# Genomics reveals broad hybridization in deeply divergent Palearctic grass and water snakes (*Natrix* spp.)

Yannis Schöneberg, Sven Winter, Oscar Arribas, Matteo Riccardo Di Nicola, Maya Master,

John Benjamin Owens, Michail Rovatsos, Wolfgang Wüster, Axel Janke, Uwe Fritz

## Molecular Phylogenetics and Evolution

## Supplementary Material 1

**Table S1**

Collection and sampling information for the studied snakes. The specimen of *Natrix natrix vulgaris* is the name-bearing neotype museum specimen for the taxon, collected in April 2001 (Fritz and Schmidtler, 2020).

|  |  |  |
| --- | --- | --- |
| Taxon | Specimen ID | Locality |
| *N. astreptophora* | MTD T 17821 | Spain: León: Puerto de las Señales |
| *N. h. helvetica* | Natrix-HFL1 | Wales: Gwynedd: Penrhyndeudraeth |
| *N. h. cetti* | MTD T 19155 | Italy: Cagliari: Sinnai near Cagliari, Sardinia |
| *N. n. natrix* | MTD T 20734 | Germany: Saxony: Liegau-Augustusbad |
| *N. n. vulgaris* | NMW 36405 | Austria: Niederösterrreich: Oberedlitz, Thaya |
| *N. maura* | MTD T 10636 | Portugal: Faro: near Alte |
| *N. tessellata* | MTD T 11214 | Hungary: Veszprém: Tihany |

MTD T: Museum of Zoology, Senckenberg Dresden (tissue collection); NMW: Natural History Museum Vienna (herpetological collection).

**Table S2**

Calibration points used for the dated phylogeny.

|  |  |  |  |
| --- | --- | --- | --- |
| Node | Fossil calibration | Age [Mya] | Reference |
| *N. h. helvetica + N. h. cetti* | Oldest *Natrix* fossil from Sardinia | 3.6 | Delfino et al. (2011) |
| *N. tessellata +* grass snake clade | Oldest record *of Natrix longivertebrata* | 16.9-16.0 | Vasilyan et al. (2022) |

**Table S3**

Comparison of selected values from the Quast report and BUSCO analysis for the different *N. n. natrix* assemblies using Raven, wtdbg2, and Flye prior to polishing.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Raven | wtdbg2 | Flye |
| Number of contigs | 5,528 | 7,098 | 12,785 |
| Total length | 1,670,916,630 | 1,655,683,932 | 1,736,742,333 |
| Largest contig | 6,821,274 | 10,755,250 | 17,021,601 |
| N50 | 634,217 | 1,050,695 | 1,581,800 |
| L50 | 713 | 443 | 295 |
| GC-content [%] | 41.41 | 41.16 | 41.5 |
| Complete BUSCOs [%] | 83.2 | 81.6 | 91 |
| Single copied BUSCOs [%] | 81.8 | 80.7 | 89.4 |
| Duplicated BUSCOs [%] | 1.4 | 0.9 | 1.6 |
| Fragmented BUSCOs [%] | 4.2 | 4.4 | 2 |
| Missing BUSCOs [%] | 12.6 | 14 | 7 |
| Total searched | 7480 | 7480 | 7480 |

**Table S4**

Comparison of selected values of the Quast report and BUSCO analysis for the two different *N. h. helvetica* assemblies using Raven and wtdbg2 prior to polishing.

|  |  |  |
| --- | --- | --- |
| Assembly | Polished Raven | Polished wtdbg2 |
| contigs | 1,215 | 10,282 |
| Largest contig | 21,110,391 | 15,563,773 |
| Total length | 1,783,919,027 | 1,878,028,187 |
| N50 | 4,622,328 | 2,110,479 |
| L50 | 105 | 229 |
| GC [%] | 41.55 | 41.61 |
| Complete BUSCOs [%] | 92.8 | 89.7 |
| Single copied BUSCOs [%] | 90.3 | 87.4 |
| Duplicated BUSCOs [%] | 2.5 | 2.3 |
| Fragmented BUSCOs [%] | 1.1 | 1.6 |
| Missing BUSCOs [%] | 6.1 | 8.7 |
| Total searched | 7480 | 7480 |

**Table S5**

Quality statistics for the final de-Novo assemblies of *N. n. natrix* and *N. h. helvetica*. Selected values of the Quast report and the results of the BUSCO analysis on the Sauropsida\_odb10 set are shown. The *N. h. helvetica* assembly is the more contiguous one.

|  |  |  |
| --- | --- | --- |
|  | *N. n. natrix* | *N. h. helvetica* |
| Number of contigs | 12,785 | 1,215 |
| Total length | 1,736,742,333 | 1,783,919,027 |
| Largest contig | 17,021,601 | 21,110,391 |
| N50 | 1,581,800 | 4,622,328 |
| L50 | 295 | 105 |
| GC-content [%] | 41.5 | 41.55 |
| Complete BUSCOs [%] | 91 | 92.8 |
| Single copied BUSCOs [%] | 89.4 | 90.3 |
| Duplicated BUSCOs [%] | 1.6 | 2.5 |
| Fragmented BUSCOs [%] | 2 | 1.1 |
| Missing BUSCOs [%] | 7 | 6.1 |
| Total searched | 7480 | 7480 |

**Table S6**

Results of repeat masking for the de-novo assemblies. The amount of repeats is given as percentage of the total assembly.

