuspended in 10% methanol and separated by flash liquid chromatography with Sephadex LH20 as a stationary phase and a gradient of 10-100% methanol as a mobile phase. Fractions were collected with a fraction collector and analyzed for lyciumin content by LC-QQQ-MS with the following LC-MS settings: LC – Phenomenex Kinetex® 2.6 µm C18 reverse phase 100 Å 150 x 3 mm LC column, LC gradient: solvent A – 0.1% formic acid, solvent B – acetonitrile (0.1% formic acid), 0.5 mL/min, 0-1 min: 5% B, 1-8 min: 5-95% B, 8-10 min: 95% B, 10-15 min: 5% B, MS – positive ion mode, Full MS: Lyciumin A/B/D - 860-920 m/z, Lyciumin C/I - 950-1010 m/z. LH20 fractions with lyciumins were combined, dried *in vacuo*, resuspended in 10% acetonitrile (0.1% trifluoroacetic acid) and subjected to preparative HPLC with a Phenomenex Kinetex® 5 µm C18 reverse phase 100 Å 150 x 21.2 mm LC column as a stationary phase for two rounds of separation. LC settings were as follows: solvent A – 0.1% trifluoroacetic acid, solvent B – acetonitrile (0.1% trifluoroacetic acid), 10 mL/min, Lyciumin A (20 mg) – 1.LC: 0-3 min: 10% B, 3-43 min: 10-50% B, 43-45 min: 50-95% B, 45-48 min: 95% B, 48-49 min: 95-10% B, 49-69 min: 10% B, 2.LC: 0-5 min: 35% B, 5-35 min: 35-50% B, 35-38 min: 50-95% B, 38-40 min: 95% B, 40-40.1 min: 95-35% B, 40.1-60 min: 35% B, Lyciumin B (13 mg) – 1.LC: 0-3 min: 20% B, 3-48 min: 20-40% B, 48-50 min: 40-95% B, 50-54 min: 95% B, 54-55 min: 95-20% B, 55-70 min: 20% B, 2.LC: 0-3 min: 30% B, 3-35 min: 30-45% B, 35-38 min: 45-95% B, 38-40 min: 95% B, 40-40.1 min: 95-30% B, 40.1-60 min: 30% B, Lyciumin C – 1.LC: 0-3 min: 10% B, 3-43 min: 10-50% B, 43-45 min: 50-95% B, 45-48 min: 95% B, 48-49 min: 95-10% B, 49-69 min: 10% B, 2.LC: 0-3 min: 40% B, 3-48 min: 40-55% B, 48-50 min: 55-95% B, 50-54 min: 95% B, 54-55 min: 95-40%, 55-70 min: 40% B, Lyciumin D (5 mg) – 1.LC: 0-3 min: 20% B, 3-48 min: 20-40% B, 48-50 min: 40-95% B, 50-54 min: 95% B, 54-55 min: 95-20% B, 55-70 min: 20% B, 2.LC: 0-3 min: 30% B, 3-48 min: 30-50% B, 48-50 min: 50-95% B, 50-54 min: 95% B, 54-55 min: 95-30% B, 55-70 min: 30% B, Lyciumin I – 1.LC: 0-3 min: 20% B, 3-48min: 20-

50% B, 48-50 min: 50-95% B, 50-54 min: 95% B, 54-55 min: 95-20% B, 55-70min: 20% B, 2.LC: 0-3 min: 25% B, 3-48 min: 25-45% B, 48-50 min: 45-95% B, 50-54 min: 95% B, 54-55 min: 95-25% B, 55-70min: 25% B. Preparative HPLC fractions with lyciumin C and lyciumin I, respectively, were combined, dried *in vacuo*, resuspended in 30% acetonitrile (0.1% trifluoroacetic acid) and subjected to semipreparative HPLC with a Phenomenex Kinetex® 5 µm C18 reverse phase 100 Å 250 x 10 mm LC column as a stationary phase. LC settings were as follows: Solvent A – 0.1% trifluoroacetic acid, solvent B – acetonitrile (0.1% trifluoroacetic acid), 1.5 mL/min, Lyciumin C (25 mg) - 0-5 min: 40% B, 5-15 min: 40-42% B, 15-17 min: 42-95% B, 17-20 min, 95% B, 20-20.1 min: 95-40% B, and lyciumin I (2.5 mg) - 0-5 min: 30% B, 5-30 min: 30-35% B, 30-32 min: 35-95% B, 32-36 min: 95% B, 36-40 min: 95-30% B, 40-60min: 30% B. For NMR analysis, lyciumin A, B, C, D and I were each dissolved in DMSO-d6. Lyciumin A was analyzed for $^{1}H$ and $^{13}C$ NMR data, lyciumin B, D and C were analyzed for $^{1}H$ NMR data. Lyciumin I was analyzed for $^{1}H$ NMR, $^{1}H$-$^{1}H$ COSY, $^{1}H$-$^{1}H$ TOCSY, HSQC, HMBC and ROESY data. NMR data was analyzed with TopSpin software (v3.5) from Bruker. Stereochemistry of crosslinked glycine α-carbons at the fourth position of lyciumins was inferred as (*R)* based on lyciumin A analysis (*8*) and same ROESY correlations of lyciumin I glycine-Hα as reported for lyciumin A (*8*). Stereochemistry of other amino acids of lyciumin I was inferred as (*L*)-amino acids because of its ribosomal biosynthesis and (*L*)-amino acid stereochemistry in all reported lyciumins (*8-10*).

**Gene expression analysis of characterized lyciumin precursors**

Gene expression of characterized lyciumin precursors was estimated by mapping raw sequencing reads to *de novo* assembled transcriptomes using RSEM (*11*). For *Solanum tuberosum*, gene expression of lyciumin precursor *PGSC0003DMG400047074* was analyzed in 16 tissue samples (NCBI SRA datasets: ERR029909, ERR029910, ERR029911, ERR029912, ERR029913, ERR029914, ERR029915, ERR029916, ERR029917, ERR029918, ERR029919, ERR029920, ERR029921, ERR029922, ERR029923, ERR029924) (*12*) by RSEM against the combined *de novo* Trinity (*13*)-assembled transcriptome of all 16 samples (Figure S24). For *Amaranthus hypochondriacus*, gene expression of lyciumin precursor *AHYPO_007393-RA* was analyzed in eight tissues and conditions (NCBI SRA: SRR1598916, SRR1598915, SRR1598914, SRR1598913, SRR1598912, SRR1598911, SRR1598910, SRR1598909) by RSEM against the combined *de novo* Trinity-assembled transcriptome of all eight samples (Figure S24). For *Chenopodium quinoa*, gene expression of lyciumin precursor *AUR62017095* was analyzed in 15 tissue samples (NCBI SRA: SRR5974430, SRR5974427, SRR5974436, SRR5974438, SRR5974437, SRR5974435, SRR5974432, SRR5974433, SRR5974425, SRR5974426, SRR5974424, SRR5974431, SRR5974428, SRR5974429, SRR5974434) against the combined *de novo* Trinity-assembled transcriptome of all 15 samples (Figure S24). For *Medicago truncatula*, gene expression of lyciumin precursor *Medtr2g081610* was assessed by the eFP bar.utoronto.ca webbrowser (*14*) of gene expression data from *Medicago truncatula* RNA-seq dataset (*15*) and displayed (Figure S24). For *Glycine max*, gene expression of lyciumin precursors *Sali3-2* and *Glyma.12G217300* was assessed by the eFP bar.utoronto.ca webbrowser (*14*) of gene expression data from *Glycine max* RNA-seq datasets (*16,17*) and displayed (Figure S24).

**Glutamine cyclotransferase co-expression assays with *LbaLycA* in *Nicotiana benthamiana***

Glutamine cyclotransferase LbaQC was characterized as the closest homolog of *Chenopodium quinoa* glutamine cyclotransferase (AUR62017096, Phytozome v12.1) by blastp search of *L. barbarum* root transcriptome on an internal Blast server (*18*). *LbaQC* was synthesized as a gBlock® (ATGGTTTCTTCTACTTCATATCTACCTACCAATCACACAAAAATGCCTCTGCTAAATCCAAGGTTTCTA
GTCATAAGCTTGATTGTTCTACTGAGCATCACCGTATTCAGAGAAGCTGAAGCATCATATAGAGTTTAC
AAAGTCAAAGTAGTCAATGAATTCCCTCACGACCCCCAAGCCTACACTCAGGGGCTTCTCTATGCAGA
AAATAATACACTCTTTGAATCAACTGGACTTTACGGACGTTCATCTGTTCGAAAAGTTGCATTGCTGGA
CGGTAAGGTTGAGAGACTTCATGAAATGGAGTCTTCTTACTTTGGAGAGGGTCTAACTCTTCTTGGTGA
GAGGTTGTTCCAACTAACATGGTTGCTGGATACAGGTTTCATATATGATCGATACAACTTCAGCAAATT
CAAAAAGTTTACTCATCACATGCAAGATGGTTGGGGATTGGCAACCGATGGGAAAGTACTTTTTGGAA
GTGATGGAACATCAACATTATATAAGATTGACCCTAAAACAATGAAAGTCATCAGAAAACAAGTTGTC
AAGTCTCAAGGGCATGAAGTGCGCTACCTGAATGAGCTGGAGTATGTGAAAGCTGAAGTCTGGGCAA

ATGTTTATGTGACTGATTGCATTGCTAGAATTTCACCAAAAGATGGCACTGTGATCGGGTGGATTCTCC
TTCAATCTCTAAGAGAAGAGTTAATATCAAGAGGATATAAGGACTTCGAGGTCCTGAATGGAATCGCA
TGGGACAGAGATGGTGACCGTATTTTTGTGACAGGGAAACTATGGCCAAAGCTCTTTGAGATCAAGTT
GCTCCCCCTCACACCGAATGATCCATTGGCTGGAGAAATCAATAACTTGTGCATCCCGAAAACCAGTT
TTCTCTTGGAAATTTAG) with a 5′-adapter (tgcccaaattcgcgaccggt) and a 3′-adapter (ctcgaggcctttaactctgg) for Gibson assembly (*19*). pEAQ-HT (*20*) was digested by AgeI and XhoI restriction enzymes and the *LbaQC* gBlock® was cloned into linearized pEAQ-HT with Gibson Assembly Master Mix (New England Biolabs). pEAQ-HT-*LbaQC* was verified by Sanger sequencing. For glutamine cyclotransferase co-expression assays, pEAQ-HT-*LbaQC* and pEAQ-HT-*LbaLycA* were transformed into *Agrobacterium tumefaciens* LBA4404 for heterologous expression as described above. For co-expression assay, leaves of three plants of *Nicotiana benthamiana* (six week old) were infiltrated with a 1:1 mixture of resuspended *A. tumefaciens* LBA4404 pEAQ-HT-*LbaQC* (OD 0.8) and *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA* (OD 0.8). For *LbaLycA* control expression without *LbaQC*, leaves of three plants of *Nicotiana benthamiana* (six week old) were infiltrated with resuspended *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA* (OD 0.4). Infiltrated plants were cultivated as described before for six days for heterologous expression. After six days, leaves of three plants of the *LbaQC-LbaLycA* co-expression and leaves of three plants of *LbaLycA* expression control were collected and freeze-dried. For comparative chemotyping of [Gln1]-lyciumin B, [Gln1]-lyciumin D, lyciumin B and lyciumin D, peptides were extracted from 0.1 g of freeze-dried tobacco leaves as described above for peptide chemotyping from the *LbaQC-LbaLycA* co-expression plants and from the *LbaLycA* expression plants. Peptide extracts were subjected to low resolution MS analysis by selected-ion monitoring (SIM) of masses of [Gln1]-lyciumins and lyciumins with the following LC-MS parameters: LC – Phenomenex Kinetex® 2.6 µm C18 reverse phase 100 Å 150 x 3 mm LC column, LC gradient: solvent A – 0.1% formic acid, solvent B – acetonitrile (0.1% formic acid), 0.5 mL/min, 0-1 min: 5% B, 1-8 min: 5-95% B, 8-10 min: 95% B, 10-15 min: 5% B, MS – positive ion mode, SIM: 896.8-897.8 m/z (lyciumin B), 899.8-900.8 m/z (lyciumin D), 913.8-914.8 m/z ([Gln1]-lyciumin B), 916.8-917.8 m/z ([Gln1]-lyciumin D). Lyciumin and [Gln1]-lyciumin ion abundance values were determined by peak area integration from each peptide SIM chromatogram in QualBrowser in the Thermo Xcalibur software package (version 3.0.63, ThermoScientific).

**LbaLycA-based lyciumin production in *Nicotiana benthamiana* in comparison to source plant extraction**

LbaLycA precursors with one, five or ten core peptide repeats of lyciumin B (QPWGVGSW) were designed manually by combination of different spacer (non-core peptide) sequences from Solanaceae lyciumin precursors with multiple core peptides with lyciumin B core peptide sequences and synthesized as gBlocks® with a 5′-adapter (tgcccaaattcgcgaccggt) and a 3′-adapter (ctcgaggcctttaactctgg) for Gibson assembly (*19*). pEAQ-HT was digested by AgeI and XhoI restriction enzymes and the *LbaLycA* gBlocks® were cloned into linearized pEAQ-HT with Gibson Assembly Master Mix (New England Biolabs). pEAQ-HT-*LycA-1x/5x/10x-QPWGVGSW* were verified by Sanger sequencing. For expression assays, pEAQ-HT-*LycA-1x-QPWGVGSW,* pEAQ-HT-*LycA-5x-QPWGVGSW* or pEAQ-HT-*LycA-10x-QPWGVGSW* were transformed into *Agrobacterium tumefaciens* LBA4404 for heterologous expression as described above. Multiple leaves of three plants of *Nicotiana benthamiana* (six week old) were infiltrated with resuspended *A. tumefaciens* LBA4404 pEAQ-HT-*LycA-1x-QPWGVGSW,* pEAQ-HT-*LycA-5x-QPWGVGSW* or pEAQ-HT-*LycA-10x-QPWGVGSW* (OD 0.8). Infiltrated plants were cultivated as described before for six days for heterologous expression. After six days, infiltrated leaves of the three plants were collected and freeze-dried. For source plant extraction, roots of three *Lycium barbarum* plants (one year old) were collected and freeze-dried. For comparative lyciumin B chemotyping, 0.1 g of freeze-dried leaves of each *N. benthamiana* plant and 0.1 g of freeze-dried roots were extracted as described above. Peptide extracts were subjected to low resolution MS analysis by selected-ion monitoring (SIM) of lyciumin B with the following LC-MS parameters: LC – Phenomenex Kinetex® 2.6 µm C18 reverse phase 100 Å 150 x 3 mm LC column, LC gradient: solvent A – 0.1% formic acid, solvent B – acetonitrile (0.1% formic acid), 0.5 mL/min, 0-1 min: 5% B, 1-8 min: 5-95% B, 8-10 min: 95% B, 10-15 min: 5% B, MS – positive ion mode, SIM: 896.8-897.8 m/z (lyciumin B). Lyciumin B ion abundance values were determined by peak area integration from each peptide SIM chromatogram in QualBrowser in the Thermo Xcalibur software package (version 3.0.63, ThermoScientific).

>LycA-1x-QPWGVGSW
ATGGAGTTGCATCACCATTACTTCTTCATACTTCTTTCTCTTGCTTTTATAGCAAGTCATGCAGCTAATTTATCTCCTGAGGTGTATTGGAAAGTCAAGCTGCCCAACACT
CCTATGCCCAGACCCATTAAGGATGCTCTACACTATTCTGAAGCCTCCGAGGGTGACGTTCACAAGTTGCGCCAACCATGGGGAGTGGGTTCGTGGTATAATACTGCTACA
AAGAAAGATGTTAATGAAAACCTCCCAGTCACCCCTTACTTTTTTGAAACAGATTTACATCAAGGGAAAAAGATGAATCTTCCATCTCTCAAAAATTATAATCCAGCTCCC
ATTTTGCCTCGCAAAGTTGCAGATTCCATCCCCTTCTCATCAGACAAGATTGAAGAAATTCTAAAGCACTTTTCCATTGATAAGGACTCAGAGGGGGCTAAAATGATCAAG
AAAACTATCAAAATGTGTGAGGAGCAAGCGGGTAATGGCGAGAAGAAATATTGTGCCACTTCCTTAGAATCAATGGTTGATTTCACCTCATCTTATCTGGGAACAAATAAT
ATTATAGCACTTTCCACTTTAGTAGAGAAGGAAACTCCAGAGGTGCAAATATATACCATCGAAGAAGTGAAAGAGAAAGCAAATGGCAAAGGCGTGATATGCCACAAAGTG
GCTTACCCGTATGCGATACATTATTGCCATAGTGTAGGAAGCACAAGGACCTTTATGGTCTCAATGGTGGGTTCTGATGGAACAAAAGTTAATGCAGTATCAGAGTGTCAT
GAGGATACTGCACCCATGAACCCTAAGGCATTGCCTTTTCAATTGCTCAACGTTAAGCCAGGAGATAAACCTATTTGCCATTTCATATTGGATGATCAGATTGCCTTAGTT
CCTTCTCAAGACGCAACTCAAGTGTCTGAAAACTAA

>LycA-5x-QPWGVGSW
ATGGAGTTGCATCACCATTACTTCTTCATACTTCTTTCTCTTGCTTTTATAGCAAGTCATGCAGCTAATTTATCTCCTGAGGTGTATTGGAAAGTCAAGCTGCCCAACACT
CCTATGCCCAGACCCATTAAGGATGCTCTACACTATTCTGAAGCCTCCGAGGGTGACGTTCACAAGTTGCGCCAACCATGGGGAGTGGGTTCGTGGTATCAAGCAGCAAAC
GAGGGTGATGTTAAAAAATTACGCCAGCCTTGGGGTGTTGGCTCTTGGTATAAGGCTGCCCCAGAGGACGAGCTTCACAAAATACGCCAGCCCTGGGGCGTCGGATCCTGG
TATCAAGCTGCCAAAGAGAATGACCTACCCAGAATGTCCCAGCCGTGGGGGGTAGGGAGTTGGTATCAGGCTGCCCCCGAGAACGAGCTTCACAAAGTACGGCAGCCATGG
GGCGTCGGAAGCTGGTACCAACCCGCAGCAGAAGGGGATTTACACAAGCTCCGATATAATACTGCTACAAAGAAAGATGTTAATGAAAACCTCCCAGTCACCCCTTACTTT
TTTGAAACAGATTTACATCAAGGGAAAAAGATGAATCTTCCATCTCTCAAAAATTATAATCCAGCTCCCATTTTGCCTCGCAAAGTTGCAGATTCCATCCCCTTCTCATCA
GACAAGATTGAAGAAATTCTAAAGCACTTTTCCATTGATAAGGACTCAGAGGGGGCTAAAATGATCAAGAAAACTATCAAAATGTGTGAGGAGCAAGCGGGTAATGGCGAG
AAGAAATATTGTGCCACTTCCTTAGAATCAATGGTTGATTTCACCTCATCTTATCTGGGAACAAATAATATTATAGCACTTTCCACTTTAGTAGAGAAGGAAACTCCAGAG
GTGCAAATATATACCATCGAAGAAGTGAAAGAGAAAGCAAATGGCAAAGGCGTGATATGCCACAAAGTGGCTTACCCGTATGCGATACATTATTGCCATAGTGTAGGAAGC
ACAAGGACCTTTATGGTCTCAATGGTGGGTTCTGATGGAACAAAAGTTAATGCAGTATCAGAGTGTCATGAGGATACTGCACCCATGAACCCTAAGGCATTGCCTTTTCAA
TTGCTCAACGTTAAGCCAGGAGATAAACCTATTTGCCATTTCATATTGGATGATCAGATTGCCTTAGTTCCTTCTCAAGACGCAACTCAAGTGTCTGAAAACTAA

>LycA-10x-QPWGVGSW
ATGGAGTTGCATCACCATTACTTCTTCATACTTCTTTCTCTTGCTTTTATAGCAAGTCATGCAGCTAATTTATCTCCTGAGGTGTATTGGAAAGTCAAGCTGCCCAACACT
CCTATGCCCAGACCCATTAAGGATGCTCTACACTATTCTGAAGCCTCCGAGGGTGACGTTCACAAGTTGCGCCAACCATGGGGAGTGGGTTCGTGGTATCAAGCAGCAAAC
GAGGGTGATGTTAAAAAATTACGCCAGCCTTGGGGTGTTGGCTCTTGGTATAAGGCTGCCCCAGAGGACGAGCTTCACAAAATACGCCAACCCTGGGGCGTCGGATCCTGG
TATCAAGCTGCCAAAGAGAATGACCTACCCAGAATGTCCCAGCCGTGGGGGGTAGGGAGTTGGTATCAGGCTGCCCCCGAGAACGAGCTTCACAAAGTACGGCAACCATGG
GGCGTCGGAAGCTGGTACCAACCCGCAGCAGAAGGGGATTTACACAAGCTCCGACAGCCTTGGGGCGTGGGGAGTTGGTACAACGACGCACCAGAGAATGAACTCCATAAG
TTCCGTCAACCCTGGGGAGTTGGATCGTGGTACAGGGCAGCTACTGAAGGGGACGTTCAAAAGCTGCGTCAGCCGTGGGGCGTCGGCTCATGGTACCAAGCAGCAAACGAG
GGAGACATTAAGAAGCTACGGCAGCCATGGGGAGTCGGCTCTTGGTACAGAGCCGCTACTGAGGGTGACGTTCAGAAGCTAAGACAGCCCTGGGGGGTCGGTAGTTGGTAT
AATACTGCTACAAAGAAAGATGTTAATGAAAACCTCCCAGTCACCCCTTACTTTTTTGAAACAGATTTACATCAAGGGAAAAAGATGAATCTTCCATCTCTCAAAAATTAT
AATCCAGCTCCCATTTTGCCTCGCAAAGTTGCAGATTCCATCCCCTTCTCATCAGACAAGATTGAAGAAATTCTAAAGCACTTTTCCATTGATAAGGACTCAGAGGGGGCT
AAAATGATCAAGAAAACTATCAAAATGTGTGAGGAGCAAGCGGGTAATGGCGAGAAGAAATATTGTGCCACTTCCTTAGAATCAATGGTTGATTTCACCTCATCTTATCTG
GGAACAAATAATATTATAGCACTTTCCACTTTAGTAGAGAAGGAAACTCCAGAGGTGCAAATATATACCATCGAAGAAGTGAAAGAGAAAGCAAATGGCAAAGGCGTGATA
TGCCACAAAGTGGCTTACCCGTATGCGATACATTATTGCCATAGTGTAGGAAGCACAAGGACCTTTATGGTCTCAATGGTGGGTTCTGATGGAACAAAAGTTAATGCAGTA
TCAGAGTGTCATGAGGATACTGCACCCATGAACCCTAAGGCATTGCCTTTTCAATTGCTCAACGTTAAGCCAGGAGATAAACCTATTTGCCATTTCATATTGGATGATCAG
ATTGCCTTAGTTCCTTCTCAAGACGCAACTCAAGTGTCTGAAAACTAA

LbaLycA-based lyciumin precursors with 1, 5 or 10 core peptides of lyciumin B (BURP domain underlined, core peptides in red):

>LycA-1x-QPWGVGSW
MELHHHYFFILLSLAFIASHAANLSPEVYWKVKLPNTPMPRPIKDALHYSEASEGDVHKLRQPWGVGSWYNTATKKDVNENLPVTPYFFETD
LHQGKKMNLPSLKNYNPAPILPRKVADSIPFSSDKIEEILKHFSIDKDSEGAKMIKKTIKMCEEQAGNGEKKYCATSLESMVDFTSSYLGTN
NIIALSTLVEKETPEVQIYTIEEVKEKANGKGVICHKVAYPYAIHYCHSVGSTRTFMVSMVGSDGTKVNAVSECHEDTAPMNPKALPFQLLN
VKPGDKPICHFILDDQIALVPSQDATQVSEN

>LycA-5x-QPWGVGSW
MELHHHYFFILLSLAFIASHAANLSPEVYWKVKLPNTPMPRPIKDALHYSEASEGDVHKLRQPWGVGSWYQAANEGDVKKLRQPWGVGSWYK
AAPEDELHKIRQPWGVGSWYQAAKENDLPRMSQPWGVGSWYQAAPENELHKVRQPWGVGSWYQPAAEGDLHKLRYNTATKKDVNENLPVTPY
FFETDLHQGKKMNLPSLKNYNPAPILPRKVADSIPFSSDKIEEILKHFSIDKDSEGAKMIKKTIKMCEEQAGNGEKKYCATSLESMVDFTSS
YLGTNNIIALSTLVEKETPEVQIYTIEEVKEKANGKGVICHKVAYPYAIHYCHSVGSTRTFMVSMVGSDGTKVNAVSECHEDTAPMNPKALP
FQLLNVKPGDKPICHFILDDQIALVPSQDATQVSEN

>LycA-10x-QPWGVGSW
MELHHHYFFILLSLAFIASHAANLSPEVYWKVKLPNTPMPRPIKDALHYSEASEGDVHKLRQPWGVGSWYQAANEGDVKKLRQPWGVGSWYK
AAPEDELHKIRQPWGVGSWYQAAKENDLPRMSQPWGVGSWYQAAPENELHKVRQPWGVGSWYQPAAEGDLHKLRQPWGVGSWYNDAPENELH
KFRQPWGVGSWYRAATEGDVQKLRQPWGVGSWYQAANEGDIKKLRQPWGVGSWYRAATEGDVQKLRQPWGVGSWYNTATKKDVNENLPVTPY
FFETDLHQGKKMNLPSLKNYNPAPILPRKVADSIPFSSDKIEEILKHFSIDKDSEGAKMIKKTIKMCEEQAGNGEKKYCATSLESMVDFTSS
YLGTNNIIALSTLVEKETPEVQIYTIEEVKEKANGKGVICHKVAYPYAIHYCHSVGSTRTFMVSMVGSDGTKVNAVSECHEDTAPMNPKALP
FQLLNVKPGDKPICHFILDDQIALVPSQDATQVSEN

**Table S1 | NMR analysis of lyciumin A from *Lycium barbarum* (600 MHz, DMSO-d6).**

| Residue | C | δ($^{13}$C) [ppm] [a] | δ($^{13}$C) [ppm] [b] | H | δ($^{1}$H) (int, m, *J*) [ppm] [c] | δ($^{1}$H) (int, m, *J*) [ppm] [d] |
|---|---|---|---|---|---|---|
| PyroGlu[1] | α | 54.8 | 54.6 | α | 4.35 (1H, m) | 4.35 (m) |
| | β | 23.9 | 23.9 | β1 | 2.28 (1H, m)/1.89 (m) | 2.29 (m)/1.9 (m) |
| | γ | 26.5 | 27.3 | γ | 2.10 (m) | 2.10 (2H, m) |
| | δ | 177.4 | 177.3 | NH | 7.9 (m) | n/a |
| Pro[2] | α | 59.5 | 59.4 | α | 4.35 (1H, m) | 4.35 (m) |
| | β | 29.1 | 29.0 | β | 2.11 (m)/1.75 (m) | 2.16 (m)/1.75 (m) |
| | γ | 24.4 | 24.2 | γ | 1.81 (m)/1.69 (m) | 1.82 (m)/ 1.68 (m) |
| | δ | 46.2 | 46.1 | δ | 3.63 (m)/3.36 (m) | 3.60 (m)/3.35 (m) |
| Tyr[3] | α | 53.9 | 53.8 | α | 4.33 (1H, m) | 4.33 (t, 7) |
| | β | 36.3 | 36.3 | β | 2.63 (2H, m) | 2.63 (2H, d, 7) |
| | C1 | 126.6 | 126.5 | H2 | 6.68 (1H, d, 8.4) | 6.63 (1H, d, 8) |
| | C2 | 129.9 | 129.8 | H3 | 6.35 (1H, d, 8.3) | 6.38 (1H, d, 8) |
| | C3 | 114.8 | 114.9 | OH | 7.73 (1H, s) | 7.71 (s, OH) |
| | C4 | 155.8 | 155.6 | H5 | 6.35 (1H, d, 8.3) | 6.38 (1H, d, 8) |
| | C5 | 114.8 | 114.9 | H6 | 6.68 (1H, d, 8.4) | 6.63 (1H, d, 8) |
| | C6 | 129.9 | 129.8 | NH | 7.94 (m) | 7.97 (br d, 7) |
| Gly[4] | α | 61.0 | 61.4 | α | 6.66 (1H, d, 8.4) | 6.67 (d, 8) |
| | | | | NH | 9.28 (1H, d, 8.4) | 9.37 (d, 8) |
| Val[5] | α | 51.7 | 56.3 | α | 4.00 (1H, t, 8.1) | 3.99 (t, 7) |
| | β | 28.9 | 29.0 | β | 2.05 (m) | 2.03 (m) |
| | γ1 | 19.6 | 19.2 | γ1 | 0.91 (3H, d, 6.7) | 0.87 (3H, d, 7) |
| | γ2 | 18.9 | 18.5 | γ2 | 0.86 (3H, d, 6.8) | 0.82 (3H, d, 7) |
| | | | | NH | 7.92 (m) | 7.92 (br d, 7) |
| Gly[6] | α | 43.1 | 43.5 | α | 4.02 (1H, m)/ 3.22 (1H, m) | 4.08 (dd, 6, 15)/ 3.23 (dd, 6, 15) |
| | | | | NH | 8.53 (1H, t, 6) | 8.68 (t, 6) |
| Ser[7] | α | 55.2 | 55.4 | α | 4.22 (1H, dd, 6, 13) | 4.11 (dd, 7, 11) |
| | β | 61.7 | 62.0 | β | 3.59 (m)/3.4 (m) | 3.59 (m)/3.49 (m) |
| | | | | NH | 7.90 (m) | 7.72 (br d, 7) |
| Trp[8] | α | 53.9 | 53.8 | α | 4.54 (1H, t) | 4.40 (t, 7) |
| | β | 28.6 | 28.3 | β | 3.30 (m)/3.16 (m) | 3.30 (m)/ 3.01 (m) |
| | C2 | 123.8 | 124.7 | H2 | 7.15 (s) | 6.91 (s) |
| | C3 | 113.2 | 111.2 | H4 | 7.53 (1H, d, 8.1) | 7.54 (d, 8) |
| | C4 | 119.1 | 119.5 | H5 | 7.08 (1H, t, 7.5) | 7.04 (t, 8) |
| | C5 | 121.3 | 121.9 | H6 | 7.18 (1H, t, 7.7) | 7.14 (t, 8) |
| | C6 | 119.1 | 118.7 | H7 | 7.37 (1H, d, 8.4) | 7.38 (d, 8) |
| | C7 | 109.4 | 109.8 | NH | 7.9 (m) | 7.83 (d, 7) |
| | C=O | 166.4 | 166.5 | | | |
| | | 169.4 | 169.2 | | | |
| | | 169.9 | 169.2 | | | |
| | | 170.6 | 170.6 | | | |
| | | 171.5 | 171.3 | | | |
| | | 171.5 | 171.3 | | | |
| | | 171.6 | 171.5 | | | |
| | | 173.0 | 175.7 | | | |

[a] $^{13}$C NMR data of isolated lyciumin A in DMSO-d6. Values were derived from $^{13}$C NMR analysis (Figure S3) [b] $^{13}$C NMR data of lyciumin A in DMSO-d6 from reference 9. [c] $^{1}$H NMR data of isolated lyciumin A in DMSO-d6. Multiplicity m (s=singlet, d=doublet, t=triplet, dd=double doublet, m=multiplet), intensity int, coupling constants *J* in Hertz (Figure S2). [d] $^{1}$H NMR data of lyciumin A in DMSO-d6 from reference 9.

**Table S2 | NMR analysis of lyciumin B from *Lycium barbarum* (600 MHz, DMSO-d6).**

| Residue | H | δ($^1$H) (int, m, *J*) [ppm] [a] | δ($^1$H) (int, m, *J*) [ppm] [b] |
|---|---|---|---|
| PyroGlu[1] | α | 4.45 (1H, dd, 7, 13) | 4.36 (m) |
| | β | 2.23 (1H, m)/1.83 (1H, m) | 2.21 (m)/1.83 (m) |
| | γ | 2.08 (2H, m) | 2.08 (2H, m) |
| | NH | 7.89 (1H) | n/a |
| Pro[2] | α | 4.34 (1H, m) | 4.36 (m) |
| | β | 2.10 (m)/1.97 (m) | 2.10 (m)/1.95 (m) |
| | γ | 1.72 (2H, m) | 1.72 (2H, m) |
| | δ | 3.63 (m)/3.33 (m) | 3.64 (dd, 5, 11)/3.11 (dd, 9, 11) |
| Trp[3] | α | 4.34 (1H, m) | 4.32 (m) |
| | β | 3.31 (m)/2.9 (1H, m) | 3.26 (m)/2.93 (m) |
| | H2 | 6.86 (s) | 6.87 (s) |
| | H4 | 7.53 (1H, d, 7.7) | 7.56 (d, 8) |
| | H5 | 7.06 (1H, t, 7.5) | 7.01 (t, 8) |
| | H6 | 7.14 (1H, t, 7.8) | 7.08 (t, 8) |
| | H7 | 7.28 (1H, t, 8) | 7.28 (t, 8) |
| | N(indole)H | 10.68 (1H, s) | 10.66 (s) |
| | NH | 7.93 (1H, d, 7.6) | 7.86 (d, 6) |
| Gly[4] | α | 6.61 (d, 8.2) | 6.67 (d, 8) |
| | NH | 9.30 (d, 8.1) | 9.35 (d, 8) |
| Val[5] | α | 4.01 (m) | 4.01 (t, 7) |
| | β | 2.08 (m) | 2.08 (m) |
| | γ1 | 0.90 (3H, d, 6.7) | 0.87 (3H, d, 7) |
| | γ2 | 0.86 (3H, d, 6.8) | 0.82 (3H, d, 7) |
| | NH | 7.97 (1H, d, 6.3) | 7.89 (br d, 7) |
| Gly[6] | α | 4.04 (m)/ 3.26 (1H, m) | 4.08 (dd, 6, 15)/3.25 (m) |
| | NH | 8.56 (1H, t, 6) | 8.55 (br s) |
| Ser[7] | α | 4.22 (1H, dd, 6.1, 12.7) | 4.18 (dd, 6, 12) |
| | β | 3.55 (m)/3.36 (m) | 3.55 (m)/3.30 (m) |
| | NH | 7.93 (1H, d, 8) | 7.69 (br d, 7) |
| Trp[8] | α | 4.53 (1H, dt, 3.3, 8.1) | 4.52 (m) |
| | β | 3.16 (1H, br s)/2.9 (1H, m) | 3.26 (m)/2.93 (br s) |
| | H2 | 7.17 (s) | 7.55 (s) |
| | H4 | 7.37 (d, 7.7) | 7.37 (d, 8) |
| | H5 | 6.86 (1H, t) | 6.86 (t, 8) |
| | H6 | 7.00 (1H, t, 7.6) | 7.00 (t, 8) |
| | H7 | 7.26 (1H, d, 8) | 7.26 (d, 8) |
| | NH | 7.71 (s) | 7.69 (br d, 7) |

[a] $^1$H NMR data of isolated lyciumin B in DMSO-d6. Multiplicity m (s=singlet, d=doublet, t=triplet, dd=double doublet, m=multiplet), intensity int, coupling constants *J* in Hertz (Figure S5). [b] $^1$H NMR data of lyciumin B in DMSO-d6 from reference 9.

**Table S3 | NMR analysis of lyciumin D from *Lycium barbarum* (600 MHz, DMSO-d6).**

| Residue | H | δ($^1$H) (int, m, *J*) [ppm] [a] | δ($^1$H) (int, m, *J*) [ppm] [b] |
|---|---|---|---|
| PyroGlu[1] | α | 4.50 (1H, m) | 4.36 (m) |
| | β | 2.27 (m)/1.88 (m) | 2.31 (m)/1.85 (m) |
| | γ | 2.10 (2H, m) | 2.10 (2H, m) |
| | NH | 8.32 (1H, d, 7.5) | n/a |
| Pro[2] | α | 4.34 (1H, m) | 4.36 (m) |
| | β | 2.00 (1H, m)/1.80 (m) | 1.97 (m)/1.80 (m) |
| | γ | 1.83 (2H, m) | 1.83 (2H, m) |
| | δ | 3.57 (m) | 3.58 (m)/3.41 (m) |
| Tyr[3] | α | 4.34 (1H, m) | 4.33 (m) |
| | β | 2.62 (2H, m) | 2.64 (m) |
| | H2 | 6.66 (1H, d, 8.4) | 6.52 (1H, d, 8) |
| | H3 | 6.33 (1H, d, 8.4) | 6.35 (1H, d, 8) |
| | OH | 9.05 (s) | 7.53 (s) |
| | H5 | 6.33 (1H, d, 8.4) | 6.35 (1H, d, 8) |
| | H6 | 6.66 (1H, d, 8.4) | 6.52 (1H, d, 8) |
| | NH | 7.90 (1H, d, 7.9) | 7.85 (d, 5) |
| Gly[4] | α | 6.65 (d, 8.5) | 6.67 (d, 9) |
| | NH | 9.38 (1H, d, 8.5) | 9.15 (d, 9) |
| Val[5] | α | 4.01 (m) | 4.04 (m) |
| | β | 2.01 (m) | 2.10 (m) |
| | γ1 | 0.87 (3H, d, 7) | 0.86 (3H, d, 6) |
| | γ2 | 0.85 (3H, d, 7) | 0.84 (3H, d, 6) |
| | NH | 7.73 (1H, m) | 7.70 (d, 8) |
| Gly[6] | α | 4.00 (m)/3.30 (m) | 4.01 (m)/3.29 (m) |
| | NH | 8.39 (1H, t, 6) | 8.30 (t, 5) |
| Ile[7] | α | 4.12 (1H, dd) | 4.07 (m) |
| | β | 1.76 (m) | 1.77 (m) |
| | γ1 | 1.48 (1H, m) | 1.49 (m) |
| | γ2 | 1.09 (1H, m) | 1.14 (m) |
| | γ(Me) | 0.86 (3H, d, 6.6) | 0.83 (3H, d, 8) |
| | δ | 0.81 (3H, t, 7.5) | 0.82 (3H, t, 8) |
| | NH | 7.39 (1H, d, 8.4) | 7.47 (d, 7) |
| Trp[8] | α | 4.28 (1H, dd, 6.6, 12.8) | 4.38 (m) |
| | β | 3.20 (m)/3.12 (m) | 3.20 (m)/3.00 (m) |
| | H2 | 6.56 (s) | 6.84 (s) |
| | H4 | 7.54 (1H, d, 7.7) | 7.54 (d, 8) |
| | H5 | 7.05 (1H) | 7.01 (t, 8) |
| | H6 | 7.09 (1H, t, 7.5) | 7.11 (t, 8) |
| | H7 | 7.18 (1H, t, 8.4) | 7.39 (t, 8) |
| | NH | 7.72 (1H, s) | 7.56 (d, 6) |

[a] $^1$H NMR data of isolated lyciumin D in DMSO-d6. Multiplicity m (s=singlet, d=doublet, t=triplet, dd=double doublet, m=multiplet), intensity int, coupling constants *J* in Hertz (Figure S7). [b] $^1$H NMR data of lyciumin D in DMSO-d6 from reference 9.

**Table S4 | Genome mining of candidate lyciumin precursor genes in plants.** Queried genome versions were from December 2017.

| Organism (Genome version) | Plant family | Database | Predicted BURP domain proteins | Predicted core peptides |
|---|---|---|---|---|
| *Aegilops tauschii* (Aet_MR_1.0) | Poaceae | NCBI (GenBank MCGU00000000.1) | 24 | |
| *Amaranthus hypochondriacus* (v1.0) | Amaranthaceae | JGI Phytozome 12.1 | 19 | QPYTVGSW, QPYTVFSW |
| *Amborella trichopoda* (v1.0) | Amborellaceae | JGI Phytozome 12.1 | 34 | |
| *Anacardium occidentale* (v0.9) | Anacardiaceae | JGI Phytozome pre-release species | 13 | |
| *Ananas comosus* (v3) | Bromeliaceae | JGI Phytozome 12.1 | 7 | |
| *Aquilegia coerulea* (v3.1) | Ranunculaceae | JGI Phytozome 12.1 | 5 | |
| *Arabidopsis halleri* (v1.1) | Brassicaceae | JGI Phytozome 12.1 | 5 | |
| *Arabidopsis lyrata* (v2.1) | Brassicaceae | JGI Phytozome 12.1 | 7 | |
| *Arabidopsis thaliana* (TAIR10) | Brassicaceae | JGI Phytozome 12.1 | 5 | |
| *Arachis duranensis* (Aradu1.1) | Fabaceae | NCBI (GenBank JQIN00000000.1) | 18 | QPYGVYTW |
| *Arachis ipaensis* (Araip1.1) | Fabaceae | NCBI (GenBank JQIO00000000.1) | 17 | QPYGVYTW |
| *Asparagus officinalis* (V1.1) | Asparagaceae | JGI Phytozome pre-release species | 7 | |
| *Beta vulgaris* (RefBeet-1.2.2) | Amaranthaceae | NCBI (GenBank AYZS00000000.2) | 18 | QPWTVYGW, QPWTVAGW, QPFTISAW, QPWTVAAW |
| *Boechera stricta* (v1.2) | Brassicaceae | JGI Phytozome 12.1 | 5 | |
| *Brachypodium distachyon* (v3.1) | Poaceae | JGI Phytozome 12.1 | 14 | |
| *Brachypodium stacei* (v1.1) | Poaceae | JGI Phytozome 12.1 | 16 | |
| *Brachypodium sylvaticum* (v1.1) | Poaceae | JGI Phytozome pre-release species | 19 | |
| *Brassica oleracea capitata* (v1.0) | Brassicaceae | JGI Phytozome 12.1 | 7 | |
| *Brassica rapa FPsc* (v1.3) | Brassicaceae | JGI Phytozome 12.1 | 9 | |
| *Cajanus cajan* (C.cajan_V1.0) | Fabaceae | NCBI (GenBank AGCT00000000.1) | 13 | |
| *Camelina sativa* (Cs) | Brassicaceae | NCBI (GenBank JFZQ00000000.1) | 17 | |
| *Capsella grandiflora* (v1.1) | Brassicaceae | JGI Phytozome 12.1 | 5 | |
| *Capsella rubella* (v1.0) | Brassicaceae | JGI Phytozome 12.1 | 5 | |
| *Capsicum annuum* (Zunla 1 Ref_v1.0) | Solanaceae | NCBI (GenBank ASJU00000000.1) | 23 | QPYGGLTW, QPWGVCLW, QPWGVGSW, QPWGVGFW |
| *Capsicum baccatum* (ASM227188v2) | Solanaceae | NCBI (GenBank MLFT00000000.2) | 21 | |
| *Capsicum chinense* (ASM227189v2) | Solanaceae | NCBI (GenBank MCIT00000000.2) | 21 | QPWGVCFW, QPWGVGSW, QPWGVGFW |
| *Carica papaya* (ASGPBv0.4) | Caricaceae | JGI Phytozome 12.1 | 11 | |
| *Chenopodium quinoa* (v1.0) | Amaranthaceae | JGI Phytozome pre-release species | 42 | QPFTVVGW, QPYTVMAW, QPYTVWGW, QPYTVMGW, QPYTVYGW, QPFTVFGW, QPYTVDGW |
| *Cicer arietinum* (v1.0) | Fabaceae | JGI Phytozome pre-release species | 15 | |
| *Citrus clementina* (v1.0) | Rutaceae | JGI Phytozome 12.1 | 11 | |
| *Citrus sinensis* (v1.1) | Rutaceae | JGI Phytozome 12.1 | 9 | |
| *Coffea arabica* (UCDv0.5) | Rubiaceae | JGI Phytozome pre-release species | 61 | |
| *Cucumis melo* (ASM31304v1) | Cucurbitaceae | NCBI (GenBank CAJI00000000.1) | 7 | |
| *Cucumis sativus* (v1.0) | Cucurbitaceae | JGI Phytozome 12.1 | 7 | |
| *Cucurbita moschata* (Cmos_1.0) | Cucurbitaceae | NCBI (GenBank NEWM00000000.1) | 10 | |
| *Daucus carota* (v2.0) | Apiaceae | JGI Phytozome 12.1 | 14 | |
| *Dichanthelium oligosanthes* (ASM163321v2) | Poaceae | NCBI (GenBank LWDX00000000.2) | 11 | |
| *Durio zibethinus* (Duzib1.0) | Malvaceae | NCBI (GenBank NSDW00000000.1) | 20 | |
| *Elaeis guineensis* (EG5) | Arecaceae | NCBI (GenBank ASJS00000000.1) | 21 | |
| *Erythranthe guttata* (Mimgu1_0) | Phrymaceae | NCBI (GenBank APLE00000000.1) | 6 | |
| *Eucalyptus grandis* (v2.0) | Myrtaceae | JGI Phytozome 12.1 | 14 | |
| *Eutrema salsugineum* (v1.0) | Brassicaceae | JGI Phytozome 12.1 | 5 | |
| *Fragaria vesca* (v1.1) | Rosaceae | JGI Phytozome 12.1 | 10 | |
| *Glycine max* (Wm82.a2.v1) | Fabaceae | JGI Phytozome 12.1 | 26 | QPFTVFAW, QPWGVGTW, QPYGVYTW |
| *Gossypium hirsutum* (v1.1) | Malvaceae | JGI Phytozome pre-release species | 27 | |
| *Gossypium raimondii* (v2.1) | Malvaceae | JGI Phytozome 12.1 | 18 | |
| *Helianthus annuus* (r1.2) | Asteraceae | JGI Phytozome pre-release species | 17 | |
| *Hevea brasiliensis* (ASM165405v1) | Euphorbiaceae | NCBI (GenBank LVXX00000000.1) | 19 | |
| *Hordeum vulgare* (r1) | Poaceae | JGI Phytozome pre-release species | 15 | |
| *Ipomoea nil* (Asagao_1.1) | Convolvulaceae | NCBI (GenBank BDFN00000000.1) | 12 | |
| *Jatropha curcas* (JatCur_1.0) | Euphorbiaceae | NCBI (GenBank AFEW00000000.1) | 32 | |
| *Juglans regia* (wgs.5d) | Juglandaceae | NCBI (GenBank LIHL00000000.1) | 13 | |
| *Kalanchoe fedtschenkoi* (v1.1) | Crassulaceae | JGI Phytozome 12.1 | 32 | |
| *Kalanchoe laxiflora* (v1.1) | Crassulaceae | JGI Phytozome 12.1 | 31 | |
| *Lactuca sativa* (V8) | Asteraceae | JGI Phytozome pre-release species | 12 | |
| *Linum usitatissimum* (v1.0) | Linaceae | JGI Phytozome 12.1 | 18 | |
| *Lupinus angustifolius* (LupAngTanjil_v1.0) | Fabaceae | NCBI (GenBank MLAU00000000.1) | 13 | |
| *Malus domestica* (v1.0) | Rosaceae | JGI Phytozome 12.1 | 28 | |
| *Manihot esculenta* (v6.1) | Euphorbiaceae | JGI Phytozome 12.1 | 31 | |
| *Marchantia polymorpha* (v3.1) | Marchantiaceae | JGI Phytozome 12.1 | 1 | |
| *Medicago truncatula* (Mt4.0v1) | Fabaceae | JGI Phytozome 12.1 | 52 | QPLLFIYW, QPYGVYFW, QPYGVYTW, QPLTTRMW, QPLITRMW, QPLTTSMW, QPITTHMW, QPFGINIW, QPFGVLTW, QPFGFFSW, QPLPAHKW, QPFRTIGW, QPLGAVKW, QPFGSLTW, QPFGVAAW, QPFGFRAW, QPFEAHTW |
| *Mimulus guttatus* (v2.0) | Phrymaceae | JGI Phytozome 12.1 | 8 | |
| *Miscanthus sinensis* (v7.1) | Poaceae | JGI Phytozome pre-release species | 23 | |
| *Morus notabilis* (ASM41409v2) | Moraceae | NCBI (GenBank ATGF00000000.1) | 10 | |
| *Musa acuminata* (v1) | Musaceae | JGI Phytozome 12.1 | 7 | |
| *Nelumbo nucifera* (Chinese Lotus 1.1) | Nelumbonaceae | NCBI (GenBank AQOG00000000.1) | 8 | |
| *Nicotiana attenuata* (v2) | Solanaceae | NCBI (GenBank MJEQ00000000.1) | 14 | QPWGVYSW |
| *Nicotiana benthamiana* (v1.0.1) | Solanaceae | Sol Genomics Network | 12 | |
| *Nicotiana sylvestris* (Nsyl) | Solanaceae | NCBI (GenBank ASAF00000000.1) | 10 | |
| *Nicotiana tabacum* (v1.0 Edwards 2017) | Solanaceae | Sol Genomics Network | 15 | |
| *Nicotiana tomentosiformis* (Ntom_v01) | Solanaceae | NCBI (GenBank ASAG00000000.1) | 10 | |
| *Olea europaea var. sylvestris* (v1.0) | Oleaceae | JGI Phytozome pre-release species | 16 | |
| *Oropetium thomaeum* (v1.0) | Poaceae | JGI Phytozome 12.1 | 4 | |
| *Oryza sativa* (v7_JGI) | Poaceae | JGI Phytozome 12.1 | 18 | |
| *Panicum hallii* (v2.0) | Poaceae | JGI Phytozome 12.1 | 13 | |
| *Panicum virgatum* (v1.1) | Poaceae | JGI Phytozome 12.1 | 36 | |
| *Petunia axillaris* (v1.6.2) | Solanaceae | Sol Genomics Network | 13 | QPYGVFAW, QPFGVFAW |
| *Petunia inflata* (v1.0.1) | Solanaceae | Sol Genomics Network | 14 | QPYGPFGW, QPFGDYVW, QPYGVFGW, QPFGVFGW, QPFGVFVW |
| *Phalaenopsis equestris* (ASM1263359v1) | Orchidaceae | NCBI (GenBank APLD00000000.1) | 7 | |
| *Phaseolus vulgaris* (v2.1) | Fabaceae | JGI Phytozome 12.1 | 11 | |
| *Phoenix dactylifera* (DPV01) | Arecaceae | NCBI (GenBank ATBV00000000.1) | 11 | |
| *Physcomitrella patens* (v3.3) | Funariaceae | JGI Phytozome 12.1 | 9 | |
| *Populus deltoides* (WV94 v2.1) | Fabaceae | JGI Phytozome pre-release species | 21 | |
| *Populus euphratica* (PopEup_1.0) | Salicaceae | NCBI (GenBank AOFL00000000.1) | 14 | |
| *Populus trichocarpa* (v3.0) | Salicaceae | JGI Phytozome 12.1 | 20 | |
| *Prunus avium* (PAV_r1.0) | Rosaceae | NCBI (GenBank BDGV00000000.1) | 19 | QPAPQLYW |
| *Prunus persica* (v2.1) | Rosaceae | JGI Phytozome 12.1 | 29 | QPAAQLYW, QPAPQLYW |
| *Pyrus x bretschneideri* (Pbr_v1.0) | Rosaceae | NCBI (GenBank AJSU00000000.1) | 21 | |
| *Raphanus sativus* (Rs1.0) | Brassicaceae | NCBI (GenBank JRUI00000000.2) | 11 | |
| *Ricinus communis* (v0.1) | Euphorbiaceae | JGI Phytozome 12.1 | 24 | |
| *Salix purpurea* (v1.0) | Salicaceae | JGI Phytozome 12.1 | 12 | |
| *Selaginella moellendorffii* (v1.0) | Selaginellaceae | JGI Phytozome 12.1 | 7 | |
| *Sesamum indicum* (v1.0) | Pedaliaceae | NCBI (GenBank APMJ00000000.1) | 9 | |
| *Setaria italica* (v2.2) | Poaceae | JGI Phytozome 12.1 | 15 | |
| *Setaria viridis* (v1.1) | Poaceae | JGI Phytozome 12.1 | 15 | |
| *Solanum lycopersicum* (iTAG2.4) | Solanaceae | JGI Phytozome 12.1 | 14 | QPWGVGAW, QPWGVYRW, QPYGVYRW, QPYGVYSW, QPWGVGSW |
| *Solanum melongena* (v2.5.1) | Solanaceae | Sol Genomics Network | 10 | QPWGVNSW, QPWGVLRW, QPWGVGSW, QPWGVLGW, QPYGVYTW |
| *Solanum pennellii* | Solanaceae | Sol Genomics Network | 13 | QPWGVGAW, QPFGVYRW, QPWGVFRW, QPWGVGSW |
| *Solanum pimpinellifolium* (LA1589) | Solanaceae | Sol Genomics Network | 13 | QPWGVGAW, QPWGVYRW, QPYGVYRW, QPYGVYSW, QPWGVGSW |
| *Solanum tuberosum* (v4.03) | Solanaceae | JGI Phytozome 12.1 | 20 | QPWGVDSW, QPYGVGVW, QPFGVGRW*, QPWGVGRW*, QPFGVVAW*, QPYGVLAW*, QPYGVSRW*, QPWGVVAW*, QPYGVFRW*, QPYGVFAW*, QPYGVDGW*, QPGVYRW*, QPWGVGAW*, QPYGVFGW*, QPFGVFGW*, QPYGVFAW*, QPWGVGSW* |
| *Sorghum bicolor* (v3.1.1) | Poaceae | JGI Phytozome 12.1 | 11 | |
| *Sphagnum fallax* (v0.5) | Sphagnaceae | JGI Phytozome 12.1 | 8 | |
| *Spinacia oleracea* (ASM200726v1) | Amaranthaceae | NCBI (GenBank LZYP00000000.1) | 16 | |
| *Spirodela polyrhiza* (v2) | Araceae | JGI Phytozome 12.1 | 11 | |
| *Tarenaya hassleriana* (ASM46358v1) | Cleomaceae | NCBI (GenBank AOUI00000000.1) | 8 | |
| *Theobroma cacao* (v1.1) | Malvaceae | JGI Phytozome 12.1 | 14 | |
| *Trifolium pratense* (v2) | Fabaceae | JGI Phytozome 12.1 | 36 | QPLGTWIW, QPFGIAAW, QPSGVYIW, QPFGINIW, QPYGVYTW |
| *Triticum aestivum* (v2.2) | Poaceae | JGI Phytozome pre-release species | 33 | |
| *Vicia faba* (VfEP_Reference-Unigene) | Fabaceae | NCBI (GenBank CSVX00000000.1) | 5 | |
| *Vigna angularis* (Vigan1.1) | Fabaceae | NCBI (GenBank JZJH00000000.1) | 13 | |
| *Vigna radiata* (release 101) | Fabaceae | NCBI (GenBank JJMO00000000.1) | 24 | |
| *Vigna unguiculata* (v1.1) | Fabaceae | JGI Phytozome pre-release species | 19 | QPATLLAW |
| *Vitis vinifera* (Genoscope.12X) | Vitaceae | JGI Phytozome 12.1 | 7 | |
| *Zea mays PH207* (v1.1) | Poaceae | JGI Phytozome 12.1 | 10 | |
| *Ziziphus jujuba* (ZizJuj_1.1) | Rhamnaceae | NCBI (GenBank JREP00000000.1) | 20 | |
| *Zostera marina* (v2.2) | Zosteraceae | JGI Phytozome 12.1 | 9 | |
| | | | | |
| Species with predicted lyciumin precursor genotype in genome | | 21/116 (18%) | | |

