|  |  |  |
| --- | --- | --- |
| Species | Average read length (bp) | Dataset Size (bp) |
| *N. tomentosiformis* | 96.7 | 1,390,138,328 |
| *N. sylvestris* | 96.7 | 1,259,648,345 |
| *N. tabacum* (SR1) | 96.7 | 2,046,389,071 |
| TR1-A (synthetic tobacco) | 96.5 | 2,318,504,049 |

**Supplementary Table 1**: Dataset size and average read length for the four Illumina runs used in this analysis following the removal of plastid sequences