|  |  |  |
| --- | --- | --- |
|  | *N. h. helvetica* | *N. n. natrix* |
| **Total Masked** | 50.87% | 50.84% |
| **Retroelements** | 25.55% | 23.41% |
| SINEs: | 1.45% | 1.33% |
| Penelope | 0.72% | 0.81% |
| LINEs: | 18.13% | 16.66% |
| CRE/SLACS | 0.00% | 0.00% |
| L2/CR1/Rex | 11.82% | 10.24% |
| R1/LOA/Jockey | 0.00% | 0.00% |
| R2/R4/NeSL | 0.77% | 0.87% |
| RTE/Bov-B | 1.53% | 1.36% |
| L1/CIN4 | 3.21% | 3.30% |
| LTR elements: | 5.97% | 5.41% |
| BEL/Pao | 0.00% | 0.02% |
| Ty1/Copia | 0.76% | 0.63% |
| Gypsy/DIRS1 | 4.44% | 4.07% |
| Retroviral | 0.77% | 0.66% |
| **DNA transposons** | 4.49% | 2.98% |
| hobo-Activator | 2.60% | 1.66% |
| Tc1-IS630-Pogo | 0.66% | 0.50% |
| En-Spm | 0.00% | 0.00% |
| MuDR-IS905 | 0.00% | 0.00% |
| PiggyBac | 0.00% | 0.01% |
| Tourist/Harbing | 0.22% | 0.00% |
| Other | 0.00% | 0.00% |
| **Rolling-circles** | 0.11% | 0.28% |
| **Unclassified** | 20.12% | 21.28% |
| **Total interspersed** | 50.17% | 47.67% |
| **Small RNA** | 0.17% | 0.01% |
| **Satellites** | 0.41% | 0.47% |
| **Simple repeats** | 0.00% | 2.07% |
| **Low complexity** | 0.00% | 0.34% |

**Table S7**

The pseudo-likelihood support values for the different number of reticulations tested in the SNAQ analysis. Adding more reticulations than three did not improve the support. Therefore, three reticulations were supported.

|  |  |
| --- | --- |
| # reticulations | -Ploglik |
| 0 | 77.7029192 |
| 1 | 46.151134 |
| 2 | 11.8880193 |
| 3 | 6.54374466 |
| 4 | 6.51174948 |
| 5 | 6.52191005 |

**Table S8**

The pseudo-likelihood support values fort he different number of reticulations tested in the PhyloNet analysis. The analysis does not support any reticulations.

|  |  |
| --- | --- |
| # reticulations | -Log Lik |
| 0 | -7488.979783 |
| 1 | -7377.272454 |
| 2 | -7479.430616 |
| 3 | -7377.068173 |
| 4 | -7377.492289 |
| 5 | -7377.278559 |
| 6 | -7377.301152 |
| 7 | -7368.977638 |
| 8 | -7377.445666 |
| 9 | -7488.601221 |
| 10 | -7377.276744 |

