Update TRENDS in Genetics Vol.23 No.9
Full text provided by www.sciencedirect.com
ScienceDirect
Genome Analysis

Mitogenomics reveals two cryptic species in *Ciona intestinalis*

Fabio Iannelli[1], Graziano Pesole[1,3], Paolo Sordino[2] and Carmela Gissi[1]

[1] Dipartimento di Scienze Biomolecolari e Biotecnologie, Università di Milano, Via Celoria 26, 20133 Milano, Italy
[2] Laboratory of Biochemistry and Molecular Biology, Stazione Zoologica ''A. Dohrn'', Villa Comunale, 80121 Naples, Italy
[3] Dipartimento di Biochimica e Biologia Molecolare ''E. Quagliariello'', Università di Bari, Via Orabona 4, 70125 Bari, Italy

Individual mitochondrial genes or genomic features are commonly used as phylogenetic markers at many taxonomic levels. We used a mitogenomics approach to demonstrate the existence of two cryptic species in the ascidian *Ciona intestinalis*, a model chordate whose status as a single species has recently been questioned. Comprehensive comparative analysis of the mitochondrial genome of the two cryptic species revealed significant differences in gene order, size and number of noncoding regions, compositional features and divergence of protein-coding genes.

Introduction

The circular mitochondrial DNA (mtDNA) of metazoans represents a rich source of genetic markers for phylogenetic, population and biogeographic studies. Sequences and genome-level features, such as gene arrangement, are both used to resolve deep-level phylogenetic relationships (reviewed in Ref. [1]), whereas single mitochondrial genes or regions are analyzed in population genetics studies [2]. In addition, the *cox1* (cytochrome oxidase subunit 1) gene has been proposed as a 'DNA barcode' for species diagnosis [3]. Several recent studies have emphasized the great resolving power of a more comprehensive, genome-level approach in the analysis of mtDNA both for population and phylogenetic purposes [1,2]. We have used a novel mitogenomics approach, based on the comparison of several sequence and higher-order mitochondrial features, to provide an unambiguous answer to the debated question of whether the ascidian *Ciona intestinalis* contains cryptic species. *C. intestinalis*, a cosmopolitan marine species, is a model chordate widely used in developmental biology and evolutionary genomics [4], thus its correct taxonomic status is fundamental for comparative studies. Recently, an unexpected genetic divergence was observed at the nuclear-genome level between *C. intestinalis* subpopulations [5], and between Californian and British individuals [6,7]. Thus, Suzuki *et al.* [7] have suggested the existence of two, morphologically almost indistinguishable, cryptic species: one denoted as type A and distributed in the Mediterranean and Pacific areas, including California; the other denoted as type B and distributed in Northern Europe, including the British coastline. Human-mediated marine transport has been invoked to explain this geographical distribution [7]. However, the two types of *C. intestinalis* hybridize in laboratory conditions [7], challenging the existence of cryptic species.

The mitogenomics approach

To solve this taxonomic controversy, we analyzed the mitochondrial genome of the two putative cryptic species. First, we determined the complete mtDNA of a type B individual of *C. intestinalis* isolated in Plymouth, UK. Next, we compared the mtDNA of the type B individual with the previously described mtDNA of a type A individual isolated in Naples [8], and used the mtDNA of the congeneric species *Ciona savignyi* [9] as outgroup. We also included in our comparative analysis the mitochondrial-like scaffolds that we identified and annotated in the whole-genome shotgun projects of *C. savignyi* (http://mendel.stanford.edu/sidowlab/ciona.html), and of a Californian type A individual of *C. intestinalis* [10] (see the supplementary material online). We examined several mitochondrial features, namely gene order, size and number of noncoding regions, base composition, codon usage and evolutionary divergence of protein-coding genes.

Gene order rearrangement and noncoding regions

The mtDNAs of *C. intestinalis* types A and B are strikingly different in gene order, and in number and size of noncoding regions (NCRs) (Figure 1). These differences were confirmed in nine additional individuals: five from type A and four from type B (see the supplementary material online).

The gene order of type B differs from that of type A in the translocation of the tRNA-Cys gene (*trnC*), thus the gene arrangement *(nd4-trnF-atp8-trnC-trnG(AGR))* observed in type A and even in *C. savignyi* mtDNA changes to (*nd4-trnC-trnF-atp8-trnG(AGR)*) in type B (Figure 1).