* - RNA-seq data complemented sequences (see Figure S26)

**Table S5 | Core peptide sequences from predicted lyciumin precursor genes.**

| Core peptide | Chemotype | Organism(s) |
|---|---|---|
| QPYTVGSW* | Lyciumin A | *Amaranthus hypochondriacus* |
| QPYTVFSW* | Lyciumin C | *Amaranthus hypochondriacus* |
| QPYGVYTW* | Lyciumin I | *Arachis duranensis, Arachis ipaensis, Glycine max, Medicago truncatula, Solanum melongena, Trifolium pratens* |
| QPWTVYGW | | *Beta vulgaris* |
| QPWTVAGW | | *Beta vulgaris* |
| QPFTISAW | | *Beta vulgaris* |
| QPWTVAAW* | Lyciumin E | *Beta vulgaris* |
| QPYGGLTW | | *Capsicum annuum* |
| QPWGVCLW | | *Capsicum annuum* |
| QPWGVGSW* | Lyciumin B | *Capsicum annuum, Capsicum chinense, Solanum lycopersicum, Solanum melongena, Solanum pennellii, Solanum pimpinellifolium, Solanum tuberosum* |
| QPWGVGFW | | *Capsicum annuum, Capsicum chinense* |
| QPWGVCFW | | *Capsicum chinense* |
| QPFTVVGW* | Lyciumin G | *Chenopodium quinoa* |
| QPYTVMAW | | *Chenopodium quinoa* |
| QPYTVWGW* | Lyciumin F | *Chenopodium quinoa* |
| QPYTVMGW | | *Chenopodium quinoa* |
| QPYTVYGW | | *Chenopodium quinoa* |
| QPFTVFGW | | *Chenopodium quinoa* |
| QPYTVDGW | | *Chenopodium quinoa* |
| QPFTVFAW | | *Glycine max* |
| QPWGVGTW* | Lyciumin H | *Glycine max* |
| QPLLFIYW | | *Medicago truncatula* |
| QPYGVYFW | | *Medicago truncatula* |
| QPLTTRMW | | *Medicago truncatula* |
| QPLTTSMW | | *Medicago truncatula* |
| QPITTHMW | | *Medicago truncatula* |
| QPFGINIW | | *Medicago truncatula, Trifolium pratense* |
| QPFGVLTW | | *Medicago truncatula* |
| QPFGFFSW | | *Medicago truncatula* |
| QPLPAHKW | | *Medicago truncatula* |
| QPFRTIGW | | *Medicago truncatula* |
| QPLGAVKW | | *Medicago truncatula* |
| QPFGSLTW | | *Medicago truncatula* |
| QPFGVAAW | | *Medicago truncatula* |
| QPFGFRAW | | *Medicago truncatula* |
| QPFEAHTW | | *Medicago truncatula* |
| QPWGVYSW | | *Nicotiana attenuata* |
| QPYGVFAW* | Lyciumin J | *Petunia axillaris, Solanum tuberosum* |
| QPFGVFAW | | *Petunia axillaris* |
| QPYGPFGW | | *Petunia inflata* |
| QPFGDYVW | | *Petunia inflata* |
| QPYGVFGW | | *Petunia inflata, Solanum tuberosum* |
| QPFGVFGW | | *Petunia inflata, Solanum tuberosum* |
| QPFGVFVW | | *Petunia inflata* |
| QPAPQLYW | | *Prunus avium, Prunus persica* |
| QPAAQLYW | | *Prunus persica* |
| QPWGVGAW* | Lyciumin K | *Solanum lycopersicum, Solanum pennellii, Solanum pimpinellifolium, Solanum tuberosum* |
| QPWGVYRW | | *Solanum lycopersicum* |
| QPYGVYRW* | Lyciumin M | *Solanum lycopersicum, Solanum pimpinellifolium, Solanum tuberosum* |
| QPYGVYSW | | *Solanum lycopersicum, Solanum pimpinellifolium* |
| QPWGVNSW | | *Solanum melongena* |
| QPWGVLRW | | *Solanum melongena* |
| QPWGVLGW | | *Solanum melongena* |
| QPFGVYRW | | *Solanum pennellii* |
| QPWGVFRW | | *Solanum pennellii* |
| QPYGVYSW | | *Solanum pimpinellifolium* |
| QPWGVDSW | | *Solanum tuberosum* |
| QPYGVGVW | | *Solanum tuberosum* |
| QPFGVGRW | | *Solanum tuberosum* |
| QPWGVGRW* | Lyciumin O | *Solanum tuberosum* |
| QPFGVVAW | | *Solanum tuberosum* |
| QPYGVLAW | | *Solanum tuberosum* |
| QPYGVSRW* | Lyciumin N | *Solanum tuberosum* |
| QPWGVVAW* | Lyciumin L | *Solanum tuberosum* |
| QPYGVFRW | | *Solanum tuberosum* |
| QPYGVFAW | | *Solanum tuberosum* |
| QPYGVDGW | | *Solanum tuberosum* |
| QPLGTWIW | | *Trifolium pratense* |
| QPFGIAAW | | *Trifolium pratense* |
| QPSGVYIW | | *Trifolium pratense* |
| QPATLLAW | | *Vigna unguiculata* |

| | |
|---|---|
| Multi-species peptides | 11 (16%) |
| Single-species peptides | 60 (84%) |
| Total core peptides | 71 |
| * Detected peptides | 13 (18%) |

**Table S6 | NMR analysis of lyciumin C from *Amaranthus hypochondriacus* (600 MHz, DMSO-d6).**

| Residue | H | δ($^1$H) (int, m, *J*) [ppm] [a] | δ($^1$H) (int, m, *J*) [ppm] [b] |
|---|---|---|---|
| PyroGlu[1] | α | 4.36 (1H, m) | 4.36 (m) |
| | β | 1.79 (m)/2.26 (1H, m) | 1.79 (m)/2.24 (m) |
| | γ | 2.10 (2H, m) | 2.09 (2H, m) |
| | NH | 8.10 (1H, s) | n/a |
| Pro[2] | α | 4.36 (1H, m) | 4.36 (m) |
| | β | 1.79 (1H, m)/2.01 (m) | 1.79 (m)/1.96 (m) |
| | γ | 1.79 (2H, m) | 1.79 (2H, m) |
| | δ | 3.38 (1H, m)/3.64 (1H, t, 7.6) | 3.41 (m)/3.73 (t, 7) |
| Tyr[3] | α | 4.36 (1H, m) | 4.36 (m) |
| | β | 2.68 (2H, m) | 2.74 (m) |
| | H2 | 6.69 (1H, d, 8.3) | 6.68 (1H, d, 8) |
| | H3 | 6.35 (1H, d, 8.4) | 6.39 (1H, d, 8) |
| | H5 | 6.35 (1H, d, 8.4) | 6.39 (1H, d, 8) |
| | H6 | 6.69 (1H, d, 8.3) | 6.68 (1H, d, 8) |
| | OH | 7.27 (1H, m) | 7.54 (s, OH) |
| | NH | 8.08 (1H, m) | 7.99 (d, 6) |
| Gly[4] | α | 6.65 (1H, d, 7.9) | 6.63 (d, 8) |
| | NH | 9.29 (1H, d, 8.1) | 9.15 (d, 8) |
| Val[5] | α | 4.31 (1H, m) | 4.31 (m) |
| | β | 1.86 (1H, m) | 1.86 (m) |
| | γ1 | 0.80 (3H, d, 7.2) | 0.78 (3H, d, 7) |
| | γ2 | 0.67 (3H, d, 7.4) | 0.63 (3H, d, 7) |
| | NH | 7.98 (1H, m) | 7.90 (m) |
| Phe[6] | α | 4.45 (1H, m) | 4.36 (m) |
| | β | 2.98 (1H, m)/3.16 (1H, m) | 3.01 (m)/3.06 (m) |
| | H2 | 7.22 (1H, m) | 7.22 (1H, m) |
| | H3 | 7.22 (1H, m) | 7.22 (1H, m) |
| | H4 | 7.22 (1H, m) | 7.22 (1H, m) |
| | H5 | 7.22 (1H, m) | 7.22 (1H, m) |
| | H6 | 7.22 (1H, m) | 7.22 (1H, m) |
| | NH | 8.15 (1H, d, 5.9) | 8.30 (d, 7) |
| Ser[7] | α | 4.09 (1H, m) | 4.16 (m) |
| | β | 3.59 (2H, m) | 3.65 (2H, m) |
| | NH | 7.94 (1H, d, 5.7) | 7.85 (d, 7) |
| Trp[8] | α | 4.60 (1H, m) | 4.42 (m) |
| | β | 3.16 (1H, m)/3.22 (1H, m) | 3.01 (m)/3.09 (m) |
| | H2 | 7.16 (1H, s) | 7.09 (s) |
| | H4 | 7.53 (1H, d, 7.9) | 7.51 (d, 8) |
| | H5 | 7.07 (1H, t, 7.5) | 7.11 (t, 8) |
| | H6 | 7.17 (1H, m) | 7.17 (t, 8) |
| | H7 | 7.41 (1H, d, 8.4) | 7.35 (d, 8) |
| | NH | 7.75 (1H, s) | 7.75 (d, 7) |

[a] $^1$H NMR data of isolated lyciumin C in DMSO-d6. Multiplicity m (s=singlet, d=doublet, t=triplet, dd=double doublet, m=multiplet), intensity int, coupling constants *J* in Hertz (Figure S12). [b] $^1$H NMR data of lyciumin C in DMSO-d6 from reference 9.

**Table S7 | NMR analysis of lyciumin I from *Glycine max* (600 MHz, DMSO-d6).**

| Residue | C | δ($^{13}$C) [ppm] [a] | H | δ($^{1}$H) (int, m, *J*) [ppm] [b] | COSY [c] | ROESY [d] | HMBC [e] |
|---|---|---|---|---|---|---|---|
| PyroGlu$^{1}$ | α | 53.6 | α | 4.35 (1H, m) | PyroGlu$^{1}$: Hβ1/2, Hγ, NH | PyroGlu$^{1}$: Hβ1/2, Hγ, NH, Pro$^{2}$: Hδ1/2 | PyroGlu$^{1}$: Cβ, Cγ, Cδ, C=O |
| | β | 23.6 | β1 | 1.85 (1H, m) | PyroGlu$^{1}$: Hα, Hβ2, Hγ | PyroGlu$^{1}$: Hα, Hβ2, Hγ | PyroGlu$^{1}$: Cα, Cγ, Cδ, C=O |
| | γ | 29.0 | β2 | 2.25 (1H, m) | PyroGlu$^{1}$: Hα, Hβ1, Hγ | PyroGlu$^{1}$: Hα, Hβ1, Hγ | PyroGlu$^{1}$: Cα, Cγ, Cδ, C=O |
| | δ | 177.2 | γ | 2.10 (2H, m) | PyroGlu$^{1}$: Hβ1/2 | PyroGlu$^{1}$: Hα, Hβ1/2 | PyroGlu$^{1}$: Cα, Cβ, Cδ |
| | C=O | 171.8 | NH | 7.75 (1H, s) | PyroGlu$^{1}$: Hα | PyroGlu$^{1}$: Hα, Hγ, Pro$^{2}$: Hδ2 | PyroGlu$^{1}$: Cα, Cβ, Cγ, Cδ |
| Pro$^{2}$ | α | 59.0 | α | 4.34 (1H, m) | Pro$^{2}$: Hβ, Hγ1/2 | Pro$^{2}$: Hβ, Hγ1/2, Hδ2 | PyroGlu$^{1}$: C=O, Pro$^{2}$: Cβ, Cγ, Cδ, C=O |
| | β | 28.4 | β | 1.77 (2H, m) | Pro$^{2}$: Hα, Hγ1/2, Hδ1/2 | Pro$^{2}$: Hα, Hγ1/2, Hδ1/2 | Pro$^{2}$: Cα, Cγ, Cδ |
| | γ | 24.4 | γ1 | 1.79 (1H, m) | Pro$^{2}$: Hα, Hβ, Hγ2, Hδ1/2 | Pro$^{2}$: Hα, Hβ, Hγ2, Hδ1/2, Val$^{5}$: Hγ1/2 | Pro$^{2}$: Cβ, Cδ |
| | δ | 46.0 | γ2 | 1.83 (1H, m) | Pro$^{2}$: Hα, Hβ, Hγ1, Hδ1/2 | Pro$^{2}$: Hα, Hβ, Hγ1, Hδ2 | Pro$^{2}$: Cβ, Cδ |
| | C=O | 171.4 | δ1 | 3.38 (1H, m) | Pro$^{2}$: Hβ, Hγ1/2, Hδ2 | PyroGlu$^{1}$: Hα, Hβ2, Pro$^{2}$: Hβ, Hγ1, Hδ2 | Pro$^{2}$: Cβ, Cγ |
| | | | δ2 | 3.58 (1H, m) | Pro$^{2}$: Hβ, Hγ1/2, Hδ1 | PyroGlu$^{1}$: Hα, NH, Pro$^{2}$: Hα, Hβ, Hγ1/2, Hδ1 | Pro$^{2}$: Cβ, Cγ |
| Tyr$^{3}$ | α | 54.5 | α | 4.35 (1H, m) | Tyr$^{3}$: Hβ1/2, NH | Tyr$^{3}$: Hβ1/2, H2/6, NH, Gly$^{4}$: NH | Pro$^{2}$: C=O, Tyr$^{3}$: Cβ, C=O |
| | β | 35.8 | β1 | 2.60 (1H, m) | Tyr$^{3}$: Hα, Hβ2 | Tyr$^{3}$: Hα, H2/6, NH | Tyr$^{3}$: Cα, C1, C2/6, C=O |
| | C1 | 126.5 | β2 | 2.70 (1H, m) | Tyr$^{3}$: Hα, Hβ1 | Tyr$^{3}$: Hα, H2/6 | Tyr$^{3}$: Cα, C1, C2/6, C=O |
| | C2 | 129.6 | H2 | 6.68 (1H, d, 8.5) | Tyr$^{3}$: H3 | Tyr$^{3}$: Hα, Hβ, H3/5 | Tyr$^{3}$: Cβ, C3/5, C4, C6 |
| | C3 | 114.4 | H3 | 6.35 (1H, d, 8.5) | Tyr$^{3}$: H2 | Tyr$^{3}$: H2/6 | Tyr$^{3}$: C1, C4, C5 |
| | C4 | 155.8 | OH | 9.18 (1H, br s) | n/a | n/a | n/a |
| | C5 | 114.4 | H5 | 6.35 (1H, d, 8.5) | Tyr$^{3}$: H6 | Tyr$^{3}$: H2/6 | Tyr$^{3}$: C1, C3, C4 |
| | C6 | 129.6 | H6 | 6.68 (1H, d, 8.5 | Tyr$^{3}$: H5 | Tyr$^{3}$: Hα, Hβ, H3/5 | Tyr$^{3}$: Cβ, C2, C3/5, C4 |
| | C=O | 171.4 | NH | 7.90 (1H, m) | Tyr$^{3}$: Hα | Tyr$^{3}$: Hα, Hβ1, Gly$^{4}$: NH | Pro$^{2}$: C=O |
| Gly$^{4}$ | α | 61.0 | α | 6.64 (1H, d, 8.3) | Gly$^{4}$: NH | Gly$^{4}$: NH, Val$^{5}$: NH, Trp$^{8}$: H2, H4 | Tyr$^{3}$: C=O, Gly$^{4}$: C=O, Trp$^{8}$: C2, C3 |
| | C=O | 166.5 | NH | 9.28 (1H, d, 8.1) | Gly$^{4}$: Hα | Tyr$^{3}$: Hα, NH, Gly$^{4}$: Hα, Trp$^{8}$: H2, H4 | Tyr$^{3}$: C=O |
| Val$^{5}$ | α | 61.4 | α | 3.68 (1H, t, 7.7) | Val$^{5}$: Hβ, NH | Val$^{5}$: Hβ, Hγ1/2, Tyr$^{6}$: NH | Gly$^{4}$: C=O, Val$^{5}$: Cβ, Cγ1/2, C=O |
| | β | 28.5 | β | 2.00 (1H, m) | Val$^{5}$: Hα, Hγ1/2 | Val$^{5}$: Hα, Hγ1/2, Tyr$^{6}$: NH, H2/6 | Val$^{5}$: Cα, Cγ1/2, C=O |
| | γ1 | 19.0 | γ1 | 0.67 (3H, d, 7.0) | Val$^{5}$: Hβ, Hγ2 | Pro$^{2}$: Hγ1, Val$^{5}$: Hα, Hβ, Hγ2, Tyr$^{6}$: H2/6, H3/5 | Val$^{5}$: Cα, Cβ, Cγ2 |
| | γ2 | 18.4 | γ2 | 0.79 (3H, d, 7.0) | Val$^{5}$: Hβ, Hγ1 | Pro$^{2}$: Hγ1, Val$^{5}$: Hα, Hβ, Hγ1, NH, Tyr$^{6}$: H2/6, H3/5, Trp$^{8}$: H4 | Val$^{5}$: Cα, Cβ, Cγ1 |
| | C=O | 170.7 | NH | 7.91 (1H, d, 6.2) | Val$^{5}$: Hα | Gly$^{4}$: Hα, Val$^{5}$: Hα, Hβ, Hγ2, Tyr$^{6}$: Hα | n/a |
| Tyr$^{6}$ | α | 55.7 | α | 4.31 (1H, m) | Tyr$^{6}$: Hβ1/2, NH | Val$^{5}$: NH, Tyr$^{6}$: Hβ1/2, H2/6, NH, Thr$^{7}$: NH | Tyr$^{6}$: Cβ, C=O |
| | β | 35.6 | β1 | 2.91 (1H, m) | Tyr$^{6}$: Hα, Hβ2 | Tyr$^{6}$: Hα, Hβ2, H2/6, NH, Thr$^{7}$: Hγ, NH | Tyr$^{6}$: Cα, C1, C2/6 |
| | C1 | 126.4 | β2 | 3.12 (1H, m) | Tyr$^{6}$: Hα, Hβ1 | Tyr$^{6}$: Hα, Hβ1, H2/6, NH Thr$^{7}$: Hγ, NH | Tyr$^{6}$: Cα, C1, C2/6 |
| | C2 | 129.7 | H2 | 7.01 (1H, d, 8.5) | Tyr$^{6}$: H3 | Val$^{5}$: Hβ, Hγ1/2, Tyr$^{6}$: Hα, Hβ1/2, H3 | Tyr$^{6}$: Cβ, C3/5, C4, C6 |
| | C3 | 114.7 | H3 | 6.63 (1H, d, 8.5) | Tyr$^{6}$: H2 | Val$^{5}$: Hγ1/2, Tyr$^{6}$: H2 | Tyr$^{6}$: C5, C4 |
| | C4 | 155.6 | OH | 9.18 (1H, br s) | n/a | n/a | n/a |
| | C5 | 114.7 | H5 | 6.63 (1H, d, 8.5) | Tyr$^{6}$: H6 | Val$^{5}$: Hγ1/2, Tyr$^{6}$: H6 | Tyr$^{6}$: C3, C4 |
| | C6 | 129.7 | H6 | 7.01 (1H, d, 8.5) | Tyr$^{6}$: H5 | Val$^{5}$: Hβ, Hγ1/2, Tyr$^{6}$: Hα, Hβ1/2, H5 | Tyr$^{6}$: Cβ, C2, C3/5, C4 |
| | C=O | 171.7 | NH | 8.02 (1H, d, 7.7) | Tyr$^{6}$: Hα | Val$^{5}$: Hα, Hβ, Tyr$^{6}$: Hα, Hβ1, Trp$^{8}$: H4 | Val$^{5}$: C=O |
| Thr$^{7}$ | α | 60.0 | α | 4.00 (1H, m) | Thr$^{7}$: Hβ, Hγ, NH | Thr$^{7}$: Hβ, Hγ, NH, Trp$^{8}$: NH | Tyr$^{6}$: C=O, Thr$^{7}$: Cβ, Cγ, C=O |
| | β | 66.5 | β | 4.02 (1H, m) | Thr$^{7}$: Hα, Hγ, NH | Tyr$^{6}$: NH, Thr$^{7}$: Hα, Hγ, NH | Thr$^{7}$: Cα, Cγ, C=O |
| | γ | 19.6 | OH | 4.84 (1H, m) | n/a | n/a | n/a |
| | C=O | 169.4 | γ | 1.10 (3H, d, 6.0) | Thr$^{7}$: Hα, Hβ | Tyr$^{6}$: Hβ1/2, H2/6, Thr$^{7}$: Hα, Hβ, NH, Trp$^{8}$: NH | Thr$^{7}$: Cα, Cβ, C=O |
| | | | NH | 7.84 (1H, d, 5.0) | Thr$^{7}$: Hα, Hβ | Tyr$^{6}$: Hα, Hβ1/2, NH, Thr$^{7}$: Hα, Hβ, Hγ, Trp$^{8}$: H2, NH | Tyr$^{6}$: Cα, C=O, Thr$^{7}$: C=O |
| Trp$^{8}$ | α | 51.4 | α | 4.57 (1H, m) | Trp$^{8}$: Hβ1/2, NH | Trp$^{8}$: Hβ1/2, H2, H7, NH | Thr$^{7}$: C=O, Trp$^{8}$: Cβ, C1, C=O |
| | β | 26.1 | β1 | 3.12 (1H, m) | Trp$^{8}$: Hα, Hβ2 | Trp$^{8}$: Hα, H7, NH | Trp$^{8}$: Cα, C1, C2, C=O |
| | C1 | 111.0 | β2 | 3.20 (1H, m) | Trp$^{8}$: Hα, Hβ1 | Trp$^{8}$: Hα, H7, NH | Trp$^{8}$: Cα, C1, C2, C=O |
| | C2 | 125.0 | H2 | 7.33 (1H, s) | Trp$^{8}$: H7 | Gly$^{4}$: Hα, NH, Thr$^{7}$: Hα, NH, Trp$^{8}$: Hα, Hβ, NH | Gly$^{4}$: Cα, Trp$^{8}$: C1, C3, C8 |
| | C3 | 135.8 | H4 | 7.40 (1H, d, 8.3) | Trp$^{8}$: H2, H4, H5, H6 | Gly$^{4}$: Hα, NH, Val$^{5}$: Hγ2, Tyr$^{6}$: NH, Trp$^{8}$: H5, H6 | Trp$^{8}$: C1, C3, C6, C8 |
| | C4 | 109.6 | H5 | 7.07 (1H, t, 7.5) | Trp$^{8}$: H4, H5, H7 | Trp$^{8}$: H4, H6, H7 | Trp$^{8}$: C4, C7, C8 |
| | C5 | 119.2 | H6 | 7.14 (1H, t, 7.5) | Trp$^{8}$: H6, H7 | Trp$^{8}$: H4, H5, H7 | Trp$^{8}$: C3, C4, C7 |
| | C6 | 121.3 | H7 | 7.52 (1H, d, 7.9) | Trp$^{8}$: H6, H7 | Trp$^{8}$: Hα, Hβ1/2, H5, H6 | Trp$^{8}$: C1, C3, C5, C6, C7, C8 |
| | C7 | 118.1 | NH | 8.10 (1H, d, 7.5) | Trp$^{8}$: Hα | Thr$^{7}$: Hα, Hγ, NH, Trp$^{8}$: Hα, Hβ1/2, H2 | Thr$^{7}$: C=O |
| | C8 | 128.0 | | | | | |
| | C=O | 172.9 | | | | | |

[a] $^{13}$C NMR data of major conformer of isolated lyciumin I in DMSO-d6. Values were derived from HSQC and HMBC NMR analyses (Figure S19 & S20), [b] $^{1}$H NMR data of major conformer of isolated lyciumin I in DMSO-d6. Multiplicity m (s=singlet, d=doublet, t=triplet, dd=double doublet, m=multiplet), intensity int, coupling constants *J* in Hertz (Figure S16), [c] $^{1}$H-$^{1}$H COSY correlations of major conformer of lyciumin I in DMSO-d6 (Figure S17), [d] ROESY correlations of major conformer of lyciumin I in DMSO-d6 (Figure S21), [e] HMBC correlations of major conformer of lyciumin I in DMSO-d6 (Figure S20)

**Table S8 | Bioinformatic analysis of lyciumin precursor genes and co-clustered glutamine cyclotransferase genes in genomes of *Beta vulgaris* and *Chenopodium quinoa* (Figure 3C).**

| Gene | Predicted function | Reference | Gene product length [aa] | Closest functional blastp homolog (organism) [Similarity/Identity, %/%] |
|---|---|---|---|---|
| *Chenopodium quinoa* (v1.0) locus | | | | |
| AUR62017095-RA (*CquBURP1*) | BURP domain lyciumin precursor | XP_021740703.1 | 619 | XP_010675925.1\|PREDICTED: BURP domain protein USPL1 [*Beta vulgaris subsp. vulgaris*] [68/57] |
| AUR62017096-RA | Glutamine cyclotransferase | XP_021740704.1 | 286 | XP_010675927.1\|PREDICTED: glutaminyl-peptide cyclotransferase isoform X1 (*Beta vulgaris subsp. vulgaris*) [80/71] |
| *Beta vulgaris* locus (RefBeet-1.2.2) | | | | |
| LOC104891851 (*BvuBURP2*) | BURP domain lyciumin precursor | XP_010675925.1 | 446 | XP_010676059.1\|PREDICTED: BURP domain protein USPL1-like (*Beta vulgaris subsp. vulgaris*) [93/89] |
| LOC104891854 | Glutamine cyclotransferase | XP_010675927.1 | 306 | XP_021771347.1\|glutaminyl-peptide cyclotransferase-like isoform X1 [*Chenopodium quinoa*] |
| LOC104891968 (*BvuBURP1*) | BURP domain lyciumin precursor | XP_010676059.1 | 404 | XP_010675925.1\|PREDICTED: BURP domain protein USPL1 (*Beta vulgaris subsp. vulgaris*) [93/89] |
| LOC104891853 | Glutamine cyclotransferase | XP_010675926.1 | 201 | XP_021740704.1\|glutaminyl-peptide cyclotransferase-like (*Chenopodium quinoa*) [73/68] |

**Table S9 | LC-MS dataset accessions (MassIVE, GNPS) (*21*)**

| Dataset | Figure | MassIVE dataset accesion |
|---|---|---|
| Amaranth_grain | Figure S11, S27 | MSV000082522 |
| beet_root | Figure S13, S27 | MSV000082523 |
| Chenopodium-quinoa-flower | Figure S14, S27 | MSV000082525 |
| Eggplant-stem | Figure S22 | MSV000082526 |
| Lycium-barbarum-root (LB-root) | Figure 1, S1, S4, S6, S29, S31 | MSV000082527 |
| Medicago_truncatula_seed | Figure S23 | MSV000082528 |
| Potato_tuber_sprout | Figure S26 | MSV000082529 |
| Russett-potato-sprout | Figure S26 | MSV000082530 |
| Soy-root | Figure S15 | MSV000082532 |
| NB-At-LBA4404-pEAQ-6d | Figure 1, 3 | MSV000082533 |
| NB-At-LBA4404-pEAQ-Sali3-2-6d | Figure 4 | MSV000082534 |
| NB-LbaLycA-6d | Figure 1, S29, S31 | MSV000082536 |
| NB-Sali3-2-QAYGVYTW-6d | Figure 4, S32 | MSV000082537 |
| NB-Sali3-2-QGVYTW-6d | Figure 4 | MSV000082538 |
| NB-Sali3-2-QPAGVYTW | Figuer S33 | MSV000082539 |
| NB-Sali3-2-QPFGFFSW | Figuer S43 | MSV000082540 |
| NB-Sali3-2-QPFGVYTW | Figure S42 | MSV000082541 |
| NB-Sali3-2-QPWGVGAW | Figure S40 | MSV000082542 |
| NB-Sali3-2-QPWGVGTW | Figuer S39 | MSV000082543 |
| NB-Sali3-2-QPWGVYSW | Figuer S44 | MSV000082544 |
| NB-Sali3-2-QPWGVYTW | Figure S42 | MSV000082545 |
| NB-Sali3-2-QPYAVYTW | Figure S34 | MSV000082546 |
| NB-Sali3-2-QPYGAYTW | Figure S35 | MSV000082547 |
| NB-Sali3-2-QPYGVATW | Figure S36 | MSV000082548 |
| NB-Sali3-2-QPYGVYAW | Figure S37 | MSV000082549 |
| NB-Sali3-2-QPYGVYFW | Figuer S45 | MSV000082550 |
| NB-Sali3-2-QPYGVYTA | Figure 4 | MSV000082552 |
| NB-Sali3-2-QPYGVYTAW | Figure 4 | MSV000082553 |
| NB-Sali3-2-QPYTVYTW | Figure S38 | MSV000082555 |
| NB-Sali3-2-QYGVYTW | Figure 4 | MSV000082556 |
| NB-StuBURP-6d | Figure S26 | MSV000082557 |

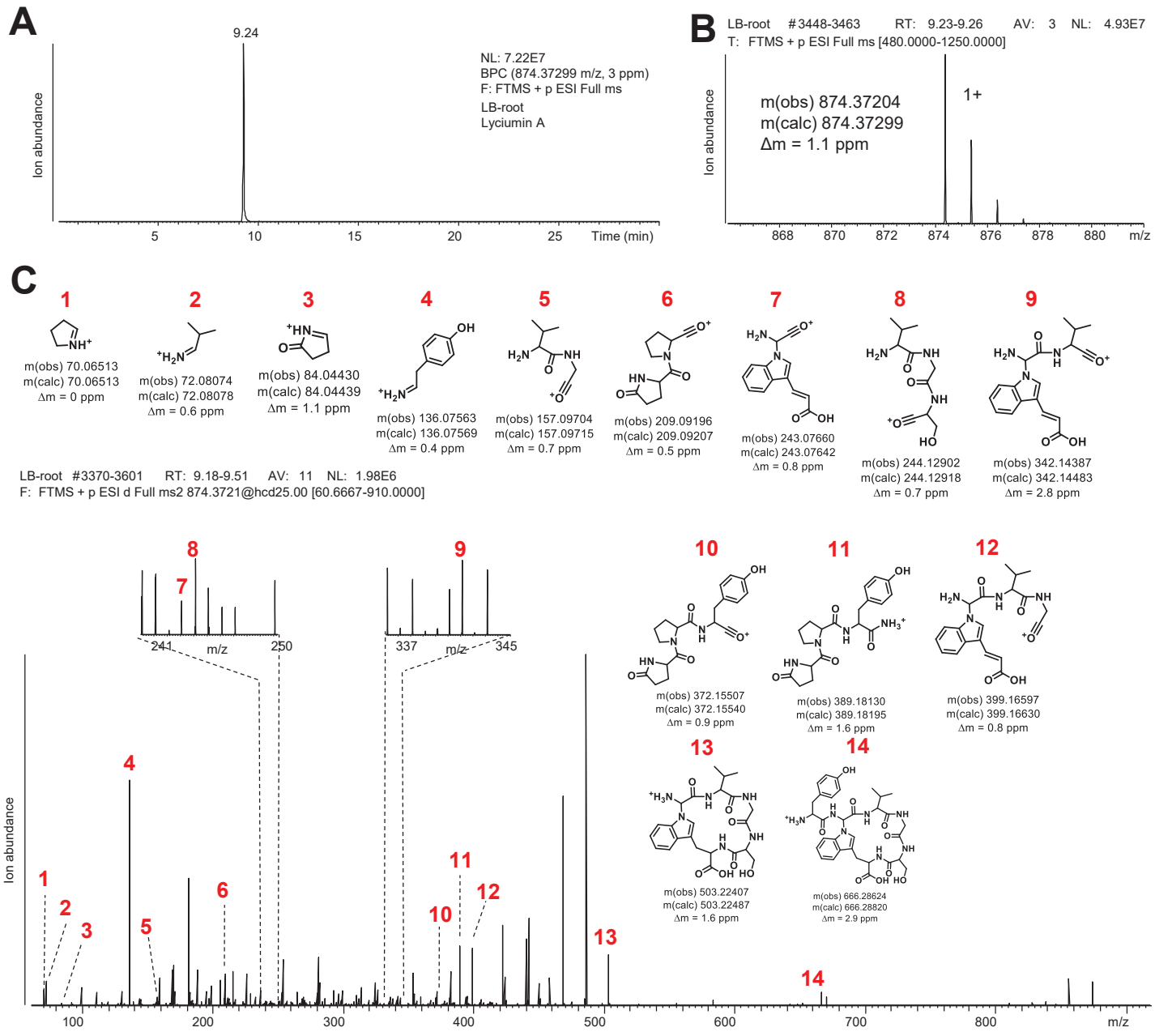



**Figure S1 | MS analysis of lyciumin A.** (**A**) LC-MS chemotyping of predicted lyciumin A in peptide extract of *Lycium barbarum* root. (**B**) MS analysis of lyciumin A. (**C**) MS/MS analysis of lyciumin A.

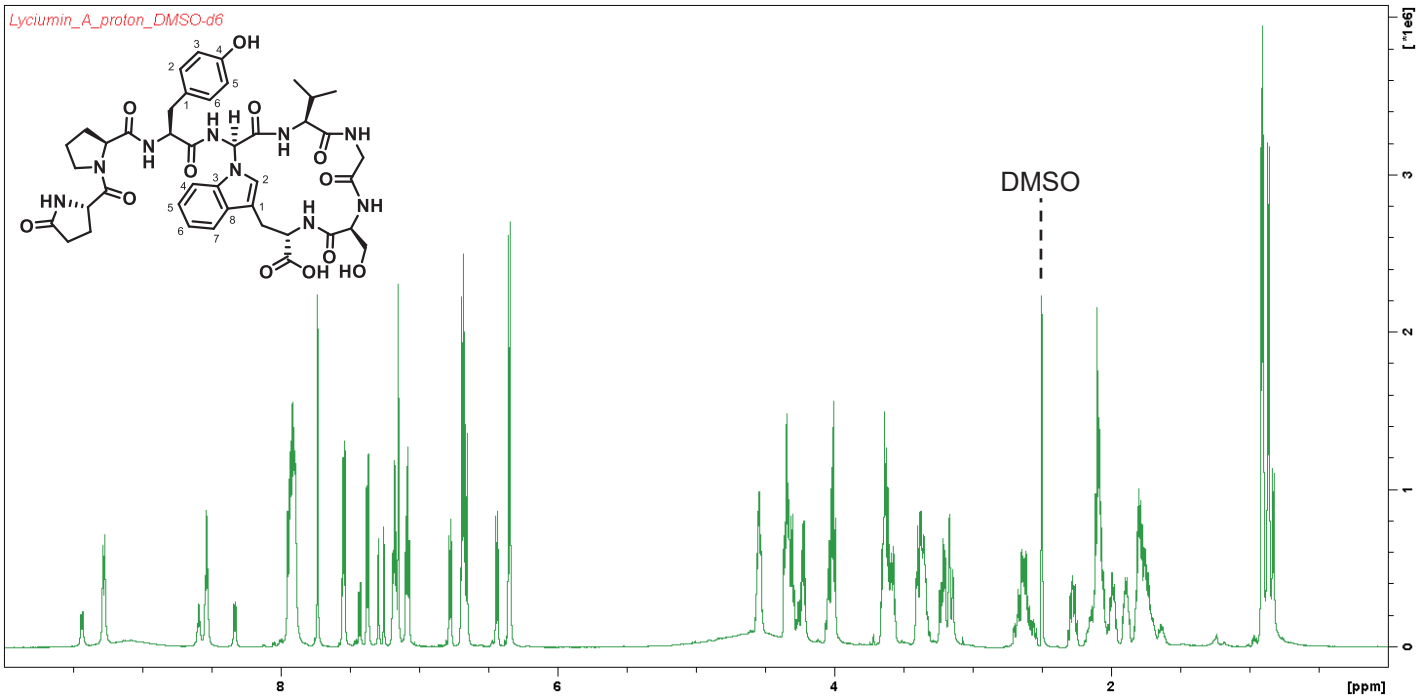



**Figure S2 | $^{1}$H NMR spectrum of lyciumin A in DMSO-d6 (600 MHz).**

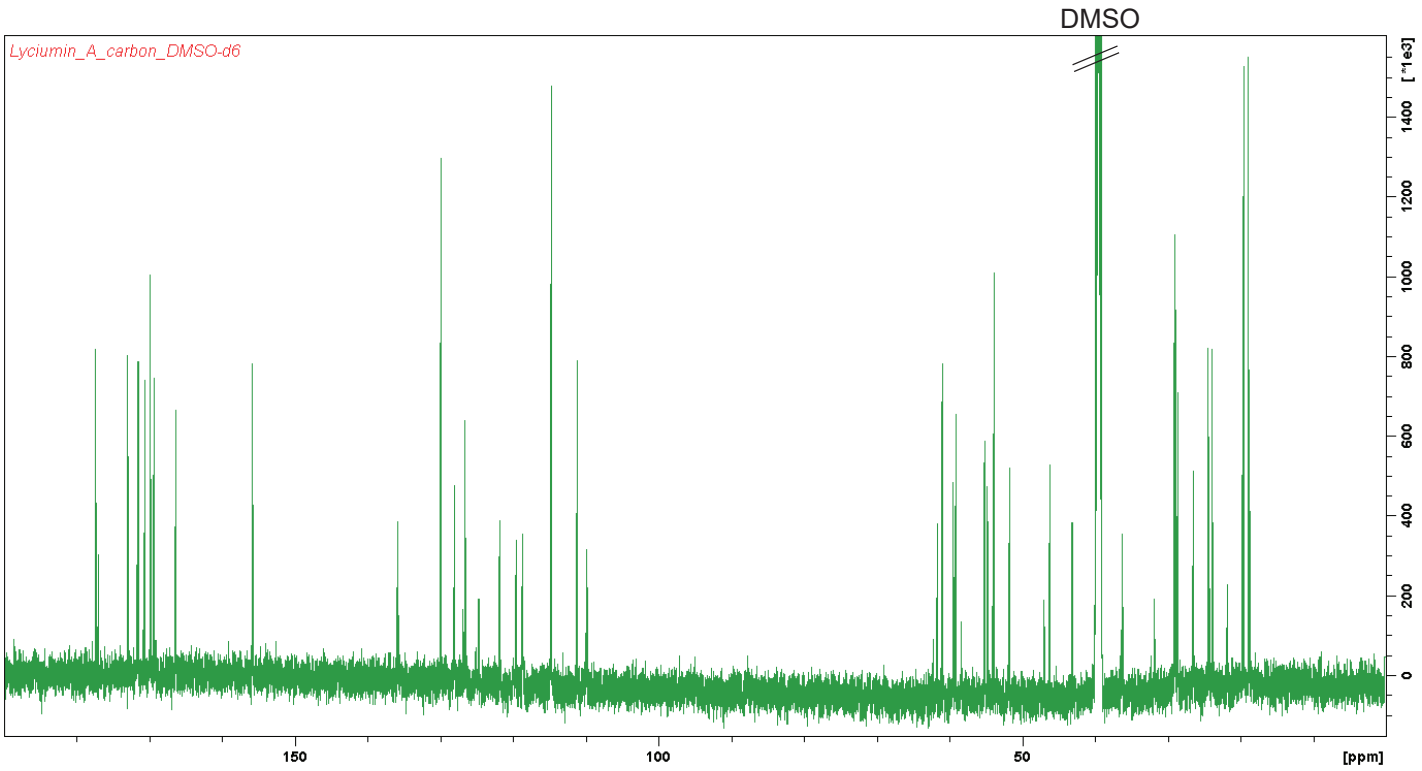



**Figure S3 | $^{13}$C NMR spectrum of lyciumin A in DMSO-d6 (600 MHz).**

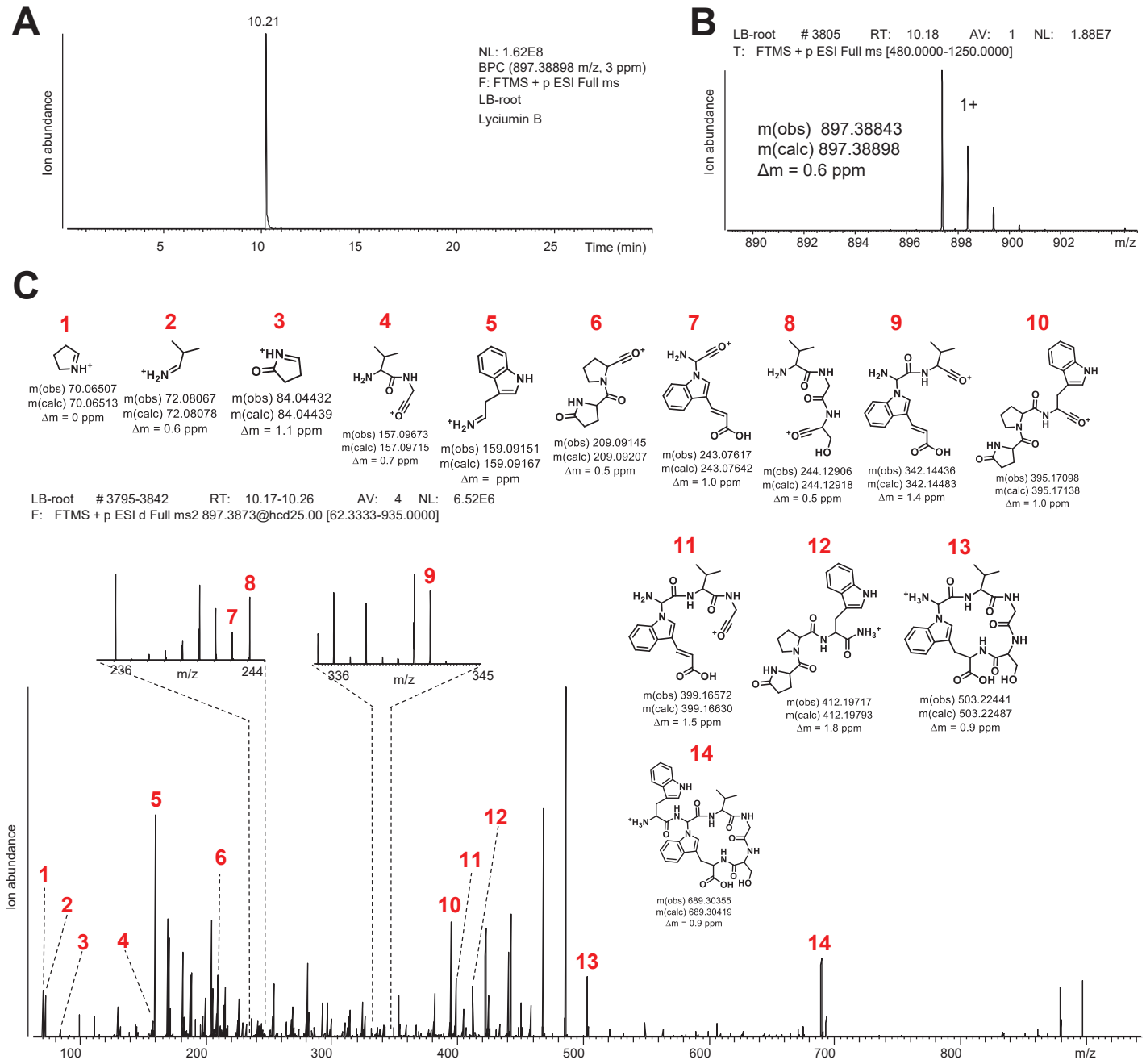



**Figure S4 | MS analysis of lyciumin B.** (**A**) LC-MS chemotyping of predicted lyciumin B in peptide extract of *Lycium barbarum* root. (**C**) MS analysis of lyciumin B. (**D**) MS/MS analysis of lyciumin B.

