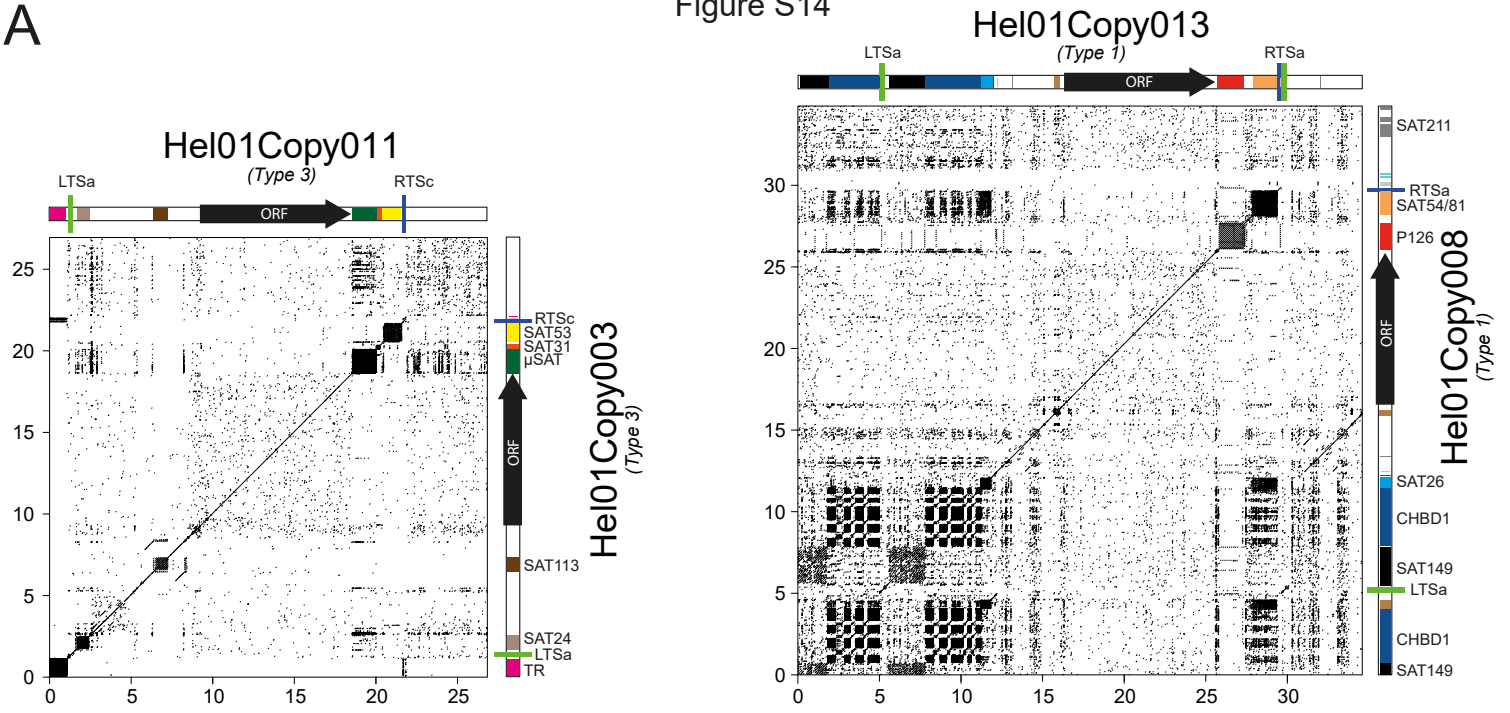


Figure S14

A



B

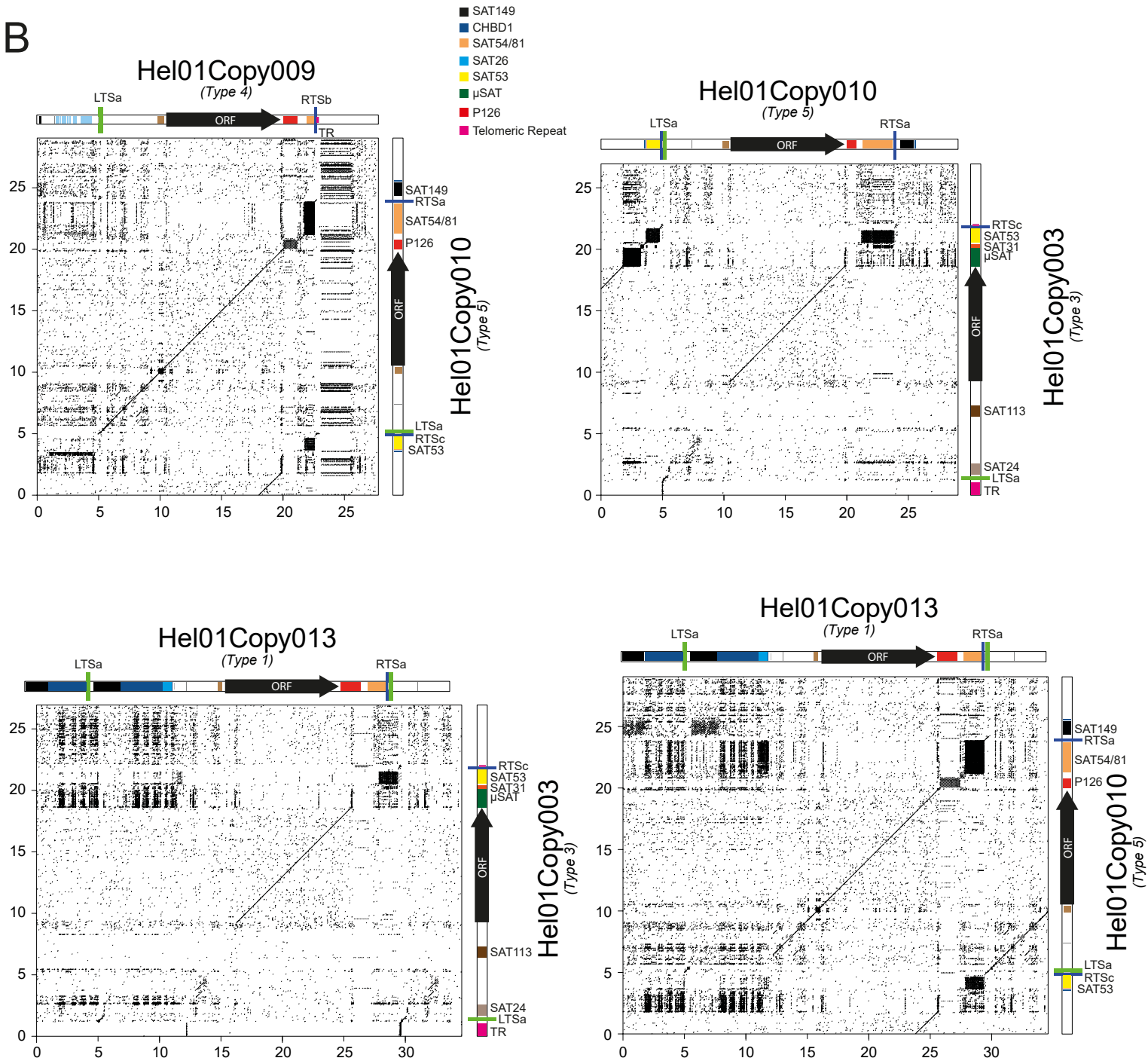


Figure S14: Nucleotide alignment of different Hel01 copies.

(A). Dot plots showing the similarity between copies of the same Hel01 Type. On the left, Hel01Copy011 and Hel01Copy003 from Type 3. On the right, Hel01Copy013 and Hel01Copy008 from Type1. **(B).** Dot plots showing similarity between copies of different Hel01 Types. Each dot corresponds to at least 70% identity over a 28 nt window. Annotations (LTS, RTS, ORF, satellites) for each copy are drawn above and to the right of each dot plot. The legend presents the different annotated satellites. Microsatellite (μ SAT) consists of (TTA/TTG) $_n$. The scale in kb is given on the left and bottom axes of each dot plot.