Mitochondrial noncoding regions, consisting of short intergenic spacers, are reduced in size and number in type B compared with type A (Figure 1). Indeed, we found only three NCRs with length $\geq$20 bp in type B, compared with six in type A. The longest NCR in the mtDNA of type A (85 bp long, between *cox3* and *trnK)* is absent in type B (Figure 1). Interestingly, in type A this longest NCR contains a 30-bp sequence that is duplicated in the NCR downstream of *nd1* [8]. Other NCRs, shorter in type B compared with type A, are located between tRNA genes, which are rearranged in *C. savignyi* compared with *C. intestinalis* (genes in gray in Figure 1). Furthermore, the first 28 bp of the small rRNA gene (*rrnS*) annotated in type

Corresponding author: Gissi, C. (carmela.gissi@unimi.it).
Available online 20 July 2007.

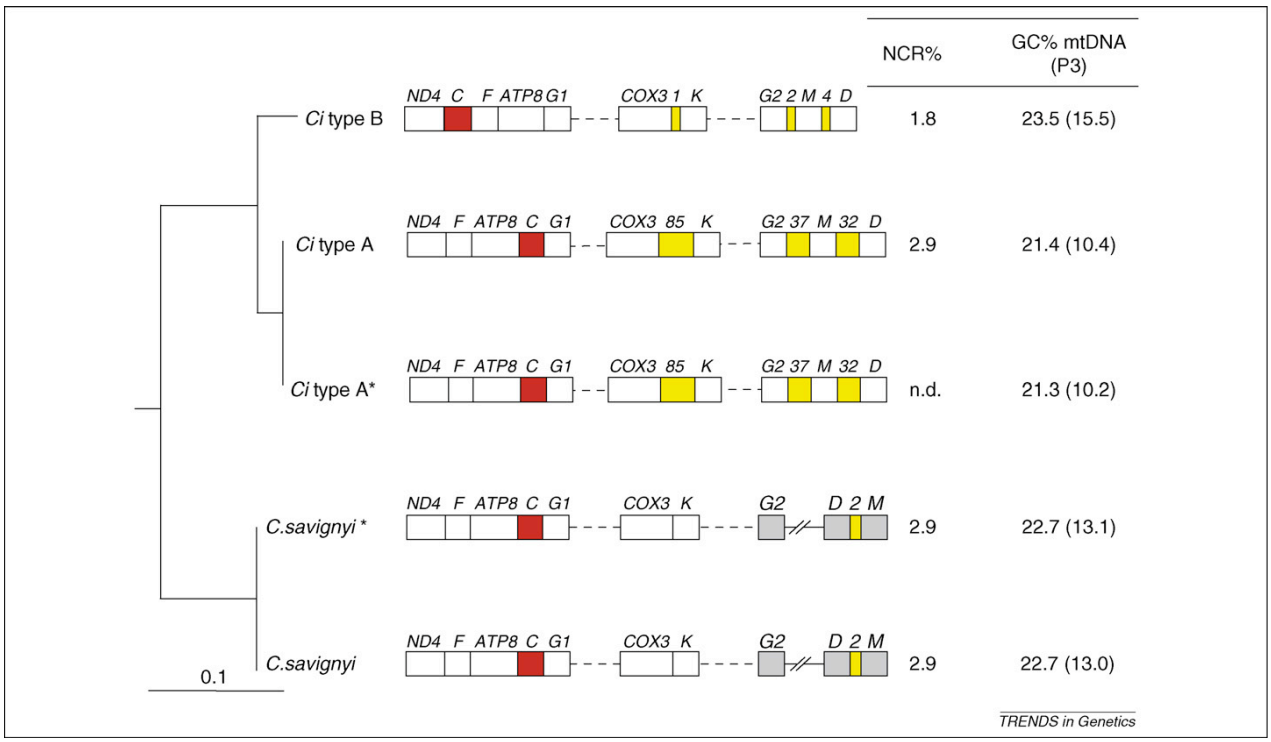

Figure 1. Differences between the mitochondrial genomes of *Ciona intestinalis* (*Ci*) type B, *C. intestinalis* type A and *C. savignyi*. Asterisks indicate mtDNAs derived from genomic scaffolds. Genome structure is shown only for regions rearranged between *C. intestinalis* types A and B, with the transposed *trnC* gene in red and noncoding regions (NCRs) with size differences in yellow (length in bp indicated above). Gray, genes rearranged in *C. savignyi* compared with *C. intestinalis*. The analyzed mtDNA sequences are: *Ciona intestinalis*, type B (EMBL accession number AM292218); *Ciona intestinalis*, type A (mtDNA, AJ517314; genomic scaffold, AABS01001113); *Ciona savignyi* (mtDNA, AB079784; genomic scaffold, AACT01048180). The *C. intestinalis* genomic scaffold is a partial mtDNA sequence (lacking a portion of *rrnS* and *nd6* genes, and the entire *trnW* gene), thus the percentage of noncoding regions was not determined (n.d.). Nonsynonymous substitutions (dN) of all protein-coding genes were calculated on the reported tree topology with the ML codon substitution method of Goldman *et al.* [14] (see the supplementary material online). The extent of differences observed in gene order, NCR% and GC content between *C. intestinalis* types A and B are incompatible with intraspecies variability. Gene abbreviations: ATP8, ATPase subunit 8; COX3, cytochrome oxidase subunit 3; ND4, NADH dehydrogenase subunit 4; C, *trnC*; D, *trnD*; F, *trnF*; G1, *trnG*(AGR); G2, *trnG(GGN)*; K, *trnK*, and M, *trnM(AUA)*. P3, third codon position.