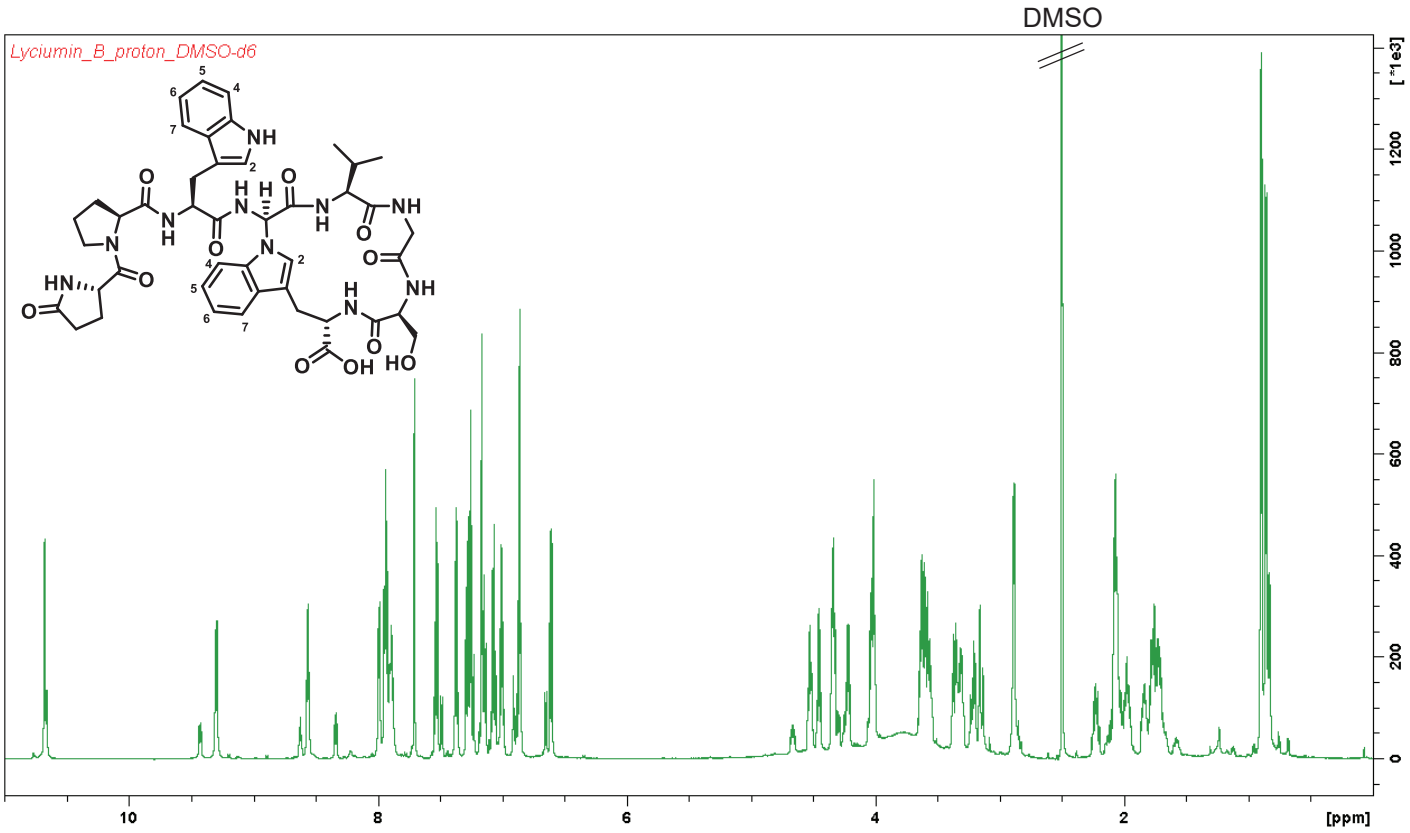



**Figure S5 | $^{1}$H NMR spectrum of lyciumin B in DMSO-d6 (600 MHz).**

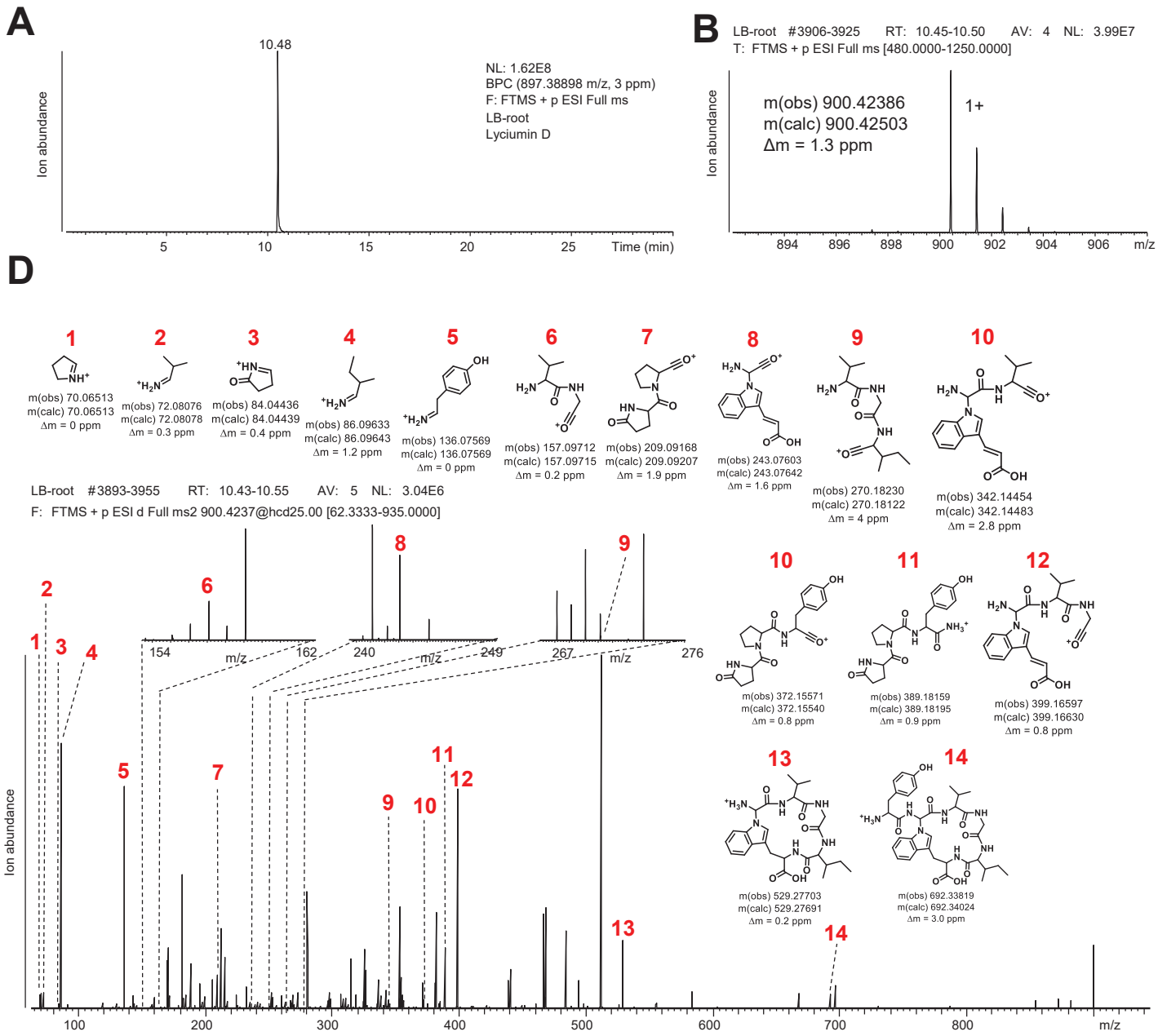



**Figure S6 | MS analysis of lyciumin D.** (**A**) Lyciumin D planar structure. (**B**) LC-MS chemotyping of predicted lyciumin D in peptide extract of *Lycium barbarum* root. (**C**) MS analysis of lyciumin D. (**D**) MS/MS analysis of lyciumin D.

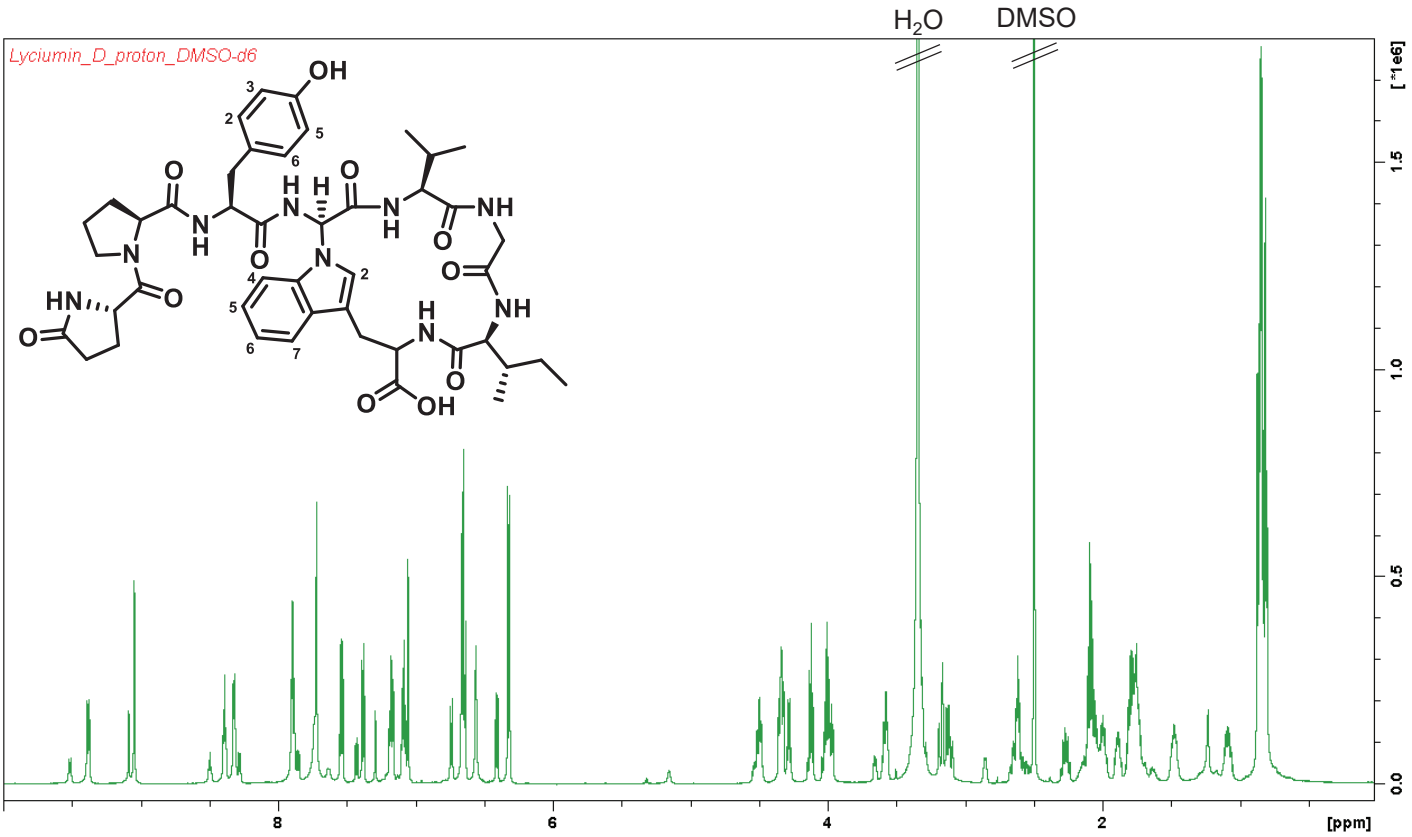



**Figure S7 | $^{1}$H NMR spectrum of lyciumin D in DMSO-d6 (600 MHz).**

>TRINITY_DN33942_c2_g1_i2|m.67 TRINITY_DN33942_c2_g1_i2|g.67 type:5prime_partial len:297
gc:universal TRINITY_DN33942_c2_g1_i2:2-892(+)
EGDVKKLPQPWGVGSWYQAANEGDVKKLRQPYGVGIWYEAANEGQVKKLRQPYGVGSWYNTATKKDVNENLPVTPYFFETDLHQGKKMNLPSL
KNYNPAPILPRKVADSIPFSSDKIEEILKHFSIDKDSEGAKMIKKTIKMCEEQAGNGEKKYCATSLESMVDFTSSYLGTNNIIALSTLVEKET
PEVQIYTIEEVKEKANGKGVICHKVAYPYAIHYCHSVGSTRTFMVSMVGSDGTKVNAVSECHEDTAPMNPKALPFQLLNVKPGDKPICHFILD
DQIALVPSQDATQVSEN
>TRINITY_DN32732_c0_g2_i1|m.98 TRINITY_DN32732_c0_g2_i1|g.98 type:internal len:130
gc:universal TRINITY_DN32732_c0_g2_i1:389-3(-)
GIKQQTRVMLKNYANHGEWVPGIKQQTRVMLKNYTNHGEWVPGIKQQTRVMLKNYPNHGEWVPGIKQQTRVMLKNYANHMELVYGMKQQTRVK
LKKLRQPYGVGSWYQAANEGDVKKIRQPWGVGSWYQ
>TRINITY_DN32732_c0_g2_i3|m.101 TRINITY_DN32732_c0_g2_i3|g.101 type:5prime_partial len:345
gc:universal TRINITY_DN32732_c0_g2_i3:1216-182(-)
WYQAANEGDVKKLRQPWGVGSWYQAANEGDVKKLHQPWGVGSWYQAANEGDVKKLPQPWGVGSWYQAANEGDVKKLRQPYGVGIWYEAANEGQ
VKKLRQPYGVGSWYNTATKKDVNENLPVTPYFFETDLRQGKKMNLPSLKNYNPAPILPRKVADSIPFSSDKIEEILKHFSIDKDSEGAKMIKK
TIKMCEEQAGNGEKKYCATSLESMVDFTSSYLGTNNIIALSTLVEKETPEVQIYTIEEVKEKANGKGVICHKVAYPYAIHYCHSVGSTRTFMV
SMVGSDGTKVNAVSECHEDTAPMNPKALPFQLLNVKPGDKPICHFILDDQIALVPSQDATQVSEN

**Figure S8 | Candidate transcripts of lyciumin precursor peptide gene from *Lycium barbarum* root transcriptome.** Predicted lyciumin core peptides are highlighted in green (lyciumin A), blue (lyciumin B) and red (lyciumin D).

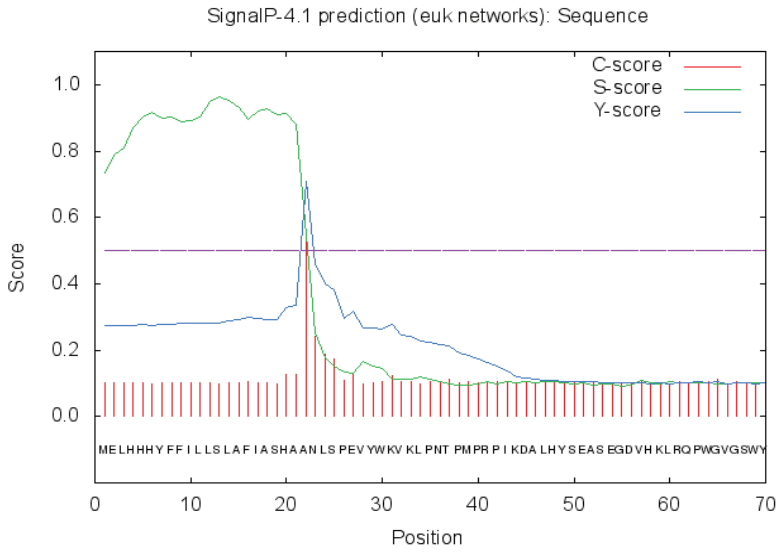



**Figure S9 | Signal peptide prediction of LbaLycA with SignalP-4.1.**

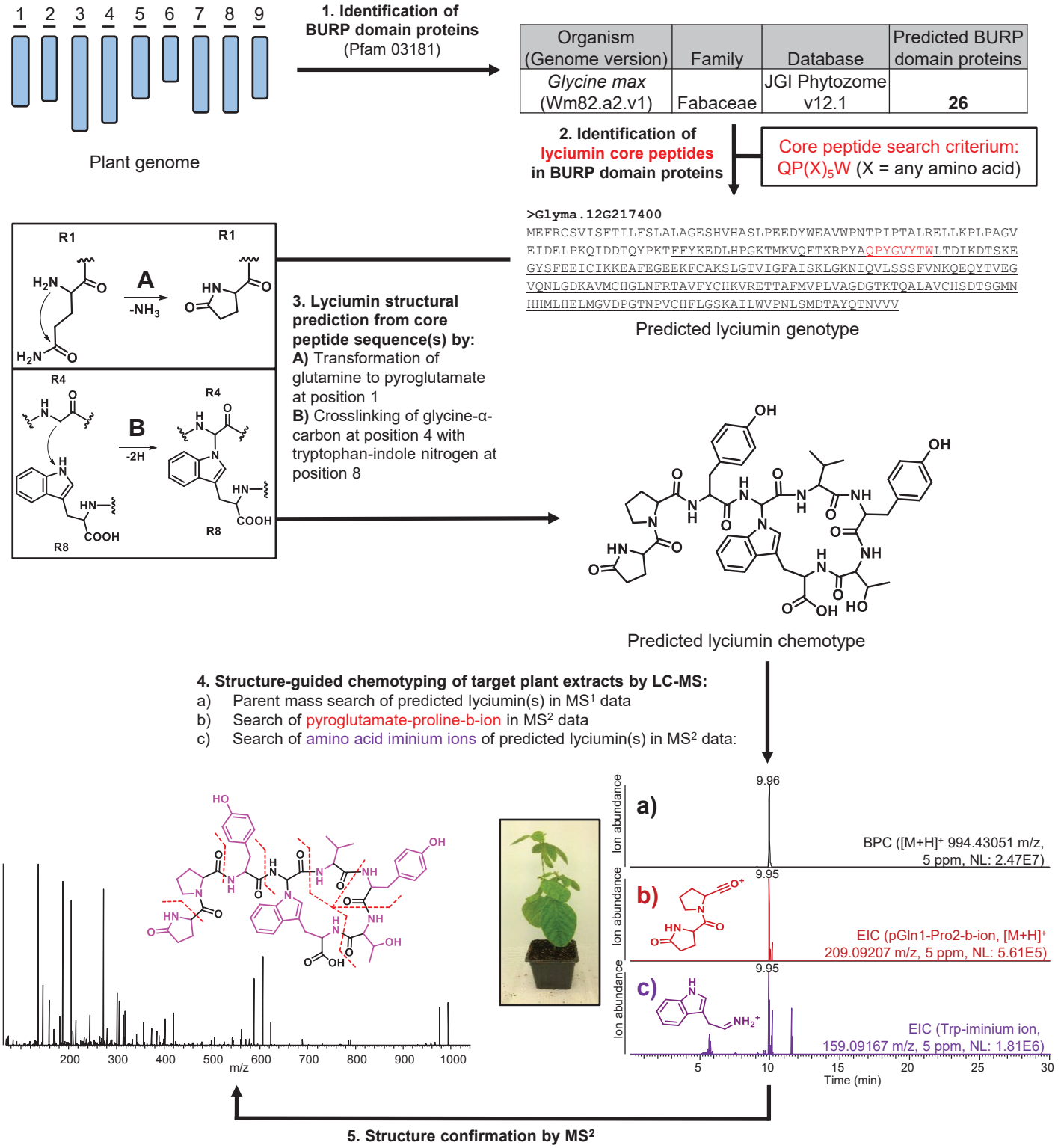



**Figure S10 | Precursor gene-guided genome mining workflow for lyciumin discovery.**

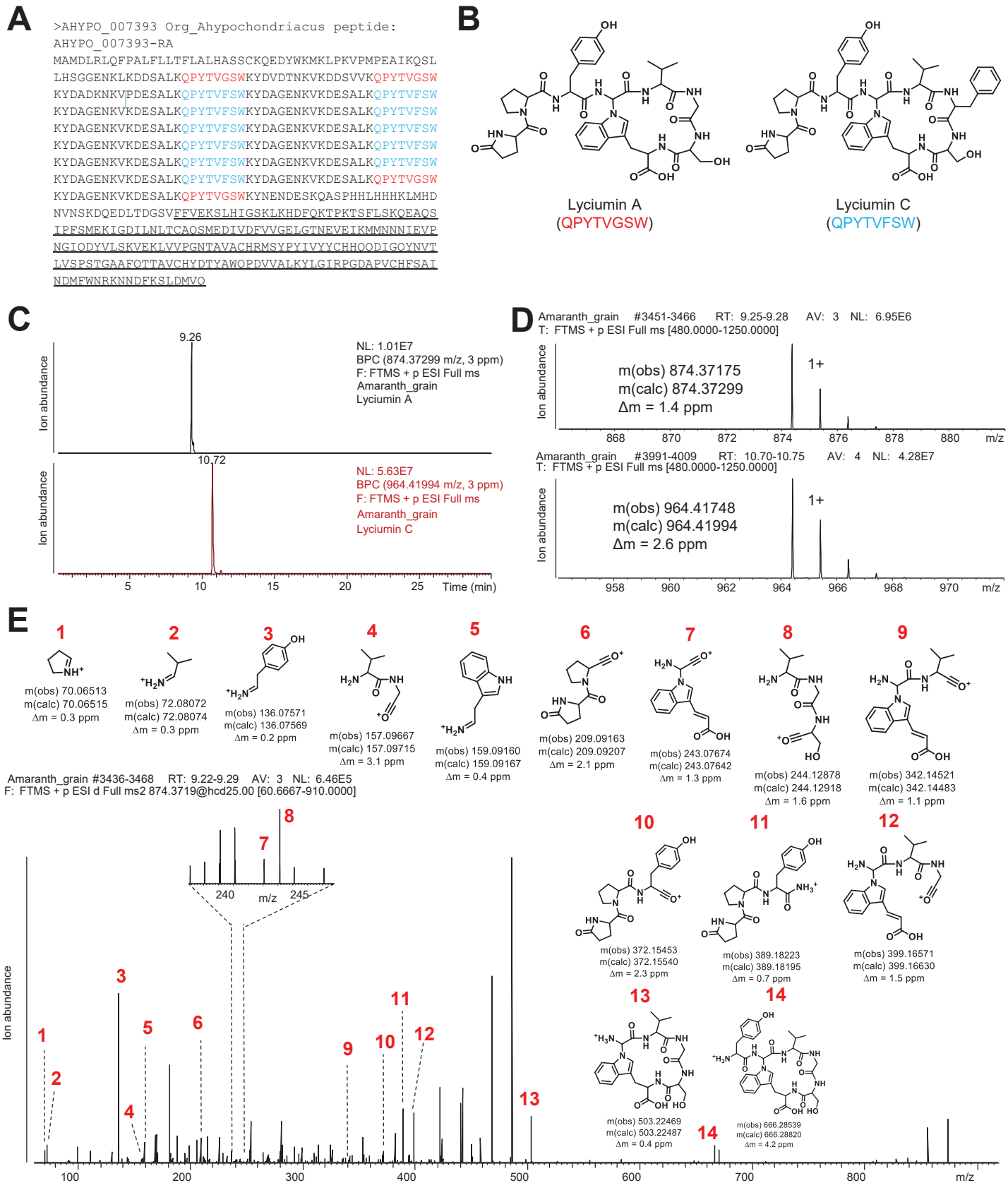



**Figure S11 | Genome mining of lyciumins from *Amaranthus hypochondriacus*.** (**A**) *Amaranthus hypochondriacus* lyciumin precursor peptide (BURP domain underlined, core peptides highlighted in blue and red). (**B**) Predicted lyciumin chemotypes. (**C**) LC-MS chemotyping of predicted lyciumin chemotypes in peptide extract of *Amaranthus hypochondriacus* seeds. (**D**) MS analysis of predicted lyciumin chemotypes of *Amaranthus hypochondriacus*. (**E**) MS/MS analysis of predicted lyciumin chemotypes of *Amaranthus hypochondriacus*.

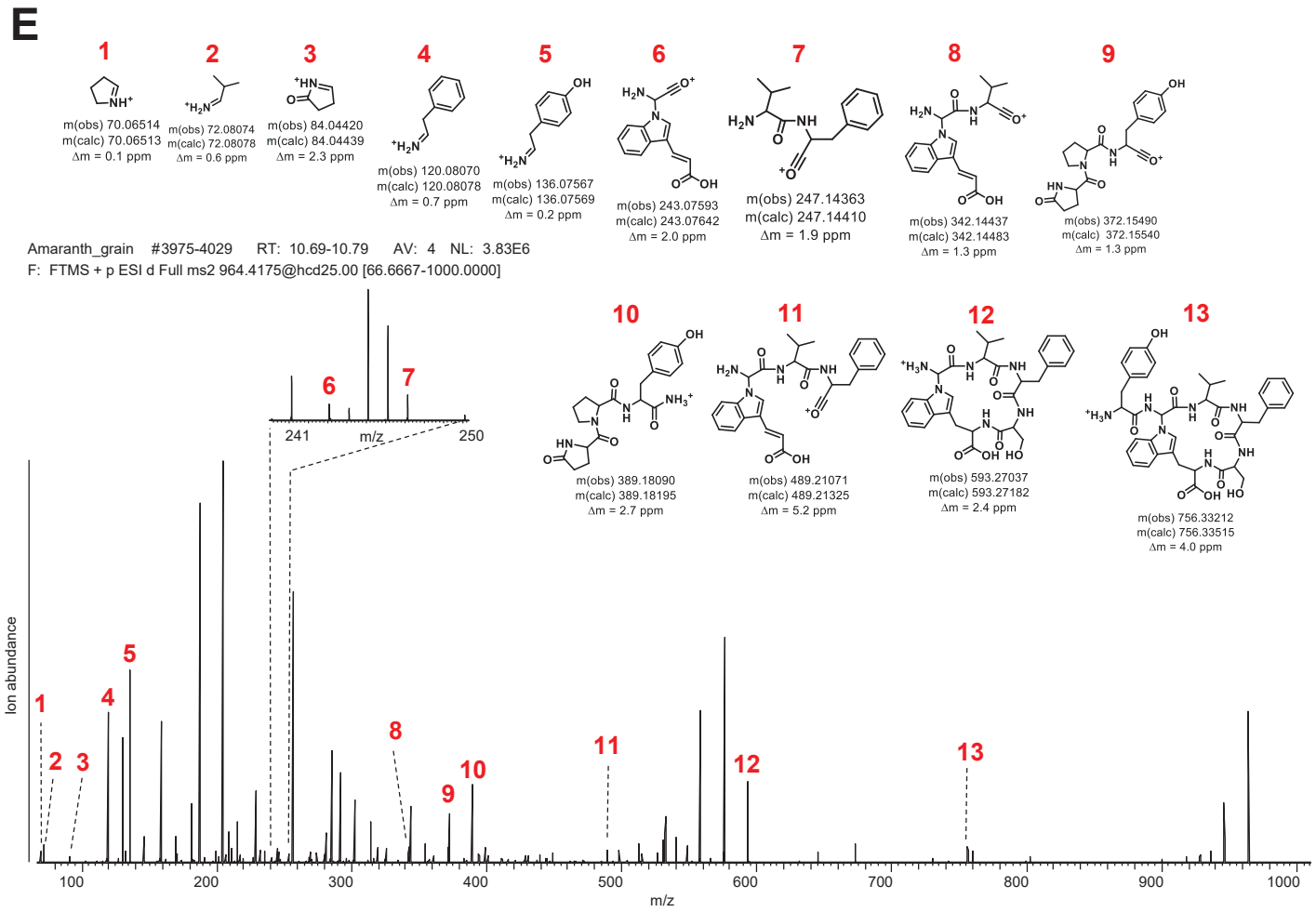



**Figure S11 | Genome mining of lyciumins from *Amaranthus hypochondriacus*.** (**E**) MS/MS analysis of predicted lyciumin chemotypes of *Amaranthus hypochondriacus*. (**F**) *Amaranthus hypochondriacus* plant.

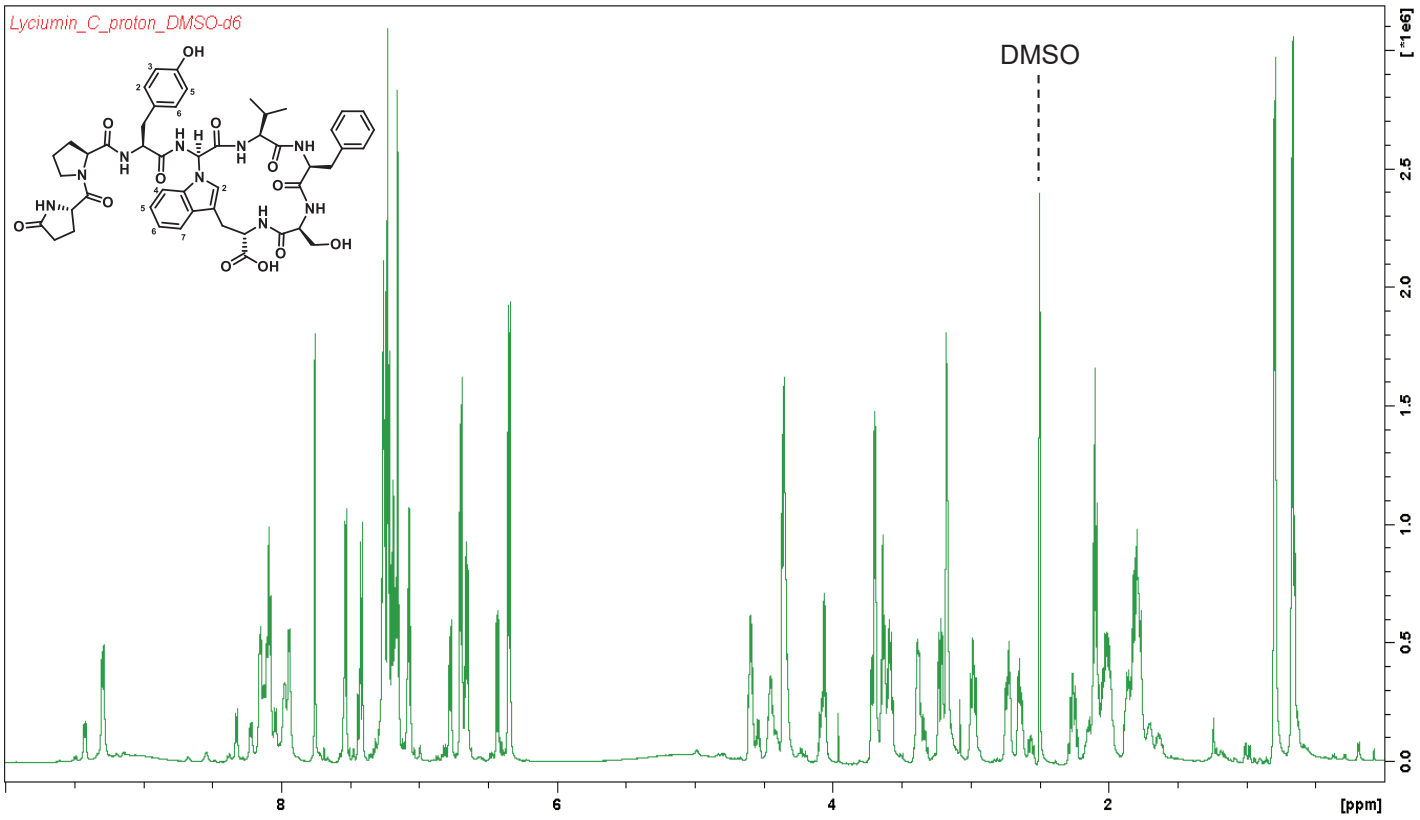



**Figure S12 | $^{1}$H NMR spectrum of lyciumin C in DMSO-d6 (600 MHz).**

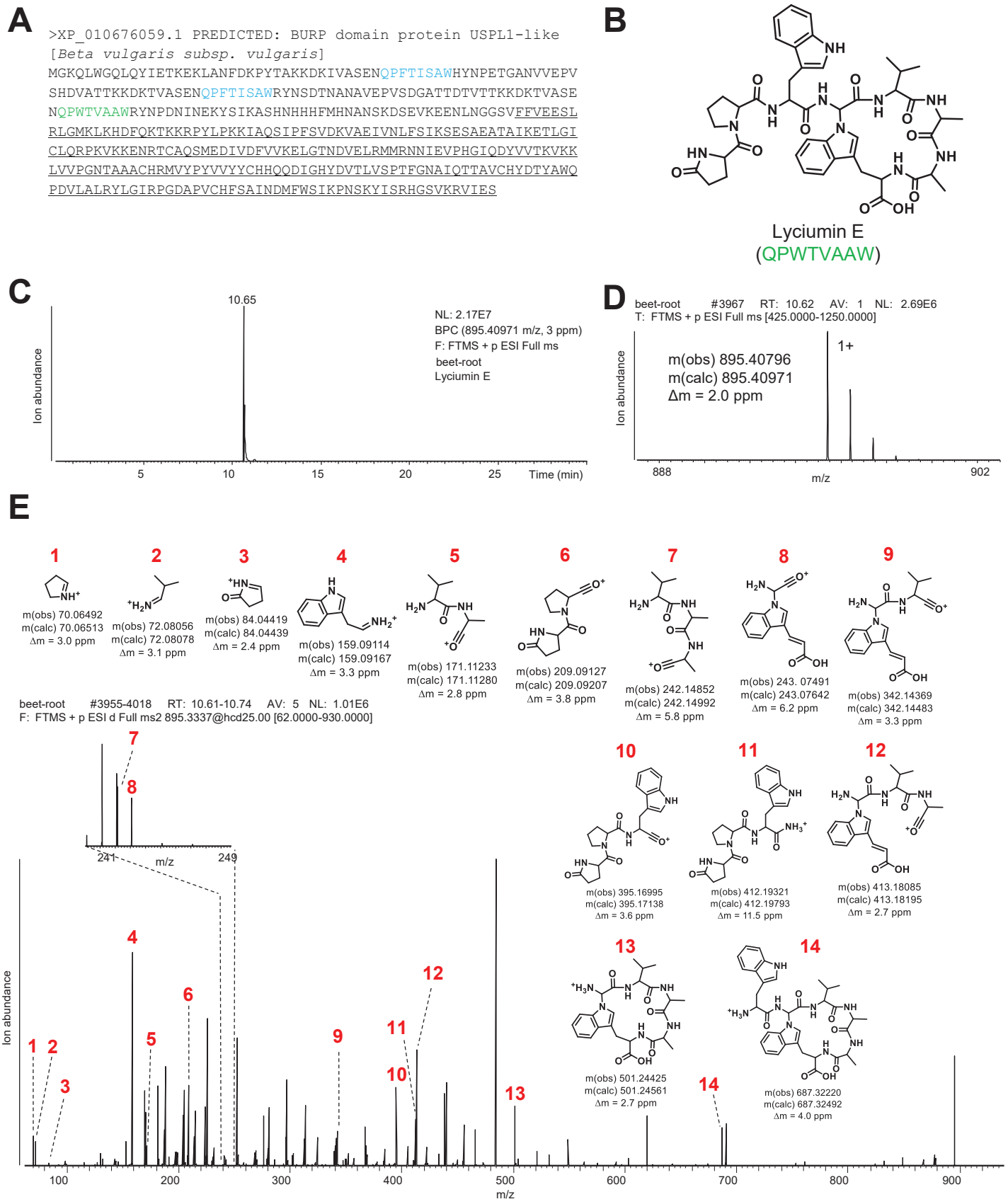



**Figure S13 | Genome mining of lyciumins from *Beta vulgaris*.** (**A**) *Beta vulgaris* lyciumin precursor peptide (BURP domain underlined, core peptides highlighted in blue and green). (**B**) Predicted lyciumin chemotype. (**C**) LC-MS chemotyping of predicted lyciumin chemotype in peptide extract of *Beta vulgaris* root. (**D**) MS analysis of predicted lyciumin chemotype of *Beta vulgaris*. (**E**) MS/MS analysis of predicted lyciumin chemotype of *Beta vulgaris*.

**F**

**Figure S13 | Genome mining of lyciumins from *Beta vulgaris*.** (**F**) *Beta vulgaris* plant.

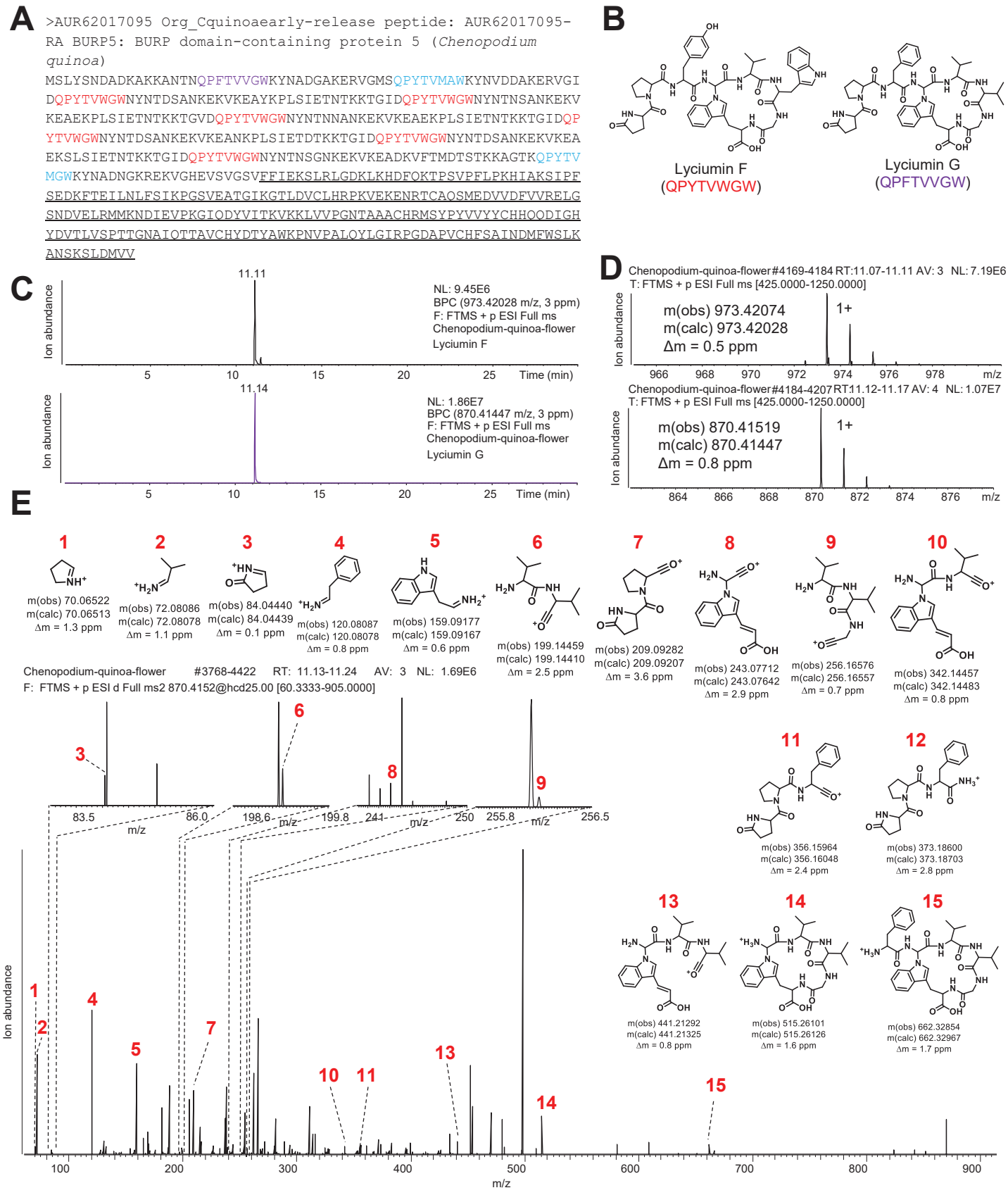
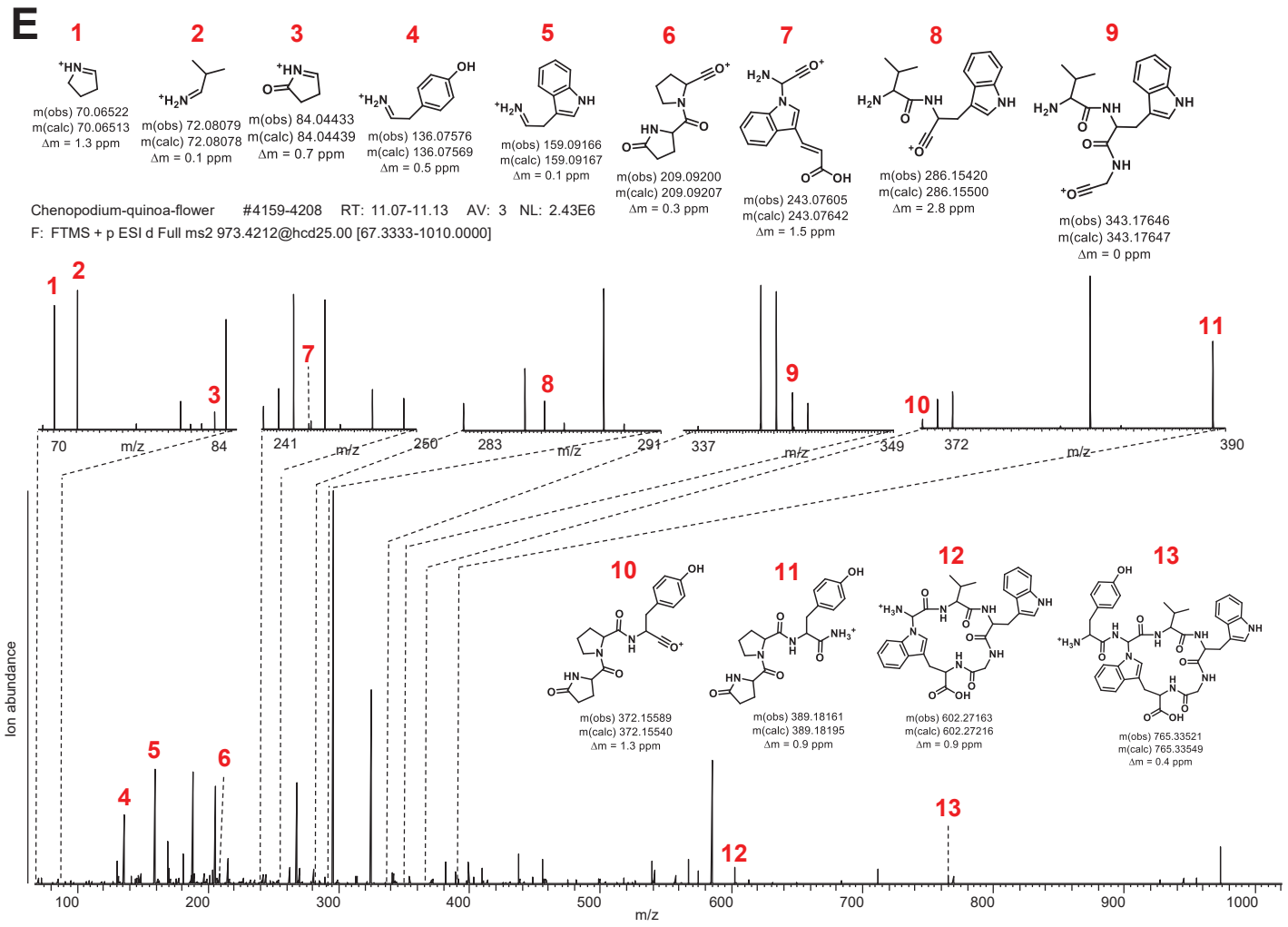



**Figure S14 | Genome mining of lyciumins from *Chenopodium quinoa*.** (**A**) *Chenopodium quinoa* lyciumin precursor peptide (BURP domain underlined, core peptides highlighted in blue, red and purple). (**B**) Predicted lyciumin chemotypes. (**C**) LC-MS chemotyping of predicted lyciumin chemotypes in peptide extract of *Chenopodium quinoa* flower. (**D**) MS analysis of predicted lyciumin chemotypes of *Chenopodium quinoa*. (**E**) MS/MS analysis of predicted lyciumin chemotypes of *Chenopodium quinoa*.




