**Supplementary Material for Saenko *et al*.**

**“The draft genome sequence of the grove snail *Cepaea nemoralis*”**

**C:\Users\suzanne.saenko\ALWOP\14.Manuscripts\1.Genome\Figure S2.tifFigure S1.** Main characteristics of the predicted protein-coding genes.

**Table S1. Estimation of genome size, repeat content, and heterozygosity by GenomeScope, based on 31-mers in Illumina reads.**

|  |  |  |
| --- | --- | --- |
| Property | min | max |
| Heterozygosity (%) | 1.425 | 1.427 |
| Genome haploid length (bp) | 3,096,453,055 | 3,097,903,799 |
| Genome repeat length (bp) | 1,721,211,761 | 1,722,018,180 |
| Genome unique length (bp) | 1,375,241,294 | 1,375,885,619 |
| Model fit (%) | 94.5 | 99.5 |
| Read error rate (%) | 0.182 | 0.182 |

**Table S3. BUSCO analysis of the *C. nemoralis* genome assembly against metazoa\_odb10 dataset (n=954).**

|  |  |
| --- | --- |
| Complete BUSCOs | 832 (87.2%) |
| Complete and single-copy BUSCOs | 709 (74.3%) |
| Complete and duplicated BUSCOs | 123 (12.9%) |
| Fragmented BUSCOs | 36 (3.8%) |
| Missing BUSCOs | 86 (9.0%) |

**Table S4. RepeatMasker analysis of the *C. nemoralis* genome.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Total genome length (bp)** | 3,490,924,950 bp | | |
| **GC level** | 41.25% | | |
| **Bases masked** | 2,667,955,916 | | 76.42% |
|  | | | |
| **Type** | **Number of elements\*** | **Length (bp)** | **Sequence (%)** |
| **Retroelements** | **3,372,460** | **1,399,286,264** | **40.08** |
| **SINEs:** | 342,120 | 46,841,651 | 1.34 |
| Penelope | 88,259 | 27,930,778 | 0.80 |
| **LINEs:** | 2,820,864 | 1,065,797,007 | 30.53 |
| RTE/Bov-B | 1,922,793 | 657,283,360 | 18.83 |
| R1/LOA/Jockey | 378,298 | 203,267,666 | 5.82 |
| L2/CR1/Rex | 91,512 | 44,303,678 | 1.27 |
| R2/R4/NeSL | 15,845 | 6,537,107 | 0.19 |
| L1/CIN4 | 567 | 106,812 | <0.01 |
| **LTR elements:** | 209,476 | 286,647,606 | 8.21 |
| Gypsy/DIRS1 | 153,922 | 278,351,598 | 7.97 |
| Retroviral | 25,408 | 1,683,511 | 0.05 |
| BEL/Pao | 2,694 | 334,881 | 0.01 |
| **DNA transposons:** | **443,363** | **157,640,040** | **4.52** |
| Tc1-IS630-Pogo | 162,488 | 58,611,608 | 1.68 |
| hobo-Activator | 147,068 | 50,294,093 | 1.44 |
| PiggyBac | 7,768 | 3,397,486 | 0.10 |
| Tourist/Harbinger | 8,622 | 746,053 | 0.02 |
| **Unclassified** | **4,400,828** | **924,623,525** | **26.49** |
| **Total interspersed repeats** | | **2,481,549,829** | **71.09** |
|  |  |  |  |
| Small RNA | 444,489 | 71,786,395 | 2.06 |
| Satellites | 32,721 | 9,713,785 | 0.28 |
| Simple repeats | 1,181,948 | 131,714,344 | 3.77 |
| Low complexity | 53,145 | 10,209,586 | 0.29 |

\* Most repeats fragmented by insertions or deletions have been counted as one element.