A are absent in the mtDNA of type B individuals. Given the similarity to the 5′ region of the downstream *trnW* gene (average identity 74%), and uncertainties in the precise definition of the rRNA boundaries, it is likely that this 28-bp sequence is a noncoding region absent in the mtDNA of type B individuals. Overall, the reduction of NCRs in type B explains the greatly compact structure of this mtDNA: 1.8% of the mtDNA of type B *C. intestinalis* consists of NCRs compared with 2.9% in type A and in *C. savignyi* (Figure 1). The differences between the NCRs of the mitochondrial genomes of the two species are particularly intriguing, because the ascidian mtDNA lacks a main NCR homologous to the control region of vertebrates, which contains regulatory elements for replication and expression of the genome [11]. Thus, differences in number and size of NCRs could provide useful indications to investigate the mechanisms of replication and transcription of these genomes.

We used the translocation of the *trnC* and the presence or absence of the 85-bp-long NCR to set up two fast PCR-based screening tests able to discriminate type A from type B individuals (see the supplementary material online).

Compositional features and sequence divergence

We found many striking differences in base composition of mtDNAs of type B and type A individuals (Figure 1). Type B is significantly more GC-rich than type A (ΔGC% = 2.1; $\chi^2 = 10.36$; $P = 0.0013$), and this difference is more evident at third codon positions (P3) (ΔGC% = 5.1; $\chi^2 = 12.85$; $P = 0.0003$) – sites almost free from selective constraints. Interestingly, the GC% of type B is greater than that of *C. savignyi* (ΔGC% = 0.75 on whole mtDNA, and 2.5 at P3). By contrast, compositional homogeneity is observed between the mtDNAs of the two type A individuals, and between the two available *C. savignyi* mtDNAs.

The compositional differences are reflected in the divergent codon usage of the two *Ciona* types. Indeed, using a chi-squared statistical test [12], we observed a significantly different codon usage strategy between type A and type B ($\chi^2 = 102.5$; $P = 5.8 \times 10^{-7}$), and between type A and the congeneric *C. savignyi* ($\chi^2 = 230.1$; $P = 2.1 \times 10^{-27}$). By contrast, the mtDNAs of the two type A individuals, and the mtDNAs of the two *C. savignyi* individuals have almost identical codon usage ($\chi^2 = 2.9$ and 0.25, respectively, with $P = 1.0$).

