

Teamed up for transcription

Christian von Mering and Peer Bork

In bacterial genomes, functionally related genes are often clustered and controlled as a unit. Such 'operons' are not normally found in animals — so why are they so abundant in one class of worm?

The genomes of animals, plants and fungi seem to be relatively disorganized. Genes appear to be randomly distributed, with only a few exceptions: repeats of similar sequences caused by gene duplications, for example, and a limited number of ancient gene clusters containing functionally related genes (such as the Hox genes that are involved in control of animal development¹). Apart from these, the average gene is generally assumed to be independent of its neighbours, and genomes are constantly rearranged and shuffled. However, in one group of animals — the nematodes (small, unsegmented worms) — neighbouring genes are occasionally assembled into regulatory units called operons². On page 851 of this issue, Blumenthal *et al.*³ now report the first whole-genome characterization of such operons in a multicellular organism, and raise intriguing questions as to how (and why) they have evolved.

Operons were first described⁴ in prokaryotes — small, single-celled organisms without a nucleus. In these organisms, genes needed for one particular process (say, synthesis of an amino acid) are often clustered in close proximity on the genome, with the same orientation on the DNA. These operons facilitate the coordinated regulation of genes, as the clustered genes are activated (that is, transcribed into messenger RNA) all in one go. Operon transcripts always code for more than one protein, and prokaryotes handle this by starting translation of the messenger RNA into amino acids separately at the beginning of each protein-coding section (Fig. 1a). In contrast, these 'polycistronic' transcripts are a problem for eukaryotes (organisms with nucleated cells, including all multicellular organisms). They generally cannot process such transcripts, and so it has long been assumed that most eukaryotes do not have operons.

Blumenthal *et al.*, however, had shown previously² that the nematode *Caenorhabditis elegans* occasionally does make polycistronic transcripts and can process them through a mechanism called *trans*-splicing. Briefly, the transcripts are split up into individual pieces, and short 'leader' sequences are simultaneously added to each piece (Fig. 1b). In *C. elegans*, *trans*-splicing is not limited to polycistronic transcripts — in fact, the majority of transcripts receive leader sequences. But the polycistronic transcripts are unique in