Figure S15

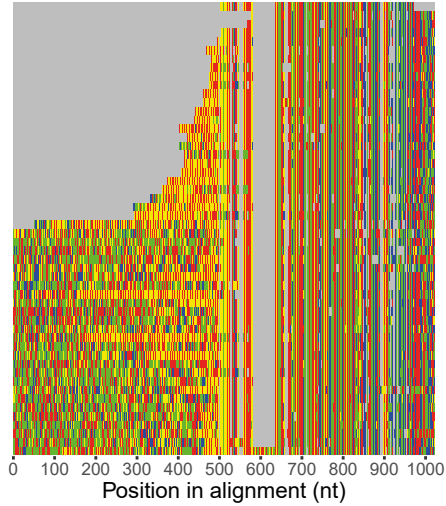
A

270 280 290 300 310 320 330 340
 RTSa CAACCCCAACCGTACCTAGTA **TAAACGAGAGAGA** **TATAACA** TAGAGAGA---AT **TGTTATAA** **CTCTCTCAGTTTA** GGGGGGGTAACCAA
 RTSb CAACCCCAACCGTACTTAGTA **TAAACGAGAGAGA** **TATAACA** TAGTGT-----AT **TGTTATAA** **CTCTCTCAGTTTA** --GAGTCAAACCAA
 RTSd CAACCCCAACCGTACTTAGTA **TAAACGAGAGAGA** **TATAACA** TAGTGT-----AT **TGTTATAA** **CTCTCTCAGTTTA** --GAGTCAAACCAA
 RTSd CAACCCCAACCGAACTAGTA **TAGACGTTAGAAATTT** **ACAAT** TGGGTGTTT **ATTGTTAAATTTCTAAAGTCTA** --TAGTCTAACCAA
 RTSe CAACCCCAACCGTA---AGTAGA **ACTT** **AGGTTGATTTT** **CTTT** CGTTTA--- **AAAGTAAATCAACCTCAGT** C-A--GAGTTATACCAA
 ..***:* ***** *.. : . :.***:~::~ *::~::~~::~: :. :.*** * .* :~::~~::~~::~