Data on substitution rates of the mitochondrial protein-coding genes consistently show a greater evolutionary rate in type B than in type A. We calculated the sequence divergence as number of synonymous (dS) and nonsynonymous (dN) substitutions per site [13,14], and compared the dN and dS of the *Ciona* genus with the corresponding estimates that were calculated for primates

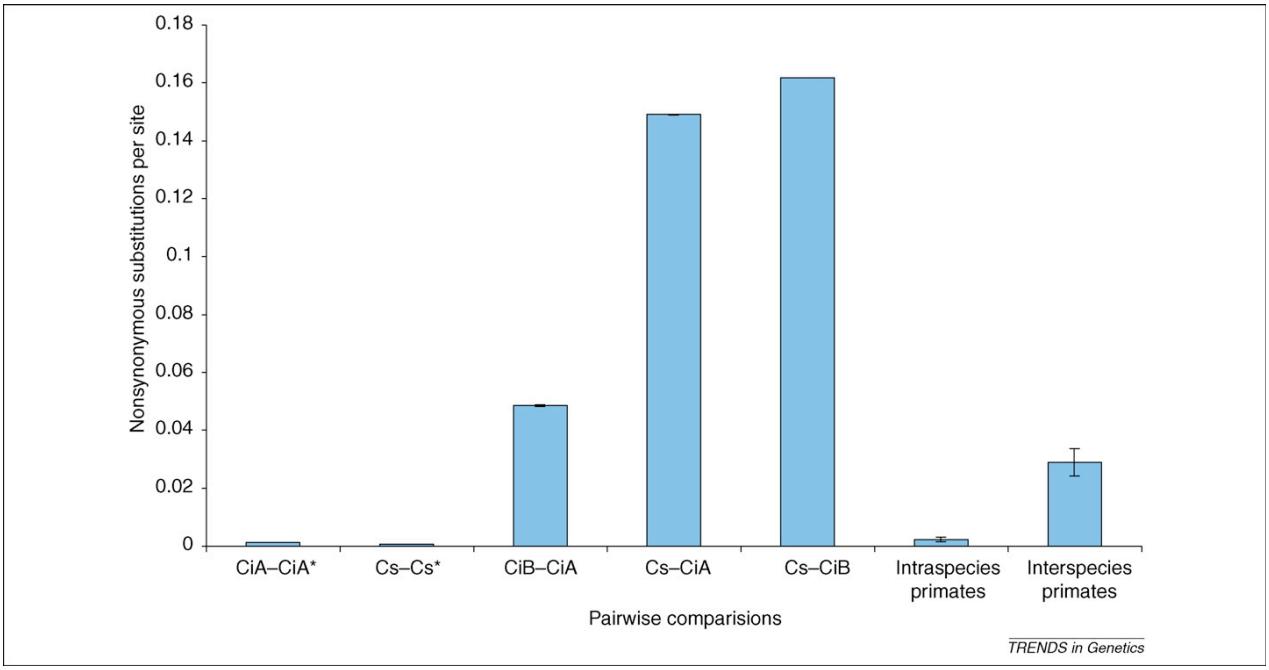

Figure 2. Comparison of pairwise nonsynonymous substitutions (dN) between *Ciona* and three species of primates. The dN rate was calculated using the CODEML program of the PAML v3.15 package [13] and the codon substitution method of Goldman and Yang [14]. The *nd6* gene was excluded from this analysis because in vertebrates it shows a different codon usage compared with all other mitochondrial genes, being the only protein-coding gene encoded by the L strand. dN values for primates are the average values estimated for three primate species, as reported in Table 2 of Hasegawa *et al.* [15]. For multiple comparisons, the standard deviation is also shown by error bars. Nonsynonymous divergence between *C. intestinalis* types A and B is significantly greater than that observed in intraspecies comparisons of ascidian (*t* test: $P < 0.05$) and primates, and comparable to that of interspecies comparisons, suggesting that type A and type B are different species. CiA, *Ciona intestinalis* type A; CiB, *Ciona intestinalis* type B; Cs, *Ciona savignyi*. Asterisks indicate mtDNAs derived from genomic scaffolds.

at intraspecies and interspecies level by Hasegawa *et al.* [15] (Figure 2 and supplementary material online). We found saturation of synonymous substitutions ($dS > 1$) in all type B versus type A comparisons (mean $dS = 3.1$), whereas dS within type A and within *C. savignyi* is not saturated and is comparable with that of primate intraspecies comparisons. The dN of type B versus type A comparisons is one or two orders of magnitude greater than in comparisons within type A or within *C. savignyi*, and is of the same order of magnitude as primate interspecies dN. Moreover, dN within type A and within *C. savignyi* is of the same order of magnitude as primate intraspecies dN comparisons. Thus, the nonsynonymous divergence between type B and type A is significantly greater than that observed in intraspecies comparisons of both primates and ascidians, and more similar to that observed in interspecies comparisons, suggesting that type B and type A are different species.