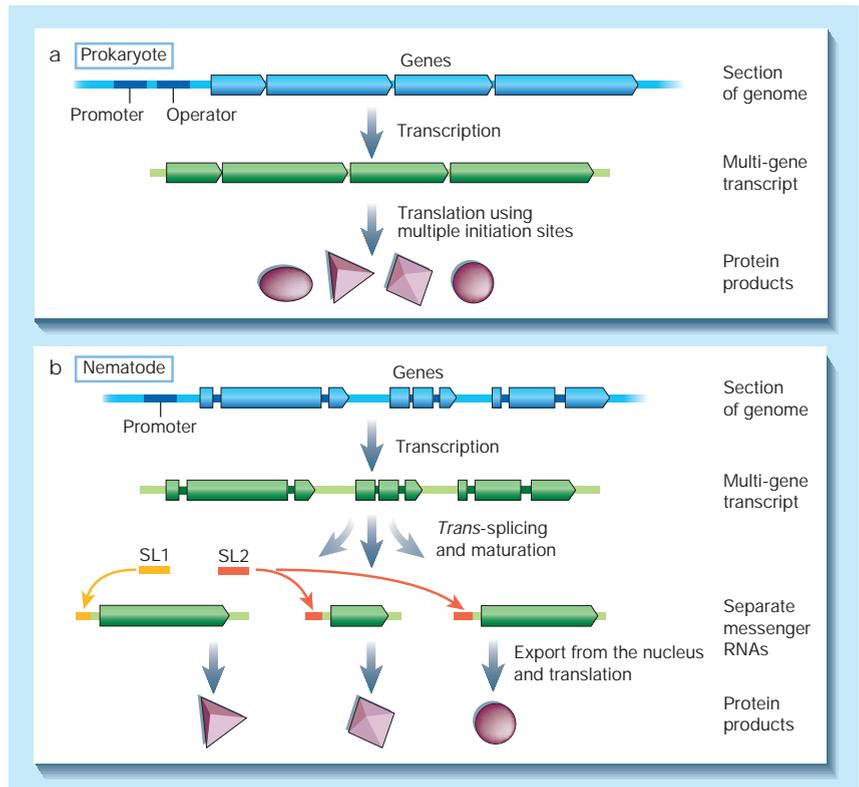


Figure 1 Operons — organizing genes for joint activation. a. A classical operon, found in prokaryotes (such as bacteria). Several genes, often functionally related, form a tight cluster on the genome. Operons are under the control of regulatory elements (promoter, operator) and factors that bind to these elements. Transcription of the cluster results in a single molecule, a multi-gene transcript of messenger RNA, which codes for several proteins and is directly translated into distinct protein products. b. An operon in the nematode *Caenorhabditis elegans*, typical of the several hundred discovered by Blumenthal *et al.*³ in the first genome-wide survey for operons in a eukaryotic organism. The initial multi-gene transcript is split up into separate messenger RNAs by *trans*-splicing, concomitant with the insertion of splice leader sequences (SL1, SL2). Although these operons fit the classical genetic definition⁴, the processing of their transcripts into proteins differs fundamentally from that of prokaryotic operons.

that processing involves a particular splice leader, called SL2, for all pieces but the first (which often has the usual leader, SL1).

To measure the extent of polycistronic transcription (and so of operons) in *C. elegans*, Blumenthal *et al.*³ initiated a systematic, genome-wide study. Reasoning that it should be possible to discover all operons by searching for SL2 transcripts and checking their arrangements on the genome, they took two independent approaches. First, they analysed *C. elegans* transcripts that had been sequenced previously (random transcripts are routinely sequenced in many model organisms). Among these, they found more

than 300 that contained SL2. For their second approach, they prepared a collection of actual transcripts from a population of worms, and enriched the sample for transcripts with an SL2 leader (this involved adding synthetic SL2-like molecules to the mixture to find the right transcripts through base-pairing). The enriched population was then compared to a normal, non-enriched fraction by measuring transcript abundances. This technically demanding approach yielded another 1,200 SL2-transcripts. Both sets of data converged well and agreed with gene clustering as observed in the genome. Together with the unlabelled

first gene of each operon, the combined list contains 2,291 genes in 881 operons. This is a surprisingly large total — the authors estimate that 13–15% of all genes in *C. elegans* (including some they might have missed) are expressed as part of an operon.

What can be learned from this first genome-wide study of eukaryotic operons? First, it nicely confirms a previous assumption²: in *C. elegans*, SL2-containing transcripts are indeed reliable operon markers (as expected, in many runs of consecutive genes on the genome, all genes but the first were found to generate SL2-transcripts). Second, the data may help researchers working with an uncharacterized gene — if it is in an operon with a gene that is better characterized, they can probably assume that both have the same transcriptional regulation. Third, a genome-wide view of operons in *C. elegans* is a step towards addressing long-standing questions such as why nematodes have operons and whether they are the only animals that do.

Blumenthal *et al.* cautiously suggest that some of the operons could serve the same purpose as their counterparts in prokaryotes: to group functionally related genes together. This clearly appears to be true for some genes³, and it would be exciting if it were the case for most of them — it would mean that the function of uncharacterized genes could be predicted from their operon context. We could not resist checking the functional annotations of these newly discovered operons (using data from the Gene Ontology Consortium⁵), but we found that only about 4% of the operons contained two or more genes annotated for the same biological process, compared to 36% for known and putative operons in the prokaryote *Escherichia coli*. This figure is low, but it could well be due to the fact that the fraction of genes with known function is smaller in *C. elegans* than in prokaryotes. Because of these limits in annotation, we cannot yet say whether most genes in operons are functionally related — but there are certainly cases where they are not.

To understand the functional significance of operons in *C. elegans*, one must consider their evolutionary history. Are they related to prokaryotic operons through common descent? *Trans*-splicing, at least, is not unique to nematodes. It was first described in single-celled eukaryotes (trypanosomes, in which entire operons have also been reported⁶), and more recently in several animals other than nematodes^{7,8}. This raises the possibility that the earliest multicellular animals possessed *trans*-splicing capability, and possibly had operons inherited from prokaryotes. However, in addition to mechanistic differences (Fig. 1), the content of nematode operons seems distinct from that of prokaryotes. Not only is the average number of genes per operon smaller (2.6 in *C. elegans* against 4.1 in *E. coli*), but we also found only a few cases where two or

more genes were in an operon together in both *C. elegans* and prokaryotes — hardly more than would be expected by chance (using the database Clusters of Orthologous Groups⁹). Furthermore, no operons have yet been found in well-studied groups such as insects, fishes and mammals, nor in plants or fungi.

