PARTIAL SEQUENCES OF THE SHRIMP Penaeus notialis MITOCHONDRIAL GENOME

Erik García Machado,[1] Mario Oliva Suárez,[1] Jean Claude Mounolou[2] and Monique Monnerot[2]

[1]*Centro de Investigaciones Marinas, University of Havana, Cuba.* [2]*Centre de Génétique Moléculaire, C.N.R.S., F-91198 Gif sur Yvette, Cedex, France.*

Introduction

Since the last past decade, animal mitochondrial DNA has received special interest for population and evolutionary biology studies (1). It has also been considered as a useful target molecule to study molecular mechanisms involved in the control of transcription and replication process (2). Gene contein of the animal mtDNA is highly conserved but in spite of this gene rearrangements have widely occur during mtDNA evolution. The more striking examples are founded in invertebrates (3, 4, 5). Most of the changes involved tRNA genes relocations, however there are a number of examples where complete segments have been inverted and have been considered as useful markers to clarified phylogenetic relationships (6).In this work we report partial sequences of the shrimp *Penaeus notialis* mtDNA. This brings new information about genome organization of a group where knowloged is limitated.

Results and Discussion

Genome Organization

A 7.9 Kb BglII fragment and a 1 Kb PCR amplified fragment of the shrimp *P. notialis* mitochondrial genome were partially sequenced. They contain the genes for the subunits I, II, III of the CO complex, the NADH deshidrogenase subunits 2, 3 and 5 and the ATPasa6 and 8 subunits, 16 tRNA genes. (phe, glu, ser(AGY), asn, arg, ala, gly, lys, leu(UUR), tyr, cys, trp, met, gln, ile and val) and the small and the large rRNA genes. The main non-coding region that have 982 nucleotides occur between the $tRNA^{ile}$ and the small rRNA genes. The $tRNA^{asp}$ was non identified probably by its potencial location up strean from the $tRNA^{lys}$ whose 3' end was not determined. All genes determined are arranged and have the same transcriptional polarities as their counterparts of *D. yakuba* and *Daphnia pulex* mtDNAs (3,7). However. in *A. franciscana*, another representant of the Crustacea, the genes for $tRNA^{ile}$ and $tRNA^{gln}$ appear translocated between the $tRNA^{gly}$ and $tRNA^{trp}$ (8).

tRNA Genes

P. notialis tRNA genes vary in size from 66bp (tyr, leu, ala and arg) to 72bp (val) with a mean value of 68, similar to that founded in other metazoan tRNAs reported to date. In contrast to *P. notialis*, there is a reduction in size of most of the homologous genes (63bp as a mean) of *A.franciscana* and *D.pulex* (7, 8) that is mainly due to the presence of smaller TΨ C stems and loops.

The Non-coding Region

The region of 982bp located between the $tRNA^{ile}$ and the small rRNA genes, lack any open reading frame longer than 57 nucleotides. It has an A+T contain (79.3 %) intermediate between other crustaceans (67 %-68 %) (7, 8) and, insects (92.8 %-96 %) (3,5). In the last third of the region (closest to the $tRNA^{ile}$ gene) there is a potencial hairping loop forming sequence, that involved a 20bp stem with an out-loop of one nucleotide, a G-T pairing, and a loop of 31 nucleotides. In this region, but closes to the small rRNA gene, there are also two perfect direct repeats of 11 nucleotides, and two imperfect repeads that will involved 30-33 nucleotides. Perfect repeats are also present in similar map positions of *D. pulex* control region. If we consider that the hairping-loop structure is in relation whit the L-strand replication initiation, as it has also been proposed for *D. yakuba* and *D. pulex* (7, 9), then in *Artemia* the translocation of such an origen between ND2 and COI (8) is a peculiarity of this specie.

rRNA Genes

The genes, for the small and large rRNA subunits are located at both sides of the $tRNA^{val}$ gene. The 5' end of the l2S gene has been deduced by sequence comparison to the homologous regions of the *D.yakuba* and *A.franciscana* mtDNAs (3,10). The beginning of the $tRNA^{va}$ gene has been taken to determine the 3' end. Thus *P. notialis* l2S-RNA gene is 859 nucleotides long, the largest one among the Arthropoda l2S genes sequenced to date. The position and 5' assigment of the l6S-RNA is only putative.

Protein Coding Genes

P. notialis and *D. pulex* protein genes are quite similar in nucleotide composition but different from *D. yakuba* where the AT content is higher. Alignements with the homologous regions of the corresponding genes of *A. franciscana, D. yakuba* and mouse (11), evidenced not unusual features in their lengths on the contrary of *A. franciscana* that exhibits a striking loss of residues, specially at the 5'end of the ND2 gene. In *P. notialis* mtDNA protein genes, as in other species (4) initiation codons will be all four ATN codon family. TAG and TAA appear as termination codons (ATPase6 and ATPase8). All other genes with the exception of ND2 (3'termini not determined) ends in TA or T.

Genetic Code and Codon Usage

P. notialis appear to use, at hight frequency, codons ending in A or T (80 %) which is much higher than the overall composition of protein genes (63 %). Codons AGA, TGA and ATA appear to have the same assigments in *P. notialis* than in *D. yakuba* mtDNA.

Remark

The gene organization as well as the tRNA structure and the length of the polypeptides differntiate the mt genome of *P.notialis* from that of *A.franciscana*. This signifies a *polyphyletism* of mtDNA within the Crustacean order.

1. Harrison Trends Ecol Evol 1989;4:6-1
2. Clayton Int.Rev.Cytol 1992;141:217-232
3. Clary and Wolstenholme J Mol Evol 1985;22:25-271.
4. Wolstenholme Int Rev Cytol 1992; 141:173-216.
5. Crozier and Crozier Genetics 1993;133:97-117.
6. Smith *et al.* J Mol Evol 1993;36:545-554.
7. Van Raay and Crease Curr.Genet. 1994; 25:66-72.
8. VálVerde *et al.* J Mol Evol 1994; 4:400-408.
9. Clary and Wolstenholme J Mol Evol 1987;25:116-125.
10. Clary and Wolstenholme Nucleic Acids Res 13:402949045;Bibb *et al.*,(1982);Cell 1985; 26:167-180.