F

**Figure S14 | Genome mining of lyciumins from *Chenopodium quinoa*.** (**E**) MS/MS analysis of predicted lyciumin chemotypes of *Chenopodium quinoa*. (**F**) *Chenopodium quinoa* plant.

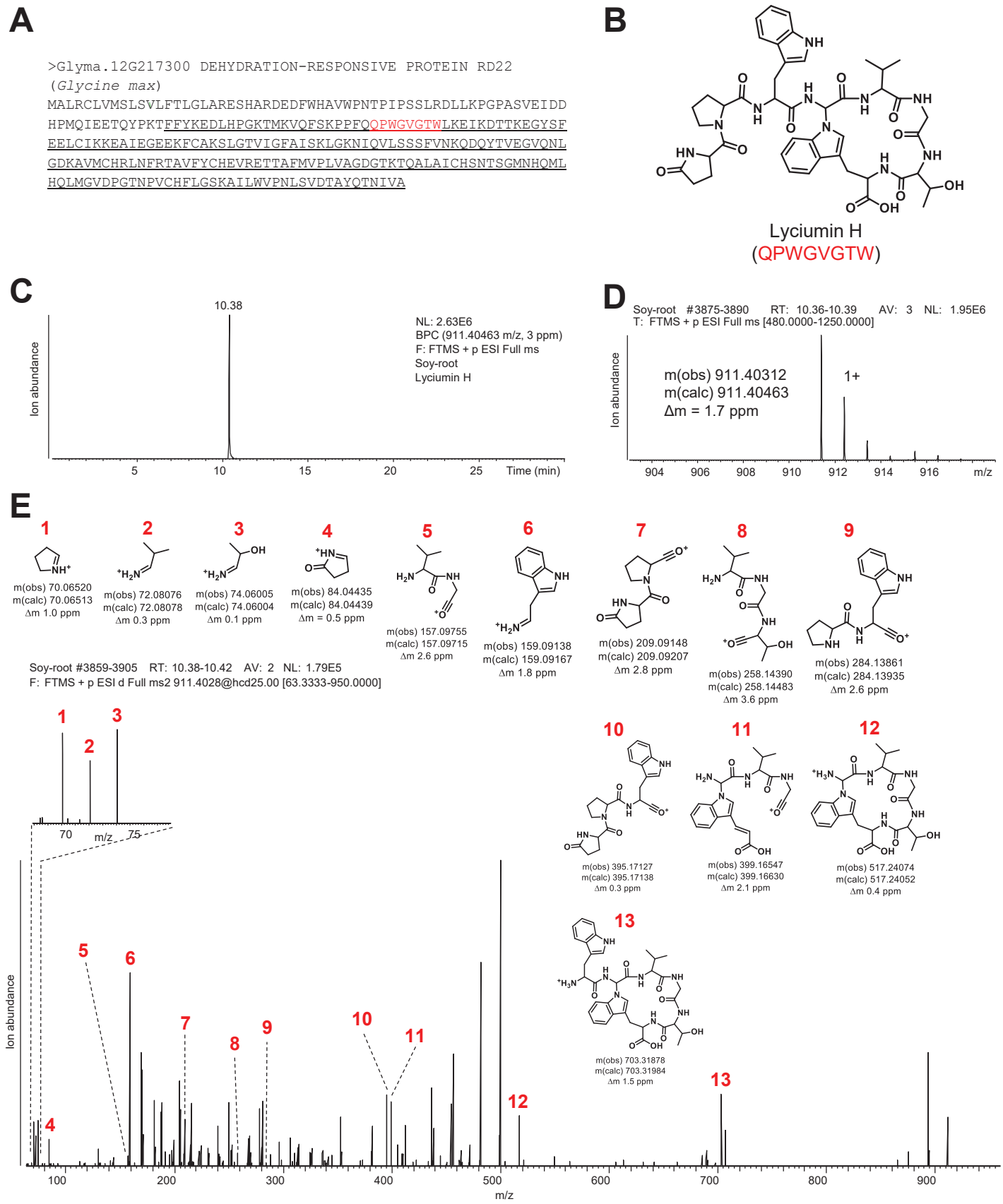



**Figure S15 | Genome mining of lyciumins from *Glycine max*.** (**A**) *Glycine max* lyciumin H precursor peptide (BURP domain underlined, core peptide highlighted in red). (**B**) Predicted lyciumin H chemotype. (**C**) LC-MS chemotyping of predicted lyciumin H chemotype in peptide extract of *Glycine max* root. (**D**) MS analysis of predicted lyciumin H chemotype of *Glycine max* root. (**E**) MS/MS analysis of predicted lyciumin H chemotype of *Glycine max* root.

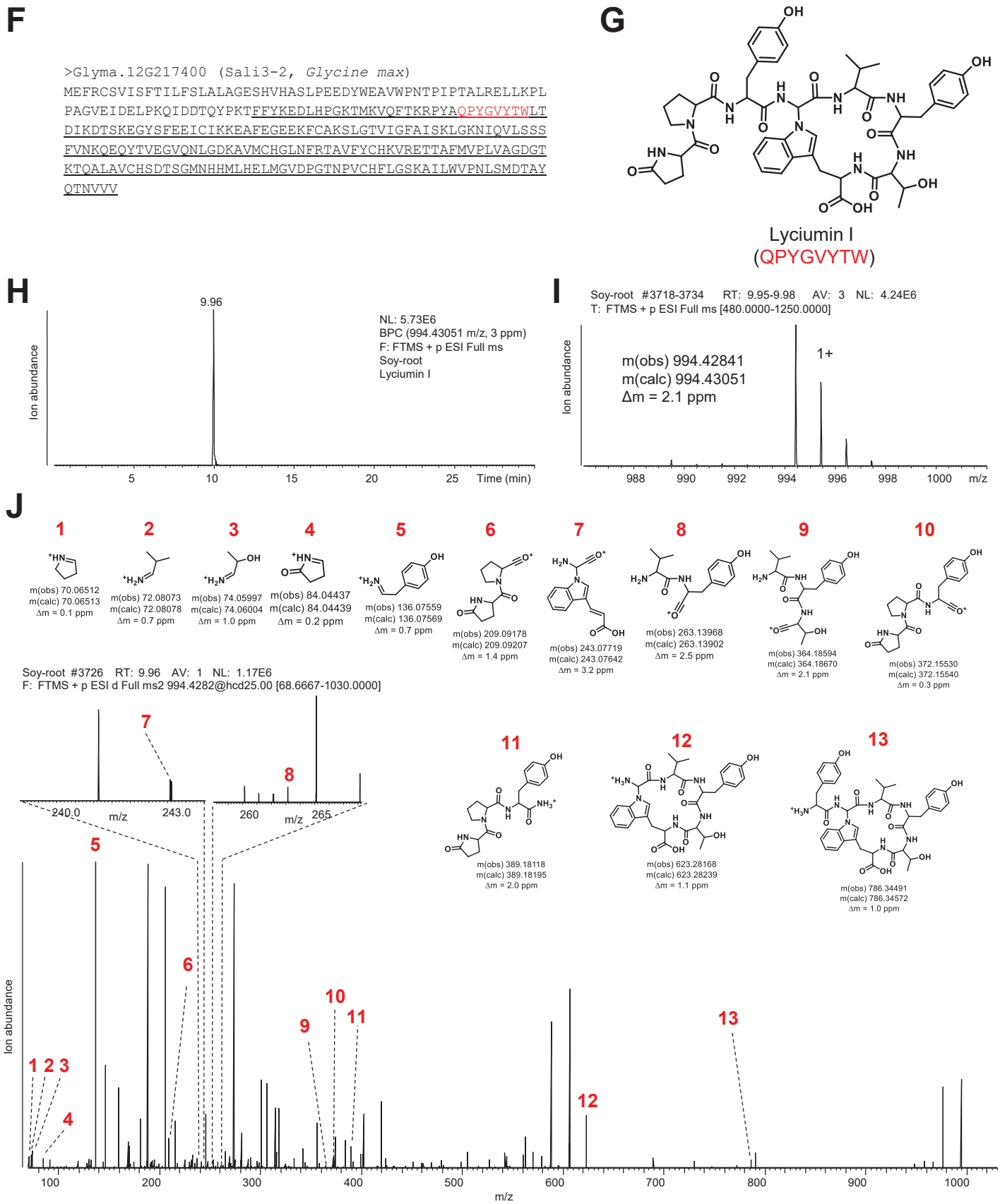



**Figure S15 | Genome mining of lyciumins from *Glycine max*.** (**F**) *Glycine max* lyciumin I precursor peptide (BURP domain underlined, core peptide highlighted in red). (**G**) Predicted lyciumin I chemotype. (**H**) LC-MS chemotyping of predicted lyciumin I chemotype in peptide extract of *Glycine max* root. (**I**) MS analysis of predicted lyciumin I chemotype of *Glycine max* root. (**J**) MS/MS analysis of predicted lyciumin I chemotype of *Glycine max* root.

**K**

**Figure S15 | Genome mining of lyciumins from *Glycine max*.** (**K**) *Glycine max* plant.

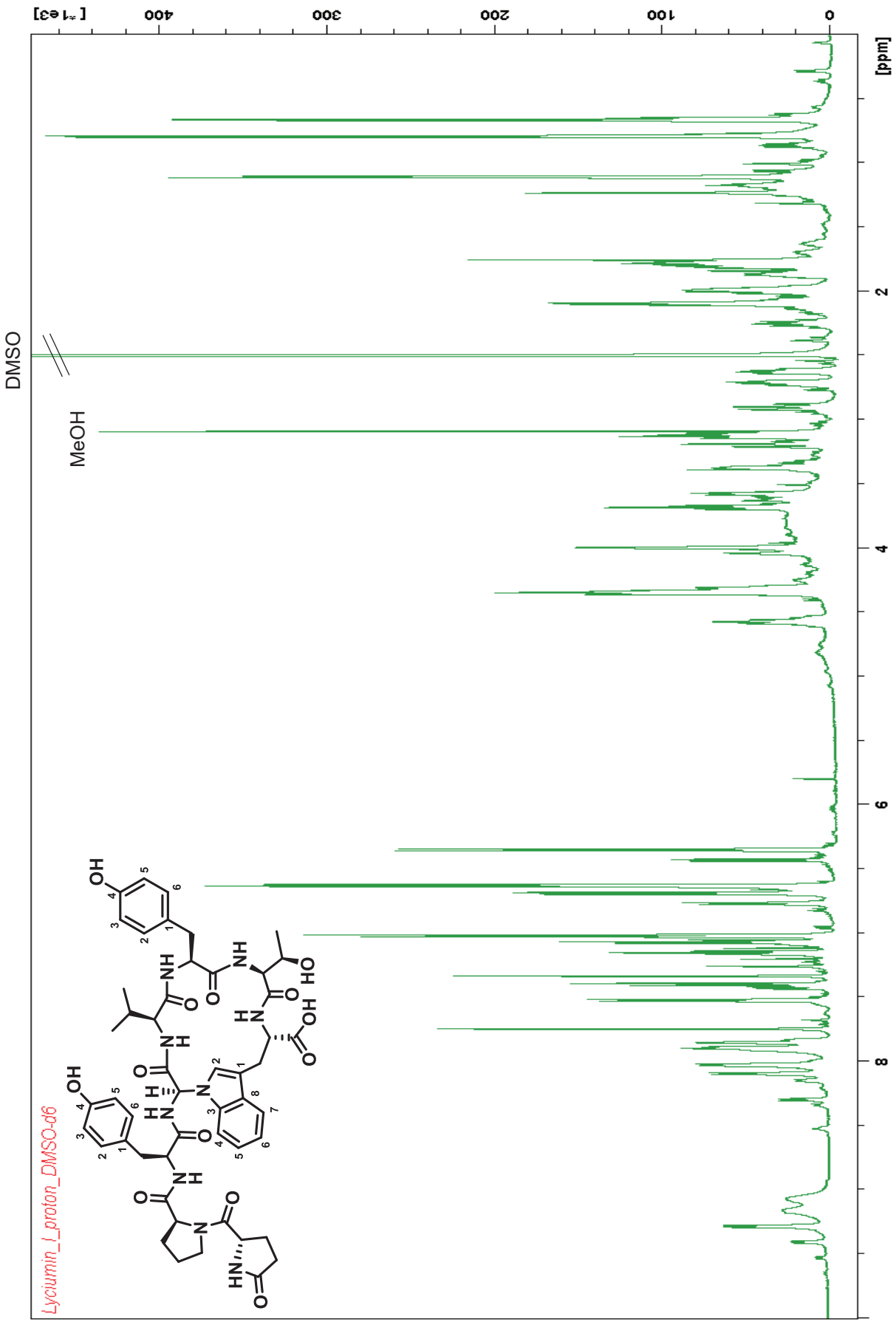



**Figure S16 | $^{1}$H NMR spectrum of lyciumin I in DMSO-d6 (600 MHz).**

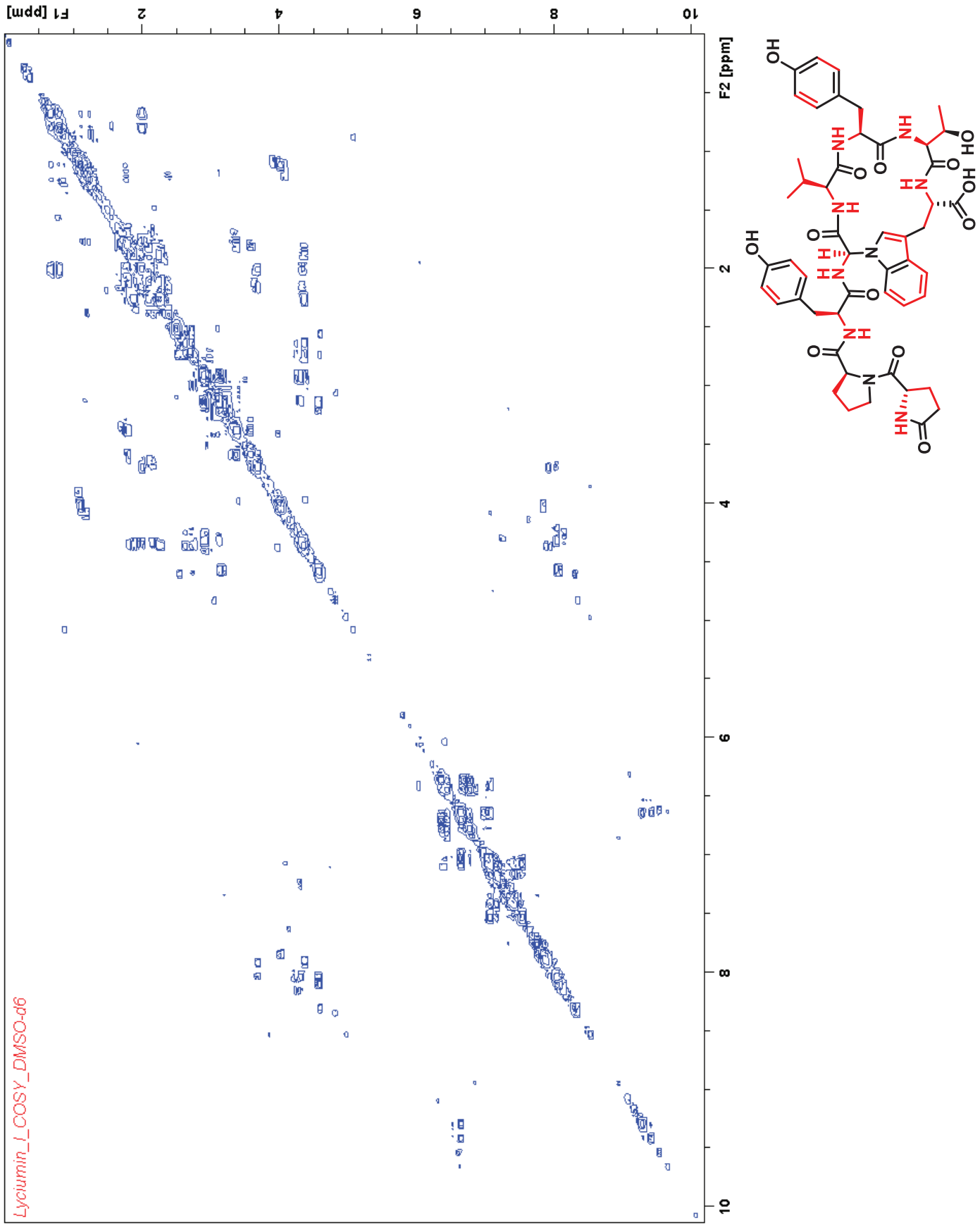



**Figure S17 | $^{1}$H-$^{1}$H COSY NMR spectrum of lyciumin I in DMSO-d6 (600 MHz).**

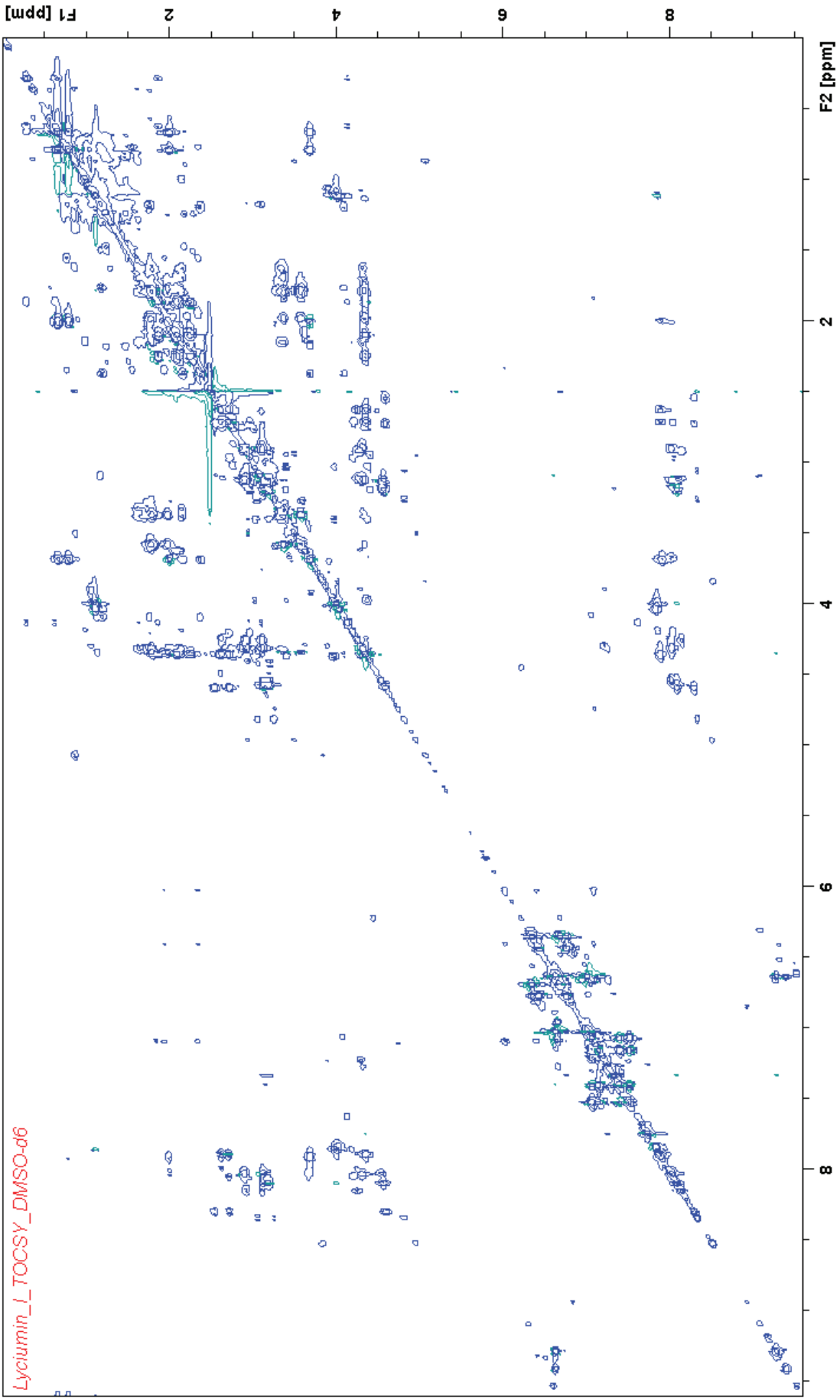



**Figure S18 | $^{1}$H-$^{1}$H TOCSY NMR spectrum of lyciumin I in DMSO-d6 (600 MHz).**

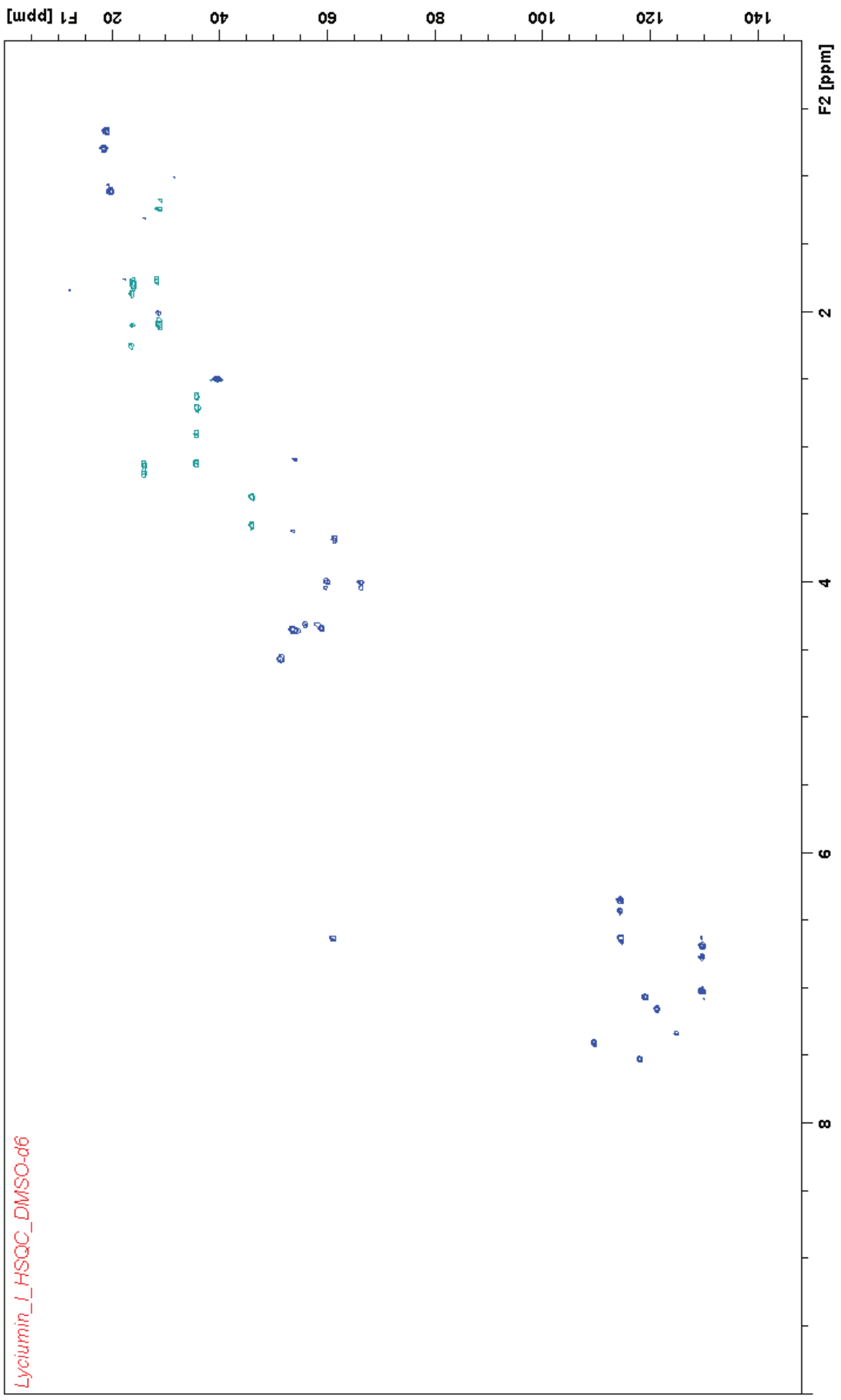



**Figure S19 | $^{1}H$-$^{13}C$ HSQC NMR spectrum of lyciumin I in DMSO-d6 (600 MHz).**

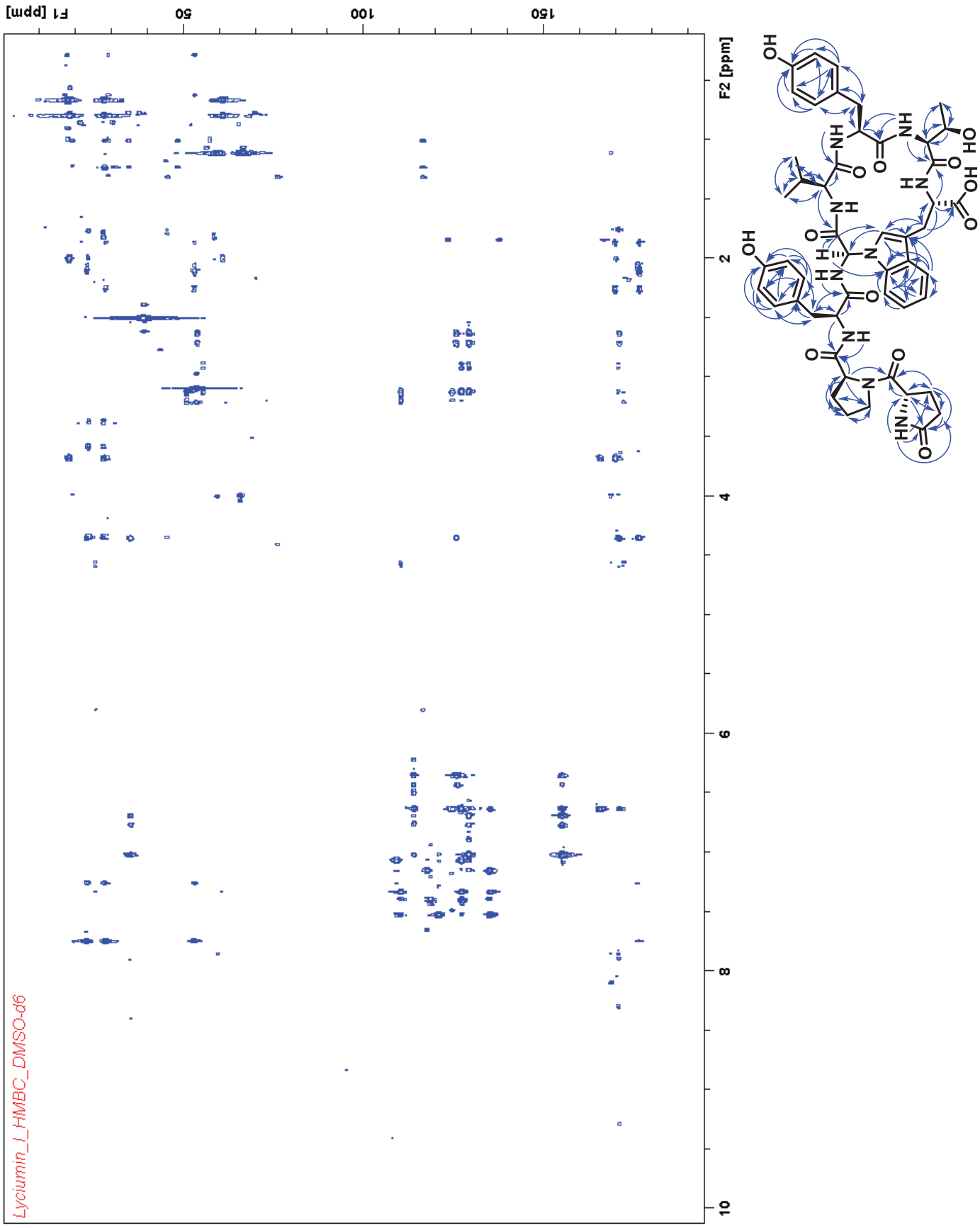



**Figure S20 | $^{1}H$-$^{13}C$ HMBC NMR spectrum of lyciumin I in DMSO-d6 (600 MHz).**

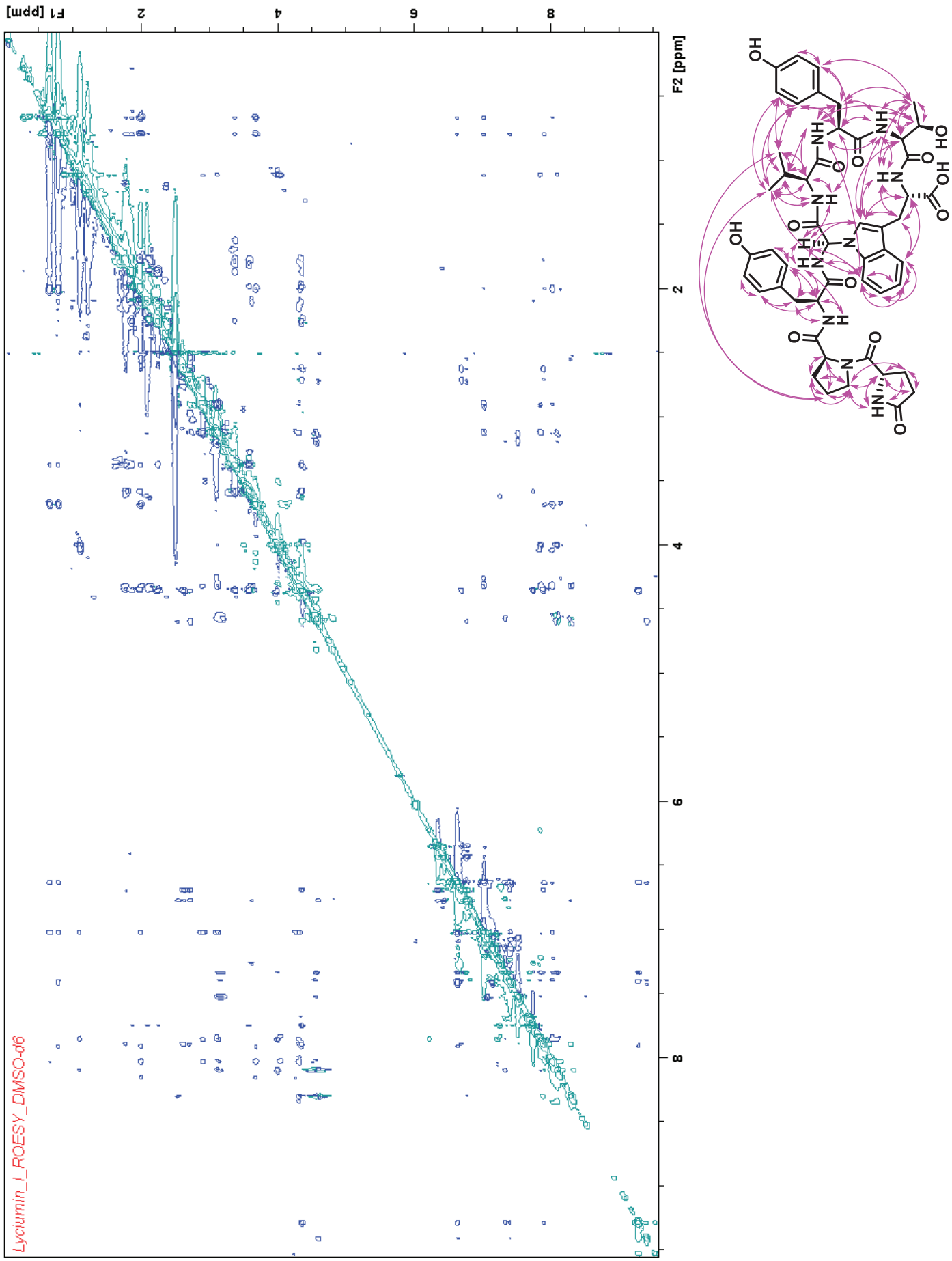



**Figure S21 | $^{1}$H-$^{1}$H ROESY NMR spectrum of lyciumin I in DMSO-d6 (600 MHz).**

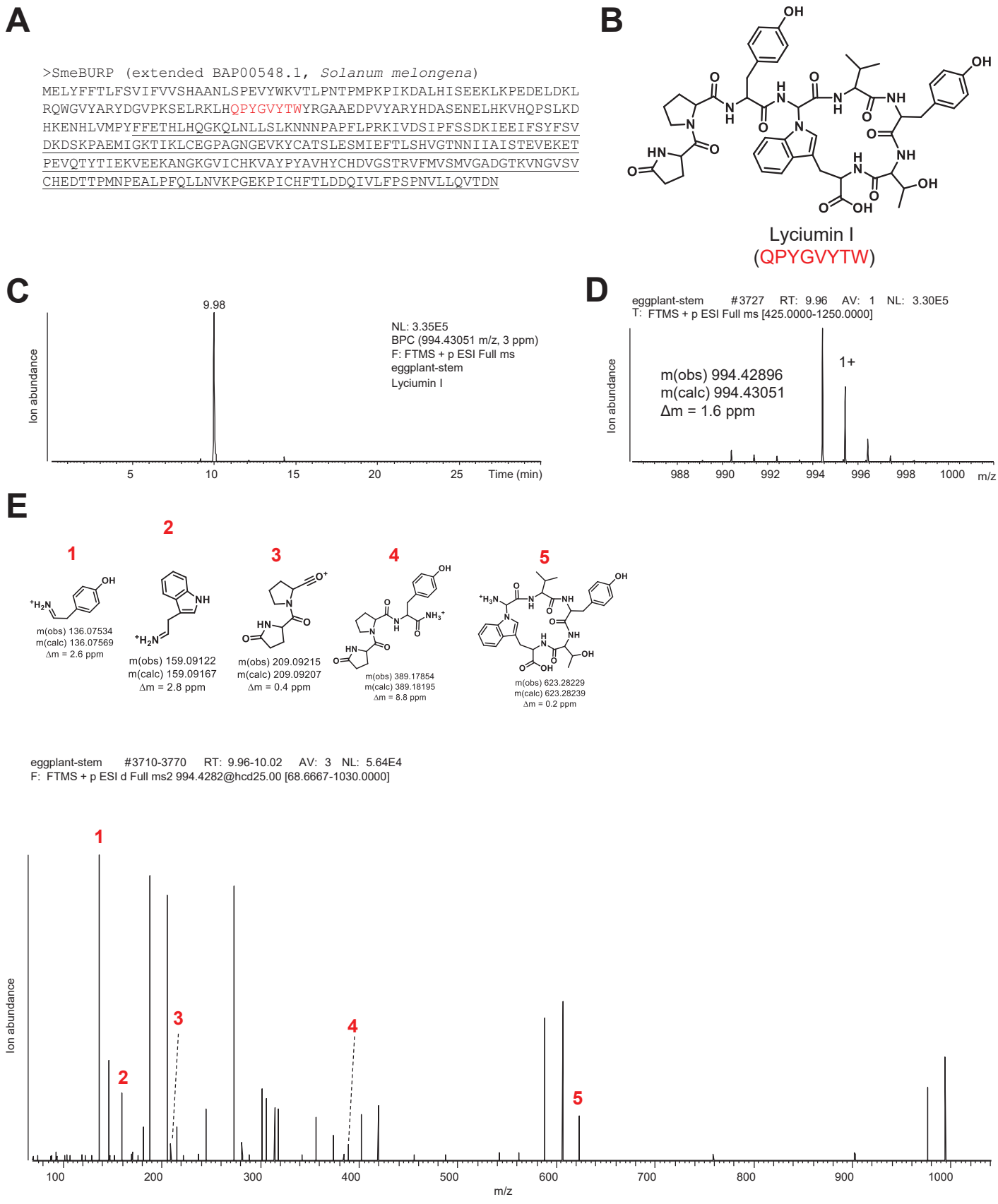


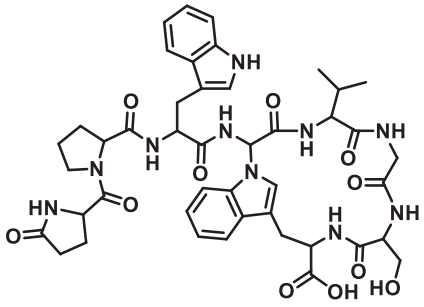

**Figure S22 | Genome mining of lyciumins from *Solanum melongena*.** (**A**) *Solanum melongena* lyciumin I precursor peptide (BURP domain underlined, core peptide highlighted in red). (**B**) Predicted lyciumin I chemotype. (**C**) LC-MS chemotyping of predicted lyciumin I chemotype in peptide extract of *Solanum melongena* stem. (**D**) MS analysis of predicted lyciumin I chemotype of *Solanum melongena* stem. (**E**) MS/MS analysis of predicted lyciumin I chemotype of *Solanum melongena* stem.

**F**

>Sme2.5_02115.1_g00002.1 unnamed protein product (*Solanum melongena*)

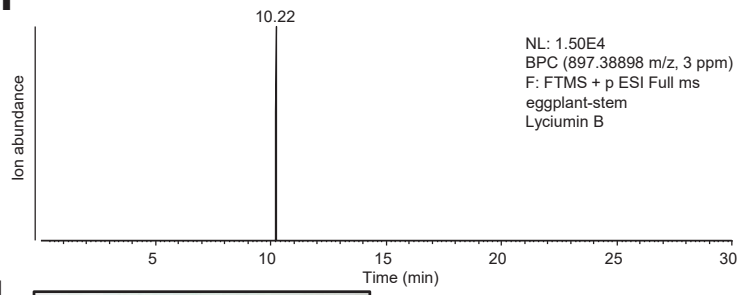
MCDVYTSQLVLVASQAANLSPEVYWKVKLPNTPMPKPIKDALHISEKTAYNGDKSTKISQ
PWGVGSWYQAAPENDLHKVRQPWGVLGWYHDAPENELHKLRQPWGVGSWYQAALGNELHK
LRQPWGVGSWYHDAPENELHKLRQPWGVGSWYQAAPENELYKVRQPWGVLGWSHVPLRMS
FTNCANRGEWAHDLHQGKAMNLLSLKNYNPAPILPRKVVDSIFFSSDKIEEILSHFSADK
DSERAEMIKKTIKMCEDPAGNGEVKHCATSLESMLDFTVSHLGTNNIIAISTEVEKETPE
VQTYTIEKVEEKANGKGVVCHKVAYPYSVHFCHDVGSTRTFMVSMVGADGTKVNAVSVCH
EDTAPMNPKALPFQLLNVKPGDKPICHFTLDDQIALFPSPNVPLQVTKN

**G**

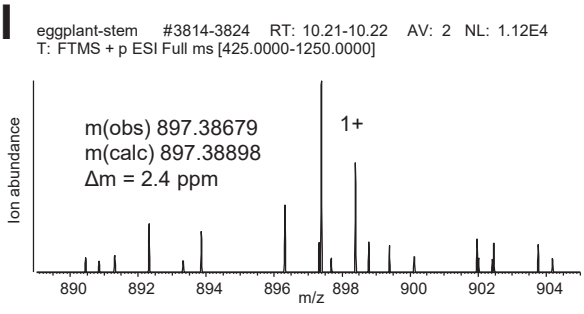


Lyciumin B
(QPWGVGSW)

**H**




**I**




**J**

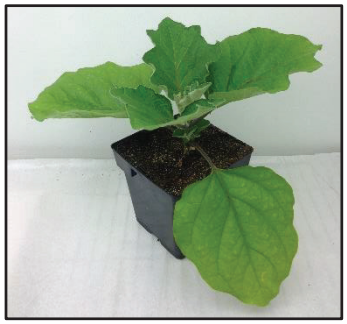


**Figure S22 | Genome mining of lyciumins from *Solanum melongena*.** (**F**) *Solanum melongena* lyciumin B precursor peptide (BURP domain underlined, core peptides highlighted in blue and red). (**G**) Predicted lyciumin B chemotype. (**H**) LC-MS chemotyping of predicted lyciumin B chemotype in peptide extract of *Solanum melongena* stem. (**I**) MS analysis of predicted lyciumin B chemotype of *Solanum melongena* stem. (**J**) *Solanum melongena* plant.

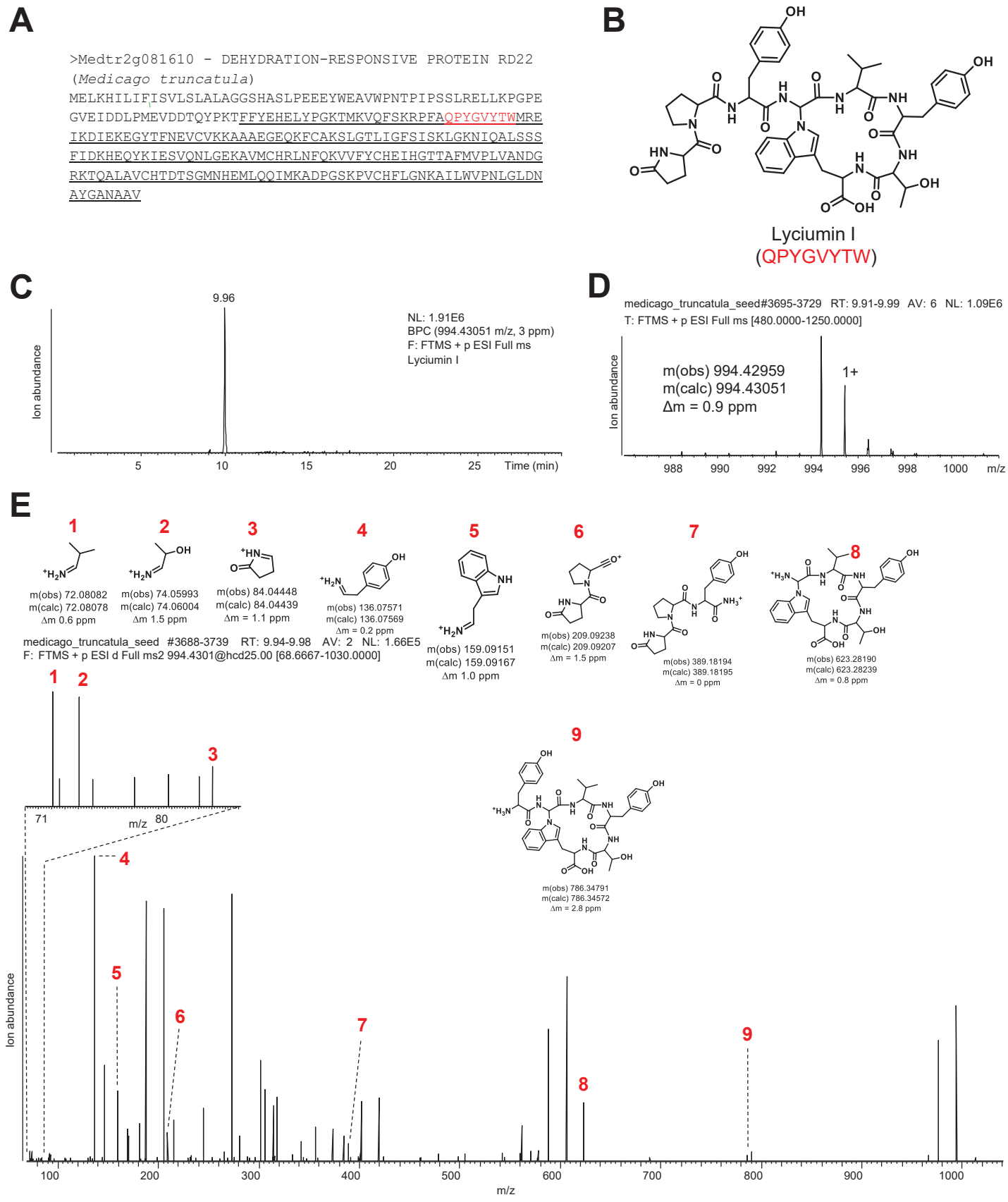



**Figure S23 | Genome mining of lyciumins from *Medicago truncatula*.** (**A**) *Medicago truncatula* lyciumin I precursor peptide (BURP domain underlined, core peptide highlighted in red). (**B**) Predicted lyciumin I chemotype. (**C**) LC-MS chemotyping of predicted lyciumin I chemotype in peptide extract of *Medicago truncatula* seeds. (**D**) MS analysis of predicted lyciumin I chemotype of *Medicago truncatula* seeds. (**E**) MS/MS analysis of predicted lyciumin I chemotype of *Medicago truncatula* seeds.