Identification of a cryptic species

Genome structural data, sequence divergence and compositional features are consistent with the existence of two cryptic species in *C. intestinalis*. No intraspecies gene-order rearrangements have been reported so far except for some bivalves, where the coexistence of two different mtDNA types is the result of doubly uniparental inheritance (DUI) [16]. In these gonochoric bivalve species, females are homoplasmic and possess only the maternally transmitted F type mtDNA, whereas males are heteroplasmic with most tissues containing both maternal (F) and paternally transmitted (M) mtDNA, and gonads containing only the M type. *C. intestinalis* is hermaphrodite; moreover, the mtDNAs of types A and B are not associated with specific tissues and both types have been obtained from ovarian tissue (see the supplementary material online), excluding the possibility that each mtDNA type could be specific to male or female gonads.

Different gene orders in congeneric species have been reported in few cases but with a wide taxonomic range, from flatworms to mollusks, arthropods and lancelets [17–21]. Although the extent of gene rearrangement is variable in these cases, there are indications of a general positive correlation between the rate of genome rearrangement and the rate of nucleotide substitution [22,23]. This situation is also observed in the *Ciona* genus where the greater dN between *C. intestinalis* type A versus *C. savignyi* corresponds to a greater number of translocated genes (six genes) [8], and the lower dN between *C. intestinalis* type A versus type B correlates with a lower number of translocated genes (one gene) (Figure 1).

Concluding remarks

The mitogenomics approach enabled us to give an unequivocal and definitive response to the hypothesis that two cryptic species exist in *C. intestinalis*; indeed, four distinct lines of evidence observed in mitochondrial genome features (gene order, number and size of noncoding regions, compositional features and sequence divergence) are all incompatible with intraspecies variability. There are many cases of distinct species differing in the mtDNA only for one of the above-mentioned genomic characteristics, thus the observed differences can be considered as independent lines

of evidence. Our conclusions are also supported by a recent comparative study between *C. intestinalis* type A and B [24], which reports the production of infertile hybrids as a result of defective gametogenesis. The fact that *C. intestinalis* encompasses at least two cryptic species represents an opportunity to unravel speciation events in chordate evolution, and has strong implications for specimen retrieval and comparative genomic studies.

Finally, our approach, based on a comprehensive comparative analysis of mtDNA, represents an accessible and powerful tool to identify cryptic species in other fast-evolving or divergent lineages.

Acknowledgements

This work was supported by the MIUR and the University of Milan (COFIN-2004 and FIRST projects to C.G.). We thank J. Bishop (MBA Plymouth), S. Shimeld (Oxford University), L. Leveque (Station Biologique de Roscoff), F. De Bernardi and C. Sotgia (Università di Milano) for kindly providing specimens; S. Groppelli for help in animal dissection; A. Boselli for technical help; D. Horner for critically reading the manuscript; and G. Zega and R. Pennati for helpful discussions.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.tig.2007.07.001.