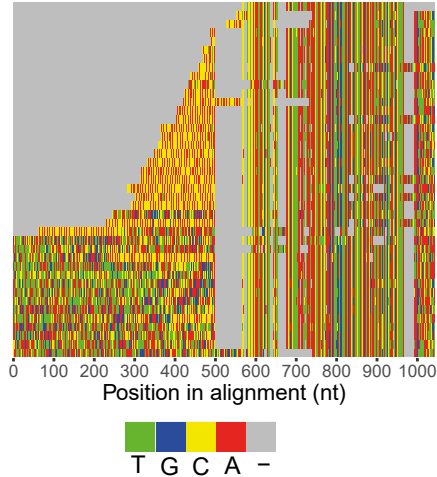
→
←

B

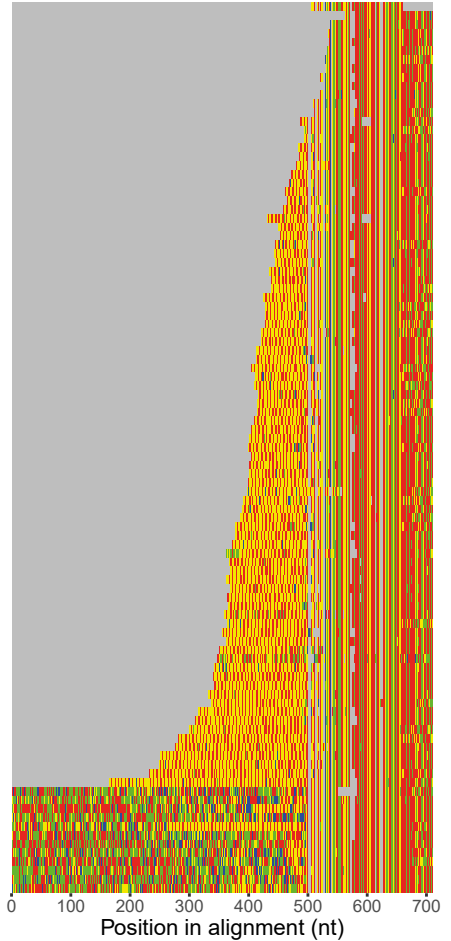
LTSa n=51



LTSb n=40

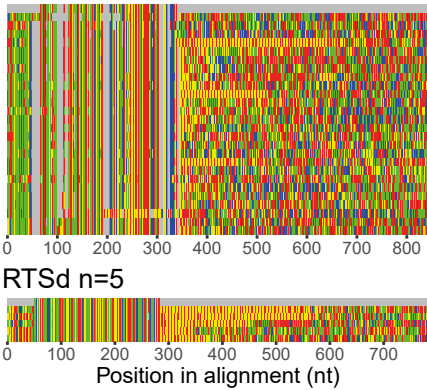


LTSc n=100

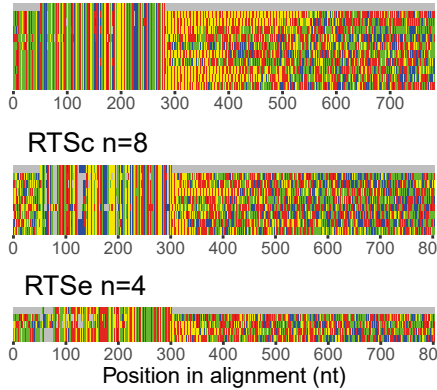


C

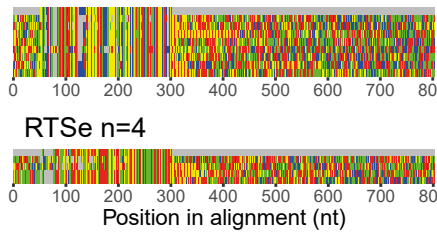
RTSa n=26



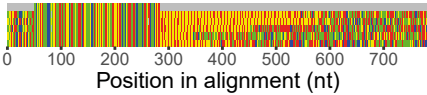
RTSb n=10



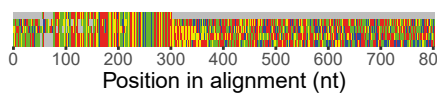
RTSc n=8



RTSd n=5

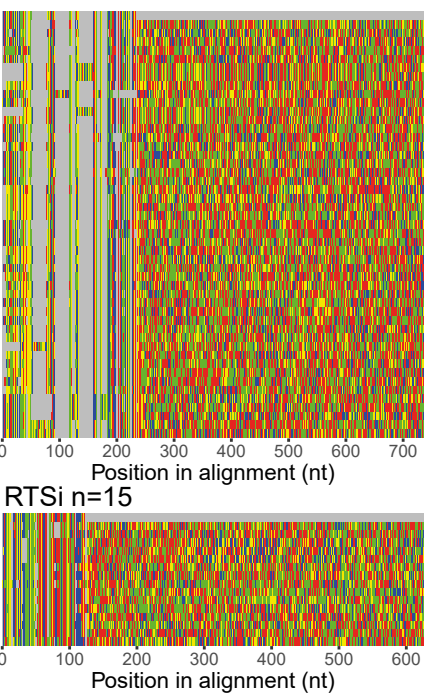


RTSe n=4

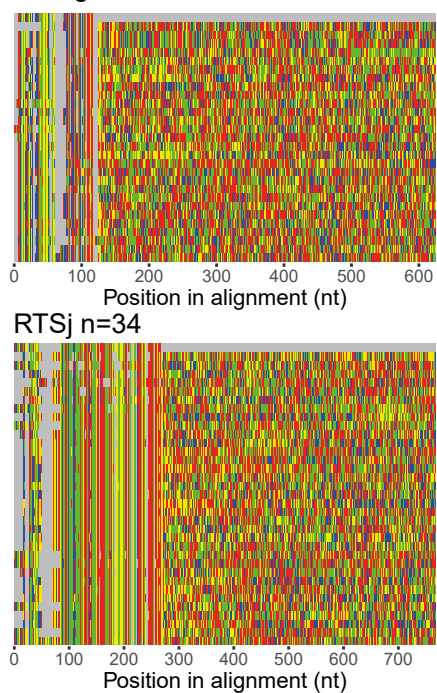


D

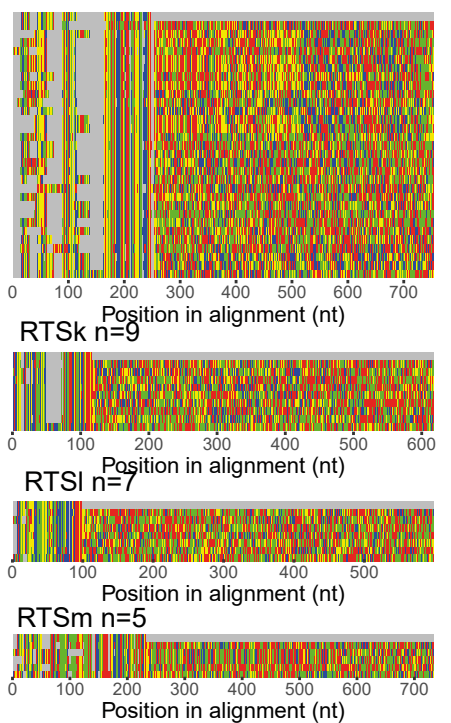
RTSf n=48



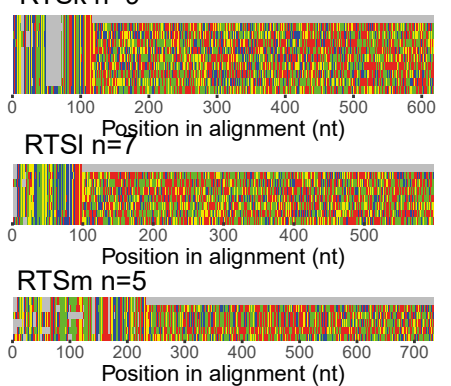
RTSg n=28



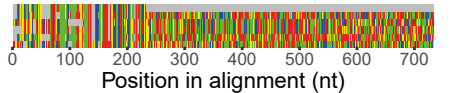
RTSh n=30



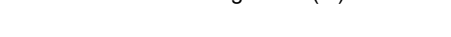
RTSi n=7



RTSj n=34



RTSk n=9



RTSl n=5

Figure S15: Structural characterisation of Helitron ends and insertion sites

(A) Alignment of different Right Terminal Sequence (RTS) consensus sequences for *Paramecium* Helitrons. The arrows beneath the alignment indicate the hairpin (positions 280-340), typically found at the 3' end. These RTS end with CCAA (positions 347-350). **(B)** Alignment of the 5' and 3' flanking sequences of 500 and 100 nucleotides, respectively, for left terminal sequence (LTS) loci. As in Figure 3B, color indicates the type of nucleotide. Note the telomeric repeats (red/yellow) before most copies of LTSa, b or c. **(C)** Alignment of RTSa to RTSe copies associated with Clade A1 Helitrons, with 5' and 3' flanking sequences of 100 and 500 nucleotides, respectively. Note telomeric repeats after most RTS copies. **(D)** Alignment of RTSf to RTSm copies associated with Clade B Helitrons, with 3' 500-nt flanking sequences.