**Table S9**

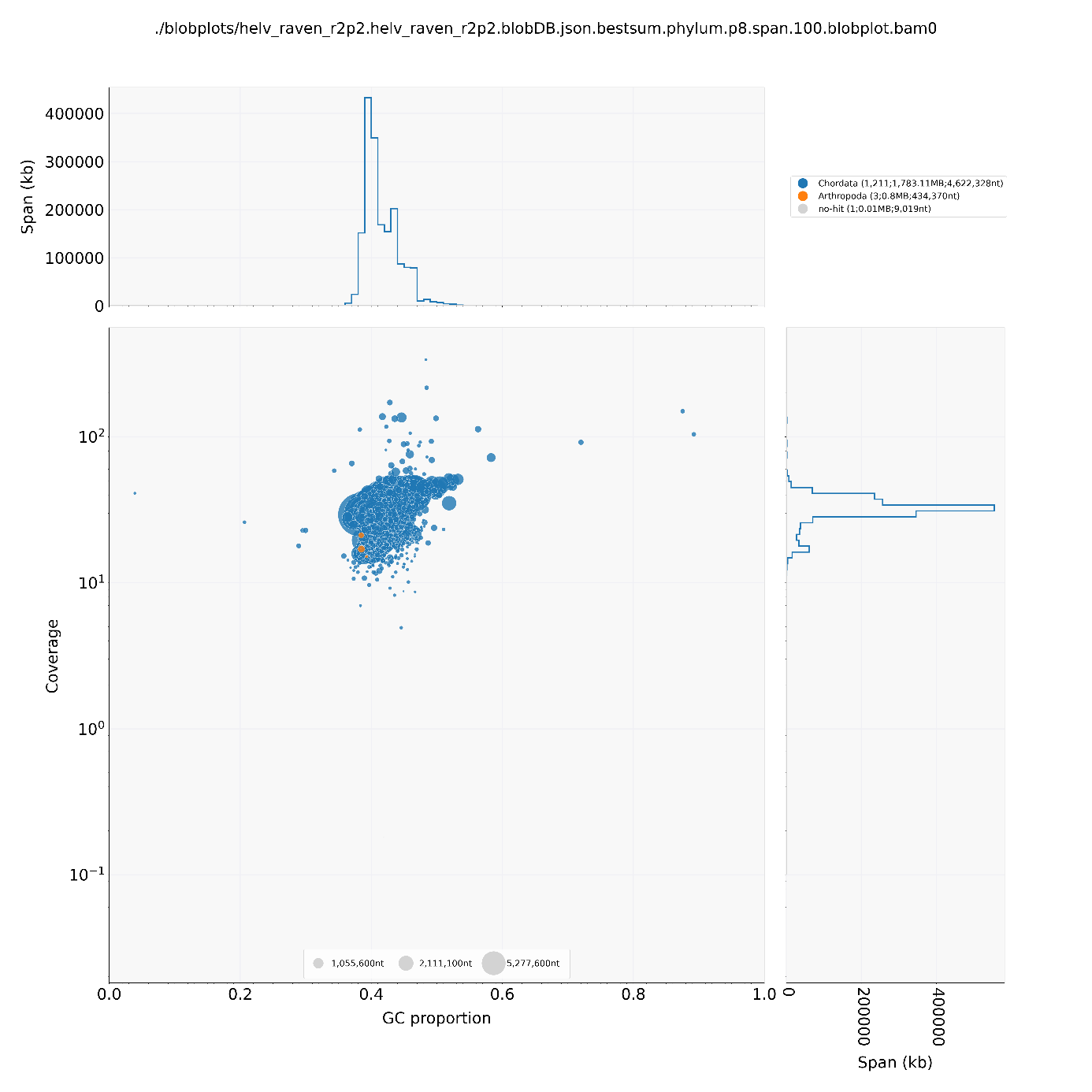
HyDe results showing only significant results. Abbreviations: *ast: Natrix astreptophora; hel: N. h. helvetica; hce: N. h. cetti; nat: N. n. natrix; nvu: N. n. natrix; mau: N. maura; tes: N. tessellata*.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| P1 | Hybrid | P2 | Zscore | Pvalue | Gamma | AAAA | AAAB | AABA | AABB | AABC | ABAA | ABAB | ABAC | ABBA | BAAA | ABBC | CABC | BACA | BCAA | ABCD |
| *ast* | *hce* | *mau* | 89.4 | 0 | 0.990 | 769221168 | 103214766 | 22049558.9 | 11445170.9 | 11992156 | 27780436.2 | 11023317.7 | 14621871.8 | 51844385.6 | 337271526 | 101384294 | 11340432.4 | 17728279.6 | 21214280.2 | 10404530 |
| *ast* | *hce* | *nat* | 848.8 | 0 | 0.830 | 815065909 | 57915051.4 | 22438379.5 | 15189129.6 | 9489637.29 | 32452121.9 | 11074881.5 | 11528912.2 | 31097356 | 400762608 | 60275173 | 9005802.54 | 18432993.1 | 28555766.1 | 7953363.98 |
| *ast* | *hce* | *nvu* | 1771.1 | 0 | 0.684 | 808839965 | 64093824.2 | 20928790 | 18367818.7 | 7678898.71 | 37853178.7 | 9127037.78 | 7934147.4 | 29148405.3 | 406362501 | 56482710.8 | 6963074.04 | 16117030.6 | 34660224.3 | 5924397.85 |
| *ast* | *hce* | *tes* | 650.1 | 0 | 0.922 | 783562063 | 88976551.3 | 21794070.6 | 13105828.1 | 10893584.3 | 31559878.5 | 10065856.8 | 12106585.9 | 45857109.2 | 355406707 | 89544361.4 | 10285848 | 17218793.4 | 24485765.3 | 9310337.28 |
| *ast* | *hel* | *hce* | 1607.3 | 0 | 0.297 | 863198152 | 31984096.8 | 16762664 | 28020453.5 | 8048746.57 | 50741953.4 | 8981342.18 | 9443010.3 | 17019756.2 | 488004973 | 33327439.3 | 8229790.8 | 16087636.9 | 55322032.8 | 7726423.45 |
| *ast* | *hel* | *mau* | 365.7 | 0 | 0.963 | 779529571 | 114987909 | 18606198.5 | 14927265.1 | 17330205.8 | 32812894.1 | 13025341.8 | 21359847.8 | 61882905.3 | 352494828 | 121980875 | 16863025.7 | 21097078.4 | 28963032.6 | 16507077.5 |
| *ast* | *hel* | *nat* | 1425.9 | 0 | 0.785 | 825836629 | 68976127.3 | 18280985.9 | 19622311.9 | 13834033.2 | 38660935.2 | 11953663 | 17457115.9 | 40013955.8 | 418527488 | 78702605.8 | 13719135.2 | 19628201.2 | 38507617.5 | 13322163.3 |
| *ast* | *hel* | *nvu* | 2162.3 | 0 | 0.691 | 819299761 | 75544185.4 | 17160307.8 | 22759748.8 | 11910790.8 | 44059584.8 | 10239725.2 | 13865944 | 38298578.6 | 423683587 | 75355476 | 11680759.2 | 17756063.4 | 44649260.3 | 11278113.8 |
| *ast* | *hel* | *tes* | 915.3 | 0 | 0.902 | 793620197 | 100984407 | 18584513.5 | 16687002 | 15848929.5 | 36799891.3 | 11945270.4 | 18709675.8 | 55774083.3 | 370964624 | 109881075 | 15673276.9 | 20326188.4 | 32680342.6 | 15263741.5 |
| *hce* | *hel* | *mau* | 330.8 | 0 | 0.976 | 787490839 | 125835307 | 16425159.3 | 10973944.8 | 12796193.5 | 24902396.9 | 9492856.73 | 14196286.5 | 68927097.7 | 364599990 | 135978930 | 12151129.1 | 16580633.3 | 20963255.6 | 11861842.1 |
| *hce* | *hel* | *nvu* | 703.4 | 0 | 0.901 | 836314546 | 77429255.6 | 17209934.6 | 13316897.6 | 10907973.2 | 27186789.6 | 10036690.9 | 12607413.4 | 39856360.6 | 452359692 | 78542732.2 | 10496984.3 | 17487416.4 | 25841820.3 | 10209301.9 |
| *mau* | *ast* | *nvu* | 751.4 | 0 | 0.138 | 740122875 | 47453500 | 27927134.7 | 38143029.6 | 13808543.1 | 85453116.3 | 13380305.8 | 16732707.2 | 17339216.6 | 306603094 | 30997665.5 | 16121539.9 | 22444269.6 | 74284258.3 | 12602309.7 |
| *mau* | *hce* | *nvu* | 841.4 | 0 | 0.114 | 748739825 | 48015307 | 22761436.6 | 46464959.7 | 15269811.8 | 97208965.7 | 12466877.4 | 18446028.2 | 16845545.9 | 320057824 | 32288577.8 | 15206460.2 | 20254608.7 | 90635303.3 | 13937019.3 |
| *mau* | *hel* | *nvu* | 995.9 | 0 | 0.136 | 757270588 | 55005571.6 | 21276603.8 | 53316860.9 | 22564732.8 | 105695882 | 16038475.5 | 27444164.8 | 21896960.8 | 331394351 | 42895768.8 | 22804800 | 26483158.3 | 105204268 | 22035370.9 |
| *mau* | *nat* | *nvu* | 288.9 | 0 | 0.037 | 738673114 | 30266692.8 | 21884477.3 | 37159377 | 9521517.31 | 80965882.9 | 7377503.79 | 5871958.29 | 8511072.56 | 300036054 | 13698496.3 | 6980216.96 | 12542915 | 66427616.5 | 3710719.94 |
| *nat* | *ast* | *mau* | 544.5 | 0 | 0.893 | 748865555 | 84042396.6 | 27207195.6 | 16122753 | 15439241.7 | 38354257.6 | 13295497 | 17159113.5 | 36933285.2 | 307711203 | 71837506.4 | 14668476.3 | 22633706 | 30036895.8 | 13403615.9 |
