Ecology & Evolution | Oct 29, 2026 | 5 min read | By Bryan White

Aptostichus stephencolberti: Reference Genome of the Californian Trapdoor Spider Spans 4.6 Billion Base Pairs

Genetics * Molecular & Cellular Biology * Biodiversity

Sequenced under the California Conservation Genomics Project, the first family-level reference genome for Euctenizidae reveals a 4.6-gigabase assembly -- roughly 50% larger than the human genome -- packaged across 52 chromosomes that lack canonical telomeric repeats and consist of 77% repetitive DNA [1, 2] .

First Reference Genome for the Mygalomorph Family Euctenizidae

In 2008, evolutionary biologist Professor Jason E. Bond described a cryptic species of coastal dune-dwelling trapdoor spider native to California and named it Aptostichus stephencolberti in honor of television satirist Stephen Colbert [2, 3] . Measuring roughly the diameter of a quarter, this sedentary mygalomorph excavates silk-lined subterranean burrows capped by camouflaged, hinged sand-and-silk trapdoors along the narrow coastal dune corridor stretching from the San Francisco Peninsula south to Point Sur and Monterey Bay [3, 4] . Despite its modest body size and restricted geographic footprint, new genomic sequencing reveals that A. stephencolberti harbors an unexpectedly massive nuclear genome .

Published in the Journal of Heredity by a multi-institutional team led by Stephen Richards, Emma Jochim, H. Bradley Shaffer, and Jason E. Bond at the University of California, Davis and UCLA, the study delivers the first chromosome-scale reference genome for the mygalomorph spider family Euctenizidae [1, 2] . Generated as part of the California Conservation Genomics Project (CCGP), the assembly combines PacBio HiFi long-read sequencing with Dovetail Omni-C proximity-ligation chromatin scaffolding to resolve a total genome size of approximately 4.6 billion base pairs (4.6 Gbp) -- roughly 50% larger than the ~3.05 to 3.2 Gbp human genome -- organized across 52 chromosomes [1, 2] .

The C-Value Enigma in Arachnids: 77% Repetitive Elements vs. 14,067 Coding Genes

Why does a quarter-sized arachnid possess 1.5 billion more DNA base pairs and more than twice as many chromosomes as Homo sapiens ? The Aptostichus stephencolberti assembly offers a textbook illustration of the eukaryotic C-value enigma -- the observation that haploid nuclear DNA content does not correlate with organismal complexity or the number of protein-coding genes [1, 5] . Using automated NCBI RefSeq eukaryotic annotation paired with multi-tissue RNA-seq transcriptomics , Richards and colleagues identified 14,067 genes in A. stephencolberti , substantially fewer than the ~19,900 protein-coding genes annotated in the human reference genome (GRCh38/T2T-CHM13) [1] .

Instead of gene duplication , the expansion of the A. stephencolberti genome is driven overwhelmingly by non-coding mobile genetic elements. Repeat masking across the 3.63 Gbp primary scaffolded assembly (551 scaffolds; scaffold N50 of 62.2 Mbp)

revealed that approximately 77% of the spider's genome consists of interspersed repetitive sequences , including retrotransposons, DNA transposons, and tandem satellite arrays [1, 2] . By comparison, repetitive elements constitute roughly 50% to 54% of the human genome. In small, geographically isolated mygalomorph populations with low effective population sizes (N e), slightly deleterious transposable element insertions face weaker purifying selection , allowing repetitive DNA to accumulate over millions of years of evolutionary divergence [1, 5] .

| Genomic & Cytogenetic Parameter | | Trapdoor Spider (Aptostichus stephencolberti) | | Human Reference (Homo sapiens T2T-CHM13) | | Western Black Widow (Latrodectus hesperus) | | Fruit Fly Model (Drosophila melanogaster) |

| Taxonomic Clade | | Araneae: Mygalomorphae (Euctenizidae) | | Chordata: Mammalia (Hominidae) | | Araneae: Araneomorphae (Theridiidae) | | Diptera: Brachycera (Drosophilidae) |

| Estimated Haploid Genome Size | | ~4.6 Gbp (3.63 Gbp primary assembly) | | ~3.05-3.20 Gbp | | ~1.60-2.10 Gbp | | ~0.18 Gbp (180 Mbp) |

| Chromosome Number | | 52 chromosomes (scaffold N50: 62.2 Mbp) | | 23 pairs (2n = 46) | | 14 pairs (2n = 28) | | 4 pairs (2n = 8) |

| Repetitive DNA Fraction | | ~77% interspersed & tandem repeats | | ~50-54% transposable & satellite elements | | ~38-45% repetitive elements | | ~20% repetitive elements |

| Annotated Gene Count | | 14,067 genes (NCBI RefSeq + RNA-seq) | | ~19,969 protein-coding genes | | ~17,500-21,000 predicted genes | | ~13,968 protein-coding genes |

| Canonical Telomeric Repeats | | Absent (lacks canonical (TTAGG) n motif) | | Present (vertebrate (TTAGGG) n array) | | Absent / non-canonical terminal structure | | Retrotransposon-maintained (HeT-A / TART) |