References

1 Boore, J.L. (2006) The use of genome-level characters for phylogenetic reconstruction. *Trends Ecol. Evol.* 21, 439–446
2 Torroni, A. *et al.* (2006) Harvesting the fruit of the human mtDNA tree. *Trends Genet.* 22, 339–345
3 Hebert, P.D. *et al.* (2003) Biological identifications through DNA barcodes. *Proc. Biol. Sci.* 270, 313–321
4 Satoh, N. (2003) The ascidian tadpole larva: comparative molecular development and genomics. *Nat. Rev. Genet.* 4, 285–295
5 Boffelli, D. *et al.* (2004) Intraspecies sequence comparisons for annotating genomes. *Genome Res.* 14, 2406–2411
6 Simmen, M.W. *et al.* (2000) Sequence analysis of transposable elements in the sea squirt, *Ciona intestinalis*. *Mol. Biol. Evol.* 17, 1685–1694
7 Suzuki, M.M. *et al.* (2005) Genomic approaches reveal unexpected genetic divergence within *Ciona intestinalis*. *J. Mol. Evol.* 61, 627–635
8 Gissi, C. *et al.* (2004) Complete mtDNA of *Ciona intestinalis* reveals extensive gene rearrangement and the presence of an *atp8* and an extra *trnM* gene in ascidians. *J. Mol. Evol.* 58, 376–389
9 Yokobori, S. *et al.* (2003) Mitochondrial genome of *Ciona savignyi* (Urochordata, Ascidiacea, Enterogona): comparison of gene arrangement and tRNA genes with *Halocynthia roretzi* mitochondrial genome. *J. Mol. Evol.* 57, 574–587
10 Dehal, P. *et al.* (2002) The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science* 298, 2157–2167
11 Saccone, C. *et al.* (1991) The main regulatory region of mammalian mitochondrial DNA: structure-function model and evolutionary pattern. *J. Mol. Evol.* 33, 83–91
12 Pesole, G. *et al.* (1988) A backtranslation method based on codon usage strategy. *Nucleic Acids Res.* 16, 1715–1728
13 Yang, Z. (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* 13, 555–556
14 Goldman, N. *et al.* (1994) A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol. Biol. Evol.* 11, 725–736
15 Hasegawa, M. *et al.* (1998) Preponderance of slightly deleterious polymorphism in mitochondrial DNA: nonsynonymous/synonymous rate ratio is much higher within species than between species. *Mol. Biol. Evol.* 15, 1499–1505
16 Zouros, E. *et al.* (1994) An unusual type of mitochondrial DNA inheritance in the blue mussel *Mytilus*. *Proc. Natl. Acad. Sci. U. S. A.* 91, 7463–7467
17 Le, T.H. *et al.* (2000) Phylogenies inferred from mitochondrial gene orders: a cautionary tale from the parasitic flatworms. *Mol. Biol. Evol.* 17, 1123–1125
18 Littlewood, D.T. *et al.* (2006) The complete mitochondrial genomes of *Schistosoma haematobium* and *Schistosoma spindale* and the evolutionary history of mitochondrial genome changes among parasitic flatworms. *Mol. Phylogenet. Evol.* 39, 452–467
19 Nohara, M. *et al.* (2005) Evolution of the mitochondrial genome in cephalochordata as inferred from complete nucleotide sequences from two *Epigonichthys* species. *J. Mol. Evol.* 60, 526–537
20 Rawlings, T.A. *et al.* (2001) A major mitochondrial gene rearrangement among closely related species. *Mol. Biol. Evol.* 18, 1604–1609
21 Shao, R. *et al.* (2006) Molecular mechanisms for the variation of mitochondrial gene content and gene arrangement among chigger mites of the genus *Leptotrombidium* (Acari: Acariformes). *J. Mol. Evol.* 63, 251–261
22 Shao, R. *et al.* (2003) Rates of gene rearrangement and nucleotide substitution are correlated in the mitochondrial genomes of insects. *Mol. Biol. Evol.* 20, 1612–1619
23 Xu, W. *et al.* (2006) The relationship between the rate of molecular evolution and the rate of genome rearrangement in animal mitochondrial genomes. *J. Mol. Evol.* 63, 375–392
24 Caputi, L. *et al.* (2007) Cryptic speciation in a model invertebrate chordate. *Proc. Natl. Acad. Sci. U. S. A.* 104, 9364–9369

0168-9525/$ – see front matter
doi:10.1016/j.tig.2007.07.001

Comparison of tryptophan biosynthetic operon regulation in different Gram-positive bacterial species

Ana Gutiérrez-Preciado[1], Charles Yanofsky[2] and Enrique Merino[1]

[1] Departmento de Microbiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos 62210, México
[2] Department of Biological Sciences, Stanford University, Stanford, CA 94305, USA

The tryptophan biosynthetic operon has been widely used as a model system for studying transcription regulation. In *Bacillus subtilis*, the *trp* operon is primarily regulated by a tryptophan-activated RNA-binding protein, TRAP. Here we show that in many other Gram-positive species the *trp* operon is regulated differently, by $tRNA^{Trp}$ sensing by the RNA-based T-box mechanism, with T-boxes arranged in tandem. Our analyses reveal an apparent relationship between *trp* operon organization and the specific regulatory mechanism(s) used.

Corresponding author: Merino, E. (merino@ibt.unam.mx).
Available online 6 June 2007.