**A**

*Amaranthus hypochondriacus*

| Gene expression (RSEM, [TPM]) in tissues/conditions | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Transcript | Lyciumin chemotype | Cotyledon | Developing embryo | Drought | Leaf | Mature seed | Root | Stem | Flower |
| AHYPO_007393-RA | Lyciumin A & C | 37338 | 2055 | 60460 | 7236 | 66366 | 77877 | 15572 | 13250 |
| | References (SRA) | SRR1598915 | SRR1598909 | SRR1598914 | SRR1598912 | SRR1598916 | SRR1598913 | SRR1598910 | SRR1598911 |

**B**

*Chenopodium quinoa*

| Gene expression (RSEM, [TPM]) in tissues/conditions | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Transcript | Lyciumin chemotype | Flower 1 | Flower 2 | Flower 3 | Leaf 1 | Leaf 2 | Leaf 3 | Seed 1 | Seed 2 | Seed 3 | Stem 1 | Stem 2 | Stem 3 | Seedling 1 | Seedling 2 | Seedling 3 |
| AUR62017095 | Lyciumin F & G | 1227.05 | 1334.44 | 1095.61 | 270.1 | 355.2 | 207.5 | 77.2 | 58.82 | 71.32 | 1196.8 | 2195.5 | 1938 | 154.87 | 183.16 | 122.36 |

| Reference (SRA) | | | | | | |
|---|---|---|---|---|---|---|
| | Flower 1 | SRR5974430 | Seed 1 | SRR5974426 | Seedling 1 | SRR5974428 |
| | Flower 2 | SRR5974427 | Seed 2 | SRR5974424 | Seedling 2 | SRR5974429 |
| | Flower 3 | SRR5974436 | Seed 3 | SRR5974431 | Seedling 3 | SRR5974434 |
| | Leaf 1 | SRR5974438 | Stem 1 | SRR5974432 | | |
| | Leaf 2 | SRR5974437 | Stem 2 | SRR5974433 | | |
| | Leaf 3 | SRR5974435 | Stem 3 | SRR5974425 | | |

**C**

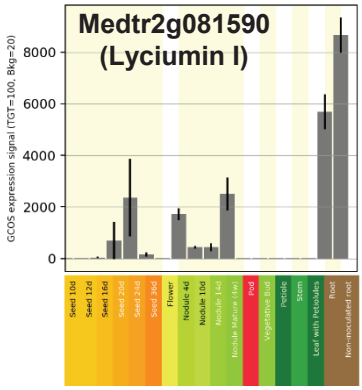
*Glycine max*

| Gene expression (bar.utoronto.ca, [RPKM]) in tissues/conditions | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Transcript | Lyciumin chemotype | Flower | Leaves | Green Pods | Nodule | Root | Root tip | Root Hair 12HAI | Roothair 12HAI mock | Root Hair 24 HAI | Roothair 24HAI mock | Root Hair 48 HAI | Roothair 48HAI mock | Root Hair 48 HAI Stripped | SAM |
| Glyma.12G217300 | Lyciumin H | 247.58 | 33.73 | 3505.92 | 1959.25 | 2072.86 | 6.48 | 508.24 | 709.11 | 720.55 | 999.91 | 782.47 | 931.26 | 6611.09 | 259.63 |
| Glyma.12G217400 (*Sali3-2*) | Lyciumin I | 330.31 | 154.83 | 681.91 | 3770.9 | 4154.42 | 1.54 | 399.73 | 513.62 | 688.9 | 1066.59 | 760.96 | 1386.26 | 1757 | 648.96 |
| | Reference | Libault, M., Farmer, A., Brechenmacher, L., *et al. Plant Physiol.* 2010, 152, 541-555. Libault, M. *et al. The Plant Journal* 2010, 63, 85-99. | | | | | | | | | | | | | |

| Transcript | Lyciumin chemotype | Flower | Young Leaf | One CM Pod | Pod Shell (10-13 DAF) | Pod Shell (14 - 17 DAF) | Nodule | Root | Seed 10 - 13 DAF | Seed 14 - 17 DAF | Seed 21 DAF | Seed 25 DAF | Seed 28 DAF | Seed 35 DAF | Seed 42 DAF |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Glyma.12G217300 | Lyciumin H | 317 | 10 | 458 | 379 | 877 | 231 | 489 | 12 | 24 | 40 | 24 | 17 | 19 | 15 |
| Glyma.12G217400 (*Sali3-2*) | Lyciumin I | 104 | 23 | 141 | 120 | 95 | 220 | 790 | 4 | 3 | 4 | 36 | 68 | 332 | 399 |
| | Reference | Severin, *et al. BMC Plant Biology* 2010, 10, 160. | | | | | | | | | | | | | |

**D**

*Medicago truncatula*




**E**

*Solanum tuberosum*

| Gene expression (RSEM, [TPM]) in tissues/conditions | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Transcript | Lyciumin chemotype | Flower | Leaf | Leaf 2 | Petiole | Root | Whole plant | Shoot apex | Stamen | Stem | Stolon | Tuber | Tuber 2 | Tuber cortex | Tuber peel | Tuber pith | Tuber sprout |
| PGSC0003DMG400047074 | Lyciumin B, K, L, M, N, O | 4.56 | 17.03 | 2.86 | 10.71 | 1190.71 | 324.87 | 474.98 | 2.26 | 118.65 | 491.75 | 718.27 | 411.05 | 91.6 | 606.46 | 38.04 | 3599.35 |
| | Reference | ERR029909 | ERR029910 | ERR029919 | ERR029911 | ERR029917 | ERR029922 | ERR029912 | ERR029918 | ERR029913 | ERR029914 | ERR029915 | ERR029916 | ERR029924 | ERR029921 | ERR029920 | ERR029923 |

**Figure S24 | Lyciumin precursor gene expression in (A) *Amaranthus hypochondriacus,* (B) *Chenopodium quinoa,* (C) *Glycine max,* (D) *Medicago truncatula*[15] and (E) *Solanum tuberosum.***

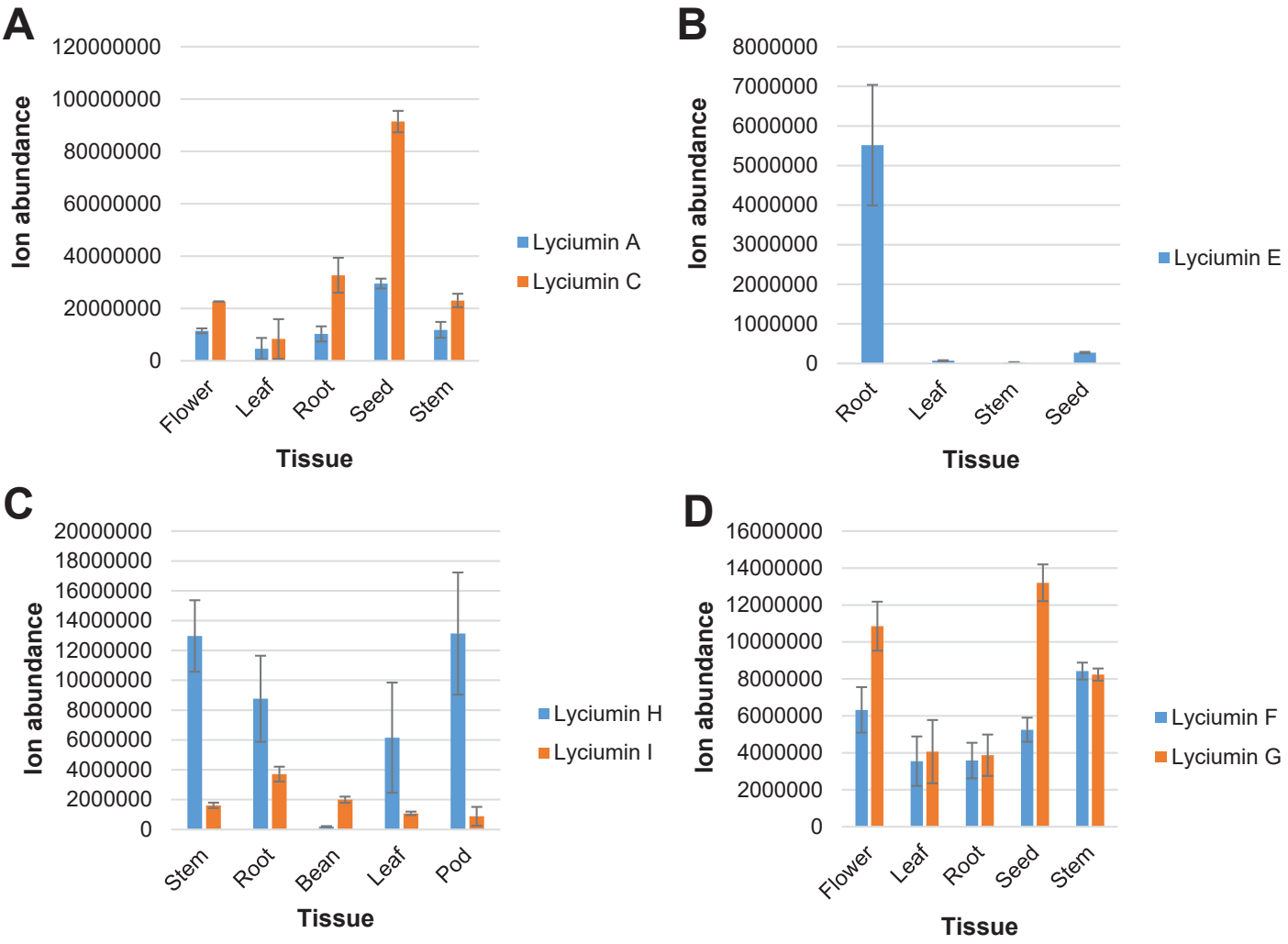



**Figure S25 | Lyciumin chemotyping in source plant tissues.** (**A**) *Amaranthus hypochondriacus*. (**B**) *Beta vulgaris*. (**C**) *Glycine max*. (**D**) *Chenopodium quinoa*.

**A** **Predicted precursor peptide from *Solanum tuberosum* genome (JGI Phytozome, v4.03)**

>PGSC0003DMG400047074 Org_Stuberosum peptide: PGSC0003DMP400069178 BURP domain-containing protein (PAC:37467747)
MELHHQYYFFTFFSVIFVVSHAANLSPEVYWRVKLPNTPMPTPIKDALHISDGIRLPLRTSFTKYANHGEWVDGIRLPFENELHKVRQPWGVDSWYQAAPENELHKVRQPYGVGVWYN
DAAKKDLNDNHPVTPYFFETDLHQGKKMNLQSLKNYNPAPILPRKVVDSIAFSSDKIEEILNHFSVDKDSERAKDIKKTIKTCEEPAGNGEVKHCATSLESMIDFTLSHLGTNNIIAM
STEVEKETPEVQAYTIEKVEEKANGKGVVCHKVAYPYAVHFCHDVGSTRTFMVSMVGADGTKVNAVSVCHEDTASMNPKALPFQLLNVKPGDKPICHFTLDDQIALFPSQNAVLQVAE
N

**B** **LbaLycA blastp hits of Trinity (v2.4) *de novo* transcriptome assembly:**

>lcl|Potato_tuber|TRINITY_DN47721_c0_g1_i1 (5' partial)
VRQPFGVGRWYNDASENELHKVRQPFGVFGWYQAASENELHKVRQPYGVFGWYNDAAKKDLNDNHPVTPYFFETDLHQGKKMNLESLKNYNPAPILPRKVVDSIAFSSDKIEEILNHF
SVDKDSERAKDIKKTIKMCEDPAGNGEVKHCATSLESMIDFTLSHLGTNNIIAISTEVEKETPEVQTYTIEKVEEKANGKGVVCHKVAYPYAVHFCHDVGSTRTFMVSMVGADGTKVN
AVSVCHEDTASMNPKALPFQLLNVKPGDKPICHFTLDDQIALFPSQNALAEN

>lcl|Potato_tuber|TRINITY_DN47721_c1_g2_i3 (5' partial)
APENELQKVRQPWGVGRWYNDAPENELQKVRQPWGVGRWYNDAPENELYKVRQPWGVGRWYNDAAKKDLNDNHPVTPYFFETDLHQGKQMNLQSLKNYNPAPILPRKVVDSIAFSSDK
IEEILNHFSVDKDSERAKDIKKTIKMCEDPAGNGEVKHCATSLESMIDFTLSHLGTNNIIAMSTEVEKETPEVQTYTIEKVEEKANGKGVVCHKVAYPYAVHFCHDVGSTRTFMVSMV
GADGTKVNAVSVCHEDTASMNPKALPFQLLNVKPGDKPICHFTLDDQIALFPSQNALAEN

>lcl|Potato_tuber|TRINITY_DN47721_c1_g2_i4 (5' partial)
MGSWYQAAPENELHKVHQPFGVVAWYNDAAKKDLNDNHPVTPYFFETDLHQGKKMNLQSLKNYNPAPILPRKVVDSIAFSSDKIEEILNHFSVDKDSERAKDIKKTIKMCEDPAGNGE
VKHCATSLESMIDFTLSHLGTNNIIAMSTEVEKETPEVQTYTIEKVEEKANGKGVVCHKVAYPYAVHFCHDVGSTRTFMVSMVGADGTKVNAVSVCHEDTASMNPKALPFQLLNVKPG
DKPICHFTLDDQIALFPSQNALAEN

>lcl|Potato_tuber|TRINITY_DN48305_c1_g2_i2 (3' partial)
MELHHQYYFFTFFSVIFVVSHAANLSPEVYWRVKLPNTPMPTPIKDALHISEKTAYNGDGNTKISQPYGVFAWYKAATENELHKVRQPYGVFAWYKAASENVLHKVRQPYG

**C** **LbaLycA hits of rnaSPAdes (v1.0) *de novo* transcriptome assembly (kmer 25,75):**

>lcl|Potato_tuber|NODE_10739_length_1460_cov_677.901_ID_21479
MELHHQYYFFTLFSVIFLVSHAANLSPEVYWRVKLPNTPMPTPIKDALHISENELHKVRQPYGVLAWYQAASENELHKVRQPYGVLALHQAASENELHKVRQPYGVLALHQAAPENEL
HKVRQPYGVYRWYQAASENELHKVRQPYGVYRWYQAASENELHKVRQPYGVYRWYQAAPENKLHKVRQPYGVSRWYNDAATKDLNDNHPVTPYFFETDLHQGKKMNLQSLKNYNPAPI
LPRKVVDSIAFSSDKIEEILNHFSVDKDSERAKDIKKTIKMCEDPAGNGEVKHCATSLESMIDFTLSHLGTNNIIAISTEVEKETPEVQTYTIEKVEEKANGKGVVCHKVAYPYAVHF
CHDVGSTRTFMVSMVGADGTKVNAVSVCHEDTASMNPKALPFQLLNVKPGDKPICHFTLDDQIALFPSQNALAEN*

>lcl|Potato_tuber|NODE_29201_length_620_cov_26.9761_ID_58407
MELLHQYYFFTLFSVIFVVSHAANLSPEVYWRVKLPNTPMPTPIKDALHISEKTAYNGDGNTKISQPWGVGAWYQDAPENELHKVRQPWGVGSWYQASPENKLHKVRQPWGVVAWYQA
ASENKLHKVRQPWGVGSWYQAAPENELHKVRQPYGVFRWYQAASENELHKVRQPWGVGSWYQDASENELHKVRQPWG

>lcl|Potato_tuber|NODE_26964_length_666_cov_44.9932_ID_53933
MELHHQYYFFTFFSVIFVVSHAANLSPEVYWRVKLPNTPMPTPIKDALHISEKTAYNGDGNTKISQPYGVFAWYQAASENELHKVRQPYGVDGWYQAASENELHKVRQPYGVFAWYKA
ITENELHKVRQPYGVFAWYKAATENELHKVRQPYGVFAWYKAASENELHKVRQPYGVFAWYNDAAKKDLNDNHPVTPYFFETDLHQGKKM

>lcl|Potato_tuber|NODE_40653_length_459_cov_9.7474_ID_81309
IKDALHISEKTAYNGDGNTKISQPWGVGAWYQDAPENELHKVRQPWGVGSWYQASPENKLHKVRQPWGVGAWYQAAPENELHKVRQPWGVGSWYQAAPENELHKVRQPYGVFRWYQAA
SENELHKVRQPWGVGSWYQDASENELHKVRQPWG

**D**

**Core peptides from BURP domain precursor peptide genes:**
Genome sequence (v4.03): QPWGVDSW, QPYGVGVW
Trinity transcriptome assembly: QPFGVGRW, QPFGVFGW, QPYGVFGW, QPWGVGRW, QPFGVVAW, **QPYGVFAW**
rnaSPAdes assembly: **QPYGVLAW**, **QPYGVYRW**, **QPYGVSRW**, **QPWGVGAW**, **QPWGVVAW**, **QPWGVGSW**, QPYGVFRW, **QPYGVFAW**, QPYGVDGW

**Figure S26 | Genome and transcriptome mining of lyciumins from *Solanum tuberosum*.** (**A**) Lyciumin precursor peptide derived from *Solanum tuberosum* genome (BURP domain underlined, core peptides highlighted in red). (**B**) Lyciumin precursor peptide transcripts derived from *Solanum tuberosum* tuber transcriptome (SRR5970148) *de novo* assembled with Trinity (v2.4, BURP domain underlined, core peptides highlighted in red). (**C**) Lyciumin precursor peptide transcripts derived from *Solanum tuberosum* tuber transcriptome (SRR5970148) *de novo* assembled with rnaSPAdes (v1.0, BURP domain underlined, core peptides highlighted in red). (**D**) Predicted lyciumin core peptides derived from genome mining and transcriptome mining of *Solanum tuberosum*. Bold core peptides indicate detected lyciumin chemotypes.

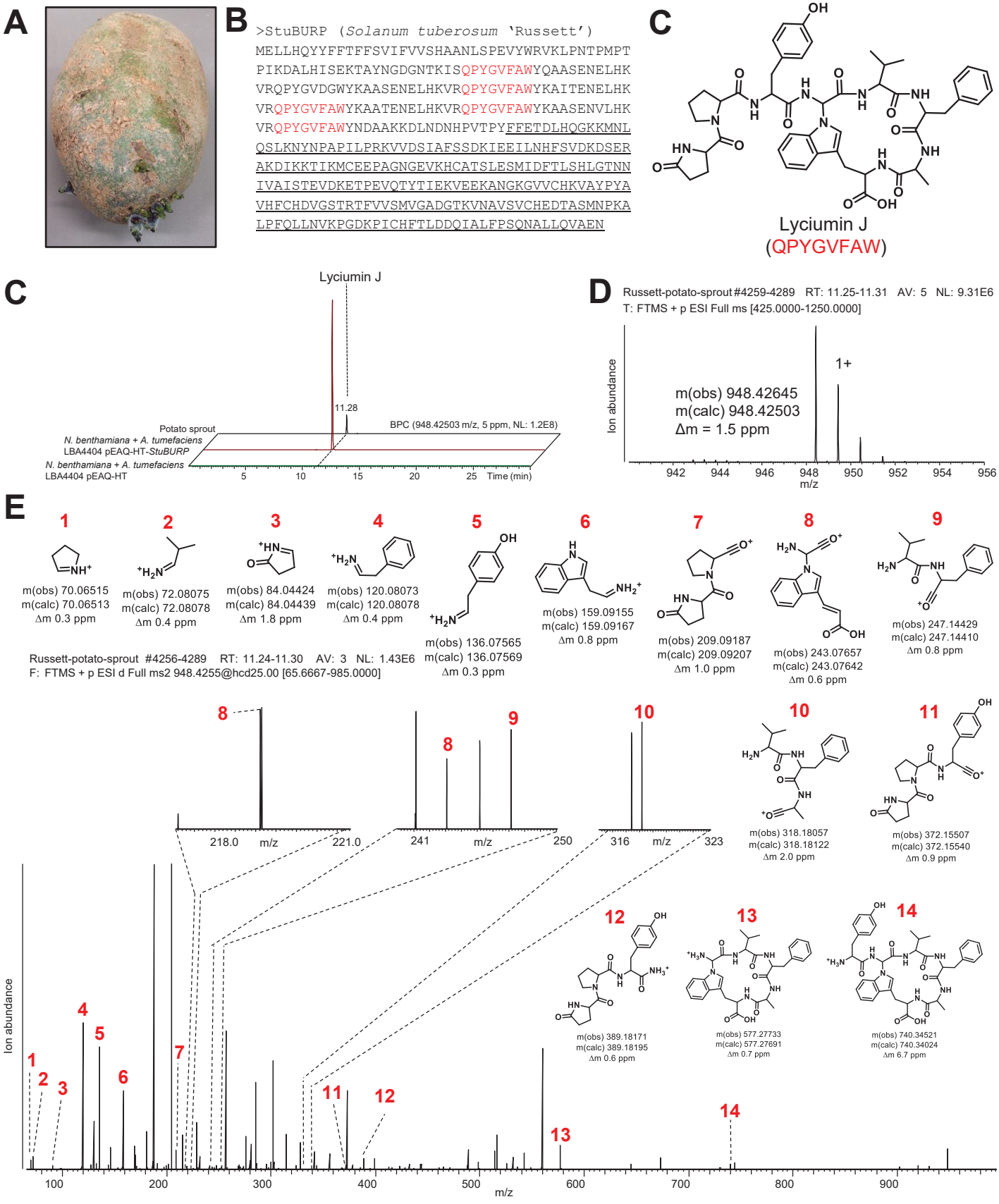



**Figure S26 | Genome and transcriptome mining of lyciumins from *Solanum tuberosum*.** (**A**) *Solanum tuberosum* 'Russett' – sprouting tuber. (**B**) *Solanum tuberosum* lyciumin J precursor peptide (BURP domain underlined, core peptide highlighted in red). (**C**) Predicted lyciumin J chemotype. (**D**) LC-MS chemotyping of predicted lyciumin J chemotype in peptide extract of *Solanum tuberosum* sprout and of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*StuBURP*. (**D**) MS analysis of predicted lyciumin J chemotype of *Solanum tuberosum* sprout. (**E**) MS/MS analysis of predicted lyciumin J chemotype of *Solanum tuberosum* sprout.

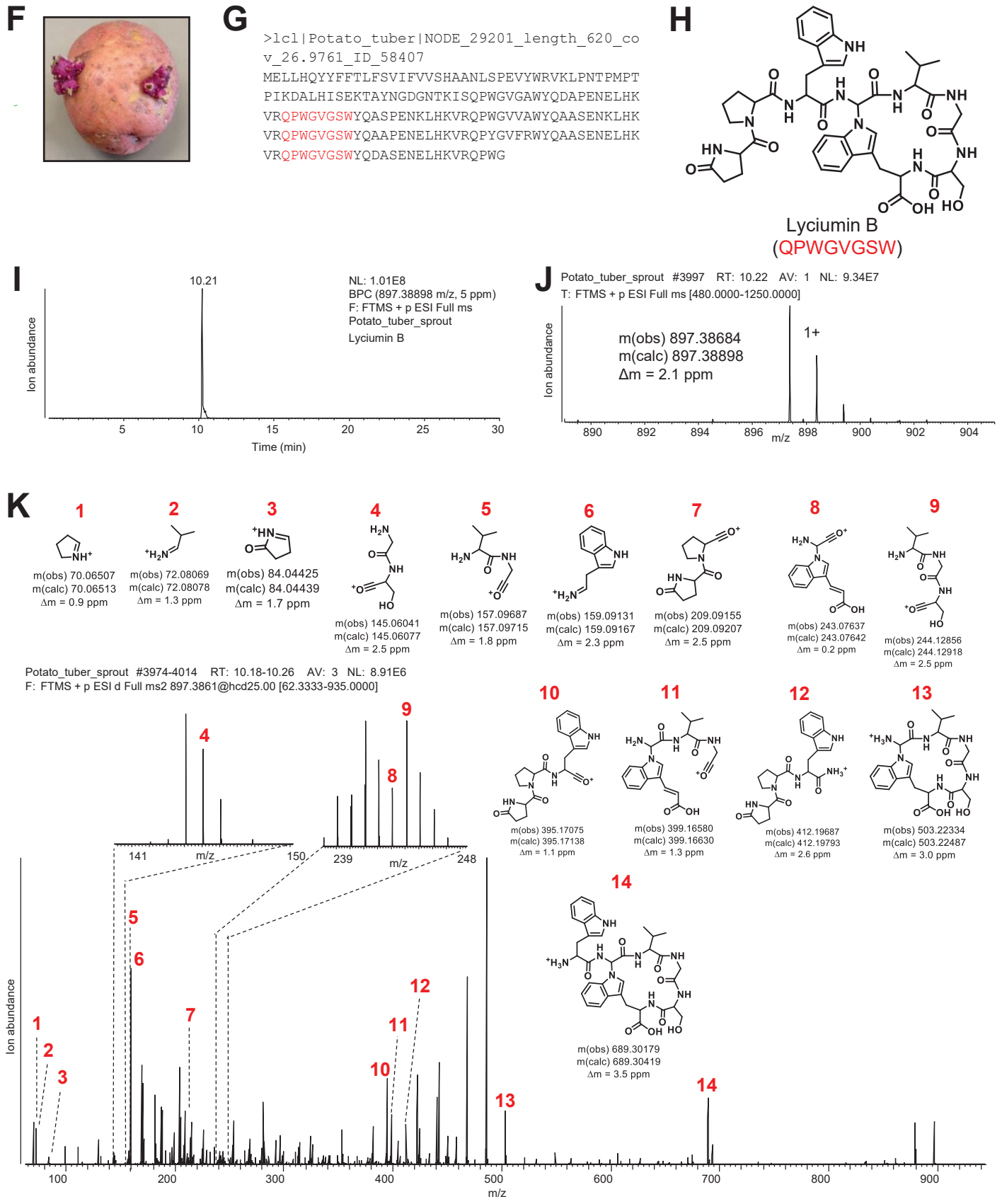



**Figure S26 | Genome and transcriptome mining of lyciumins from *Solanum tuberosum*.** (**F**) *Solanum tuberosum* 'Red potato' - sprouting tuber. (**G**) *Solanum tuberosum* lyciumin B precursor peptide transcript (BURP domain underlined, core peptides highlighted in red). (**H**) Predicted lyciumin B chemotype. (**I**) LC-MS chemotyping of predicted lyciumin B chemotype in peptide extract of *Solanum tuberosum* sprout. (**J**) MS analysis of predicted lyciumin B chemotype of *Solanum tuberosum* sprout. (**K**) MS/MS analysis of predicted lyciumin B chemotype of *Solanum tuberosum* sprout.

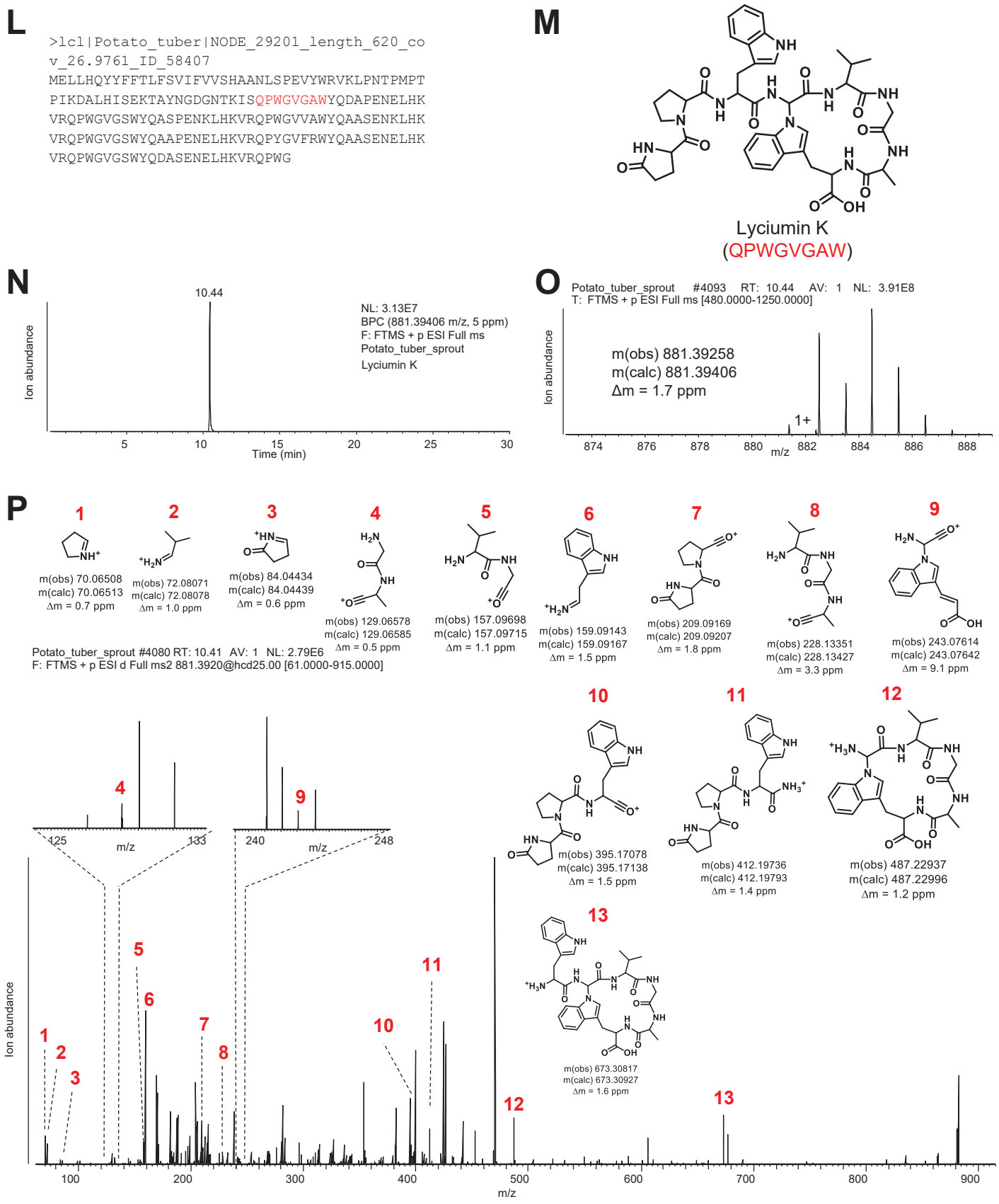
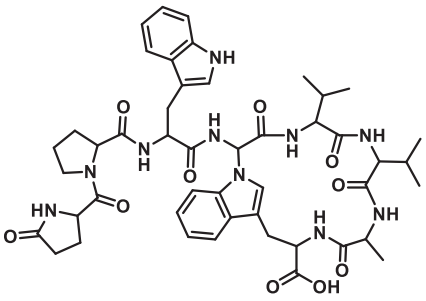



**Figure S26 | Genome and transcriptome mining of lyciumins from *Solanum tuberosum*.** (**L**) *Solanum tuberosum* lyciumin K precursor peptide transcript (BURP domain underlined, core peptides highlighted in red). (**M**) Predicted lyciumin K chemotype. (**N**) LC-MS chemotyping of predicted lyciumin K chemotype in peptide extract of *Solanum tuberosum* sprout. (**O**) MS analysis of predicted lyciumin K chemotype of *Solanum tuberosum* sprout. (**P**) MS/MS analysis of predicted lyciumin K chemotype of *Solanum tuberosum* sprout.

**Q**

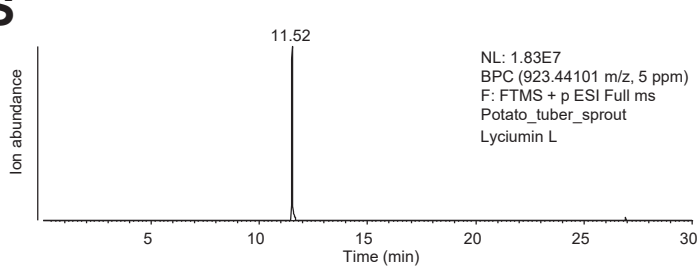
>lcl|Potato_tuber|NODE_29201_length_620_cov_26.9761_ID_58407
MELLHQYYFFTLFSVIFVVSHAANLSPEVYWRVKLPNTPMPT
PIKDALHISEKTAYNGDGNTKISQPWGVGAWYQDAPENELHK
VRQPWGVGSWYQASPENKLHKVRQPWGVVAWYQAASENKLHK
VRQPWGVGSWYQAAPENELHKVRQPYGVFRWYQAASENELHK
VRQPWGVGSWYQDASENELHKVRQPWG

**R**

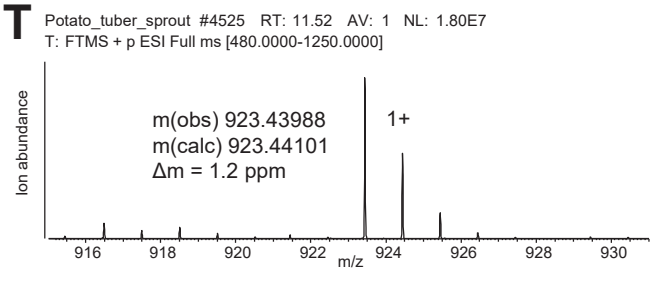


Lyciumin L
(QPWGVVAW)

**S**




**T**




**U**

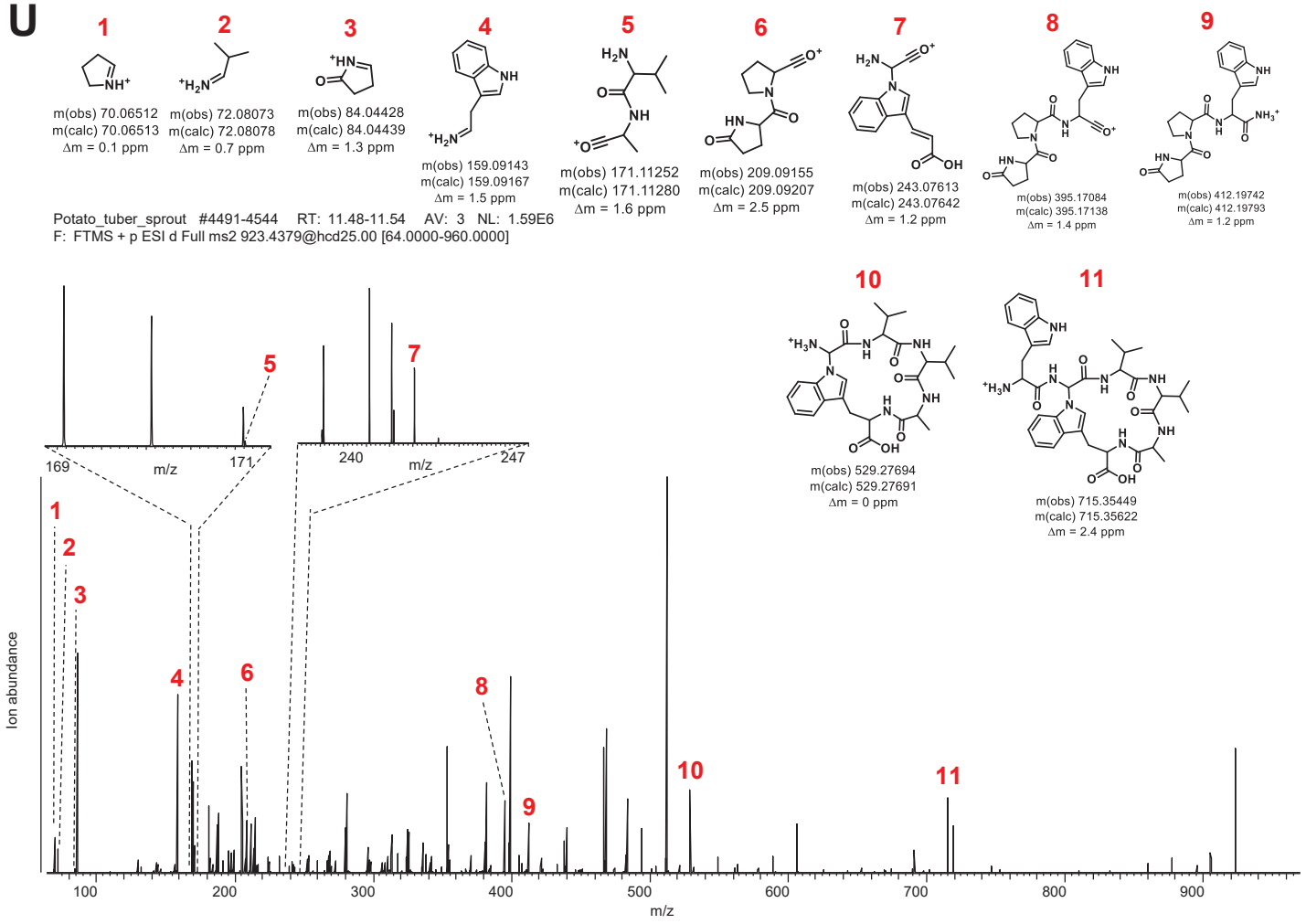
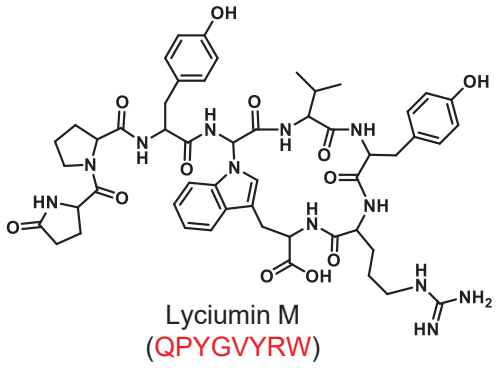



**Figure S26 | Genome and transcriptome mining of lyciumins from *Solanum tuberosum*.** (**Q**) *Solanum tuberosum* lyciumin L precursor peptide transcript (BURP domain underlined, core peptides highlighted in red). (**R**) Predicted lyciumin L chemotype. (**S**) LC-MS chemotyping of predicted lyciumin L chemotype in peptide extract of *Solanum tuberosum* sprout. (**T**) MS analysis of predicted lyciumin L chemotype of *Solanum tuberosum* sprout. (**U**) MS/MS analysis of predicted lyciumin L chemotype of *Solanum tuberosum* sprout.

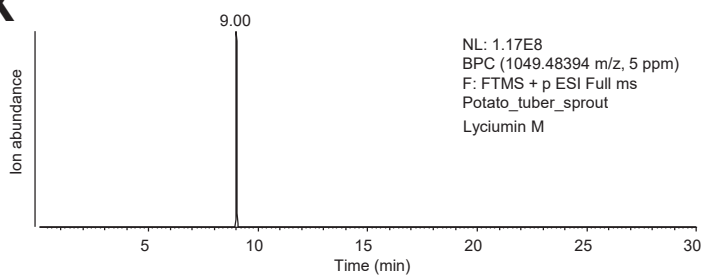
**V**

>lcl|Potato_tuber|NODE_10739_length_1460_cov_677.901_ID_21479
MELHHQYYFFTLFSVIFLVSHAANLSPEVYWRVKLPNTPMPTPIKDALHISEN
ELHKVRQPYGVLAWYQAASENELHKVRQPYGVLALHQAASENELHKVRQPYGV
LALHQAAPENELHKVRQPYGVYRWYQAASENELHKVRQPYGVYRWYQAASENE
LHKVRQPYGVYRWYQAAPENKLHKVRQPYGVSRWYNDAATKDLNDNHPVTPYF
FETDLHQGKKMNLQSLKNYNPAPILPRKVVDSIAFSSDKIEEILNHFSVDKDS
ERAKDIKKTIKMCEDPAGNGEVKHCATSLESMIDFTLSHLGTNNIIAISTEVE
KETPEVQTYTIEKVEEKANGKGVVCHKVAYPYAVHFCHDVGSTRTFMVSMVGA
DGTKVNAVSVCHEDTASMNPKALPFQLLNVKPGDKPICHFTLDDQIALFPSQN
ALAEN

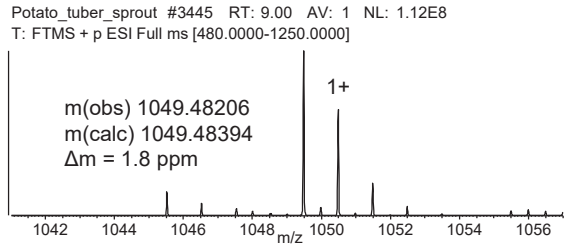
**W**




**X**




**Y**




**Z**

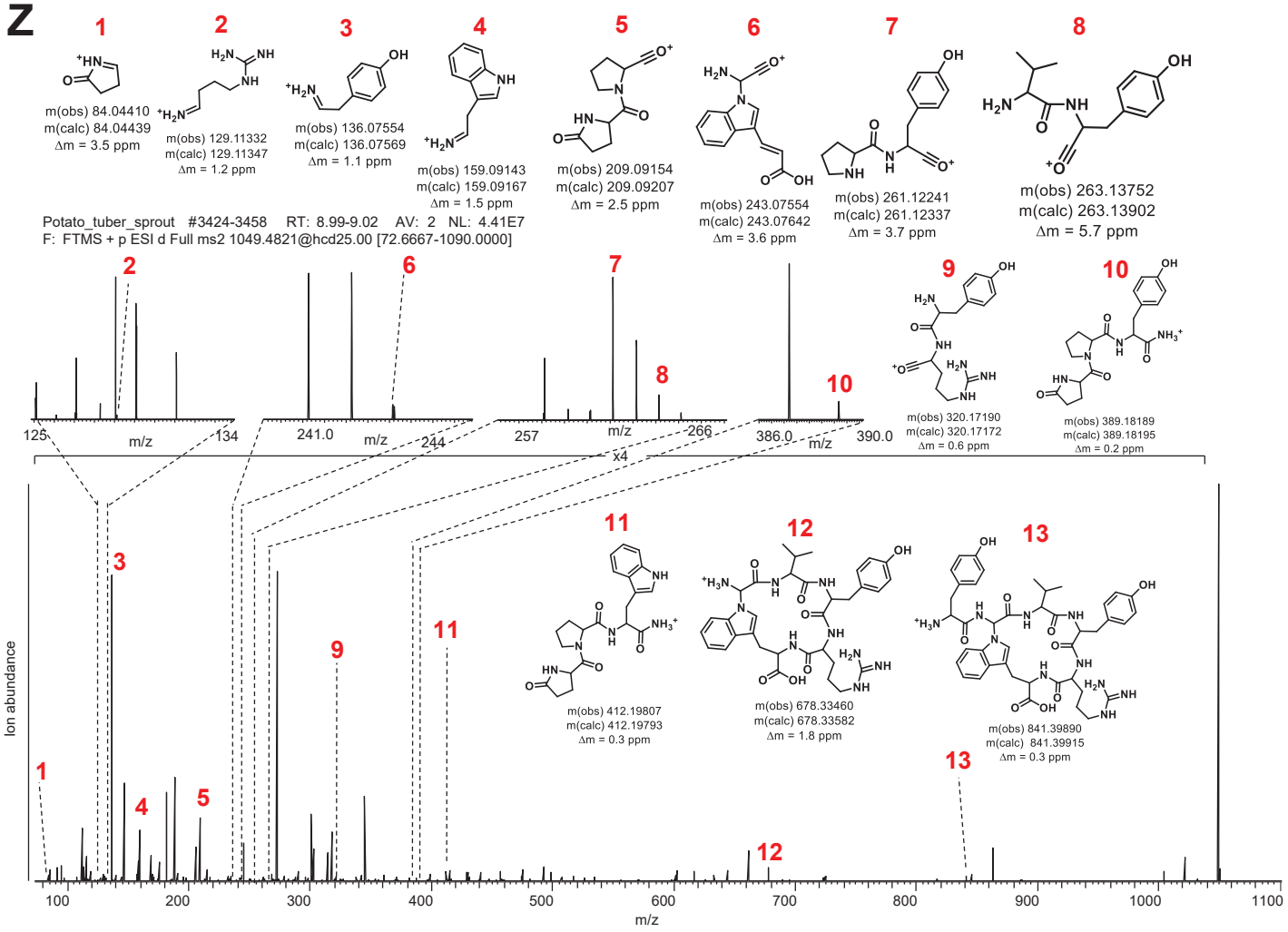



**Figure S26 | Genome and transcriptome mining of lyciumins from *Solanum tuberosum*.** (**V**) *Solanum tuberosum* lyciumin M precursor peptide transcript (BURP domain underlined, core peptides highlighted in red). (**W**) Predicted lyciumin M chemotype. (**X**) LC-MS chemotyping of predicted lyciumin M chemotype in peptide extract of *Solanum tuberosum* sprout. (**Y**) MS analysis of predicted lyciumin M chemotype of *Solanum tuberosum* sprout. (**Z**) MS/MS analysis of predicted lyciumin M chemotype of *Solanum tuberosum* sprout.

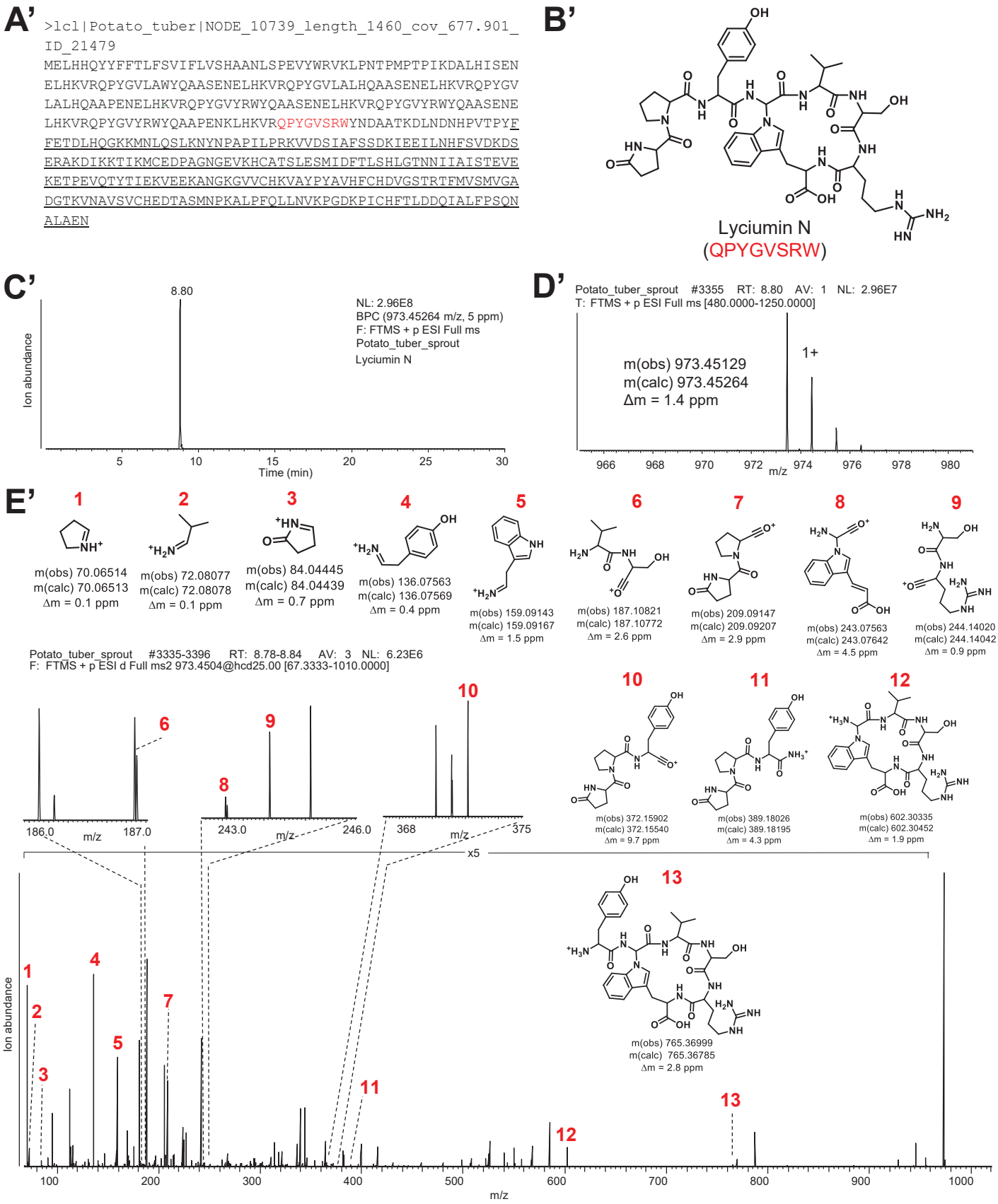


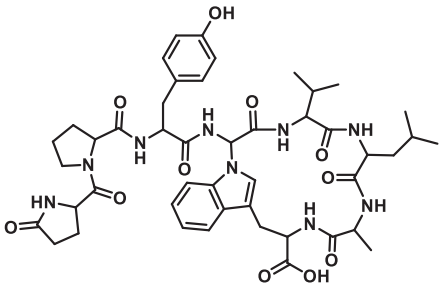

**Figure S26 | Genome and transcriptome mining of lyciumins from *Solanum tuberosum*.** (**A'**) *Solanum tuberosum* lyciumin N precursor peptide transcript (BURP domain underlined, core peptides highlighted in red). (**B'**) Predicted lyciumin N chemotype. (**C'**) LC-MS chemotyping of predicted lyciumin N chemotype in peptide extract of *Solanum tuberosum* sprout. (**D'**) MS analysis of predicted lyciumin N chemotype of *Solanum tuberosum* sprout. (**E'**) MS/MS analysis of predicted lyciumin N chemotype of *Solanum tuberosum* sprout.

