More than expected: Operon turnover in three *Caenorhabditis* species

Christian Rödelsperger and Christoph Dieterich*
Max Planck Institute for Developmental Biology, Department of Evolutionary Biology,
Spemannstraße 35-37, 72072 Tübingen, Germany

Operons are polycistronic gene units that are transcribed from one shared upstream promoter. They are quite common in bacteria but only a small number of eukaryotes including *C. elegans* have been reported to contain operon structures. Unlike in bacteria, nematodes (e.g. *C. elegans*) use the mechanism of trans-splicing to seperate individual gene transcripts. Transplicing removes intron-like sequences, the so-called outrons, at the 5-end of a pre-mRNA and in between genes of an operon. The mechanism is the same as in *cis*-splicing except that the 5 splice site is missing. In this case the splice site is provided by a small nuclear ribonucleoprotein particle (snRNP), which binds to a branching point on the outron and thus releases a mini exon of 22 nucleotides, the so-called splice leader (SL). About 70 % of *C.elegans* genes are trans-spliced and $\sim$ 15 % are found in operons. Usually, the first gene in an operon is SL1 trans-spliced. All downstream genes are predominantly trans-spliced to SL2, although some downstream genes have been reported to be trans-spliced to SL1.

We analyzed gene order relations in three *Caenorhabditis* species: *C.elegans, C.briggsae and C.remanei* with a novel approach (partial gene order alignment with SYNTENATOR). To this end, we use the published *Caenorhabditis* operons as annotated in Wormbase release 150. Having identified operon genes and their orthologs, we implemented a classification scheme (Figure 1) to distinguish three possible scenarios: operons that are clearly conserved, operons that may be conserved and clearly disrupted operons. We also discuss lineage-specific disruptions and highlight those where downstream genes have an embryonic lethal phenotype in *C.elegans*. Finally, nucleotide substitution rates are compared to test differences in selective pressure for genes that belong to operons and other genes. Of the 1054 operons in WS150,

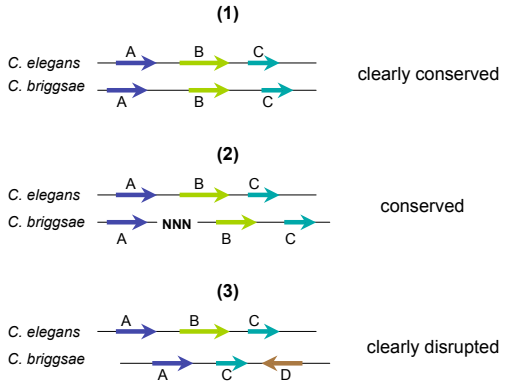

Figure 1: Classification example of *C. elegans* operons consisting of three genes. (1) clearly conserved (2) potentially conserved due to assembly gap (3) clearly disrupted.

906 could be aligned at least partially to *C. briggsae*. 770 operons could be identified as conserved and 95 were found to be clearly disrupted, meaning that the disruption cannot be due to gaps in the assembly. For the *C. elegans* and *C. remanei* genome a total of 845 operons could be partially aligned. The number of clearly disrupted operons was 74 with 720 conserved operons.
28 operons of the clearly disrupted operons are disrupted in both species and 674 operons were found to be conserved in both genomes. These findings clearly indicate that operon turnover is higher than previously reported by Stein et al. (2003) and challenges our views on the evolvability of operons.

*christoph.dieterich@tuebingen.mpg.de