| *nat* | *ast* | *tes* | 452.5 | 0 | 0.884 | 762327440 | 71090893.5 | 28972523.7 | 16511763.5 | 14814389.2 | 39384832.3 | 14129092.8 | 16824201.6 | 32287562.9 | 324194777 | 61530619.6 | 15509085.3 | 24184635.6 | 31054399.9 | 13053136.4 |
| *nat* | *hce* | *mau* | 596.1 | 0 | 0.895 | 755358800 | 91408910.7 | 24561757.6 | 17772567.1 | 17962096.6 | 41437500.9 | 14525197.2 | 22311027.3 | 42273837 | 317477140 | 82506032.4 | 17783896.6 | 23180259.2 | 34065717.8 | 16631521.1 |
| *nat* | *hce* | *tes* | 999.2 | 0 | 0.804 | 770217052 | 76696467.3 | 24566591.8 | 19328752.7 | 16837037.3 | 44742734.7 | 13907728 | 20059255.6 | 36178025.3 | 335991947 | 70524360.6 | 16706962.4 | 23057977.3 | 37191365.6 | 15576940.2 |
| *nat* | *hel* | *hce* | 765.5 | 0 | 0.092 | 837086607 | 27519925.1 | 17152305.5 | 44660133.6 | 10526590.4 | 76592988.9 | 9712428.32 | 12406419.6 | 13267508.1 | 437884284 | 25832388.1 | 10822809.8 | 17072720.2 | 88021722.9 | 10219495.5 |
| *nat* | *hel* | *mau* | 780.0 | 0 | 0.869 | 764131484 | 100003814 | 22793324.1 | 22671862.6 | 25478422.5 | 48144021.4 | 17944919 | 31227113.2 | 49248583.7 | 329106536 | 97216675.7 | 25113484.6 | 29142687 | 44406536.5 | 24684153.8 |
| *nat* | *hel* | *nvu* | 449.8 | 0 | 0.260 | 815004438 | 49511618.1 | 23003214.1 | 27110618.3 | 21959930.4 | 48342184 | 19646585.8 | 30456577.1 | 22264582.2 | 410361893 | 44088767.9 | 21837994.3 | 28713357.1 | 53584293.2 | 21471834.7 |
| *nat* | *hel* | *tes* | 1178.6 | 0 | 0.784 | 778860539 | 85371077.5 | 22877109.7 | 24260763.8 | 24090450.3 | 51501773.7 | 17138655.2 | 28960233 | 42963845.7 | 348114446 | 84782190 | 24022076.5 | 28565448.2 | 47784102.8 | 23544288 |
| *nat* | *nvu* | *hce* | 212.5 | 0 | 0.154 | 802821055 | 20560874.6 | 36700131.5 | 15844895.9 | 2174718.51 | 35918602.5 | 10727101.4 | 1937002.14 | 11662115.1 | 388976564 | 14206409.6 | 8974808.24 | 20816266.7 | 30711288.8 | 1375052.24 |
| *nat* | *nvu* | *tes* | 30.4 | 0 | 0.994 | 753460972 | 68237160.6 | 32256838.2 | 9747340.41 | 7700907.1 | 26039658.6 | 9613764.68 | 7913948.15 | 32787864.5 | 319157026 | 57853304.8 | 11622733.1 | 16994105.3 | 17480014.7 | 5755068.29 |
| *tes* | *ast* | *mau* | 1299.6 | 0 | 0.620 | 726978614 | 74199914.3 | 29639558.6 | 24245813.3 | 21285957.4 | 60217835.6 | 16267446.7 | 23960800.9 | 29282134.1 | 284155185 | 56722555.6 | 19818910.5 | 26749792.6 | 45399748 | 18418545.5 |
| *tes* | *ast* | *nvu* | 415.4 | 0 | 0.111 | 754001798 | 48205430.6 | 31575257.5 | 33071855.8 | 13607051.8 | 72222996.5 | 15051106.1 | 16253851.3 | 17301225.6 | 323786454 | 32156901.4 | 14456226.1 | 25772531.1 | 64117521.4 | 12204033.6 |
| *tes* | *hce* | *mau* | 1343.1 | 0 | 0.657 | 733103066 | 81588792.7 | 27264313 | 25927426.4 | 23458770 | 63571114.4 | 17528580.1 | 28762505.6 | 33582032.5 | 292587762 | 65538706.4 | 23408051.3 | 28301898.1 | 50433978.5 | 22250288.5 |
| *tes* | *hce* | *nvu* | 1010.6 | 0 | 0.162 | 765554520 | 49358217.8 | 22965420.4 | 39741650.2 | 14603537.3 | 80534208 | 12483646.9 | 17504643.7 | 17771164.7 | 340435853 | 34163954.4 | 14495153.3 | 20517841.5 | 77400284.9 | 13301162.9 |
| *tes* | *hel* | *hce* | 376.7 | 0 | 0.032 | 805060362 | 25428629.6 | 16978012.4 | 61429271.5 | 11830047.2 | 108371040 | 9639129.53 | 13834672.1 | 11341333.3 | 387240640 | 21956486.7 | 12186921.2 | 16867218.7 | 121151716 | 11524806.6 |
| *tes* | *hel* | *mau* | 1428.4 | 0 | 0.643 | 741419753 | 88778922.5 | 26007198.3 | 31423121.4 | 31949141.6 | 70788413 | 21544576.6 | 38653039.7 | 39299680.9 | 302705284 | 77664103 | 31649633.1 | 35345968.5 | 61854955.7 | 31164977 |
| *tes* | *hel* | *nvu* | 1185.8 | 0 | 0.186 | 773947733 | 56413241.4 | 21571797.1 | 46413454.7 | 21874645.1 | 89113084.6 | 15875479.5 | 26478183.5 | 22864392.4 | 352248172 | 45042845.4 | 21820959.5 | 26311382 | 91536527.6 | 21307264.9 |
| *tes* | *nat* | *mau* | 975.1 | 0 | 0.674 | 717105434 | 67016147.9 | 32867891.2 | 18658368.3 | 14128293.2 | 52300446.2 | 13506685.4 | 15082927 | 24136536.2 | 271425027 | 45412725.1 | 12952626.8 | 22307529.6 | 33557315.3 | 10591899.5 |
| *tes* | *nvu* | *mau* | 740.2 | 0 | 0.728 | 709529435 | 70355900.5 | 42428390.1 | 19654908.3 | 15744921.7 | 53479626.5 | 15636351.8 | 17807944 | 26395642.5 | 271451640 | 49689869 | 15466671.4 | 26103824.3 | 36199693.7 | 12851696.1 |