**F’**

>lcl|Potato_tuber|NODE_10739_length_1460_cov_677.901_ID_21479

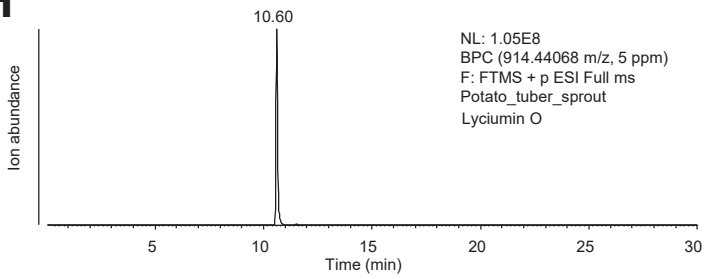
MELHHQYYFFTLFSVIFLVSHAANLSPEVYWRVKLPNTPMPTPIKDALHISEN
ELHKVRQPYGVLAWYQAASENELHKVRQPYGVLALHQAASENELHKVRQPYGV
LALHQAAPENELHKVRQPYGVYRWYQAASENELHKVRQPYGVYRWYQAASENE
LHKVRQPYGVYRWYQAAPENKLHKVRQPYGVSRWYNDAATKDLNDNHPVTPYF
FETDLHQGKKMNLQSLKNYNPAPILPRKVVDSIAFSSDKIEEILNHFSVDKDS
ERAKDIKKTIKMCEDPAGNGEVKHCATSLESMIDFTLSHLGTNNIIAISTEVE
KETPEVQTYTIEKVEEKANGKGVVCHKVAYPYAVHFCHDVGSTRTFMVSMVGA
DGTKVNAVSVCHEDTASMNPKALPFQLLNVKPGDKPICHFTLDDQIALFPSQN
ALAEN

**G’**

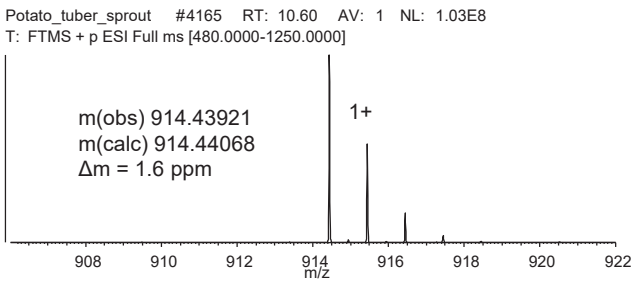


Lyciumin O
(QPYGVLAW)

**H’**




**I’**




**J’**

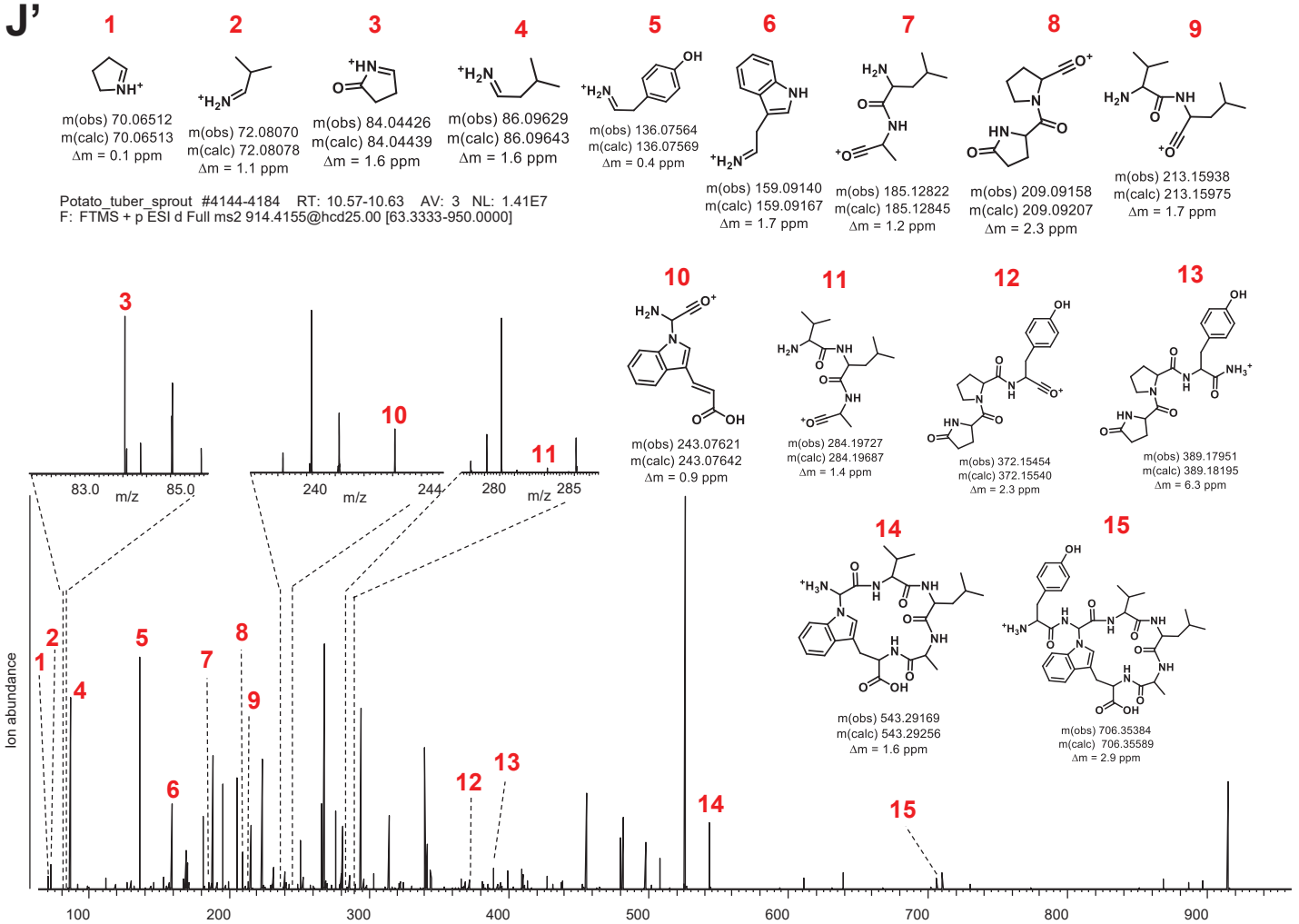



**Figure S26 | Genome and transcriptome mining of lyciumins from *Solanum tuberosum*.** (**F’**) *Solanum tuberosum* lyciumin O precursor peptide transcript (BURP domain underlined, core peptides highlighted in red). (**G’**) Predicted lyciumin O chemotype. (**H’**) LC-MS chemotyping of predicted lyciumin O chemotype in peptide extract of *Solanum tuberosum* sprout. (**I’**) MS analysis of predicted lyciumin O chemotype of *Solanum tuberosum* sprout. (**J’**) MS/MS analysis of predicted lyciumin O chemotype of *Solanum tuberosum* sprout.

**K’**

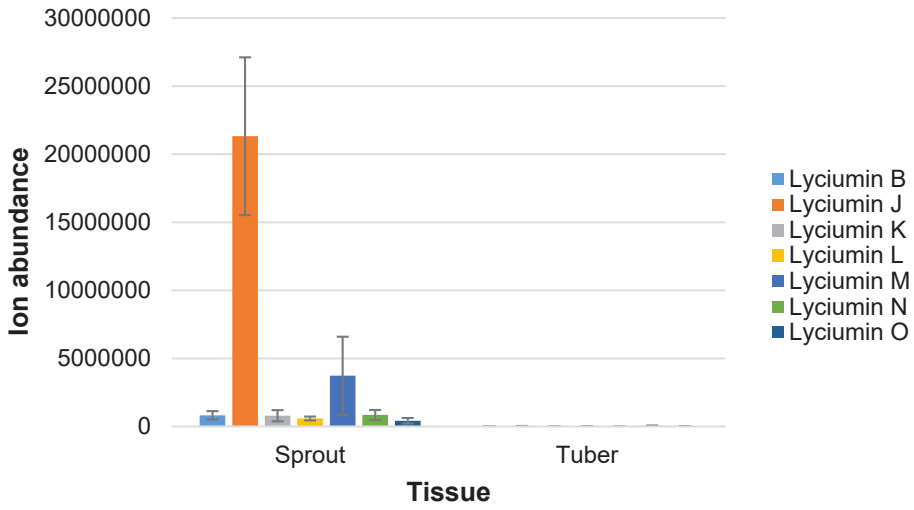



**Figure S26 | Genome and transcriptome mining of lyciumins from *Solanum tuberosum*.** (**K’**) Chemotyping of lyciumins in sprout and tuber tissues of *Solanum tuberosum* (‘red potato’).

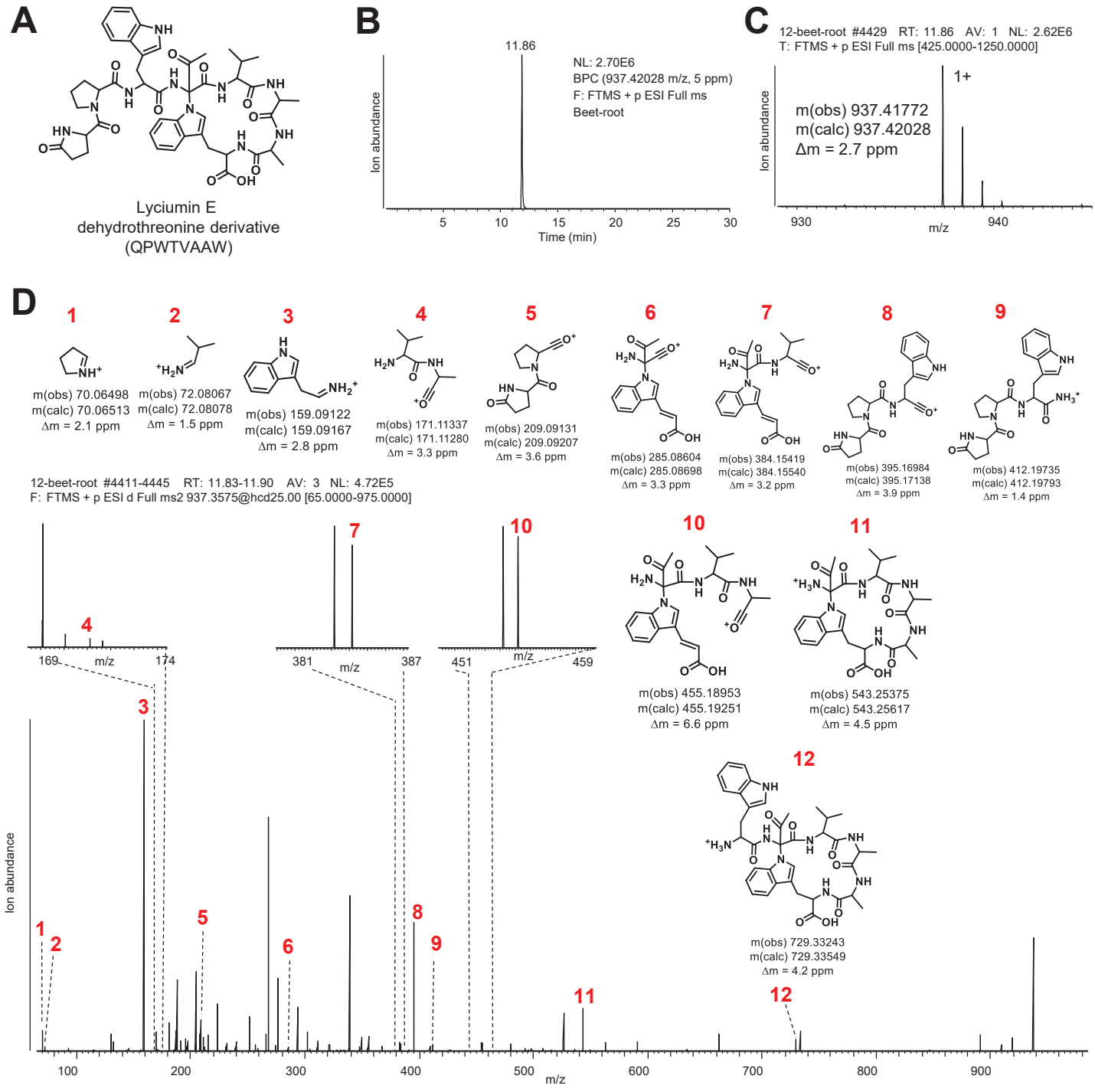



**Figure S27 | Detection of candidate dehydrothreonine-lyciumin derivatives in Amaranthaceae.** (**A**) Predicted *Beta vulgaris* lyciumin E dehydrothreonine derivative chemotype (core peptide). (**B**) LC-MS chemotyping of predicted lyciumin E dehydrothreonine derivative chemotype in peptide extract of *Beta vulgaris* root. (**C**) MS analysis of predicted lyciumin E dehydrothreonine derivative chemotype of *Beta vulgaris*. (**D**) MS/MS analysis of predicted lyciumin E dehydrothreonine derivative chemotype of *Beta vulgaris*.

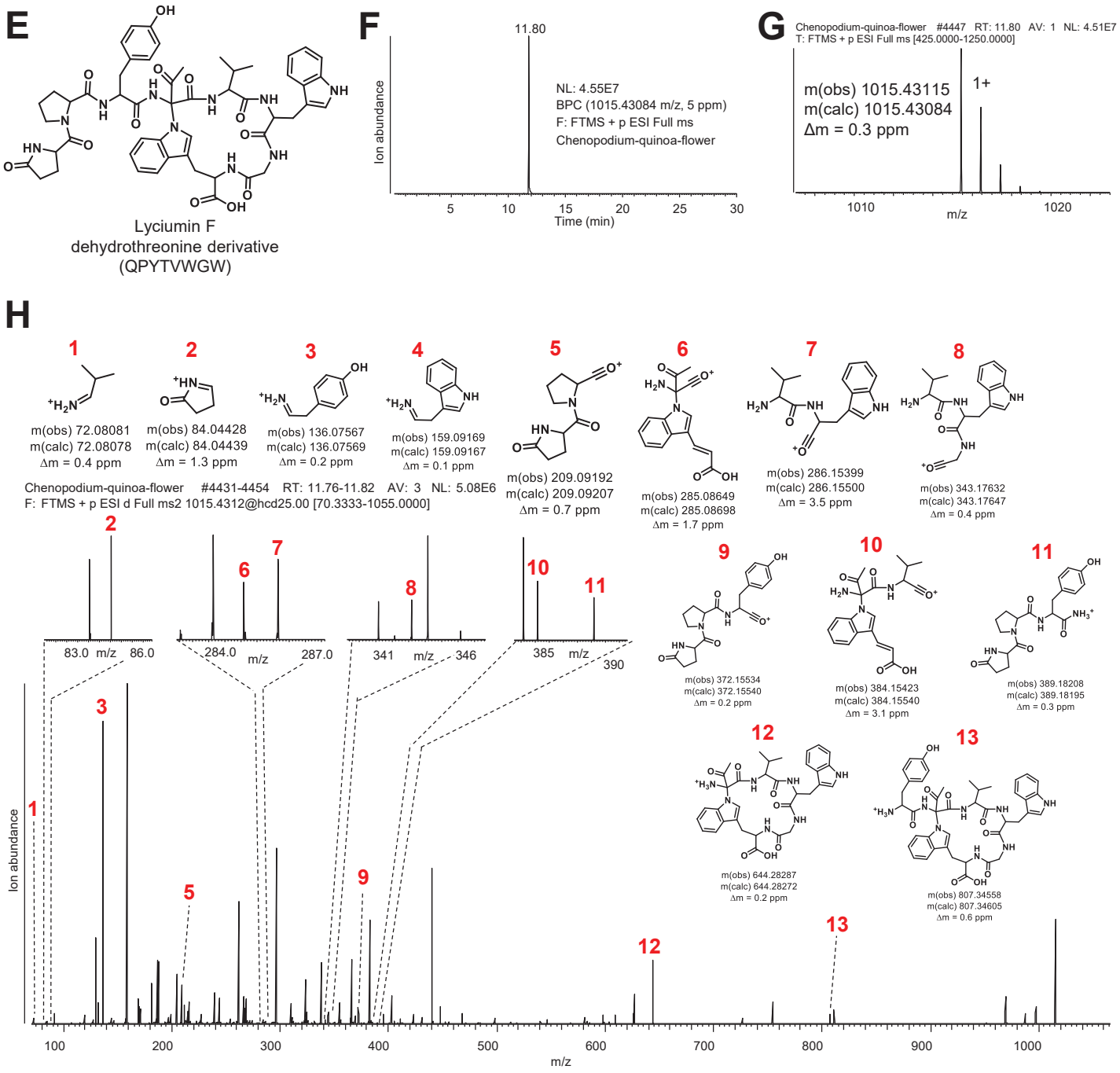



**Figure S27 | Detection of candidate dehydrothreonine-lyciumin derivatives in Amaranthaceae.** (**E**) Predicted *Chenopodium quinoa* lyciumin F dehydrothreonine derivative chemotype (core peptide). (**F**) LC-MS chemotyping of predicted lyciumin F dehydrothreonine derivative chemotype in peptide extract of *Chenopodium quinoa* flower. (**G**) MS analysis of predicted lyciumin F dehydrothreonine derivative chemotype of *Chenopodium quinoa*. (**H**) MS/MS analysis of predicted lyciumin F dehydrothreonine derivative chemotype of *Chenopodium quinoa*.

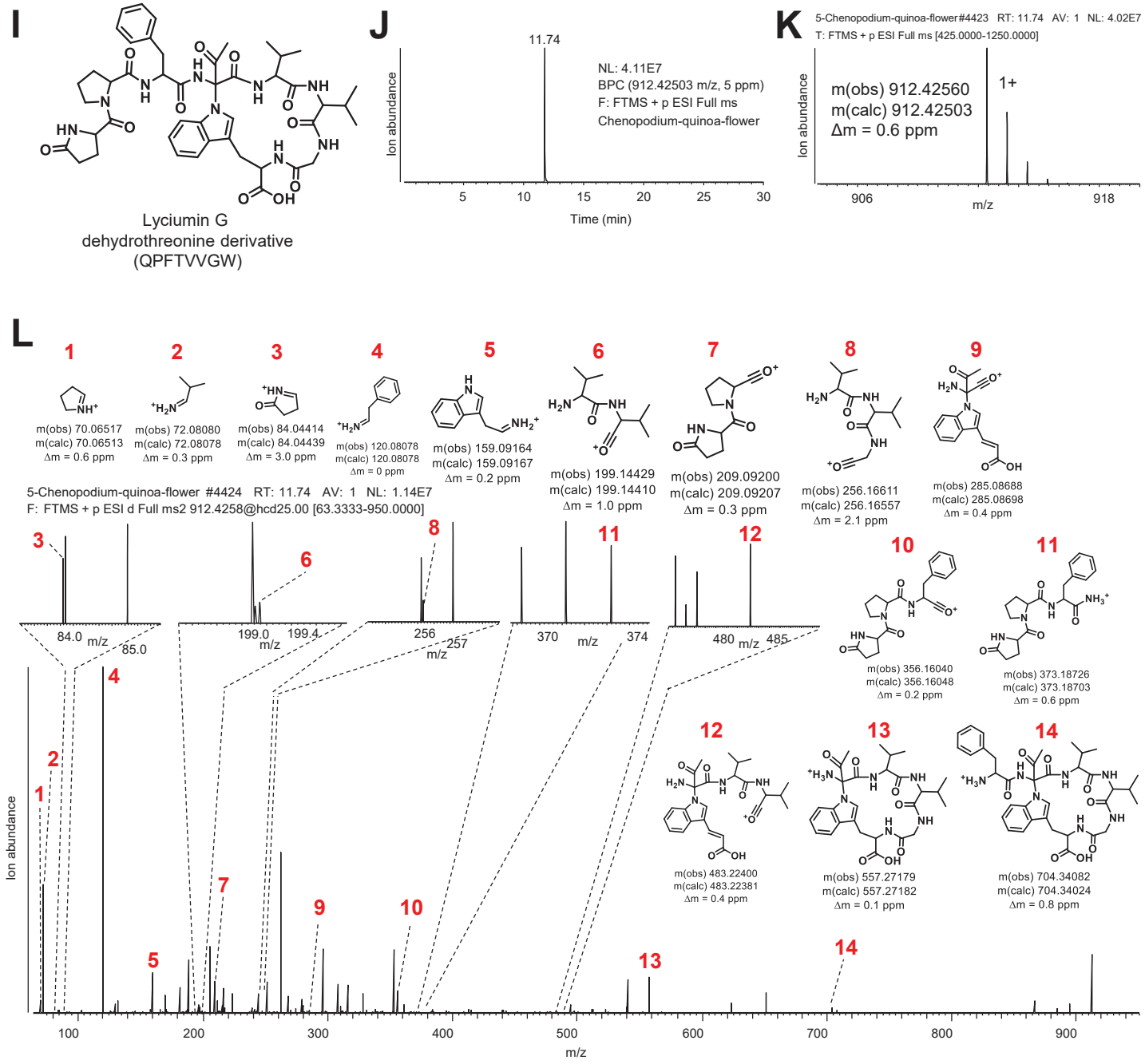



**Figure S27 | Detection of candidate dehydrothreonine-lyciumin derivatives in Amaranthaceae.** (**I**) Predicted *Chenopodium quinoa* lyciumin G dehydrothreonine derivative chemotype (core peptide). (**J**) LC-MS chemotyping of predicted lyciumin G dehydrothreonine derivative chemotype in peptide extract of *Chenopodium quinoa* flower. (**K**) MS analysis of predicted lyciumin G dehydrothreonine derivative chemotype of *Chenopodium quinoa*. (**L**) MS/MS analysis of predicted lyciumin G dehydrothreonine derivative chemotype of *Chenopodium quinoa*.

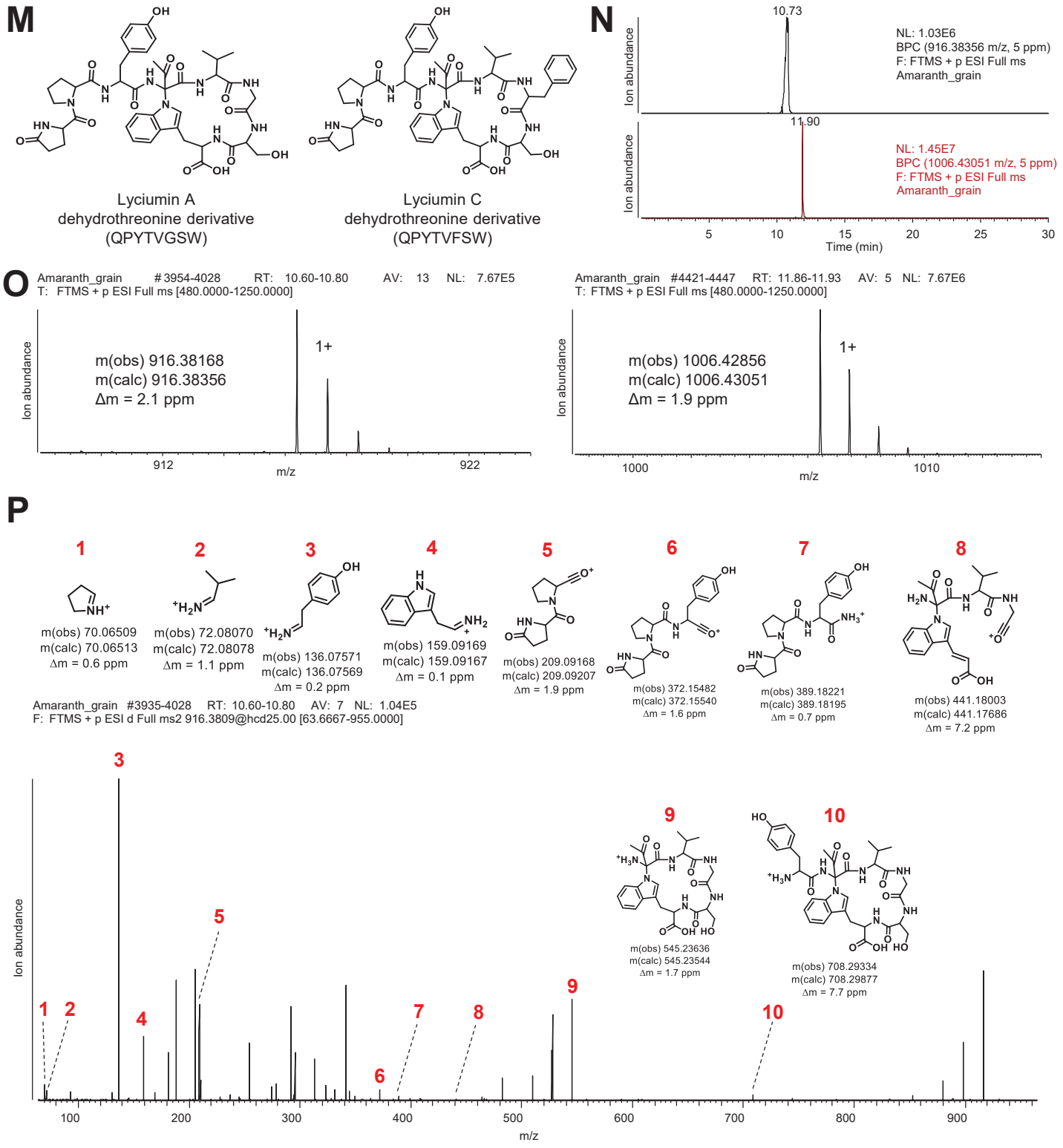



**Figure S27 | Detection of candidate dehydrothreonine-lyciumin derivatives in Amaranthaceae.** (**M**) Predicted *Amaranthus hypochondriacus* lyciumin A and C dehydrothreonine derivatives chemotypes (core peptide). (**N**) LC-MS chemotyping of predicted lyciumin A and C dehydrothreonine derivative chemotypes in peptide extract of *Amaranthus hypochondriacus* seeds. (**O**) MS analysis of predicted lyciumin A and C dehydrothreonine derivative chemotypes of *Amaranthus hypochondriacus*. (**P**) MS/MS analysis of predicted lyciumin A and C dehydrothreonine derivative chemotypes of *Amaranthus hypochondriacus*.

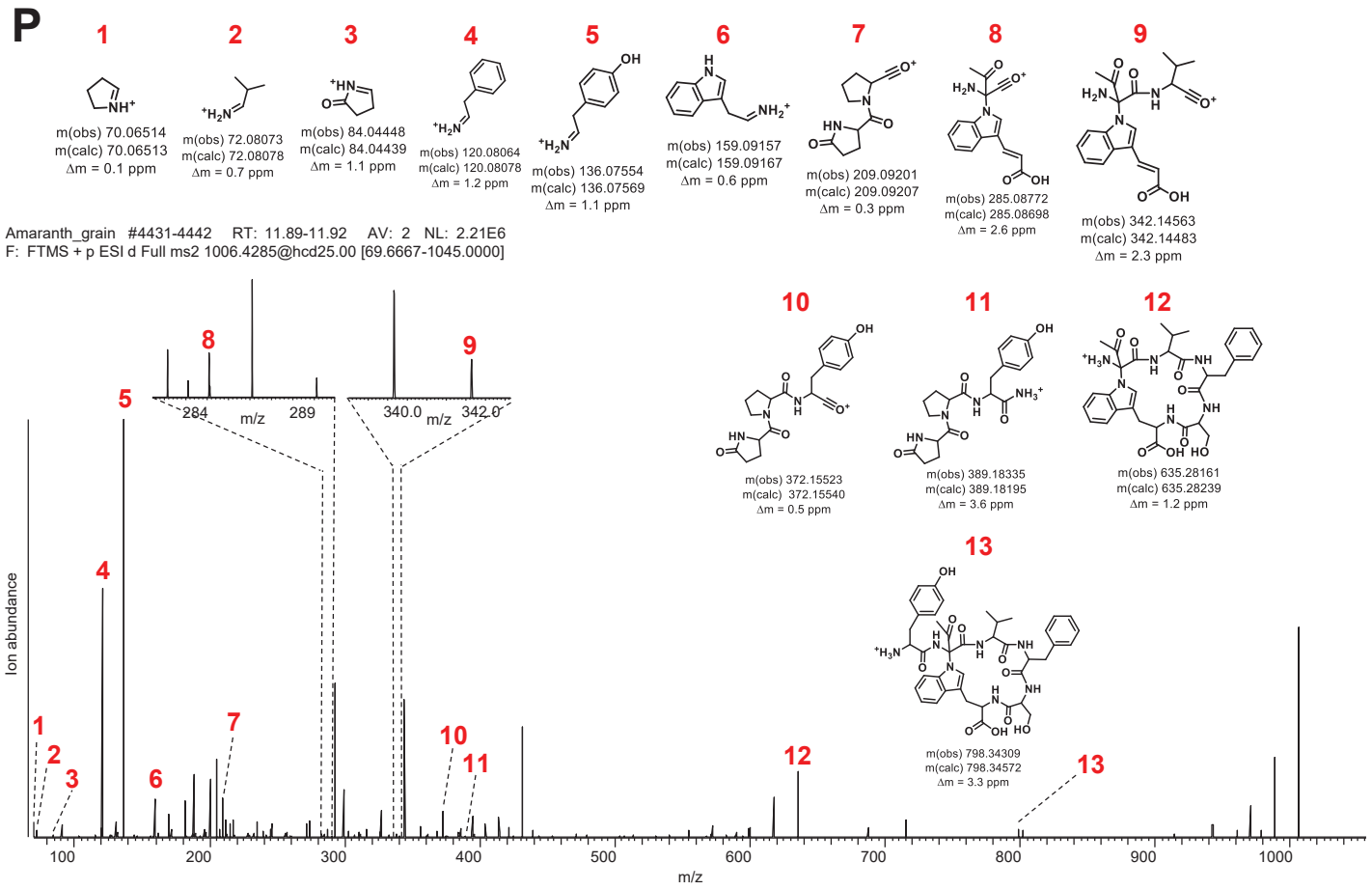



**Figure S27 | Detection of candidate dehydrothreonine-lyciumin derivatives in Amaranthaceae.** (**P**) MS/MS analysis of predicted lyciumin A and C dehydrothreonine derivative chemotypes of *Amaranthus hypochondriacus*.

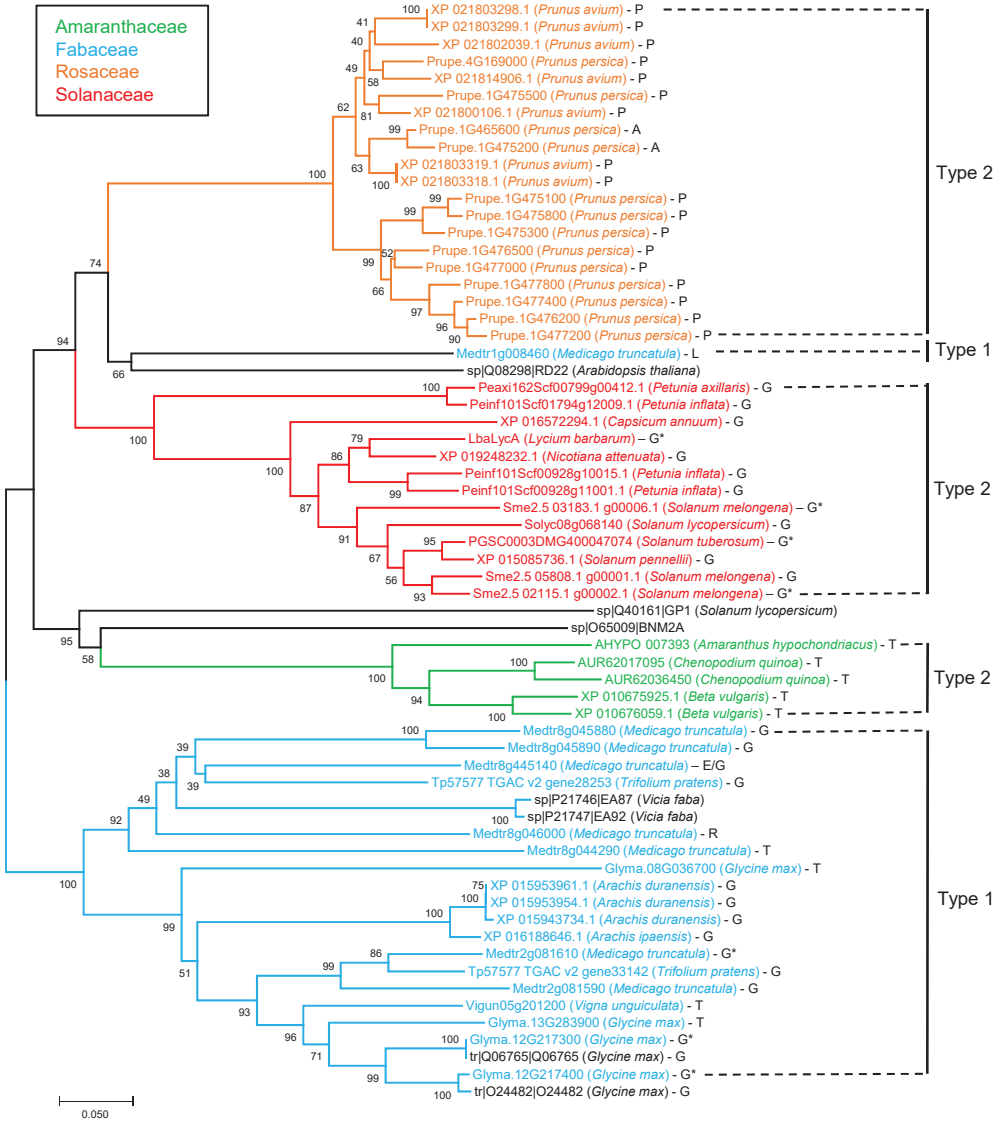



**Figure S28 | Phylogenetic analysis of predicted and characterized lyciumin precursors from analyzed plant genomes.** A neighbor-joining phylogenetic tree of the BURP-domain sequences of lyciumin precursors predicted from plant genomes and founding members of BURP domain protein family generated with 2000 bootstrap generations using the p-distance method. The scale measures evolutionary distances in substitutions per amino acid. Precursors with characterized chemotypes are noted with an asterick. Predicted cyclization sites at the fourth core pepitde position of precursor peptides are noted after species name as capital letter (Abbreviations: A – alanine, G – glycine, L – leucine, P – proline, R – arginine, T – threonine).

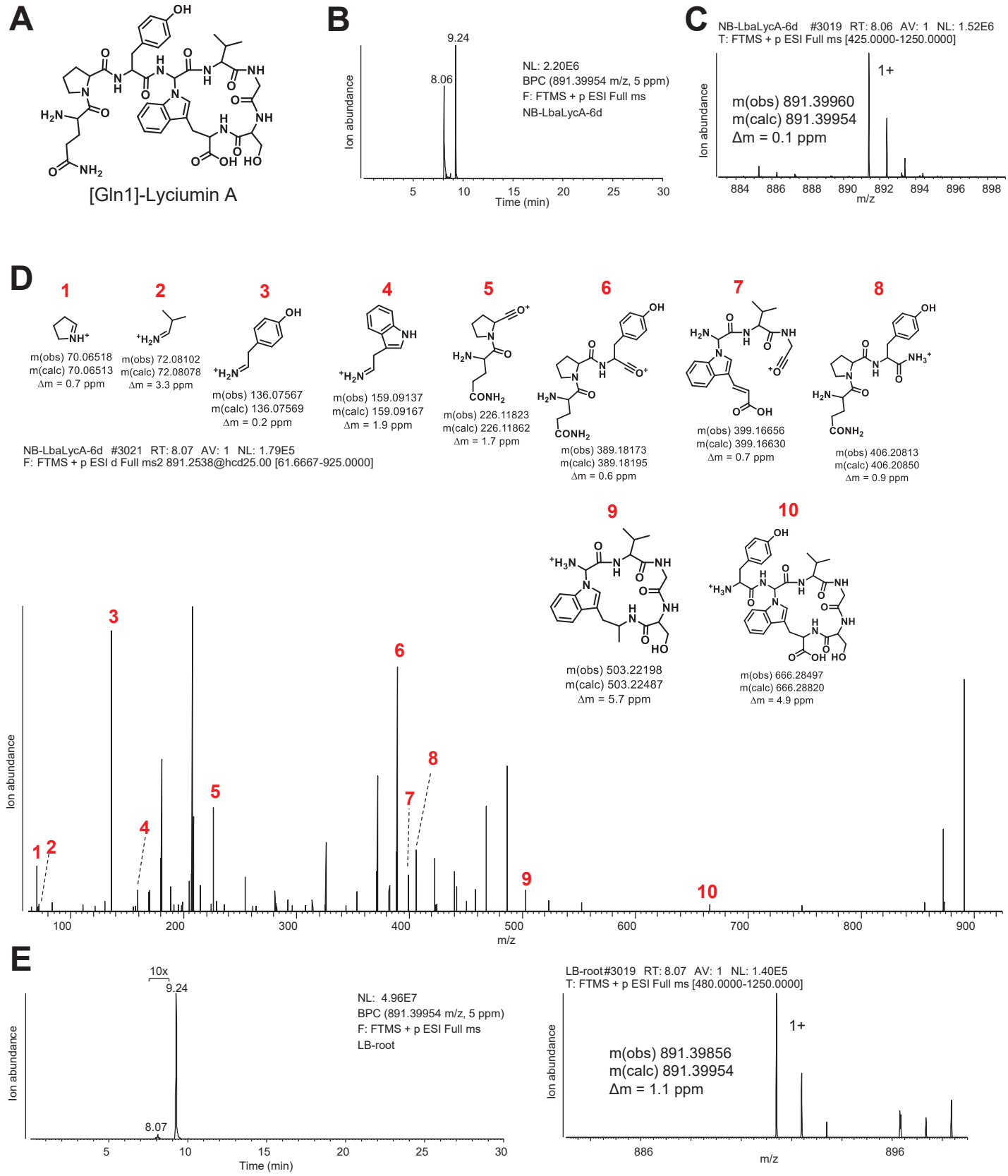



**Figure S29 | Detection of [Gln1]-lyciumins after heterologous expression of *LbaLycA* in *Nicotiana benthamiana*.** (**A**) [Gln1]-lyciumin A chemotype. (**B**) LC-MS chemotyping of [Gln1]-lyciumin A chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**C**) MS analysis of [Gln1]-lyciumin A chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**D**) MS/MS analysis [Gln1]-lyciumin A chemotype in peptide extract of *N. benthamiana* leaves six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**E**) LC-MS chemotyping of [Gln1]-lyciumin A chemotype in peptide extract of *Lycium barbarum* root.

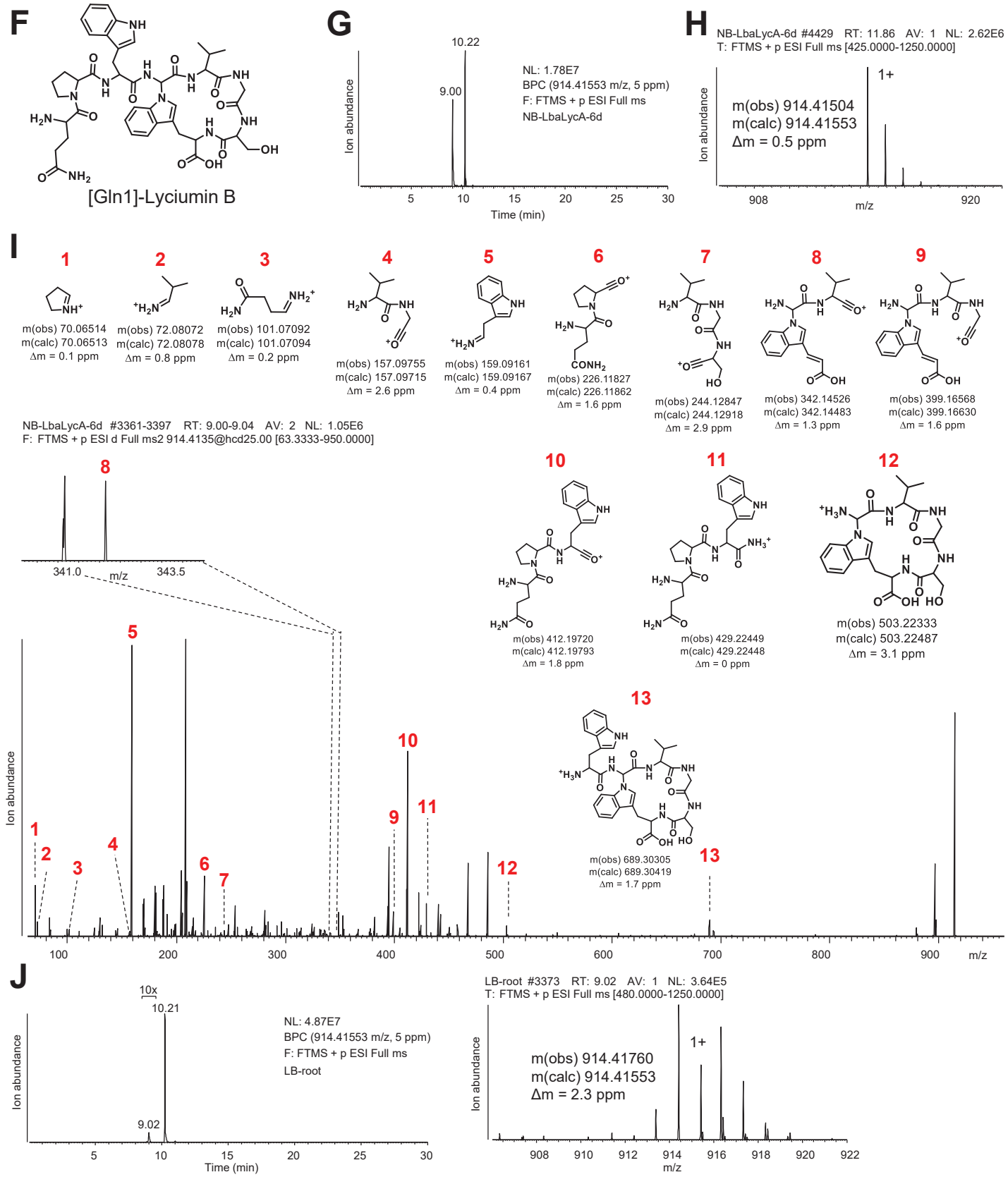



**Figure S29 | Detection of [Gln1]-lyciumins after heterologous expression of *LbaLycA* in *Nicotiana benthamiana*.** (**F**) [Gln1]-lyciumin B chemotype. (**G**) LC-MS chemotyping of [Gln1]-lyciumin B chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**H**) MS analysis of [Gln1]-lyciumin B chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**I**) MS/MS analysis [Gln1]-lyciumin B chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**J**) LC-MS chemotyping of [Gln1]-lyciumin B chemotype in peptide extract of *Lycium barbarum* root.

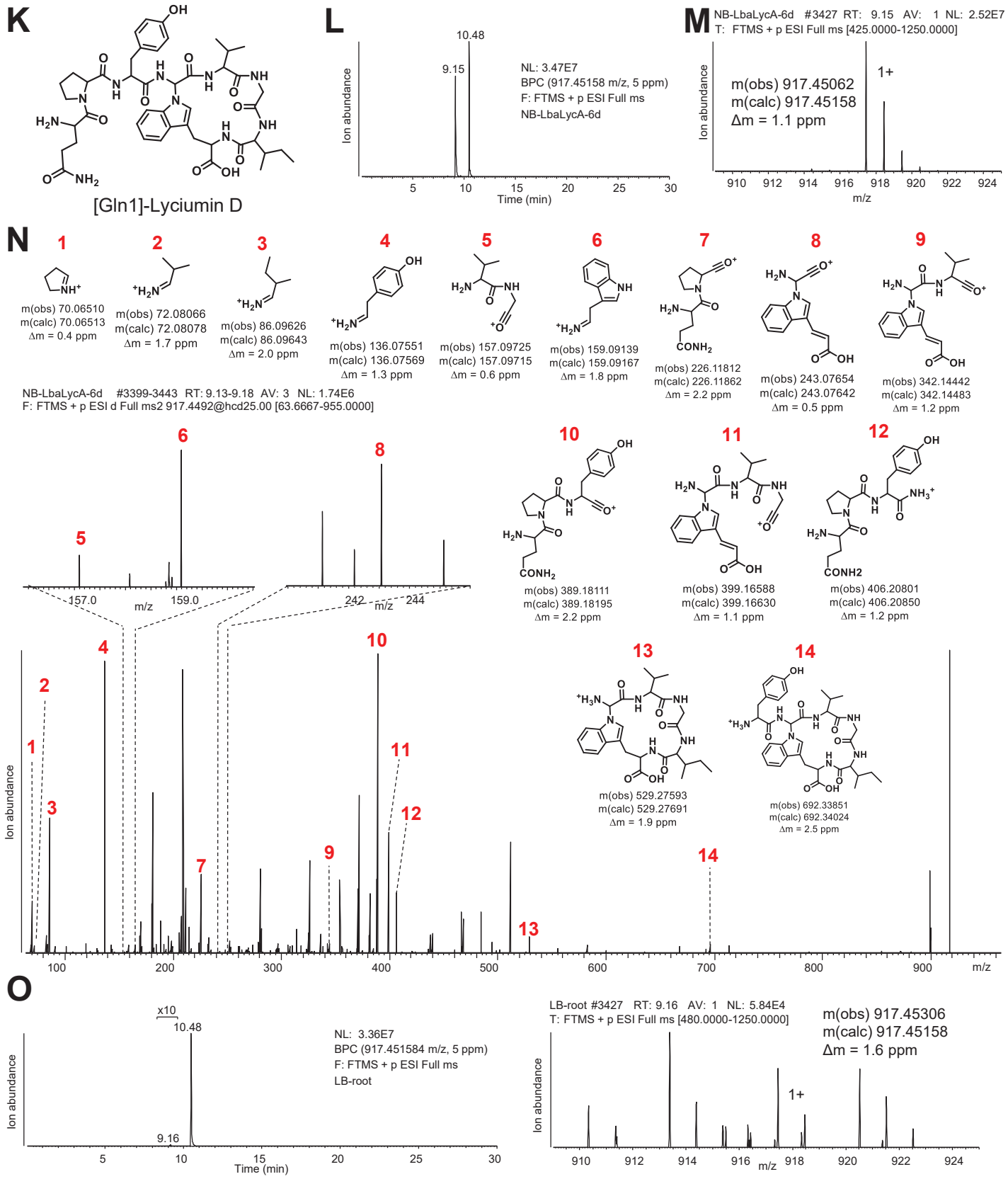



**Figure S29 | Detection of [Gln1]-lyciumins after heterologous expression of *LbaLycA* in *Nicotiana benthamiana*.** (**K**) [Gln1]-lyciumin D chemotype. (**L**) LC-MS chemotyping of [Gln1]-lyciumin D chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**M**) MS analysis of [Gln1]-lyciumin D chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**N**) MS/MS analysis [Gln1]-lyciumin D chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**O**) LC-MS chemotyping of [Gln1]-lyciumin D chemotype in peptide extract of *Lycium barbarum* root.

