Table S1. Primers used for assembly validation of IR/SC junction

|  |  |  |  |
| --- | --- | --- | --- |
| Sequence name | Sequence (5’-3’) | Target size(bp) | Target junction |
| A-f | CTATTATCGCCTTTCGCA | 700 | IRa/LSC |
| A-r | TTAGTTAGAGCAGCGGGT |  |  |
| B-f | TGGCTTGGATTGGTATTAG | 699 | IRb/SSC |
| B-r | GGGCGGAAACATAAAAGA |  |  |
| C-f | GAGTAAAGTAAAGGGAAGG | 361 | IRb/LSC |
| C-r | CTGTAGTAAGAGGAGTAGT |  |  |
| D-f | AAATATGTAGGAGGTGGG | 258 | IRa/SSC |
| D-r | CGAGTGAATGGAAAGGAA |  |  |

Table-S2. Codon–anticodon recognition patterns and codon usage for the A. spinosa chloroplast genomes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Amino Acid** | **Codon** | **Count** | **RSCU** | **Recognition pattern (tRNA)** |
| Phe | UUU | 1058 | 1.25 |  |
|  | UUC | 630 | 0.75 | trnF-GAA |
| Leu | UUA | 741 | 1.72 | trnL-UAA |
|  | UUG | 539 | 1.25 | trnL-CAA |
|  | CUU | 562 | 1.30 |  |
|  | CUC | 225 | 0.52 |  |
|  | CUA | 344 | 0.80 | trnL-UAG |
|  | CUG | 175 | 0.41 |  |
| Ile | AUU | 1078 | 1.48 |  |
|  | AUC | 529 | 0.73 | trnI-GAU |
|  | AUA | 580 | 0.80 | trnI-UAU |
| Met | AUG | 564 | 1.00 | trn(f)M-CAU |
| Val | GUU | 530 | 1.47 |  |
|  | GUC | 190 | 0.53 | trnV-GAC |
|  | GUA | 509 | 1.41 | trnV-UAC |
|  | GUG | 217 | 0.60 |  |
| Ser | UCU | 667 | 1.68 |  |
|  | UCC | 371 | 0.93 | trnS-GGA |
|  | UCA | 489 | 1.23 | trnS-UGA |
|  | UCG | 245 | 0.62 |  |
| Pro | CCU | 413 | 1.52 |  |
|  | CCC | 219 | 0.81 |  |
|  | CCA | 298 | 1.10 | trnP-UGG |
|  | CCG | 155 | 0.57 |  |
| Thr | ACU | 492 | 1.48 |  |
|  | ACC | 278 | 0.83 | trnT-GGU |
|  | ACA | 399 | 1.20 | trnT-UGU |
|  | ACG | 163 | 0.49 |  |
| Ala | GCU | 550 | 1.71 |  |
|  | GCC | 217 | 0.67 |  |
|  | GCA | 368 | 1.14 | trnA-UGC |
|  | GCG | 155 | 0.48 |  |
| Tyr | UAU | 865 | 1.53 |  |
|  | UAC | 269 | 0.47 | trnY-GUA |
| TER | UAA | 302 | 1.20 |  |
|  | UAG | 184 | 0.73 |  |
| His | CAU | 506 | 1.50 |  |
|  | CAC | 168 | 0.50 | trnH-GUG |
| Gln | CAA | 689 | 1.46 | trnQ-UUG |
|  | CAG | 253 | 0.54 |  |
| Asn | AAU | 999 | 1.48 |  |
|  | AAC | 349 | 0.52 | trnN-GUU |
| Lys | AAA | 1103 | 1.47 | trnK-UUU |
|  | AAG | 396 | 0.53 |  |
| Asp | GAU | 804 | 1.59 |  |
|  | GAC | 207 | 0.41 |  |
| Glu | GAA | 965 | 1.47 |  |
|  | GAG | 350 | 0.53 |  |
| Cys | UGU | 299 | 1.33 |  |
|  | UGC | 151 | 0.67 | trnC-GCA |
| TER | UGA | 272 | 1.08 |  |
| Trp | UGG | 480 | 1.00 | trnW-CCA |
| Arg | CGU | 326 | 1.06 | trnR-ACG |
|  | CGC | 104 | 0.34 |  |
|  | CGA | 383 | 1.24 |  |
|  | CGG | 149 | 0.48 |  |
| Ser | AGU | 417 | 1.05 |  |
|  | AGC | 193 | 0.49 | trnS-GCU |
| Arg | AGA | 634 | 2.05 | trnR-UCU |
|  | AGG | 256 | 0.83 |  |
| Gly | GGU | 551 | 1.22 |  |
|  | GGC | 214 | 0.47 |  |
|  | GGA | 684 | 1.52 |  |
|  | GGG | 356 | 0.79 |  |