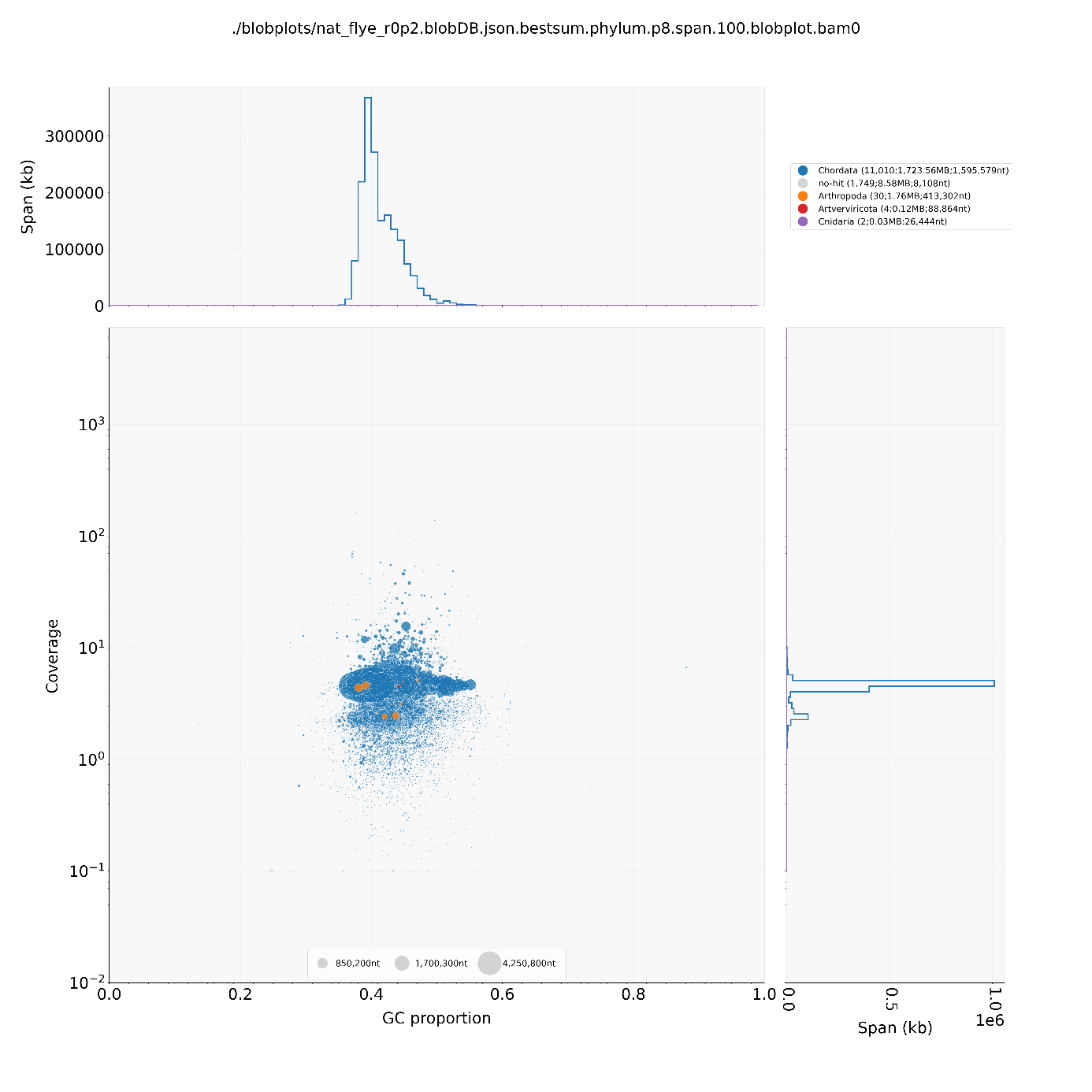
**Table S10**

Colubrid and natricid genomes used for comparison to our de novo assemblies available on NCBI with their NCBI accession numbers and assembly lengths.

|  |  |  |  |
| --- | --- | --- | --- |
| Accession number | Taxon | Total length [bp] | Publication |
| GCA\_022577455.1 | *Arizona elegans* | 1,842,534,626 | — |
| GCA\_019457695.1 | *Chrysopelea ornata* | 1,334,265,196 | — |
| GCF\_001185365.1 | *Pantherophis guttatus* | 1,706,961,544 | Ullate-Agote et al., 2014. Int. J. Dev. Biol. 58, 881–888. https://doi.org/10.1387/ijdb.150060at |
| GCA\_012654085.1 | *Pantherophis obsoletus* | 1,692,461,844 | — |
| GCA\_019677565.1 | *Pituophis catenifer pumilus* | 1,516,097,738 | — |
| GCA\_012654045.1 | *Ptyas mucosa* | 1,721,466,642 | — |
| GCF\_009769535.1 | *Thamnophis elegans* | 1,672,190,305 | — |
| GCF\_001077635.2 | *Thamnophis sirtalis* | 1,424,897,867 | Perry et al., 2018. Genome Biol. Evol. 10, 2110–2129. https://doi.org/10.1093/gbe/evy157 |

