**Table S5** Types of different soybean *GmCOL* proteins and their haplotype distribution

|  |  |  |  |
| --- | --- | --- | --- |
| Gene | Types of protein | Amino acids in mutation sites | Haplotype distribution |
| *GmCOL2* | P0 |  | *Hap1, Hap5, Hap4, Hap10* |
|  | P1 | D216 | *Hap2, Hap3, Hap6, Hap11* |
|  | P2 | I116 | *Hap8* |
|  | P3 | I116 L176 | *Hap9* |
|  | P4 | A103 N104 D217 | *Hap7* |
| *GmCOL4* | P0 |  | *Hap1, Hap2, Hap3, Hap4* |
|  | P1 | R104 | *Hap5* |
| *GmCOL5* | P0 |  | *Hap1* |
|  | P1 | E167 | *Hap2, Hap4, Hap5* |
|  | P2 | E167 L265 | *Hap3* |
| *GmCOL6* | P0 |  | *Hap1, Hap2* |
|  | P1 | \*Del153 | *Hap1, Hap2* |
|  | P 2 | \*Del153 \*Del166-185 L220 R295 | *Hap3, Hap5* |
|  | P3 | \*Del79 \*Del153 | *Hap1* |
|  | P4 | L220 P323 Q404 | *Hap4* |
|  | P5 | P323 | *Hap2* |
|  | P6 | P98 L99 Stop codon104 | *Hap1* |
| *GmCOL8* | P0 |  | *Hap1, Hap2* |
| *GmCOL9* | P0 |  | *Hap1, Hap2, Hap3, Hap4, Hap5, Hap6, Hap7, Hap8, Hap9, Hap10, Hap11* |
| *GmCOL10* | P0 |  | *Hap1* |
|  | P1 | S364 | *Hap2, Hap3, Hap4* |
| *GmCOL13* | P0 |  | *Hap1, Hap2, Hap3* |

**Table S5** Continued

|  |  |  |  |
| --- | --- | --- | --- |
| Gene | Types of protein | Amino acids in mutation sites | Haplotype distribution |
| *GmCOL14* | P0 |  | *Hap1* |
|  | P1 | Stop codon59 | *Hap1* |
|  | P2 | T66 | *Hap2* |
|  | P3 | Stop codon110 | *Hap1* |
|  | P4 | \*Del291 | *Hap1* |
|  | P5 | P292 | *Hap3* |
| *GmCOL15* | P0 |  | *Hap1, Hap2, Hap3* |
|  | P1 | Q328 | *Hap4* |
| *GmCOL16* | P0 |  | *Hap1, Hap2, Hap4, Hap6, Hap9* |
|  | P1 | A101 | *Hap3, Hap5, Hap7, Hap8* |
| *GmCOL19* | P0 |  | *Hap1, Hap5* |
|  | P1 | D102 | *Hap2, Hap3, Hap4* |
| *GmCOL20* | P0 |  | *Hap1, Hap6, Hap7, Hap8* |
|  | P1 | E161 | *Hap2, Hap3, Hap9* |
|  | P2 | Q319 | *Hap4, Hap5* |
|  | P3 | E161 | *Hap10* |
|  | P4 | V103 E161 Q319 | *Hap2* |
|  | P5 | Y276 Q319 | *Hap4* |
| *GmCOL22* | P0 |  | *Hap1, Hap4, Hap5, Hap6* |
|  | P1 | \*Del76 | *Hap2* |
|  | P2 | A274 | *Hap3* |
|  | P3 | Stop codon155 | *Hap1, Hap4* |
|  | P4 | I157 | *Hap1* |
|  | P5 | Stop codon229 | *Hap1* |

Table S5 Continued

|  |  |  |  |
| --- | --- | --- | --- |
| **Gene** | **Types of protein** | **Amino acids in mutation sites** | **Haplotype distribution** |
| *GmCOL22* | P6 | Stop codon229 | *Hap1* |
|  | P7 | \*Del79-80 | *Hap4* |
| *GmCOL23* | P0 |  | *Hap1* |
|  | P1 | Q104 Q105 Y109 | *Hap2, Hap3, Hap5* |
|  | P2 | Y107 G352 | *Hap6, Hap9* |
|  | P3 | T104 Y108 | *Hap8* |
|  | P4 | Y107 G352\*Del418 | *Hap4* |
|  | P5 | Y107 Stop codon434 | *Hap10* |
|  | P6 | \*Del418 | *Hap1* |
|  | P7 | Y107 | *Hap3* |
|  | P8 | K431 Stop codon434 | *Hap7* |
|  | P9 | G354 | *Hap4* |
| *GmCOL24* | P0 |  | *Hap1* |
|  | P1 | D75 M92 P93 R112 F174 | *Hap2, Hap3* |
|  | P2 | D75 M92 P93 R112 | *Hap5* |
|  | P3 | F174 | *Hap4* |
| *GmCOL25* | P0 |  | *Hap1, Hap4* |
|  | P1 | L73 I84 Y108 Stop codon152 | *Hap2, Hap5* |
|  | P2 | I165 | *Hap3* |
| *GmCOL26* | P0 |  | *Hap1* |
|  | P1 | A8 H42 K55 F83 K95 A129 | *Hap2* |
|  | P2 | A8 T25 M60 E127 A129 | *Hap3* |
|  | P3 | A8 T39 C42 I59 A129 | *Hap4* |

Table S5 Continued

|  |  |  |  |
| --- | --- | --- | --- |
| **Gene** | **Types of protein** | **Amino acids in mutation sites** | **Haplotype distribution** |
| *GmCOL28* | P0 |  | *Hap1* |
|  | P1 | A88 | *Hap2* |
|  | P2 | S4 | *Hap3* |

Note: 1. A103, subscript 103 means which position of soybean *GmCOL* protein sequence this amino acid is located in; others are the same as A103

2. \*Del153 denotes deletion of nucleotide from the position of 153 bp of amino acid sites, others are the same as Del153