From all of this it seems that operons in *C. elegans* were a lineage-specific invention, perhaps facilitated by an existing capability for *trans*-splicing. A possible evolutionary scenario is suggested by the finding that genes with similar transcription profiles may have a tendency to cluster in eukaryotic genomes^{10,11}. It is conceivable that such clusters were among the first to be joined into operons. This might have been promoted by a selective pressure for a small genome (*C. elegans* is known for its compact genome). Later, in the continuing processes of genome shuffling, segmental duplications and gene loss, more transcriptionally or functionally related genes could have come close together by chance, forming additional operons.

Whatever the precise course of events,

Blumenthal *et al.*³ have provided us with a key study. The operons they have identified will be the basis for much further investigation into genome evolution and gene regulation in animals. ■

Christian von Mering and Peer Bork are at the European Molecular Biology Laboratory, Meyerhofstrasse 1, 69012 Heidelberg, and the Max-Delbrück Centre for Molecular Medicine, Berlin, Germany.
e-mail: bork@embl-heidelberg.de

1. Ferrier, D. E. & Holland, P. W. *Nature Rev. Genet.* **2**, 33–38 (2001).
2. Spieth, J., Brooke, G., Kuersten, S., Lea, K. & Blumenthal, T. *Cell* **73**, 521–532 (1993).
3. Blumenthal, T. *et al.* *Nature* **417**, 851–854 (2002).
4. Jacob, F. & Monod, J. *J. Mol. Biol.* **3**, 318–356 (1961).
5. The Gene Ontology Consortium *Genome Res.* **11**, 1425–1433 (2001).
6. Gao, G., Nara, T., Nakajima-Shimada, J. & Aoki, T. *J. Mol. Biol.* **285**, 149–161 (1999).
7. Stover, N. A. & Steele, R. E. *Proc. Natl Acad. Sci. USA* **98**, 5693–5698 (2001).
8. Vandenberghe, A. E., Meedel, T. H. & Hastings, K. E. *Genes Dev.* **15**, 294–303 (2001).
9. Tatusov, R. L. *et al.* *Nucleic Acids Res.* **29**, 22–28 (2001).
10. Cohen, B. A., Mitra, R. D., Hughes, J. D. & Church, G. M. *Nature Genet.* **26**, 183–186 (2000).
11. Lercher, M. J., Urrutia, A. O. & Hurst, L. D. *Nature Genet.* **31**, 180–183 (2002).

Astronomy

Nebulous explanation

Mark Claussen

The planetary nebulae of gas and dust that are formed from red-giant stars are usually far from spherical in shape. Observations of the gas distribution in one red giant caught in the act of transition show why.

Planetary nebulae are glowing remnants of gas, left over from the dying stages of Sun-like stars. Images from the Hubble Space Telescope have provided astronomers with exquisitely detailed images of these beautiful objects. But a long-standing mystery is why most (but not all) planetary nebulae are not spherical, although the distributions of gas and dust in their immediate predecessors are. Now Imai *et al.*¹ (page 829 of this issue), using a ground-based interferometric technique, provide evidence that a well-collimated outflow, or jet, of gas is emanating from a red-giant star that is in transition to becoming a planetary nebula. Their observations also suggest that the jet may be precessing — wobbling around on its axis like a toy top does as it spins. Their findings should lead to a better understanding of how planetary nebulae are shaped.

The Sun is a fairly common type of star, and will remain substantially unchanged over the next several billion years. Nuclear fusion reactions (converting hydrogen into helium) in the core of the Sun provide energy that supports the outer gaseous layers as gravity tries to force them inwards. But after many billions of years, when all the hydrogen in the core of such a star has been converted into helium, the core contracts under gravity,

heating up and creating pressure that causes the outer layers of the star to expand and cool. The star becomes a red giant; when the Sun reaches that stage, its outer layers will expand as far as the Earth's orbit.

As the expansion progresses, the outer layers of the star become unstable and the star pulsates, changing its total energy output on a regular basis (the cycle time is usually around a year). The pulsation causes matter to be lost from the star as a very slow 'wind' (with a velocity of about 10 km s⁻¹), into a surrounding shell. These pulsational changes last several tens of thousands of years. Eventually, the core of the star collapses to form a white dwarf with a surface temperature of 30,000 K — hot enough to ionize the gas in the shell that surrounds the dying star. The ionized gas is what we see as a planetary nebula.

This model of the later stages of stellar evolution works just fine until the last part. The slow mass loss is, as far as we can tell, generally spherically symmetrical. So the distribution of any gas that is expelled in this manner should have spherical symmetry. And yet the ionized gas distribution for most planetary nebulae is far from spherical (Fig. 1); many have a flattened, even two-lobed, structure^{2–5}.