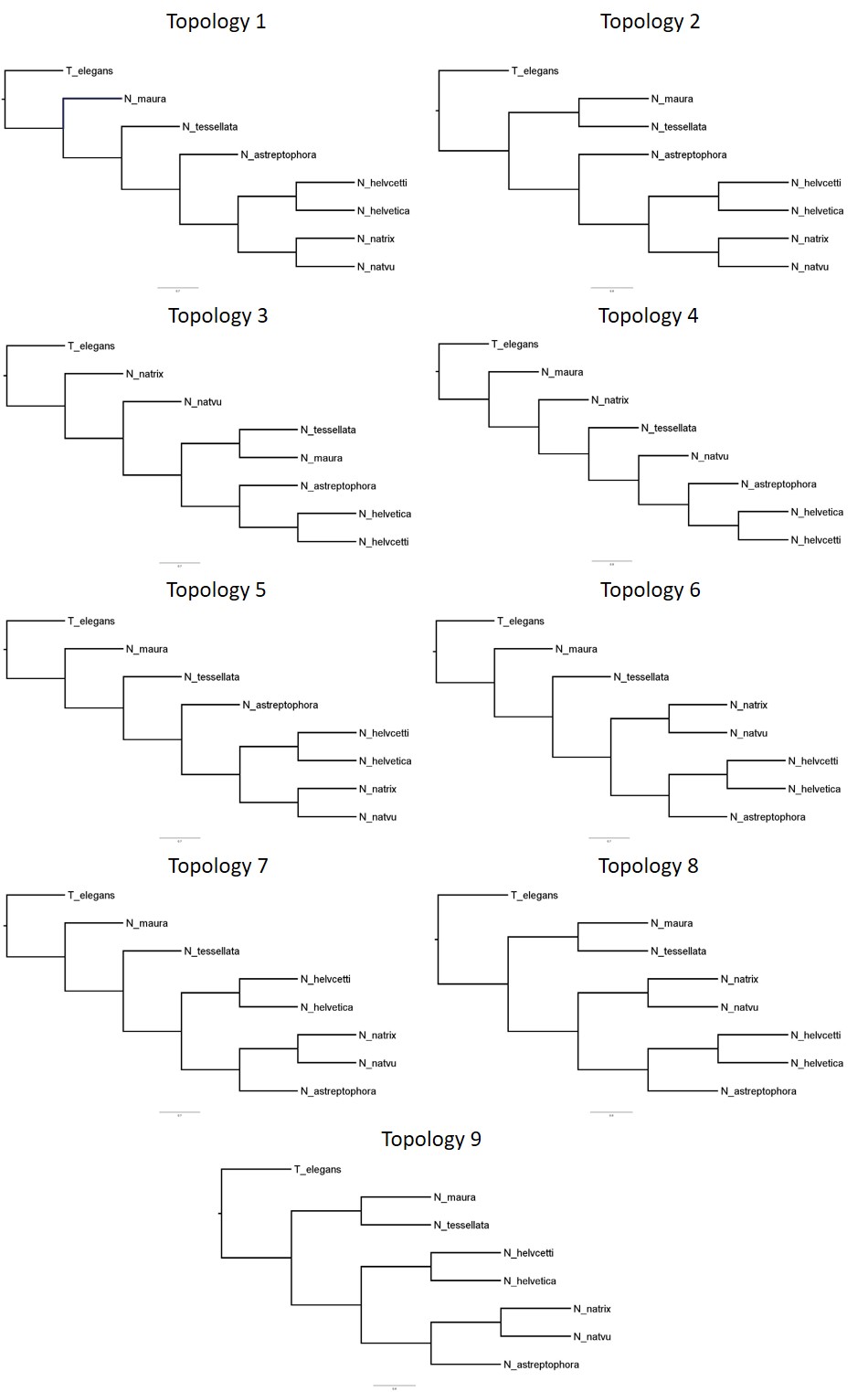
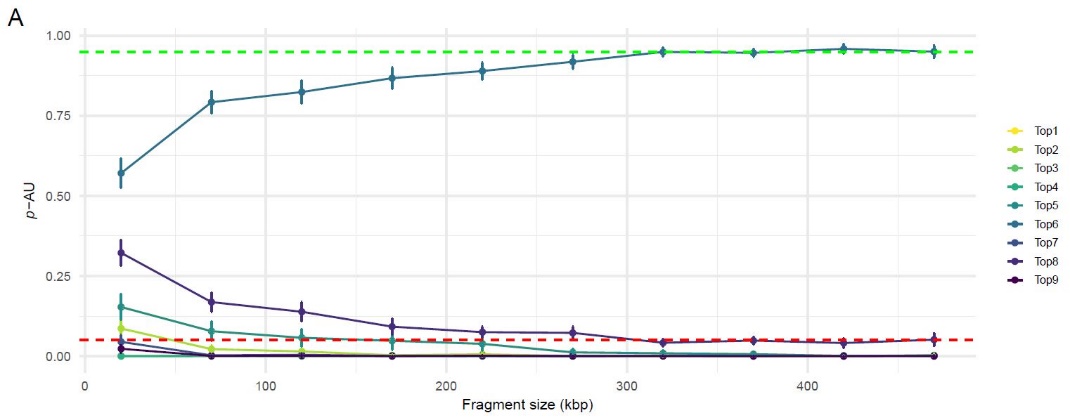


**Fig. S1.** Blobtools plot of the de-novo assembly of *Natrix helvetica helvetica*. It shows only negligible signal for contamination, which could result from small contig length and a sampling bias in NCBI. Furthermore, the putatively contaminated contigs (red) are located in the main cloud. Therefore, the assembly was not filtered.

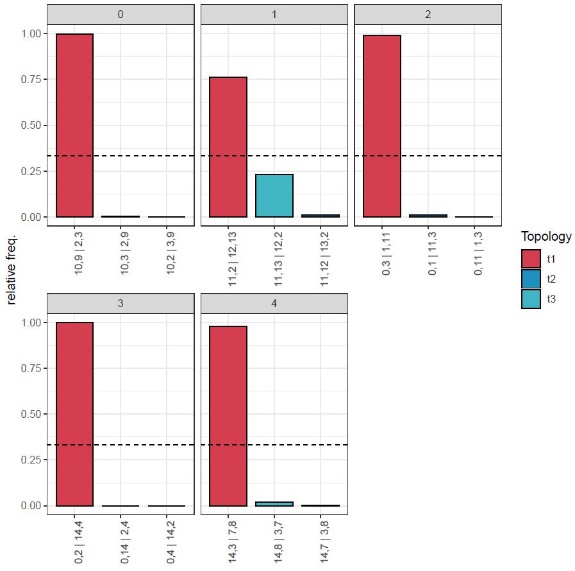


**Fig. S2.** Blobtools plot of the de-novo assembly of *Natrix natrix natrix*. It shows only negligible signal for contamination, which could result from small contig length and a sampling bias in NCBI. Furthermore, the putatively contaminated contigs (red) are located in the main cloud. Therefore, the assembly was not filtered.