Table S5. Long-Repeats distribution within A. spinosa cp genome.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Repeat length | Start | Match direction | Repeat length | Start2 | Distance | e-value |
| 30 | 57759 | F | 30 | 89259 | -1 | 5.54e-07 |
| 30 | 114661 | F | 30 | 116885 | -2 | 2.41e-05 |
| 30 | 4685 | F | 30 | 4727 | -3 | 6.75e-04 |
| 30 | 65506 | F | 30 | 65548 | -3 | 6.75e-04 |
| 30 | 149170 | F | 30 | 149176 | -3 | 6.75e-04 |
| 30 | 110140 | P | 30 | 149271 | -1 | 5.54e-07 |
| 30 | 106570 | P | 30 | 106570 | -2 | 2.41e-05 |
| 30 | 127468 | P | 30 | 127468 | -2 | 2.41e-05 |
| 30 | 4685 | P | 30 | 65506 | -3 | 6.75e-04 |
| 30 | 4727 | P | 30 | 65548 | -3 | 6.75e-04 |
| 30 | 14210 | P | 30 | 111848 | -3 | 6.75e-04 |
| 30 | 110140 | P | 30 | 120047 | -3 | 6.75e-04 |
| 31 | 42013 | C | 31 | 85378 | -3 | 1.87e-04 |
| 31 | 157161 | F | 31 | 157183 | -3 | 1.87e-04 |
| 31 | 76762 | P | 31 | 102871 | -3 | 1.87e-04 |
| 32 | 120050 | F | 32 | 149274 | -3 | 5.15e-05 |
| 32 | 76525 | P | 32 | 124945 | -2 | 1.72e-06 |
| 33 | 23266 | F | 33 | 23298 | -1 | 9.52e-09 |
| 33 | 46932 | F | 33 | 46964 | -1 | 9.52e-09 |
| 33 | 152088 | F | 33 | 152107 | -3 | 1.42e-05 |
| 33 | 154426 | F | 33 | 154453 | -3 | 1.42e-05 |
| 33 | 23266 | P | 33 | 46932 | -1 | 9.52e-09 |
| 33 | 23298 | P | 33 | 46964 | -1 | 9.52e-09 |
| 33 | 123220 | R | 33 | 123220 | -2 | 4.57e-07 |
| 35 | 149171 | R | 35 | 149171 | -2 | 3.22e-08 |
| 36 | 102819 | P | 36 | 102819 | -2 | 8.52e-09 |
| 37 | 108539 | F | 37 | 108557 | -3 | 7.88e-08 |
| 39 | 34079 | F | 39 | 111851 | -2 | 1.57e-10 |
| 39 | 56026 | F | 39 | 111851 | -2 | 1.57e-10 |
| 39 | 14198 | P | 39 | 111851 | -2 | 1.57e-10 |
| 41 | 114650 | F | 41 | 116874 | -3 | 4.22e-10 |
| 42 | 34078 | F | 42 | 56025 | 0 | 3.67e-16 |
| 42 | 14196 | P | 42 | 34078 | 0 | 3.67e-16 |
| 42 | 77992 | P | 42 | 77992 | 0 | 3.67e-16 |
| 48 | 80434 | P | 48 | 80434 | 0 | 8.96e-20 |
| 49 | 151696 | P | 49 | 151696 | -3 | 1.11e-14 |
| 50 | 12505 | P | 50 | 89208 | 0 | 5.60e-21 |
| 60 | 147606 | P | 60 | 147606 | -2 | 8.50e-23 |
| 79 | 57708 | F | 79 | 89208 | -1 | 4.60e-36 |
| 79 | 12476 | P | 79 | 89208 | -1 | 4.60e-36 |
| 87 | 139400 | F | 87 | 150791 | 0 | 2.96e-43 |
| 109 | 70263 | P | 109 | 158739 | 0 | 1.68e-56 |
| 135 | 25839 | P | 135 | 44292 | 0 | 3.74e-72 |
| 25836 | 0 | P | 25836 | 44427 | 0 | 0.00e+00 |