Figure S16

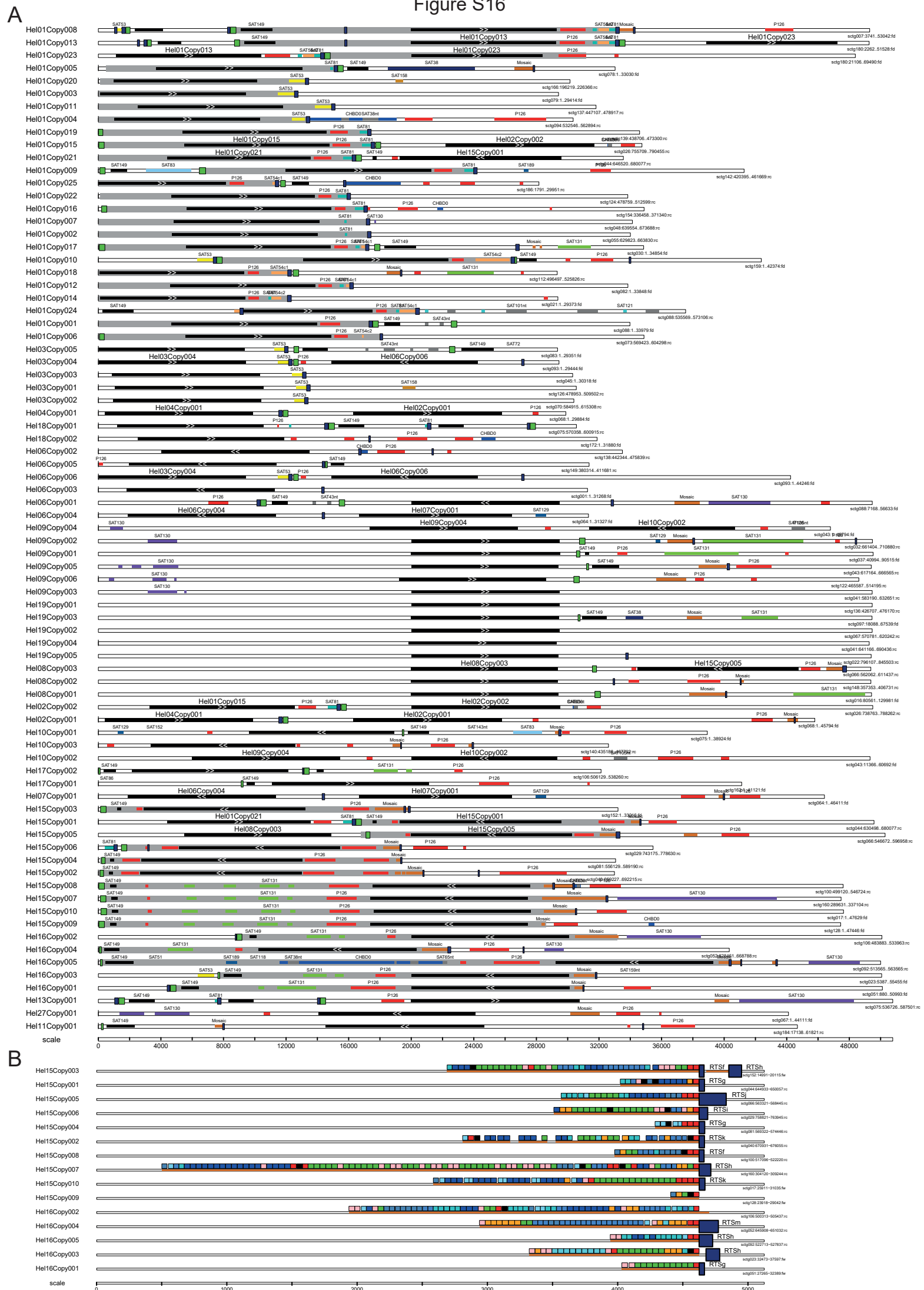


Figure S16: Annotations of helitron copies with full-length ORFs

(A) Scale drawings showing all supercontig regions with Helitrons containing full-length ORFs. The closest extremity of the supercontig is shown on the left and the genomic position is indicated on the right with forward (fd) and reverse complement (rc) orientation. The RepHel ORF (black box) is oriented with ">>", LTS (lime green box) and RTS (navy blue box) sequences delimiting the putative Helitron element (gray rectangles). Satellite DNA is colored as in *Additional file 6: Data S3-S4*. Nucleotide scale is indicated at the bottom. **(B)** Zoom of Hel15/Hel16 copies to show mosaic satellite repeat structure and the presence or absence of a putative RTS (navy blue). Each "color" of the mosaic satellite represents a distinct ~44 nt repeat unit.