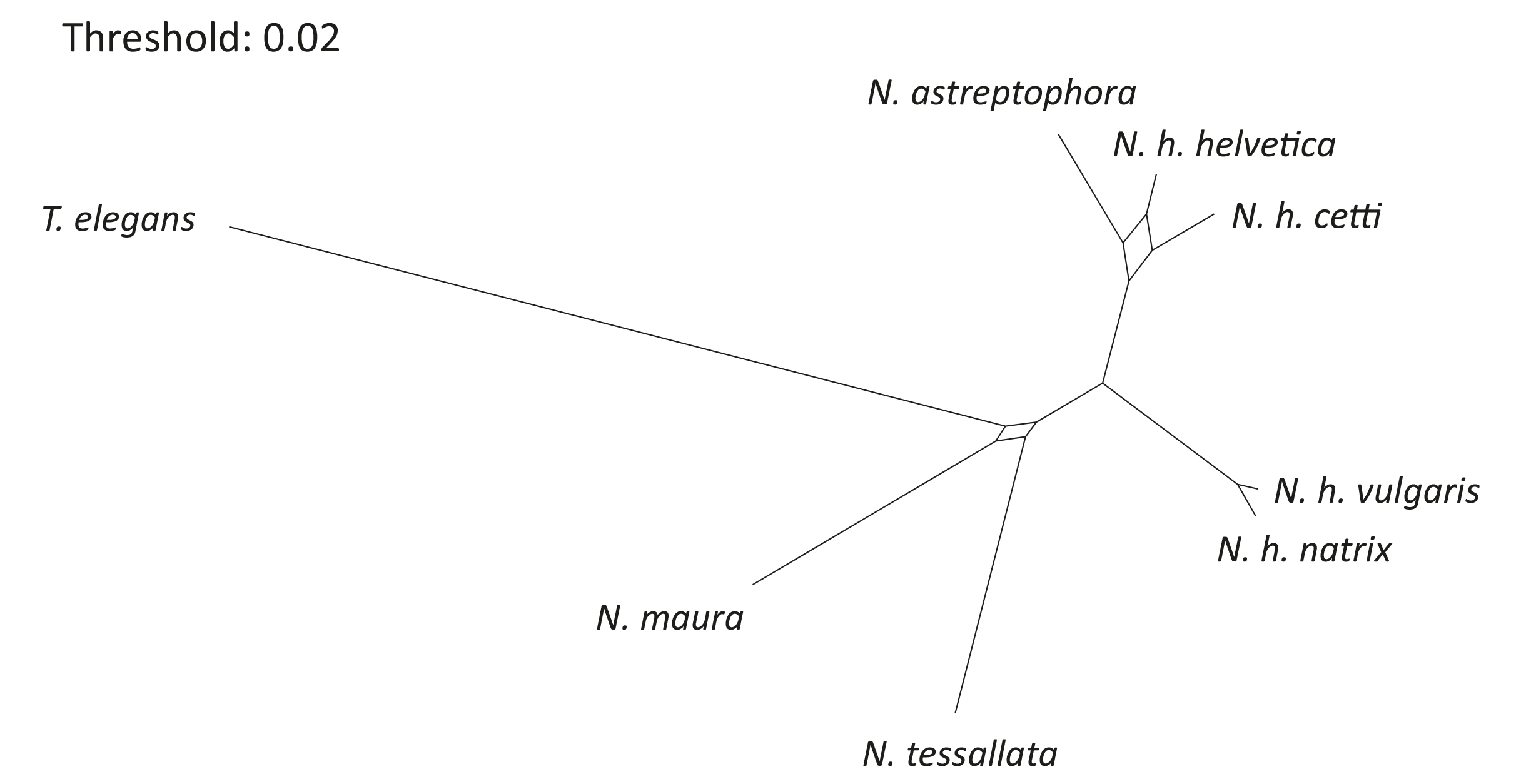
**Fig. S3.** Results of the AU test of nine different topologies for different fragment sizes between 20 kb and 470 kb. This analysis tests whether one of the given topologies is supported significantly (95 % p-AU) over the others. In this case, topology 6 (same as Fig. 2) was supported significantly at a fragment length of 320 kb or higher. Therefore, we chose this length as fragment length for the genome fragment analysis.

