

Homology searches for group I introns in archaea

* downloaded all archaeal sequences in GenBank (6.7 Gb as of Sept 2017)

* searched archaeal sequences with all GISSD models + RF00028 with default cmsearch parameters and with --anytrunc

* 95 non-overlapping hits with $E < 0.01$ corresponding to 39 group I intron candidates(\$) (12 IA3 and 27 IB4)

* 30/39 introns have at least one hit with $E < 1E-10$

* 36 within LSU rRNA, 3 within SSU rRNA

* All IA3s are in one of two LSU insertion positions:

- LSU/2593 (N=10)
- LSU/2500 (N=2)

* All IB4s are in one of two LSU insertion positions and one SSU position:

- LSU/1931 (N=15)
- LSU/1923 (N=9)
- SSU/1498 (N=3)

(§) as determined via manual sequence analysis by Tom Jones