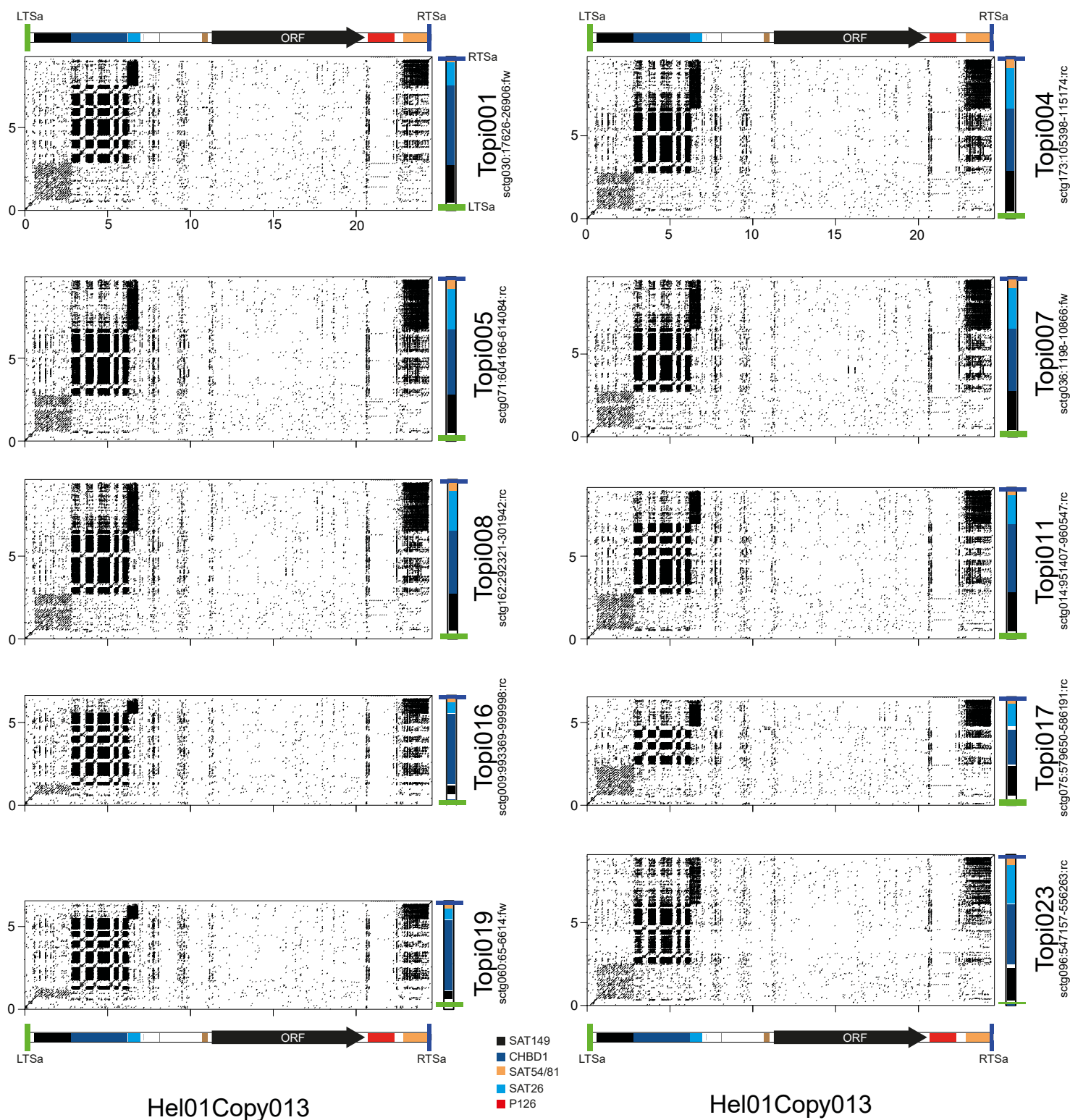


Figure S17: Structure of Hel01 ORF-less copies.

Dot plots as in Additional file 9:SupFigure14, illustrating 10 examples of Hel01 copies with internal deletions that encompass the ORF (y-axis, labeled with “Topi” prefix) with complete Hel01Copy013 (x-axis).

