**Table S1. Knowledgebase annotation and *ab initio* prediction of assembly genes**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Gene model\*** | **1H** | **2H** | **3H** | **4H** | **5H** | **6H** | **7H** | **Un** |
| GMAP alone | **6229** | **7091** | **7356** | **5364** | **7457** | **5862** | **7239** | **6476** |
| Augustus alone | 42711 | 50696 | 50345 | 50157 | 46721 | 45736 | 50500 | 46842 |
| Union both | 44784 | 52916 | 52661 | 51842 | 48990 | 47610 | 53079 | 49690 |
| GMAP in union | 5525 | 5769 | 6266 | 4700 | 6244 | 5051 | 6321 | 5784 |
| Augustus in union | 41828 | 49931 | 49297 | 49352 | 45767 | 44928 | 49593 | 46247 |
| Overlap union | **2569** | **2784** | **2902** | **2210** | **3021** | **2369** | **2835** | **2341** |

\*Genes of the two annotations were compared and merged using gffcompare (v0.11.2) (Pertea *et al*, 2020). When the transcrits have the same exact intron chain, they are combined and are thus reported only once in the union set.

**Reference:**

Pertea. G., and M. Pertea, 2020 GFF Utilities: GffRead and GffCompare. F1000Research 9:304