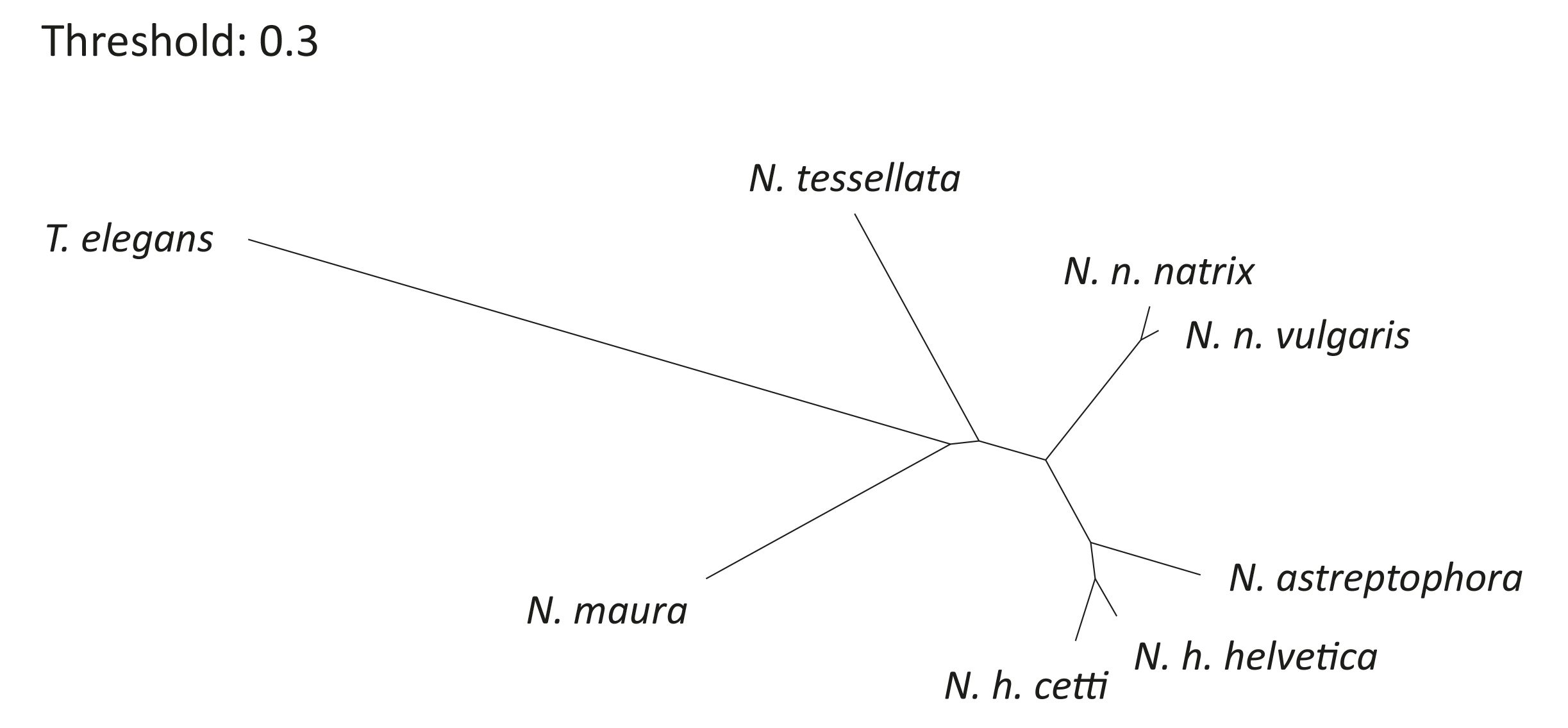
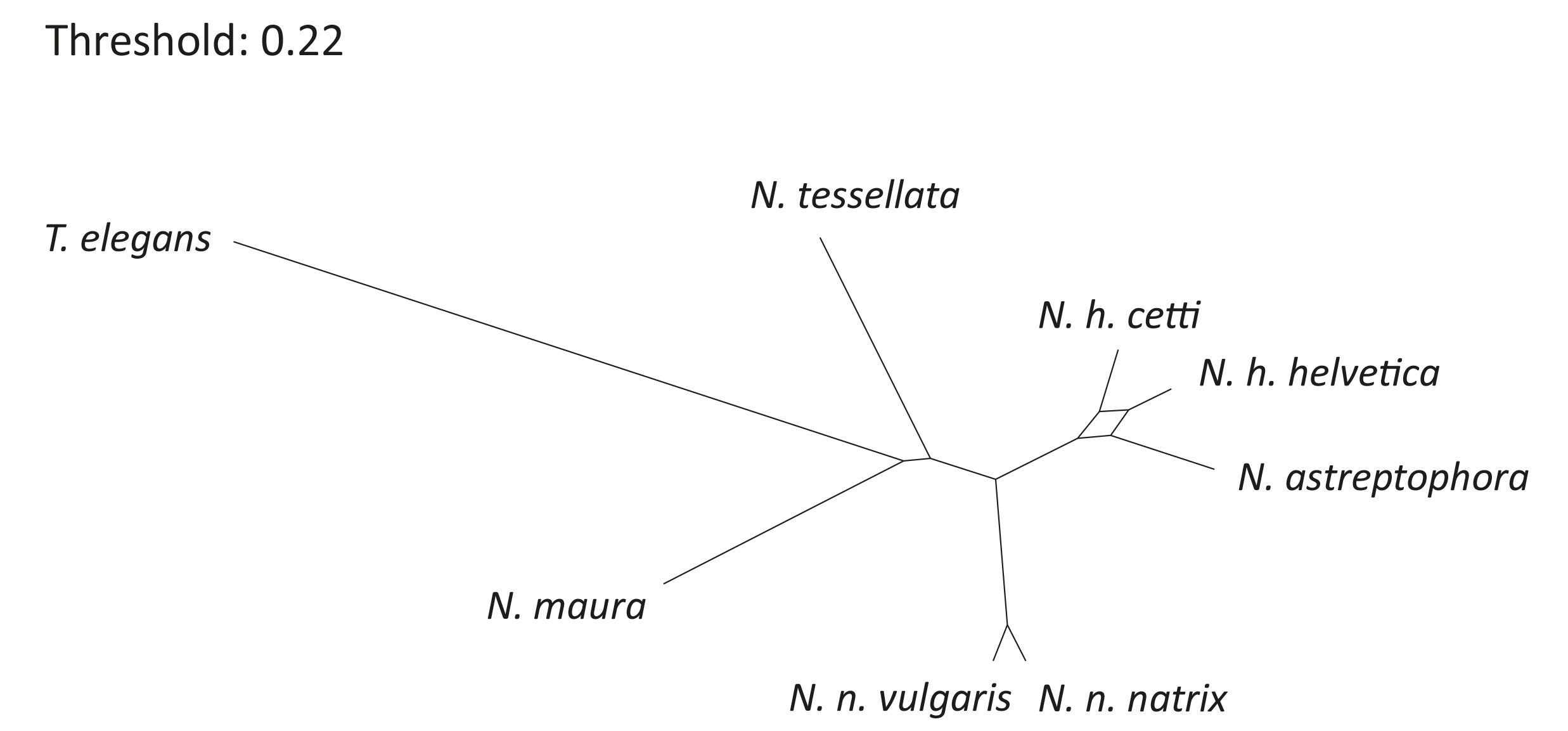


**Fig. S4.** Tested topologies in pAU test (Fig. S3). All nine topologies were tested to choose the best genome fragment length. At a length of ~320 kb, topology 6 was significantly supported (Fig. S3).



**Fig. S5.** Left: Quartet Frequencies in the genome fragments for the Astral species tree. These statistics show the frequencies of each possible topology for focal quartets. The numbers of the quartets given on the x-axis of each graph refers to the branch numbers given on the right. Right: Tree topology of the species tree by Astral showing the branch numeration used for Quartet Frequency analysis. *N natvu* stands for *N. n. vulgaris*, and *N helvcetti* for *N. h. cetti*.



**Fig. S6*.*** Phylogenetic networks based on 860 genome fragments of 320 kb length, showing phylogenetic conflict among the species at different thresholds. The shown thresholds are the lowest ones before a reticulation disappears. The networks show conflict among *N. h. helvetica*, *N. h. cetti* and *N. astreptophora* as between *N. maura* and *N. tessellata*.