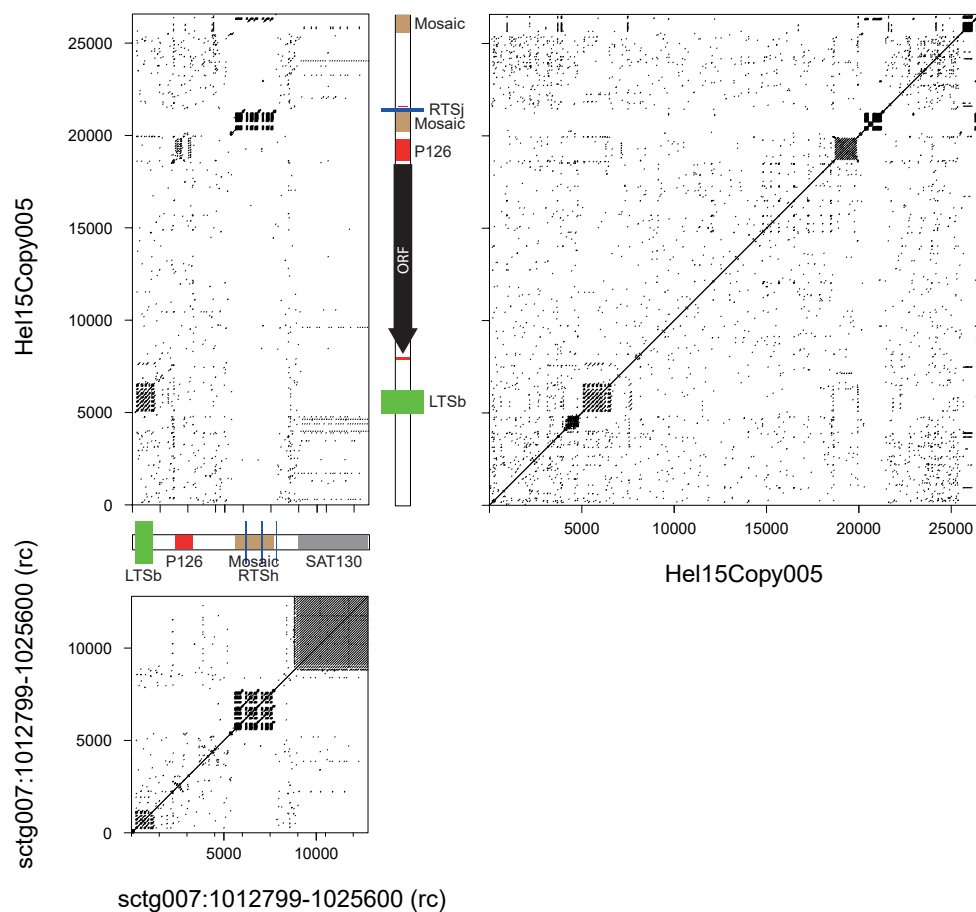


Figure S18: Structure of Hel15/Hel16 ORF-less copies.

As in Additional file 9: SupFigures 16-17. (A) Comparison of Hel15Copy005 with itself, an ORF-less locus (right end of sctg007) and of the ORF-less locus with itself.

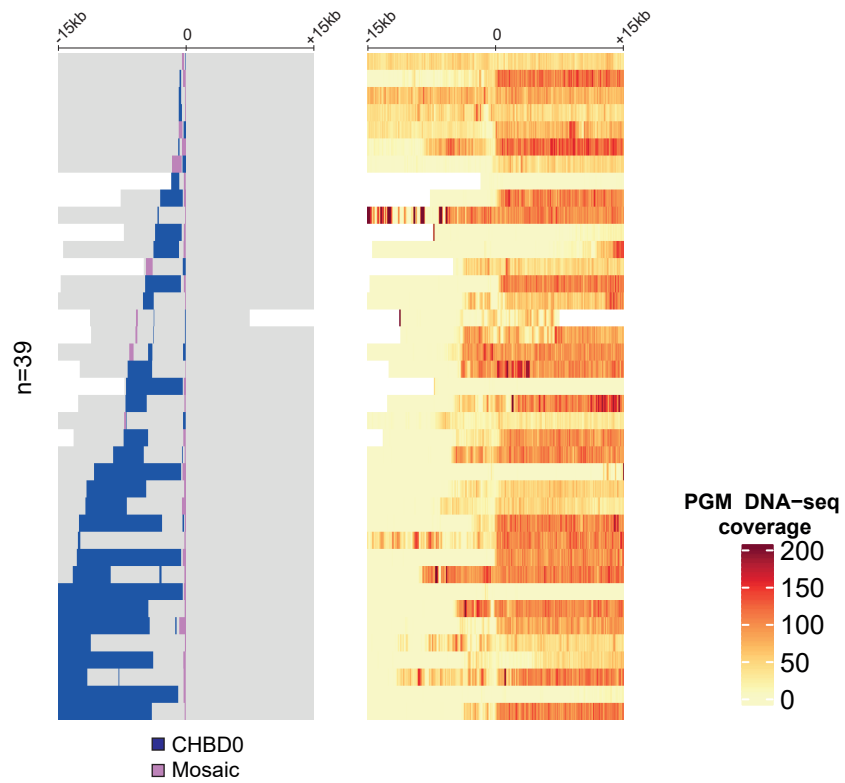


Figure S19: Loci that contain both CHBD0 and Mosaic satellite.

As in Figure 8 C-D, heatmap for 39 loci that contain CHBD0 (blue) and Mosaic (orchid) separated by less than 15 kb centered in a 30 kb window.