**A**

>LbaQC
MPLLNPRFLVISLIVLLSITVFREAEASYRVYKVKVVNEFPHDPQAYTQGLLYAENNTLFESTGLYGRSSVRKVALLDGKVERLHEMESSYFGEGLTLLGE
RLFQLTWLLDTGFIYDRYNFSKFKKFTHHMQDGWGLATDGKVLFGSDGTSTLYKIDPKTMKVIRKQVVKSQGHEVRYLNELEYVKAEVWANVYVTDCIARI
SPKDGTVIGWILLQSLREELISRGYKDFEVLNGIAWDRDGDRIFVTGKLWPKLFEIKLLPLTPNDPLAGEINNLCIPKTSFLLEI

**B**

| Gene | Size [aa] | RSEM [TPM] | Transcript rank (Total transcript #) | Similarity/Identity [%/%] to precursor co-localized *Chenopodium quinoa* QC (AUR62017096-RA) |
|---|---|---|---|---|
| *LbaQC* | 287 | 42 | 3365 (100355) | 72/61 |
| *LbaLycA* | 543 | 13397 | 13 (100355) | - |

**Figure S30 | Characterization of *Lycium barbarum* glutamine cyclotransferase (LbaQC).** (**A**) LbaQC sequence with predicted secretory pathway signaling peptide underlined (SignalP v4.1). (**B**) Bioinformatic analysis of candidate lyciumin-glutamine cyclotransferase LbaQC from root transcriptome of *Lycium barbarum*.

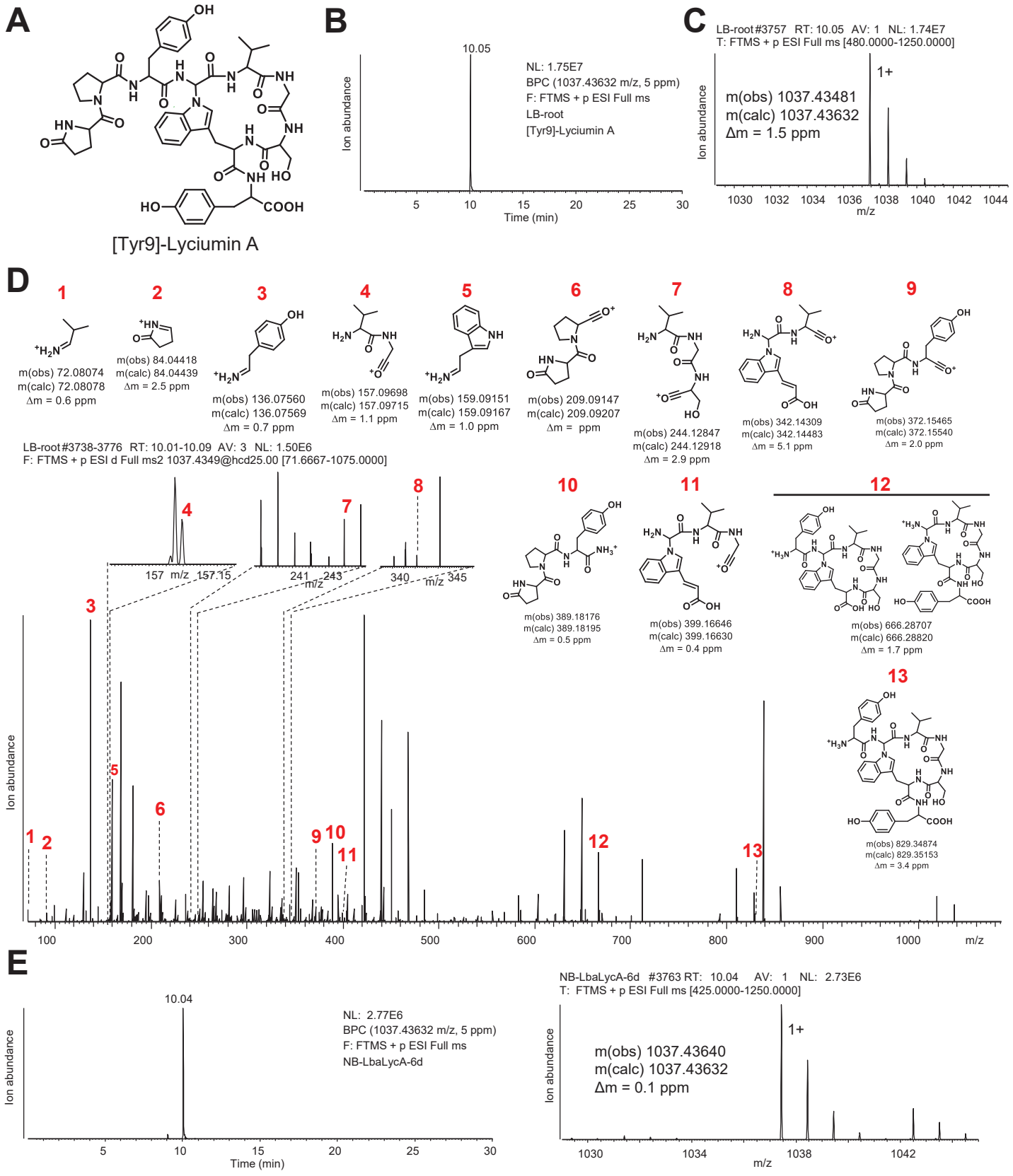



**Figure S31 | Detection of C-terminally extended lyciumins after *LbaLycA* heterologous expression in *Nicotiana benthamiana*.** (**A**) [Tyr9]-lyciumin A chemotype. (**B**) LC-MS chemotyping of [Tyr9]-lyciumin A chemotype in peptide extract of *Lycium barbarum* root. (**C**) MS analysis of [Tyr9]-lyciumin A chemotype in peptide extract of *Lycium barbarum* root. (**D**) MS/MS analysis [Tyr9]-lyciumin A chemotype in peptide extract of *Lycium barbarum* root. (**E**) LC-MS chemotyping of [Tyr9]-lyciumin A chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*.

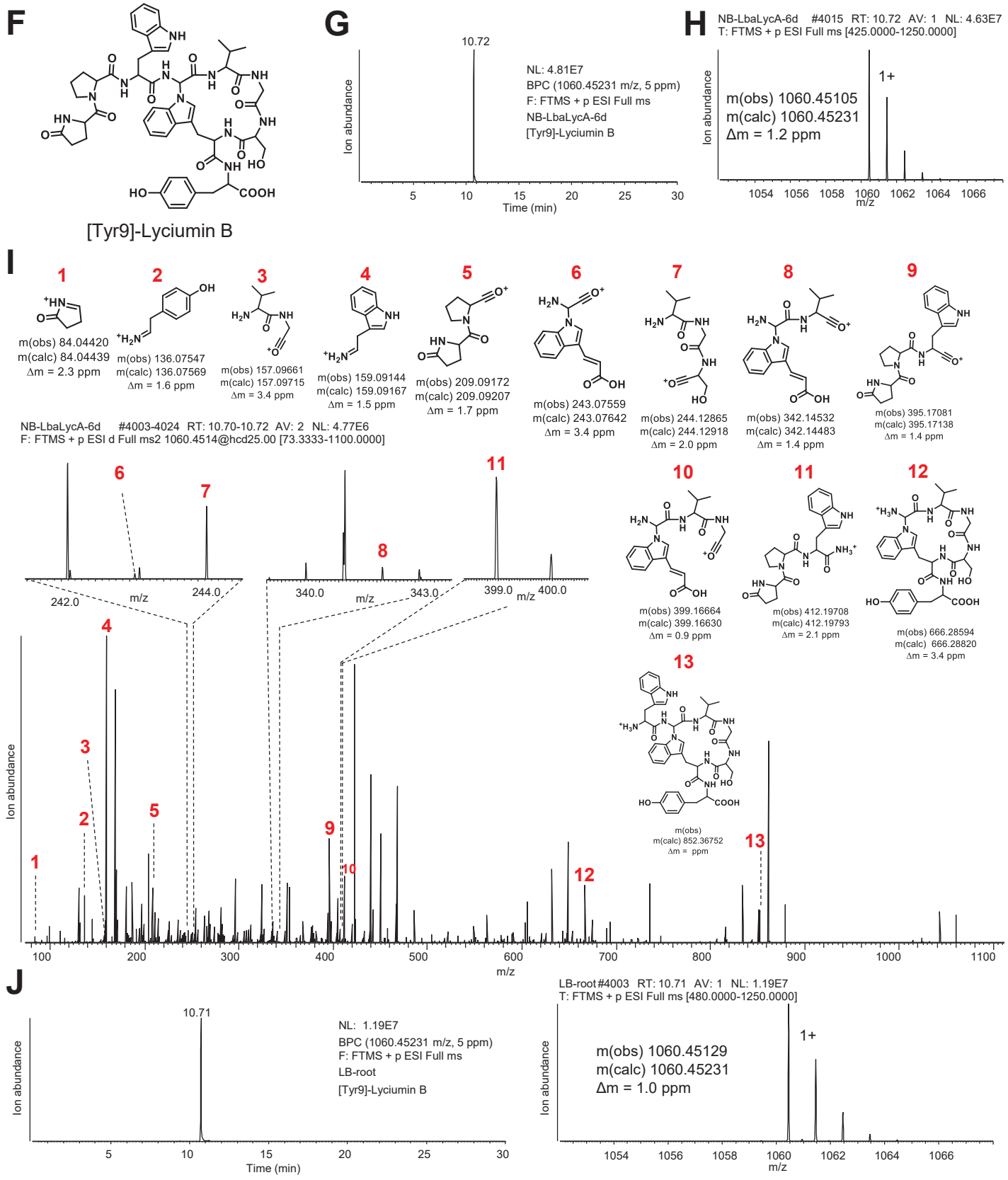



**Figure S31 | Detection of C-terminally extended lyciumins after *LbaLycA* heterologous expression in *Nicotiana benthamiana*.** (**F**) [Tyr9]-lyciumin B chemotype. (**G**) LC-MS chemotyping of [Tyr9]-lyciumin B chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**H**) MS analysis of [Tyr9]-lyciumin B chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**I**) MS/MS analysis [Tyr9]-lyciumin B chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**J**) LC-MS chemotyping of [Tyr9]-lyciumin B chemotype in peptide extract of *Lycium barbarum* root.

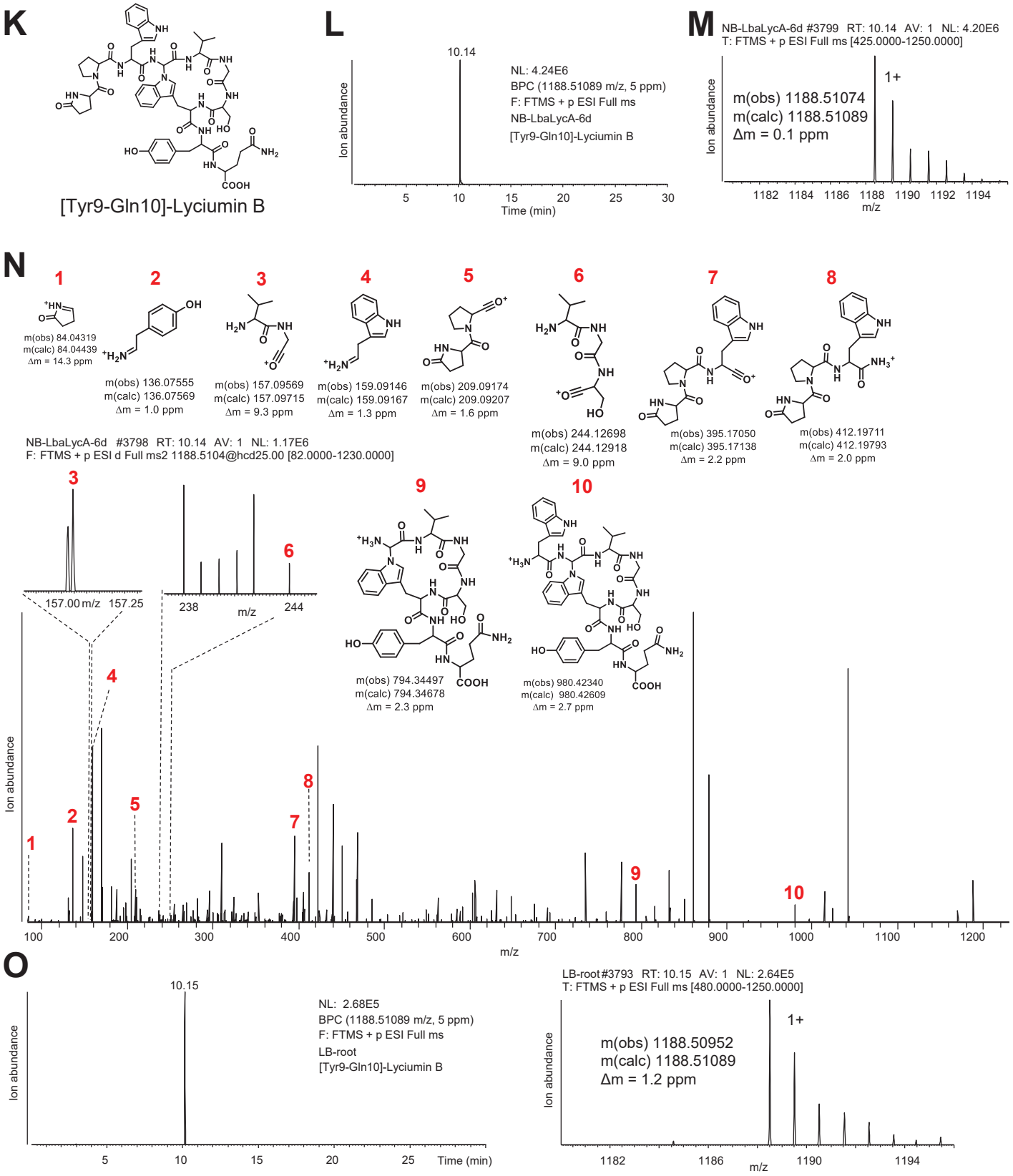



**Figure S31 | Detection of C-terminally extended lyciumins after *LbaLycA* heterologous expression in *Nicotiana benthamiana*.** (**K**) [Tyr9-Gln10]-lyciumin B chemotype. (**L**) LC-MS chemotyping of [Tyr9-Gln10]-lyciumin B chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**M**) MS analysis of [Tyr9-Gln10]-lyciumin B chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**N**) MS/MS analysis [Tyr9-Gln10]-lyciumin B chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**O**) LC-MS chemotyping of [Tyr9-Gln10]-lyciumin B chemotype in peptide extract of *Lycium barbarum* root.

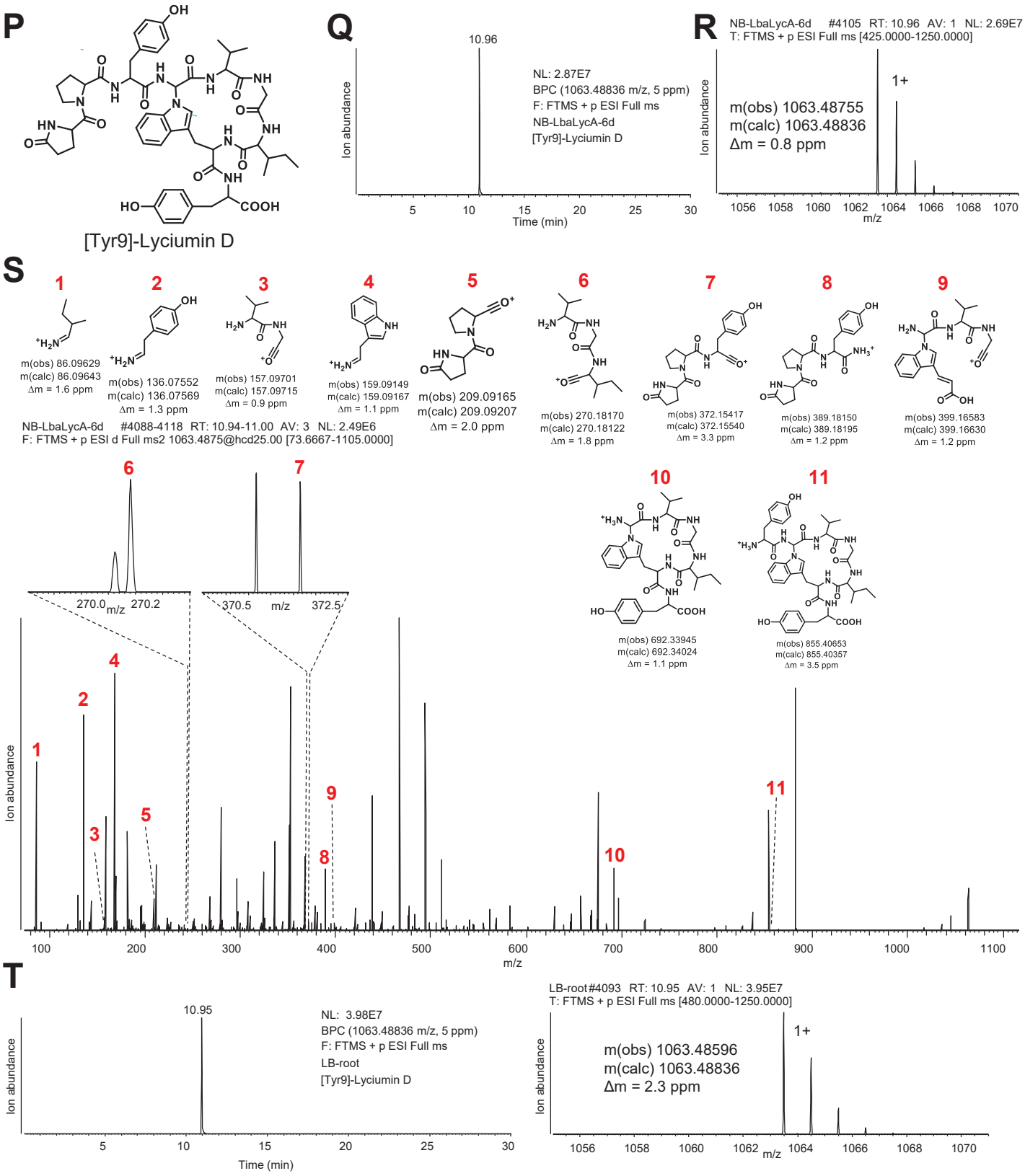



**Figure S31 | Detection of C-terminally extended lyciumins after *LbaLycA* heterologous expression in *Nicotiana benthamiana*.** (**P**) [Tyr9]-lyciumin D chemotype. (**Q**) LC-MS chemotyping of [Tyr9]-lyciumin D chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**R**) MS analysis of [Tyr9]-lyciumin D chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**S**) MS/MS analysis [Tyr9]-lyciumin D chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**T**) LC-MS chemotyping of [Tyr9]-lyciumin D chemotype in peptide extract of *Lycium barbarum* root.

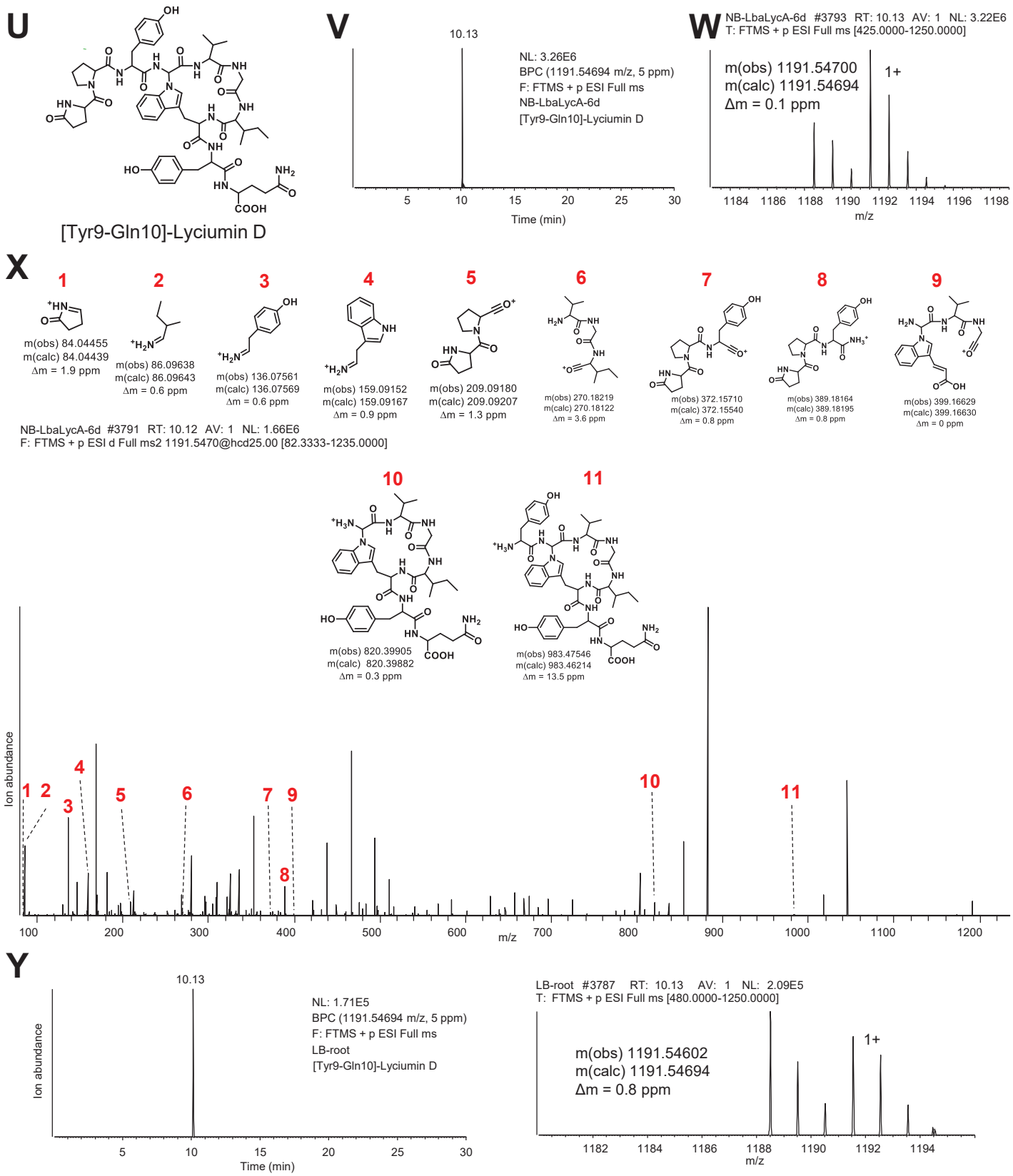



**Figure S31 | Detection of C-terminally extended lyciumins after *LbaLycA* heterologous expression in *Nicotiana benthamiana*.** (**U**) [Tyr9-Gln10]-lyciumin D chemotype. (**V**) LC-MS chemotyping of [Tyr9-Gln10]-lyciumin D chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**W**) MS analysis of [Tyr9-Gln10]-lyciumin D chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**X**) MS/MS analysis [Tyr9-Gln10]-lyciumin D chemotype in peptide extract of *N. benthamiana* six days after infiltration with *A. tumefaciens* LBA4404 pEAQ-HT-*LbaLycA*. (**Y**) LC-MS chemotyping of [Tyr9-Gln10]-lyciumin D chemotype in peptide extract of *Lycium barbarum* root.

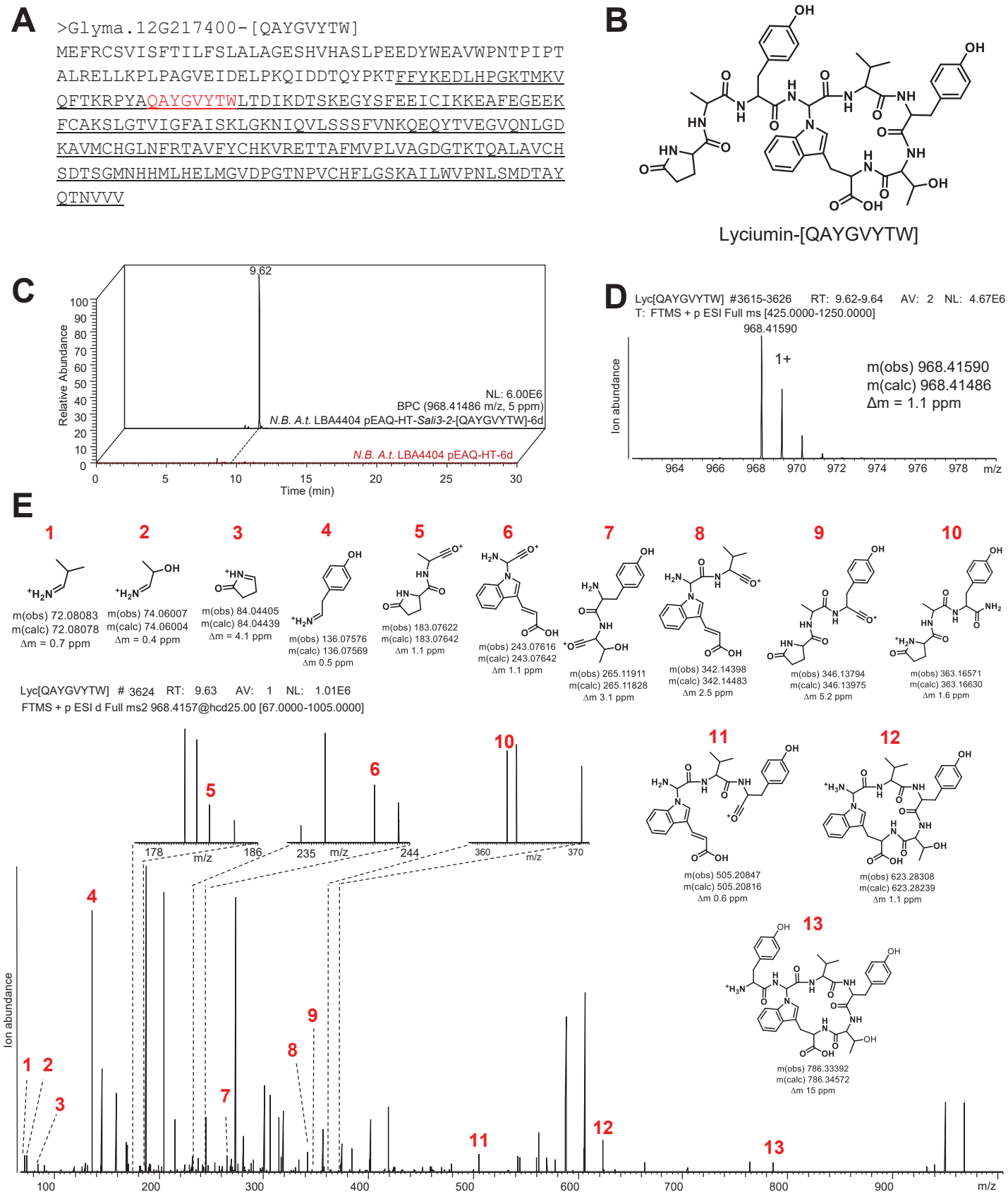



**Figure S32 | Heterologous expression of *Sali3-2*-[QAYGVYTW] in *Nicotiana benthamiana*.** (**A**) Sali3-2-[QAYGVYTW] lyciumin precursor peptide (BURP domain underlined, core peptide highlighted in red). (**B**) Predicted [QAYGVYTW]-lyciumin chemotype. (**C**) LC-MS chemotyping of predicted [QAYGVYTW]-lyciumin in peptide extract of *N. benthamiana* leaves infiltrated with *A. tumefaciens* LBA4404 pEAQ-HT-*Sali3-2*-[QAYGVYTW] for six days. (**D**) MS analysis of predicted [QAYGVYTW]-lyciumin chemotype. (**E**) MS/MS analysis of predicted [QAYGVYTW]-lyciumin chemotype.

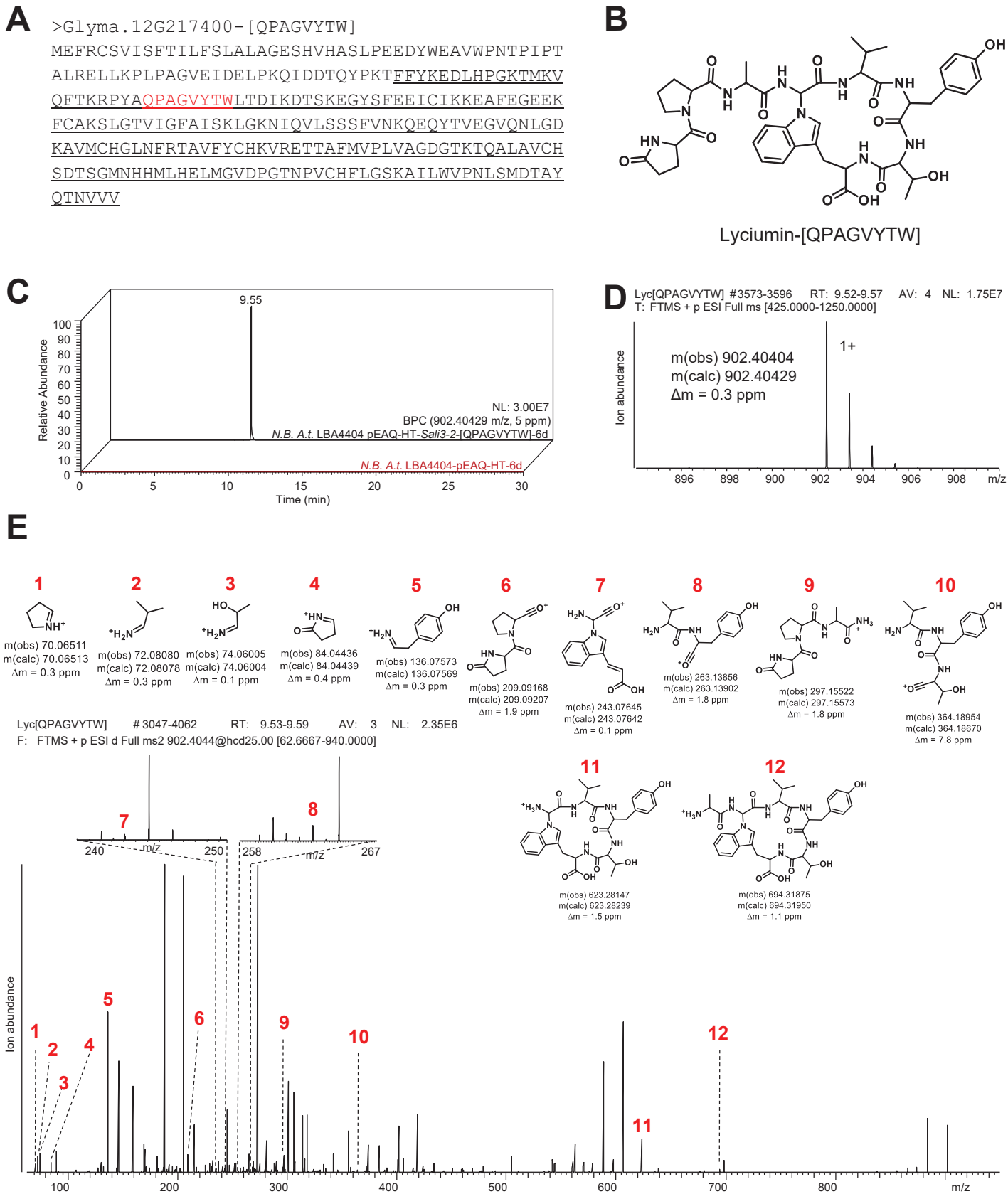
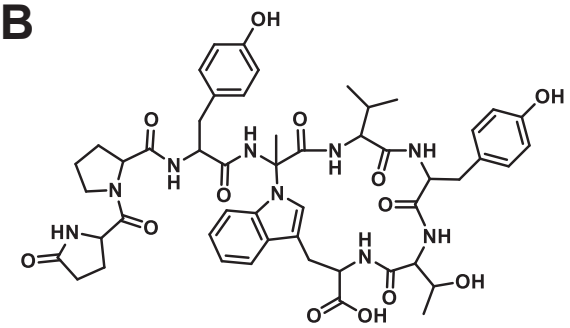



**Figure S33 | Heterologous expression of *Sali3-2*-[QPAGVYTW] in *Nicotiana benthamiana*.** (**A**) Sali3-2-[QPAGVYTW] lyciumin precursor peptide (BURP domain underlined, core peptide highlighted in red). (**B**) Predicted [QPAGVYTW]-lyciumin chemotype. (**C**) LC-MS chemotyping of predicted [QPAGVYTW]-lyciumin in peptide extract of *N. benthamiana* leaves infiltrated with *A. tumefaciens* LBA4404 pEAQ-HT-*Sali3-2*-[QPAGVYTW] for six days. (**D**) MS analysis of predicted [QPAGVYTW]-lyciumin chemotype. (**E**) MS/MS analysis of predicted [QPAGVYTW]-lyciumin chemotype.

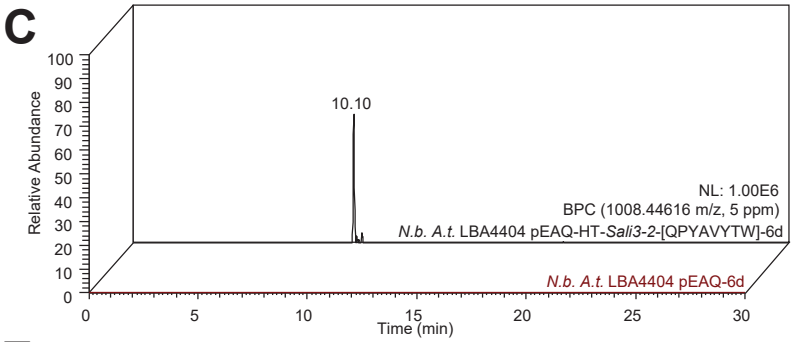
**A**

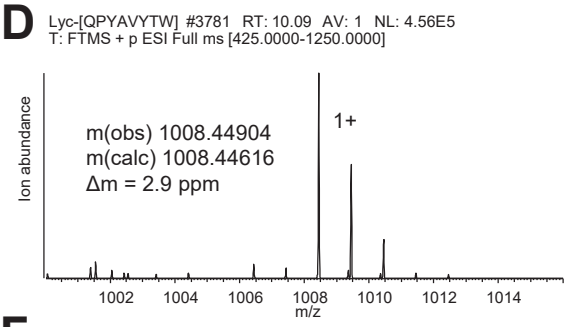
>Glyma.12G217400-[QPYAVYTW]

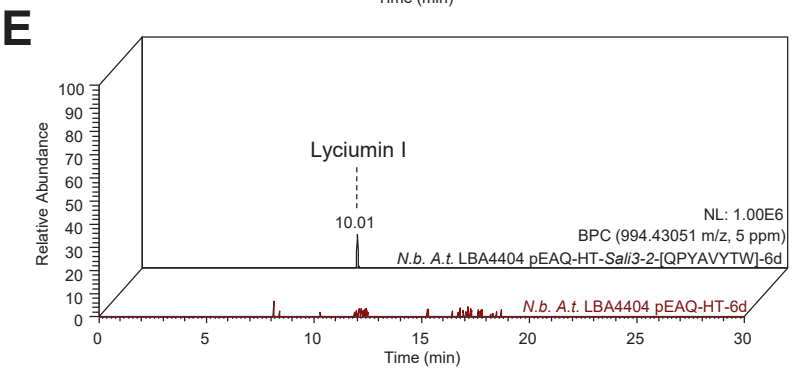
MEFRCSVISFTILFSLALAGESHVHASLPEEDYWEAVWPNTPIPT
ALRELLKPLPAGVEIDELPKQIDDTQYPKTFFYKEDLHPGKTMKV
QFTKRPYAQPYAVYTWLTDIKDTSKEGYSFEEICIKKEAFEGEEK
FCAKSLGTVIGFAISKLGKNIQVLSSSFVNKQEQYTVEGVQNLGD
KAVMCHGLNFRTAVFYCHKVRETTAFMVPLVAGDGTKTQALAVCH
SDTSGMNHHMLHELMGVDPGTNPVCHFLGSKAILWVPNLSMDTAY
QTNVVV




Lyciumin-[QPYAVYTW]



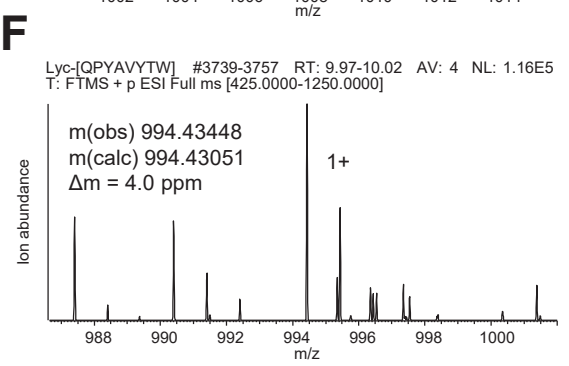




**Figure S34 | Heterologous expression of *Sali3-2*-[QPYAVYTW] in *Nicotiana benthamiana*.** (**A**) Sali3-2-[QPYAVYTW] lyciumin precursor peptide (BURP domain underlined, core peptide highlighted in red). (**B**) Predicted [QPYAVYTW]-lyciumin chemotype. (**C**) LC-MS chemotyping of predicted [QPYAVYTW]-lyciumin in peptide extract of *N. benthamiana* leaves infiltrated with *A. tumefaciens* LBA4404 pEAQ-HT-*Sali3-2*-[QPYAVYTW] for six days. (**D**) MS analysis of predicted [QPYAVYTW]-lyciumin chemotype. (**E**) LC-MS chemotyping of lyciumin I in peptide extract of *N. benthamiana* leaves infiltrated with *A. tumefaciens* LBA4404 pEAQ-HT-*Sali3-2*-[QPYAVYTW] for six days. (**F**) MS analysis of predicted lyciumin I chemotype.

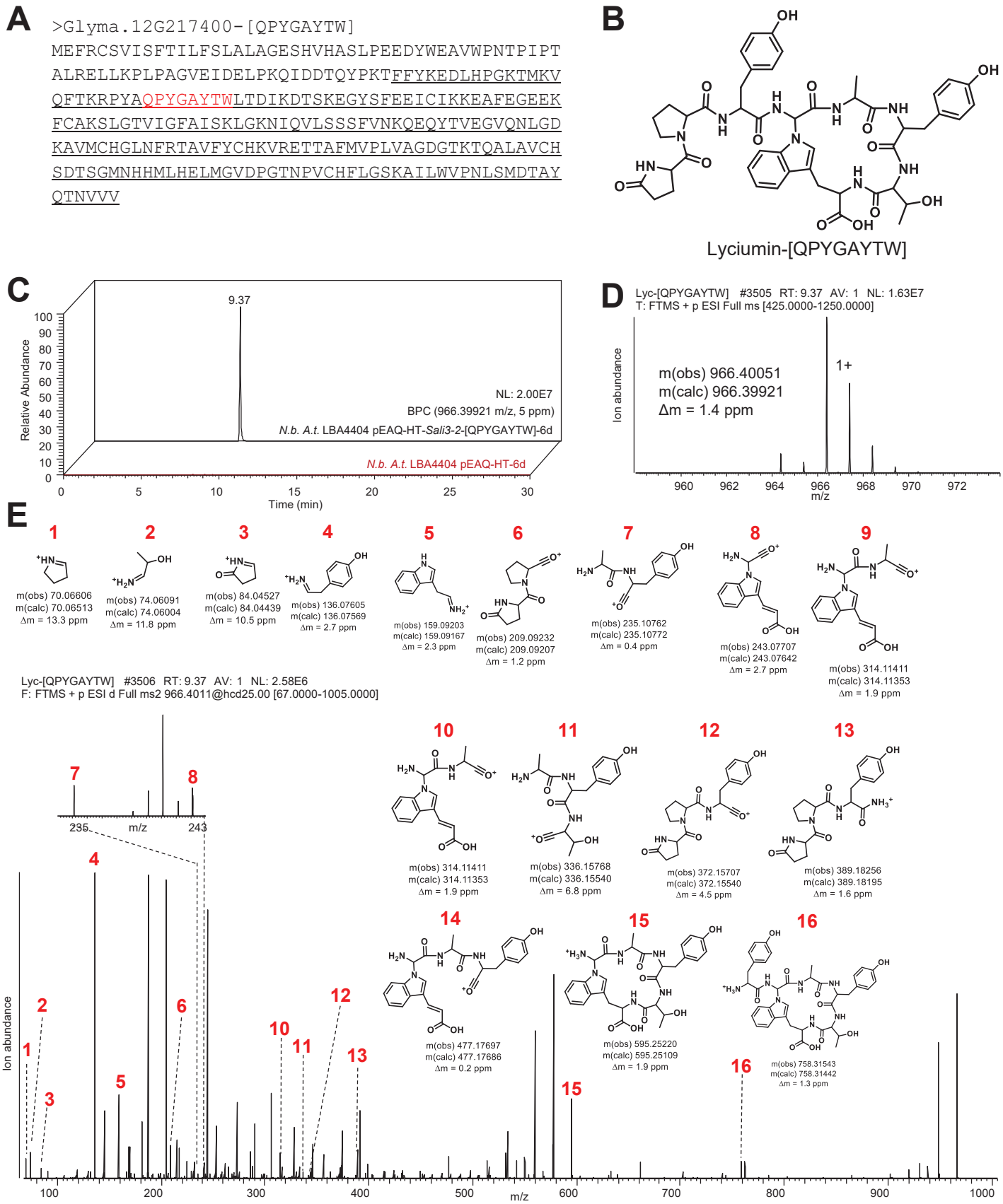



**Figure S35 | Heterologous expression of *Sali3-2*-[QPYGAYTW] in *Nicotiana benthamiana*.** (**A**) Sali3-2-[QPYGAYTW] lyciumin precursor peptide (BURP domain underlined, core peptide highlighted in red). (**B**) Predicted [QPYGAYTW]-lyciumin chemotype. (**C**) LC-MS chemotyping of predicted [QPYGAYTW]-lyciumin in peptide extract of *N. benthamiana* leaves infiltrated with *A. tumefaciens* LBA4404 pEAQ-HT-*Sali3-2*-[QPYGAYTW] for six days. (**D**) MS analysis of predicted [QPYGAYTW]-lyciumin chemotype. (**E**) MS/MS analysis of predicted [QPYGAYTW]-lyciumin chemotype.

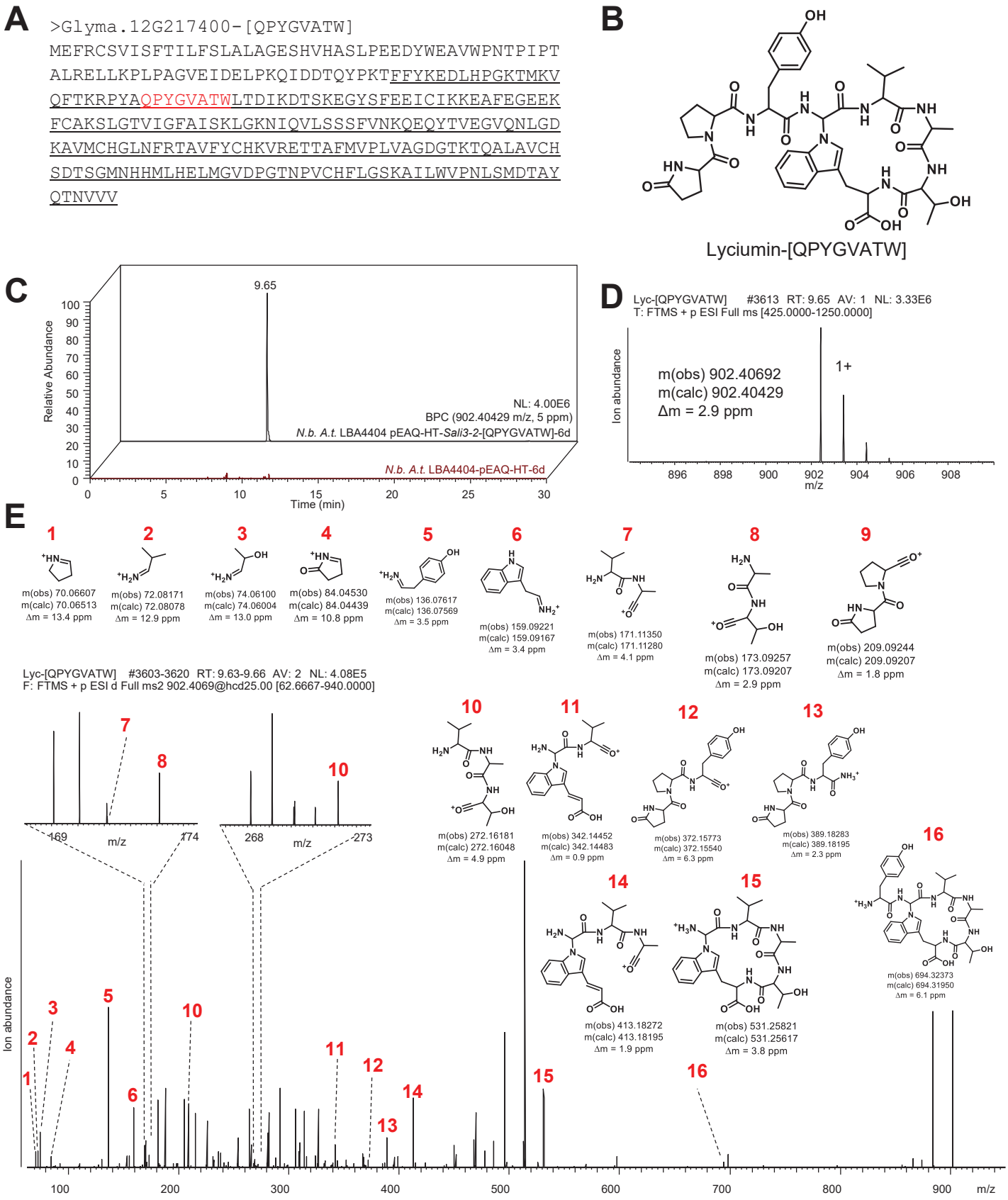



**Figure S36 | Heterologous expression of *Sali3-2*-[QPYGVATW] in *Nicotiana benthamiana*.** (**A**) Sali3-2-[QPYGVATW] lyciumin precursor peptide (BURP domain underlined, core peptide highlighted in red). (**B**) Predicted [QPYGVATW]-lyciumin chemotype. (**C**) LC-MS chemotyping of predicted [QPYGVATW]-lyciumin in peptide extract of *N. benthamiana* leaves infiltrated with *A. tumefaciens* LBA4404 pEAQ-HT-*Sali3-2*-[QPYGVATW] for six days. (**D**) MS analysis of predicted [QPYGVATW]-lyciumin chemotype. (**E**) MS/MS analysis of predicted [QPYGVATW]-lyciumin chemotype.