Absence of Canonical Telomeric Repeats & Accelerated Chromosomal Evolution

One of the most striking cytogenetic findings in the Aptostichus stephencolberti genome is the complete absence of canonical arthropod telomeric repeats [1, 2] . In most vertebrates and insects, chromosome ends are capped by short, tandemly repeated hexanucleotide or pentanucleotide motifs -- such as vertebrate (TTAGGG) n or ancestral arthropod (TTAGG) n -- synthesized by the ribonucleoprotein enzyme telomerase to prevent chromosomal end-to-end fusion and exonucleolytic degradation during DNA replication [1, 5] .

Scanning the chromosome-level Omni-C scaffolds of A. stephencolberti confirmed that, like several recently sequenced araneomorph spiders, this basal mygalomorph lineage lacks recognizable telomerase-generated terminal repeat arrays [1, 2] . This finding reinforces the hypothesis that the loss of canonical (TTAGG) n telomeres occurred deep in the ancestral stem of all spiders (order Araneae) more than 300 million years ago [1, 5] . Without conventional telomeric caps, spiders appear to rely on alternative lengthening of telomeres (ALT), terminal satellite recombination, or

specialized retrotransposon capping mechanisms -- a structural plasticity that may help explain why spider karyotypes undergo rapid chromosomal fission, fusion, and sex-chromosome turnover across families [1, 2] .

Conservation Genomics of Fragmented California Coastal Dune Ecosystems

Beyond its value for comparative arachnid phylogenomics, the Aptostichus stephencolberti reference genome serves an urgent conservation mission under the California Conservation Genomics Project [1, 4] . Unlike web-building araneomorph spiders that disperse tens of kilometers on wind currents via silk "ballooning," A. stephencolberti juveniles walk only a few meters from their maternal burrow before establishing lifelong subterranean residences [3, 4] . This extreme philopatry produces deep phylogeographic structure and severe genetic isolation between adjacent dune pockets along the Central California coast [3] .

Today, those coastal foredunes face compounding anthropogenic pressures, including urban coastal development, recreational trampling, encroaching sea-level rise, and habitat stabilization by invasive succulent mats such as highway iceplant (Carpobrotus edulis), which binds loose sand and prevents trapdoor burrow excavation [1, 2] . By anchoring whole-genome resequencing across disjunct dune populations from San Francisco to Big Sur, the chromosome-level assembly enables conservation geneticists to quantify inbreeding coefficients (F ROH), map adaptive loci, and identify genetically distinct management units before localized coastal populations are lost [1, 4] .

References & Primary Sources

* [1] Richards, S., Jochim, E., Newton, L., Starrett, J., Sahasrabudhe, R., Nguyen, O., Beraut, E., Fairbairn, C., Sacco, S., Seligmann, W., Escalona, M., Cooper, R. D., Shaffer, H. B., Toffelmier, E., & Bond, J. E. (2026). Reference genome of the Californian trapdoor spider Aptostichus stephencolberti Bond 2008 (Araneae: Mygalomorphae: Euctenizidae). Journal of Heredity, esag069. -- https://doi.org/10.1093/jhered/esag069

* [2] University of California, Davis News / Phys.org (2026). A spider named for Stephen Colbert has a genome about 50% larger than ours. -- https://www.ucdavis.edu/news/spider-named-stephen-colbert-has-genome-about-50-larger-ours

* [3] Bond, J. E., & Stockman, A. K. (2008). An integrative method for delimiting cohesion species: finding the population-species interface in a group of Californian trapdoor spiders with extreme genetic divergence and geographic structuring. Systematic Biology, 57(4), 628-646. -- https://doi.org/10.1080/10635150802302443

* [4] Bond, J. E. (2012). Phylogenetic treatment and taxonomic revision of the trapdoor spider genus Aptostichus Simon (Araneae, Mygalomorphae, Euctenizidae). ZooKeys, 252, 1-209. -- https://doi.org/10.3897/zookeys.252.3588

* [5] Shaffer, H. B., Toffelmier, E., Corbett-Detig, R. B., Escalona, M., Erickson, B., Fiedler, P., Gold, M., Harrigan, R. J., Hodges, S., Luckau, T. K., Miller, C., Oliveira, D. R., Shaffer, K. E., Shapiro, B., Sork, V. L., & Wang, I. J. (2022). Landscape genomics to enable conservation actions: The California Conservation

Genomics Project. Journal of Heredity, 113(6), 577-588. -- https://doi.org/10.1093/jhered/esac020