



Figure S3. (A) Schematic multiple alignment of 23S rRNA (*rrnL*) genes of diverse *Chlamydomonas* species and *Polyotma uvella*. Stacked black rectangles represent conserved regions across all sequences. The identity profile was calculated using a 15 nt sliding window (dark blue bars represent identical sites, light blue is for incomplete identity and red highlights low identity). Light gray blocks highlight regions of the *P. uvella* *rrnL* gene similar to at least one of the *Chlamydomonas* analyzed sequences. Dark gray blocks represent AT-rich (>80%) “insertions” in the *P. uvella* *rrnL* gene with no evident similarity to Chlamydomonadalean *rrnL* genes. The red coloured box indicates the only intron (type I) predicted by RNAweasel in the *P. uvella* plastid scaffolds. GC content (150 nt sliding window size) in the rRNA-coding genes *rrnL* (B) and *rrnS* (C) of *Polyotma uvella*. Light gray blocks indicate conserved regions of relative high GC content and dark gray blocks correspond to insertions of high AT content.