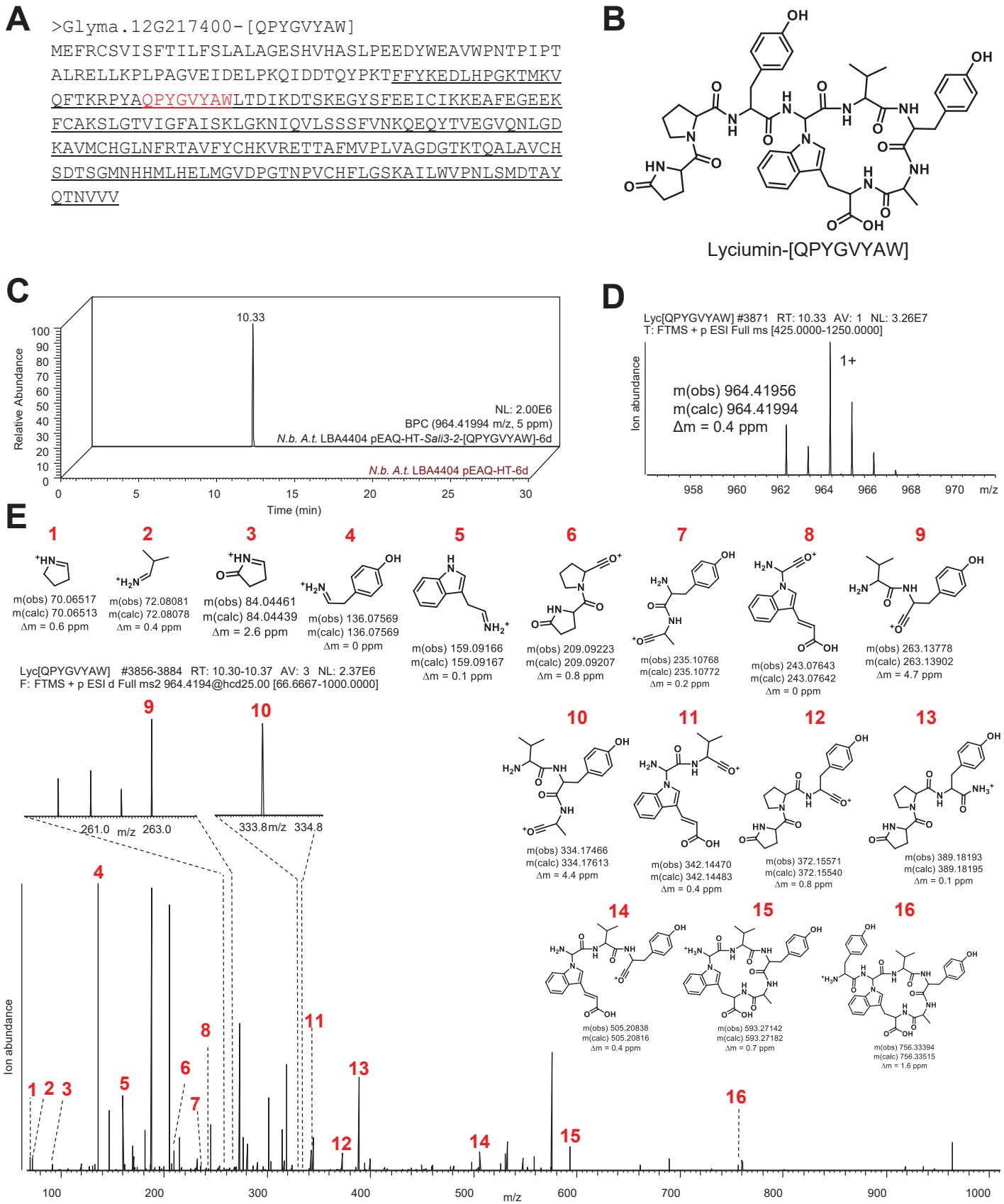



**Figure S37 | Heterologous expression of *Sali3-2*-[QPYGVYAW] in *Nicotiana benthamiana*.** (**A**) Sali3-2-[QPYGVYAW] lyciumin precursor peptide (BURP domain underlined, core peptide highlighted in red). (**B**) Predicted [QPYGVYAW]-lyciumin chemotype. (**C**) LC-MS chemotyping of predicted [QPYGVYAW]-lyciumin in peptide extract of *N. benthamiana* leaves infiltrated with *A. tumefaciens* LBA4404 pEAQ-HT-*Sali3-2*-[QPYGVYAW] for six days. (**D**) MS analysis of predicted [QPYGVYAW]-lyciumin chemotype. (**E**) MS/MS analysis of predicted [QPYGVYAW]-lyciumin chemotype.

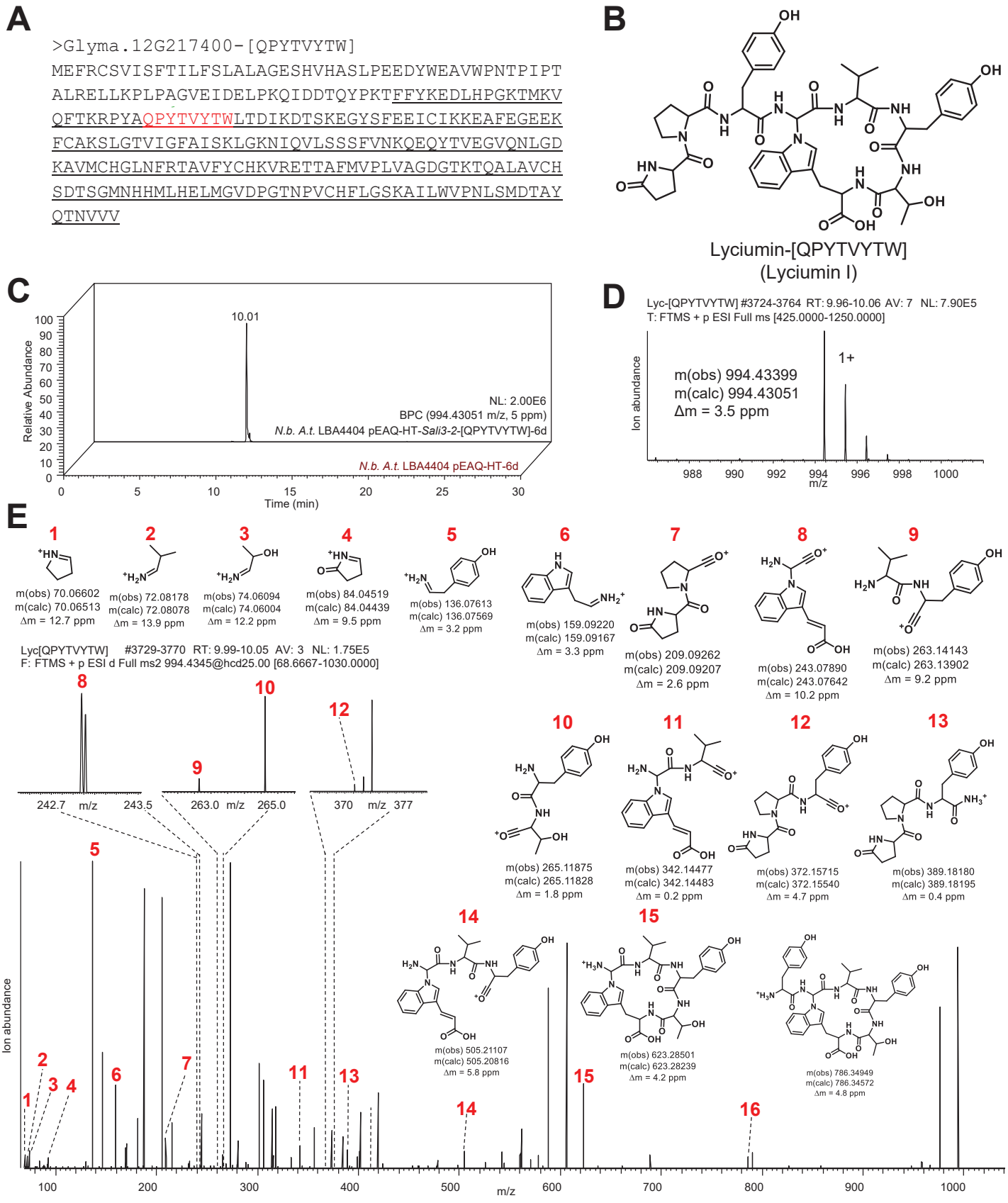



**Figure S38 | Heterologous expression of *Sali3-2*-[QPYTVYTW] in *Nicotiana benthamiana*.** (**A**) Sali3-2-[QPYTVYTW] lyciumin precursor peptide (BURP domain underlined, core peptide highlighted in red). (**B**) Predicted [QPYTVYTW]-lyciumin chemotype. (**C**) LC-MS chemotyping of predicted [QPYTVYTW]-lyciumin in peptide extract of *N. benthamiana* leaves infiltrated with *A. tumefaciens* LBA4404 pEAQ-HT-*Sali3-2*-[QPYTVYTW] for six days. (**D**) MS analysis of predicted [QPYTVYTW]-lyciumin chemotype. (**E**) MS/MS analysis of predicted [QPYTVYTW]-lyciumin chemotype.

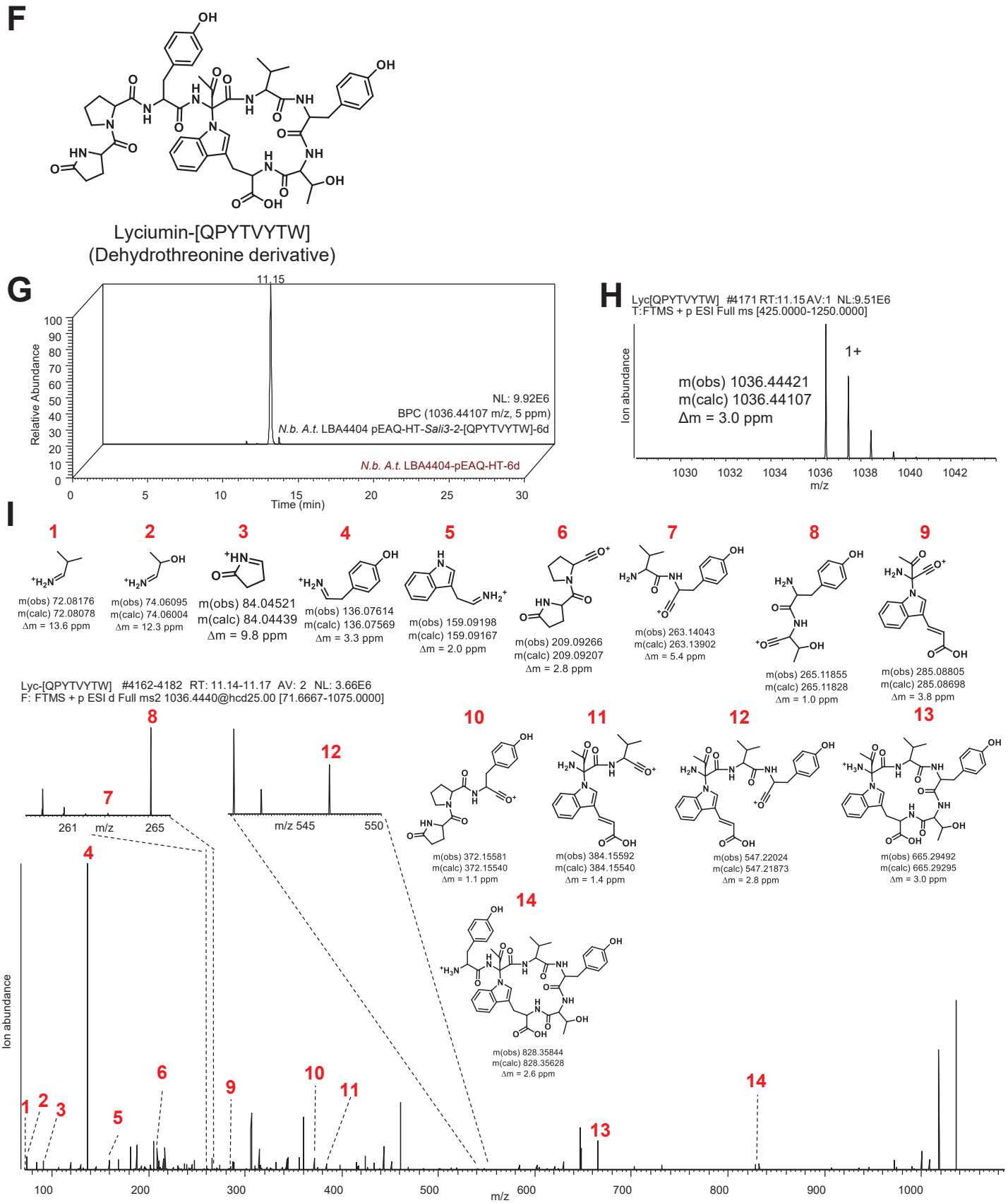



**Figure S38 | Heterologous expression of *Sali3-2*-[QPYTVYTW] in *Nicotiana benthamiana*.** (**F**) Predicted [QPYTVYTW]-dehydrothreonine lyciumin chemotype. (**G**) LC-MS chemotyping of predicted [QPYTVYTW]-dehydrothreonine lyciumin in peptide extract of *N. benthamiana* leaves infiltrated with *A. tumefaciens* LBA4404 pEAQ-HT-*Sali3-2*-[QPYTVYTW] for six days. (**H**) MS analysis of predicted [QPYTVYTW]-dehydrothreonine lyciumin chemotype. (**I**) MS/MS analysis of predicted [QPYTVYTW]-dehydrothreonine lyciumin chemotype.

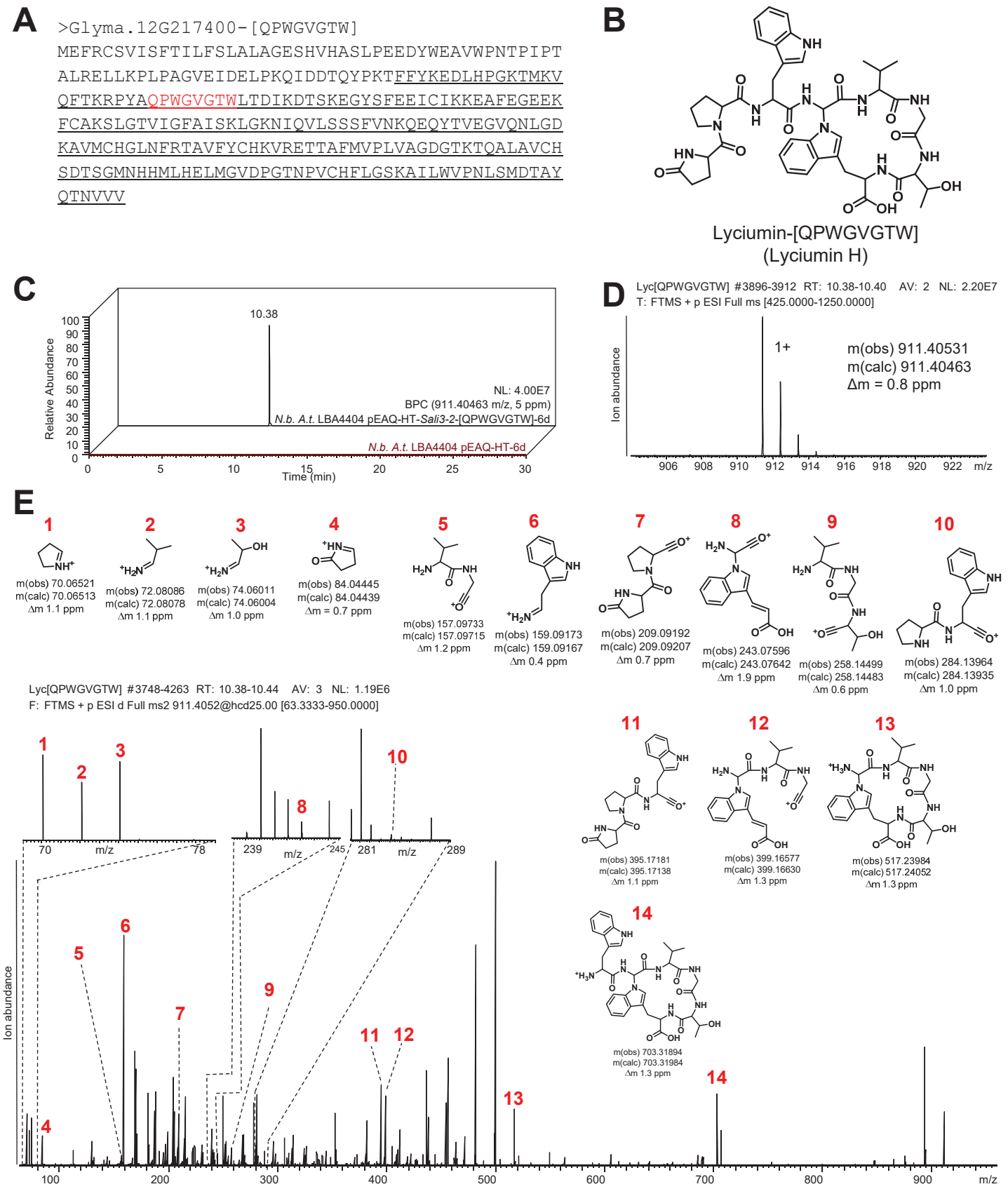



**Figure S39 | Heterologous expression of *Sali3-2*-[QPWGVGTW] in *Nicotiana benthamiana*.** (**A**) Sali3-2-[QPWGVGTW] lyciumin precursor peptide (BURP domain underlined, core peptide highlighted in red). (**B**) Predicted [QPWGVGTW]-lyciumin chemotype. (**C**) LC-MS chemotyping of predicted [QPWGVGTW]-lyciumin in peptide extract of *N. benthamiana* leaves infiltrated with *A. tumefaciens* LBA4404 pEAQ-HT-*Sali3-2*-[QPWGVGTW] for six days. (**D**) MS analysis of predicted [QPWGVGTW]-lyciumin chemotype. (**E**) MS/MS analysis of predicted [QPWGVGTW]-lyciumin chemotype.

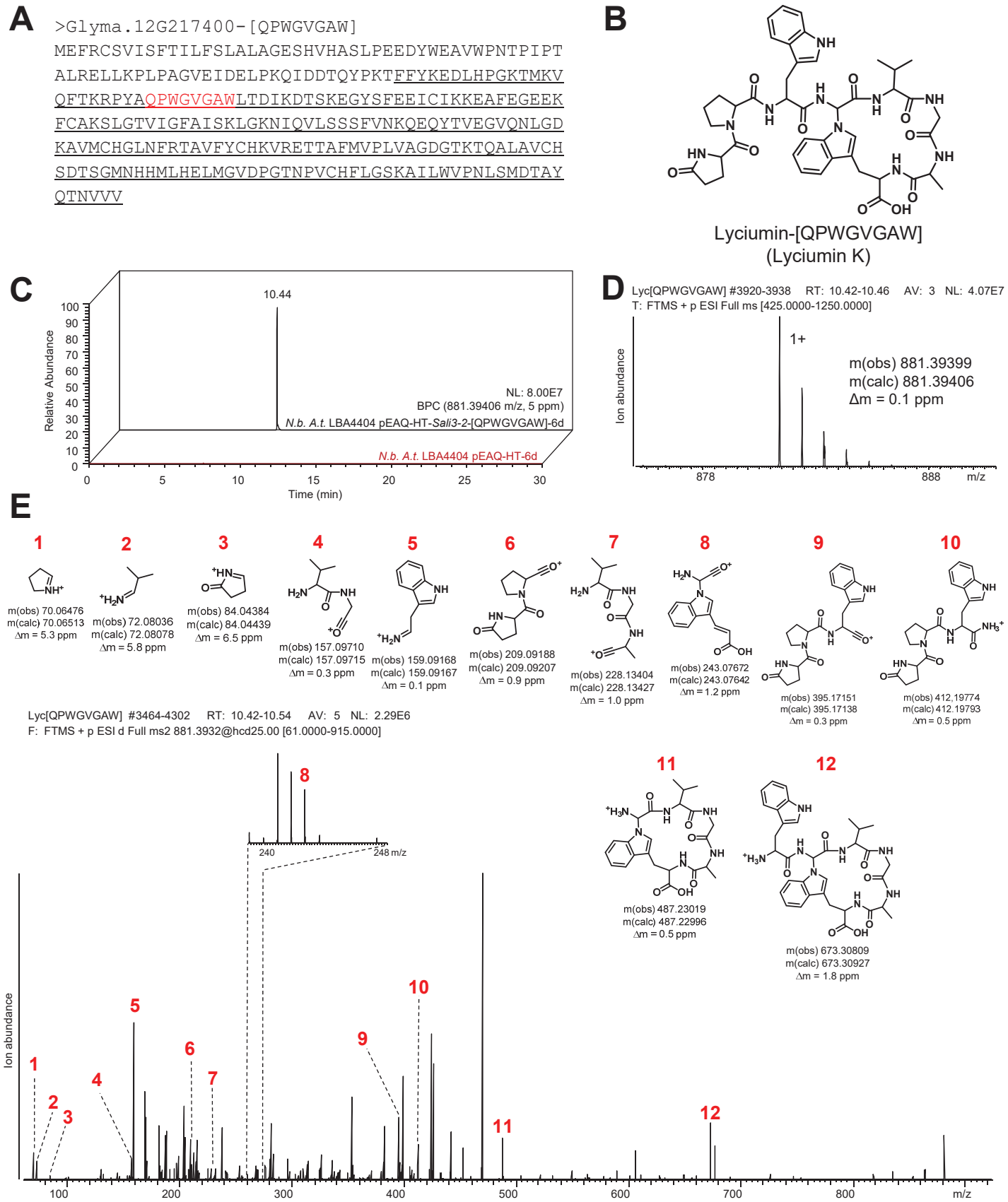



**Figure S40 | Heterologous expression of *Sali3-2*-[QPWGVGAW] in *Nicotiana benthamiana*.** (**A**) Sali3-2-[QPWGVGAW] lyciumin precursor peptide (BURP domain underlined, core peptide highlighted in red). (**B**) Predicted [QPWGVGAW]-lyciumin chemotype. (**C**) LC-MS chemotyping of predicted [QPWGVGAW]-lyciumin in peptide extract of *N. benthamiana* leaves infiltrated with *A. tumefaciens* LBA4404 pEAQ-HT-*Sali3-2*-[QPWGVGAW] for six days. (**D**) MS analysis of predicted [QPWGVGAW]-lyciumin chemotype. (**E**) MS/MS analysis of predicted [QPWGVGAW]-lyciumin chemotype.

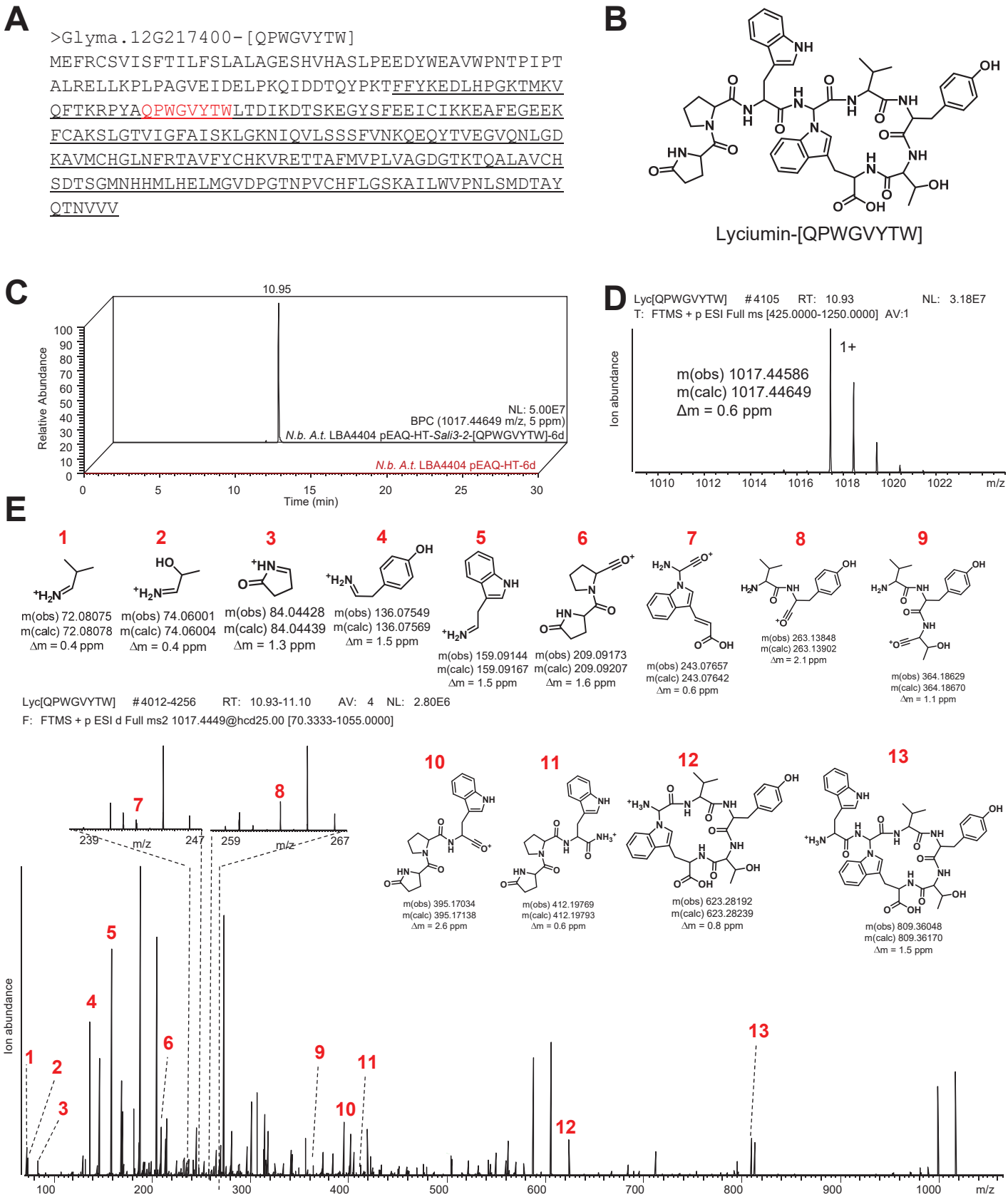



**Figure S41 | Heterologous expression of *Sali3-2*-[QPWGVYTW] in *Nicotiana benthamiana*.** (**A**) Sali3-2-[QPWGVYTW] lyciumin precursor peptide (BURP domain underlined, core peptide highlighted in red). (**B**) Predicted [QPWGVYTW]-lyciumin chemotype. (**C**) LC-MS chemotyping of predicted [QPWGVYTW]-lyciumin in peptide extract of *N. benthamiana* leaves infiltrated with *A. tumefaciens* LBA4404 pEAQ-HT-*Sali3-2*-[QPWGVYTW] for six days. (**D**) MS analysis of predicted [QPWGVYTW]-lyciumin chemotype. (**E**) MS/MS analysis of predicted [QPWGVYTW]-lyciumin chemotype.

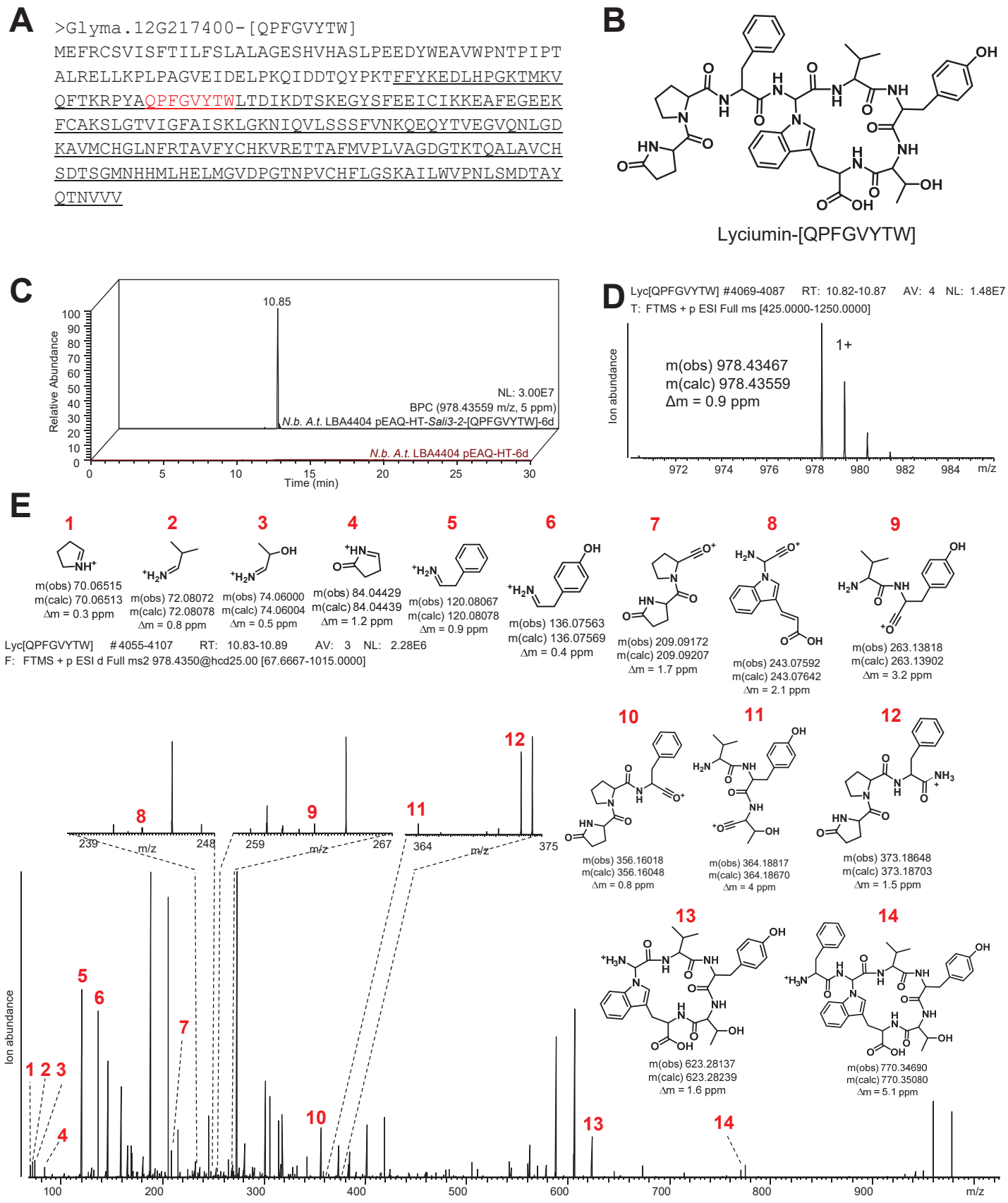



**Figure S42 | Heterologous expression of *Sali3-2*-[QPFGVYTW] in *Nicotiana benthamiana*.** (**A**) Sali3-2-[QPFGVYTW] lyciumin precursor peptide (BURP domain underlined, core peptide highlighted in red). (**B**) Predicted [QPFGVYTW]-lyciumin chemotype. (**C**) LC-MS chemotyping of predicted [QPFGVYTW]-lyciumin in peptide extract of *N. benthamiana* leaves infiltrated with *A. tumefaciens* LBA4404 pEAQ-HT-*Sali3-2*-[QPFGVYTW] for six days. (**D**) MS analysis of predicted [QPFGVYTW]-lyciumin chemotype. (**E**) MS/MS analysis of predicted [QPFGVYTW]-lyciumin chemotype.

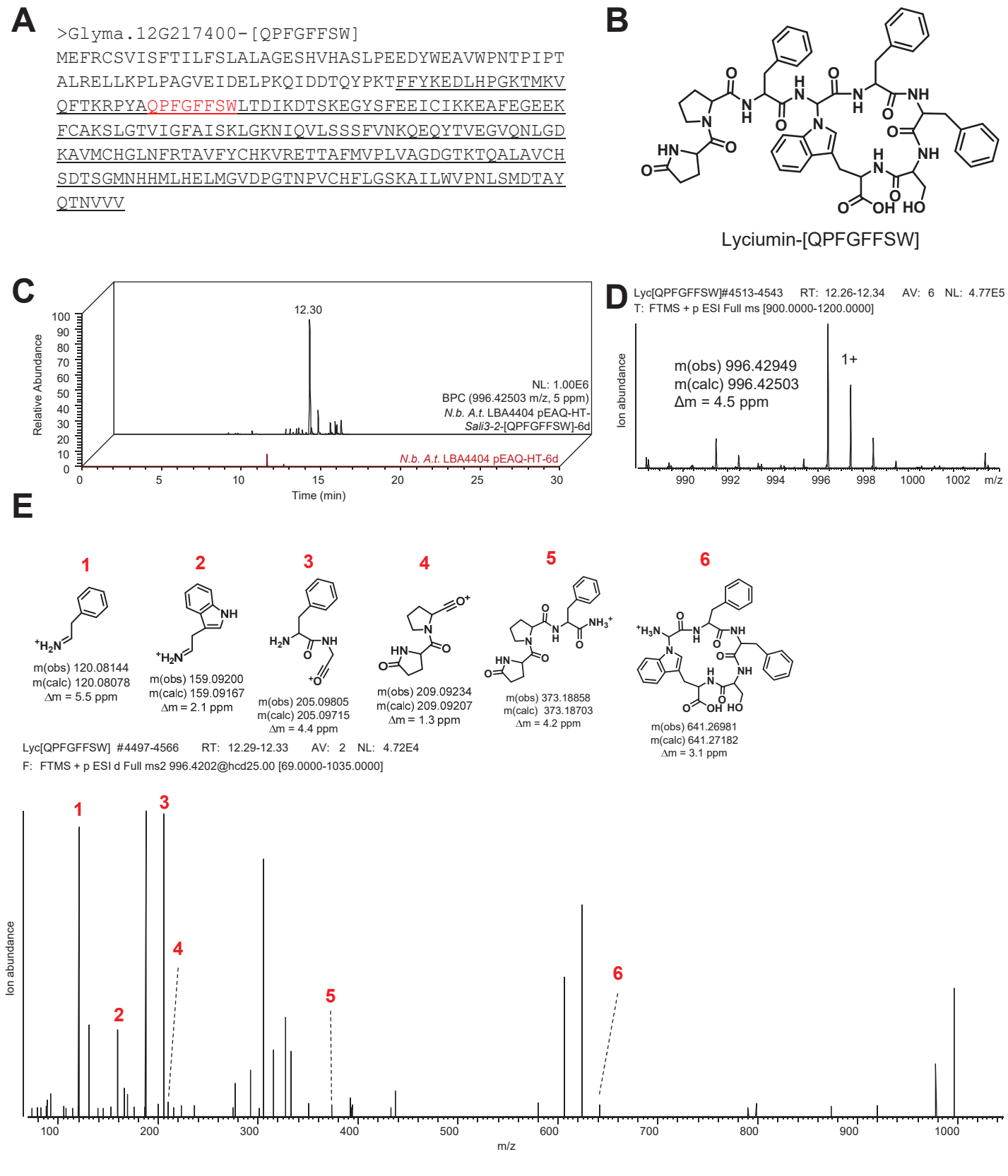



**Figure S43 | Heterologous expression of *Sali3-2*-[QPFGFFSW] in *Nicotiana benthamiana*.** (**A**) Sali3-2-[QPFGFFSW] lyciumin precursor peptide (BURP domain underlined, core peptide highlighted in red). (**B**) Predicted [QPFGFFSW]-lyciumin chemotype. (**C**) LC-MS chemotyping of predicted [QPFGFFSW]-lyciumin in peptide extract of *N. benthamiana* leaves infiltrated with *A. tumefaciens* LBA4404 pEAQ-HT-*Sali3-2*-[QPFGFFSW] for six days. (**D**) MS analysis of predicted [QPFGFFSW]-lyciumin chemotype. (**E**) MS/MS analysis of predicted [QPFGFFSW]-lyciumin chemotype.

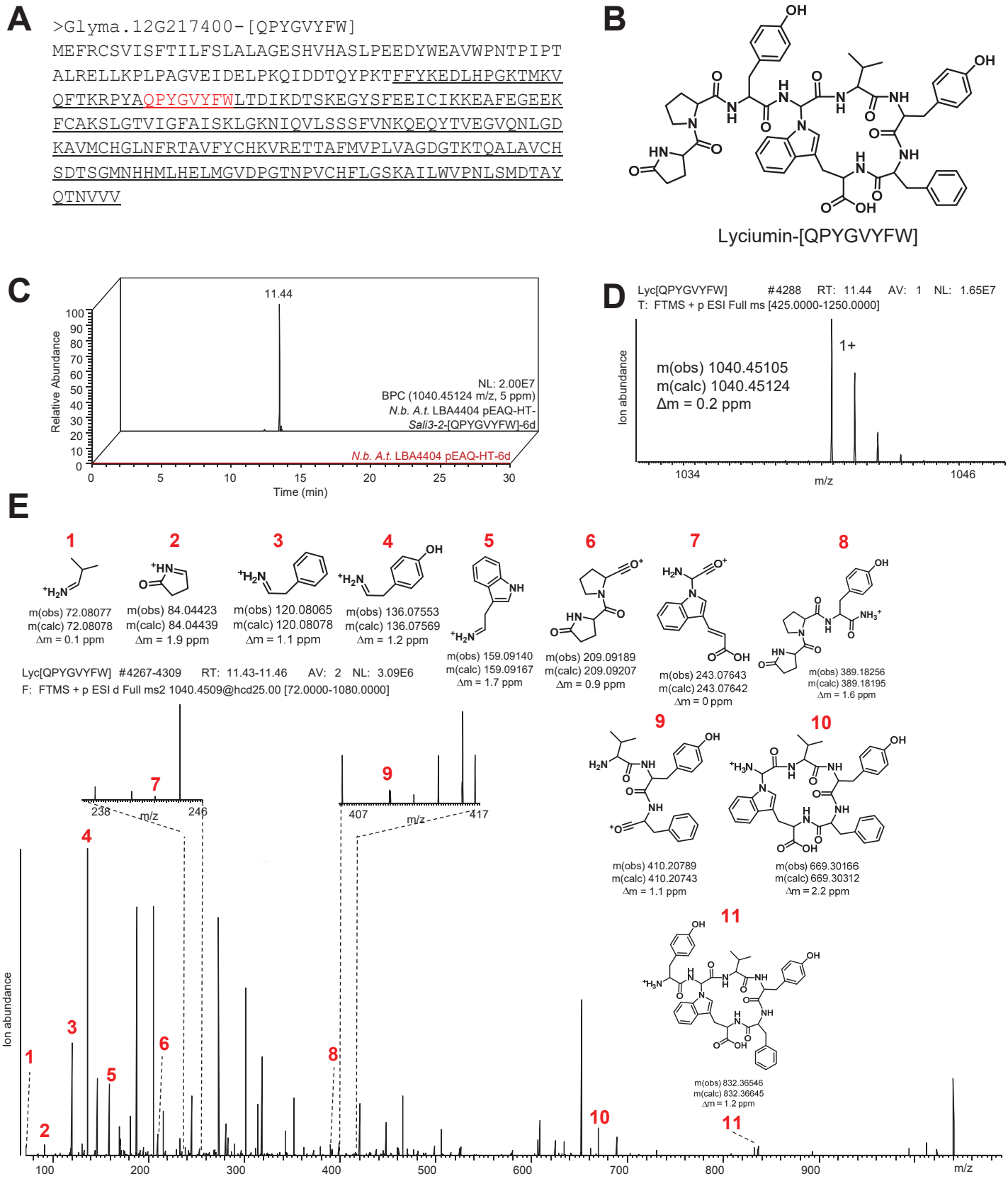



**Figure S44 | Heterologous expression of *Sali3-2*-[QPYGVYFW] in *Nicotiana benthamiana*.** (**A**) Sali3-2-[QPYGVYFW] lyciumin precursor peptide (BURP domain underlined, core peptide highlighted in red). (**B**) Predicted [QPYGVYFW]-lyciumin chemotype. (**C**) LC-MS chemotyping of predicted [QPYGVYFW]-lyciumin in peptide extract of *N. benthamiana* leaves infiltrated with *A. tumefaciens* LBA4404 pEAQ-HT-*Sali3-2*-[QPYGVYFW] for six days. (**D**) MS analysis of predicted [QPYGVYFW]-lyciumin chemotype. (**E**) MS/MS analysis of predicted [QPYGVYFW]-lyciumin chemotype.

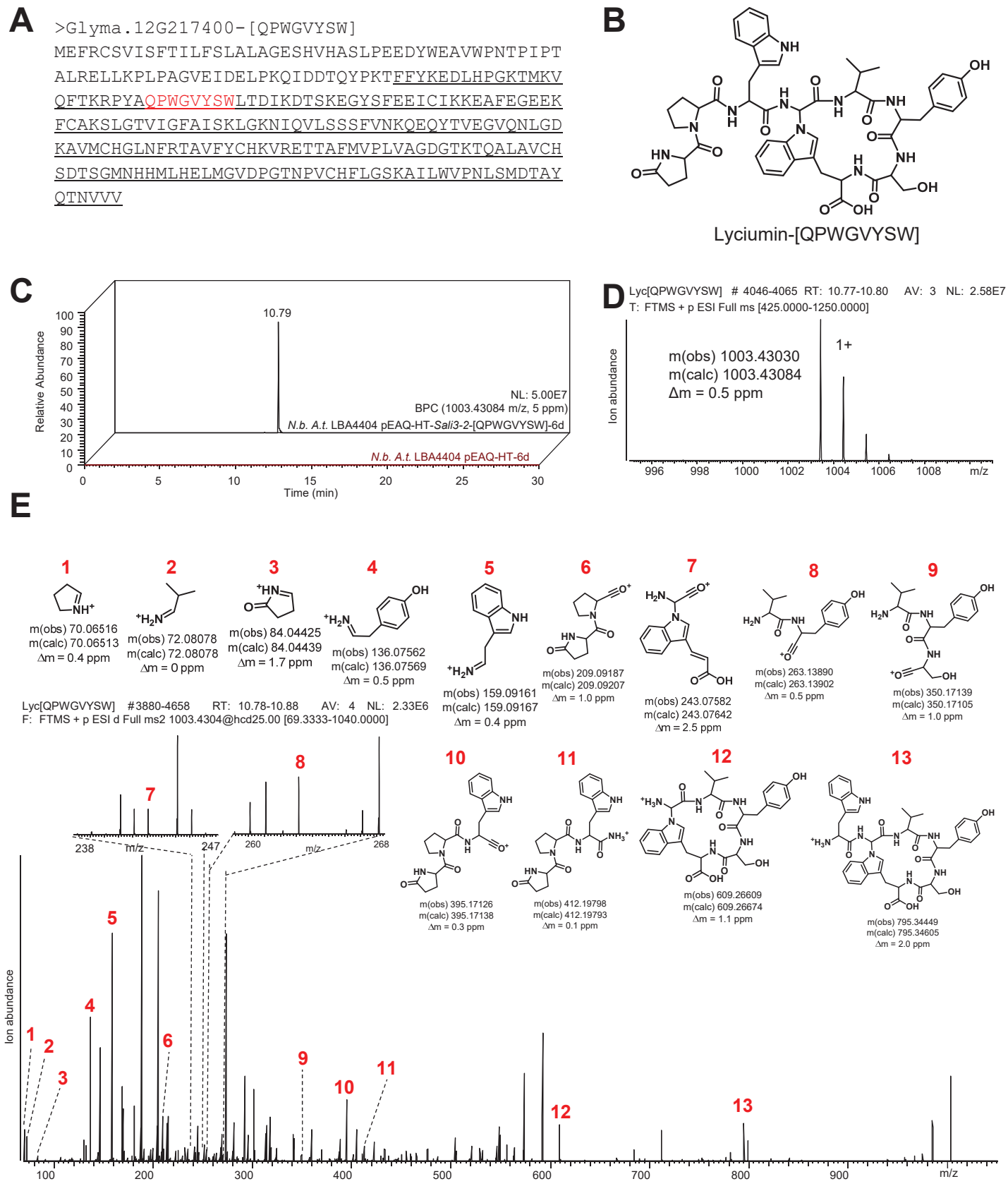



**Figure S45 | Heterologous expression of *Sali3-2*-[QPWGVYSW] in *Nicotiana benthamiana*.** (**A**) Sali3-2-[QPWGVYSW] lyciumin precursor peptide (BURP domain underlined, core peptide highlighted in red). (**B**) Predicted [QPWGVYSW]-lyciumin chemotype. (**C**) LC-MS chemotyping of predicted [QPWGVYSW]-lyciumin in peptide extract of *N. benthamiana* leaves infiltrated with *A. tumefaciens* LBA4404 pEAQ-HT-*Sali3-2*-[QPWGVYSW] for six days. (**D**) MS analysis of predicted [QPWGVYSW]-lyciumin chemotype. (**E**) MS/MS analysis of predicted [QPWGVYSW]-lyciumin chemotype.

**References**


1. Bassüner R *et al.* (1988) Abundant embryonic mRNA in field bean (Vicia faba L.) codes for a new class of seed proteins: cDNA cloning and characterization of the primary translation product. *Plant Mol Biol* 11(3): pp 321–334.
2. Yamaguchi-Shinozaki K, Shinozaki K (1993) The plant hormone abscisic acid mediates the drought-induced expression but not the seed-specific expression of rd22, a gene responsive to dehydration stress in Arabidopsis thaliana. *Mol Gen Genet* 238(1-2): pp 17–25.
3. Zheng L, Heupel RC, DellaPenna D (1992) The b Subunit of Tomato Fruit Polygalacturonase Isoenzyme 1: Isolation, Characterization, and Identification of Unique Structural Features. *Plant Cell* 4(9): pp 1147-1156.
4. Boutilier KA *et al.* (1994) Expression of the BnmNAP subfamily of napin genes coincides with the induction of Brassica microspore embryogenesis. *Plant Mol Biol* 26(6): pp 1711–1723.
5. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput *Nucleic Acids Res* 32(5): pp 1792-1797.
6. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets *Mol Biol Evol* 33, pp 1870-1874.
7. Nei M, Kumar S (2000) Molecular Evolution and Phylogenetics. Oxford University Press, New York.
8. Morita H, Yoshida N, Takeya K, Itokawa H, Shirota O (1996) Configurational and conformational analyses of a cyclic octapeptide, lyciumin A, from Lycium chinense Mill. *Tetrahedron* 52(8): pp 2795-2802.
9. Yahara S *et al.* (1993) Cyclic peptides, acyclic diterpene glycosides and other compounds from Lycium chinense Mill. *Chem Pharm Bull* 41(4): pp 703-709.
10. Morita H, Suzuki H, Kobayashi J (2004) Celogenamide A, a New Cyclic Peptide from the Seeds of Celosia argentea. *J Nat Prod* 67(9): pp 1628–1630.
11. Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform* 12(1): pp 323.
12. Potato Genome Sequencing Consortium (2011) Genome sequence and analysis of the tuber crop potato. *Nature* 475(7355): pp 189-195.
13. Grabherr MG *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29(7): pp 644-652.
14. Waese J *et al.* (2017) ePlant: Visualizing and Exploring Multiple Levels of Data for Hypothesis Generation in Plant Biology. *Plant Cell* 29: pp 1806-1821.
15. Benedito VA *et al.* (2008) A gene expression atlas of the model legume Medicago truncatula. *Plant J* 55(3): pp 504-513.
16. Libault M *et al.* (2010) An integrated transcriptome atlas of the crop model Glycine max, and its use in comparative analyses in plants. *Plant J* 63(1): pp 86-99.
17. Severin AJ *et al.* (2010) RNA-Seq Atlas of Glycine max: a guide to the soybean transcriptome. *BMC Plant Biol* 10(1): pp 160.
18. Priyam A *et al.* (2015) Sequenceserver: a modern graphical user interface for custom BLAST databases. *Biorxiv* 033142.
19. Gibson DG *et al.* (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6(5): pp 343-345.
20. Sainsbury F, Thuenemann EC, Lomonossoff GP (2009) pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnol J* 7(7): pp 682-693.
21. Wang MX *et al.* (2016) Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat Biotechnol* 34(8): pp 828-837.