Atti Accademia Nazionale dei Lincei

Classe Scienze Fisiche Matematiche Naturali

Rendiconti

Daniela Bellincampi, Francesco Amaldi

Duplication-divergence evolution of the procaryotic genome suggested by DNA reassociation kinetics of E.coli DNA

Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti, Serie 8, Vol. **65** (1978), n.5, p. 222–226.

Accademia Nazionale dei Lincei

<http://www.bdim.eu/item?id=RLINA_1978_8_65_5_222_0>

Biologia molecolare. — *Duplication-divergence evolution of the procaryotic genome suggested by* DNA *reassociation kinetics of* E.coli DNA. Nota di DANIELA BELLINCAMPI (*) e FRANCESCO AMALDI (*), presentata (**) dal Corrisp. E. MARRÈ.

RIASSUNTO. — Da studi delle cinetiche di riassociazione del DNA di *E.coli*, abbiamo dimostrato che il genoma batterico comprende delle sequenze ripetute con diverso grado di divergenza: alcune molto simili, capaci cioè di riassociare bene tra loro, ed altre più o meno eterogenee.

Queste evidenze sperimentali sono in accordo con le ipotesi di Zipkas e Riley (1975) e di Masters (1975) secondo i quali il genoma di *E.coli* si sarebbe originato per successive duplicazioni di un genoma primitivo. La conseguente divergenza evolutiva avrebbe introdotto una certa eterogeneità tra sequenze inizialmente identiche.

It is generally accepted that a polyphasic pattern of reassociation kinetics is characteristic of eucaryotic DNA, the procaryotic DNA exhibiting a more simple kinetics. In other words, the eucaryotic genome includes medium repetitive, highly repetitive and palindromic sequences, besides unique ones (Britten and Kohne, 1968; Davidson and Britten, 1973), while procaryotic DNA would be made up only of unique sequences. It is reasonable, though, to expect the presence of DNA fractions reassociating faster than bulk DNA in procaryotes also: repeated DNA sequences as genes for rRNA which are present in only a few copies in *E.coli* (Attardi and Amaldi, 1970) and palindromes within genes coding for RNA molecules with a complex secondary structure as in eucaryotes (Davidson *et al.*, 1973). On the other hand, although a polyphasic pattern of DNA reassociation might be expected, it has not been observed to date in the procaryotes.

We present here a study of the reassociation kinetics of *E.coli* B DNA showing the presence of sequences reassociating faster than bulk DNA. An example of a reassociation experiment on *E.coli* DNA sheared to a fragment size of about 300 nucleotides is shown in Fig. 1 (closed circles). The presence of a repetitive fraction (about 10 %) reassociating about ten times faster than bulk DNA, and a very fast reassociating DNA fraction (about 5 %) is evident. These results cannot be due to the presence of a virus or a plasmid since, using the method of Telford *et al.* (Telford *et al.*, 1977), we have checked our *E.coli* strain, which on the other hand had been chosen because known to contain no such structures. Moreover, the observed pattern of interspersion of unique and repeated sequences (see below) rules out this type of artifact. Another possible artifact, namely that some of the DNA was not sheared and thus renatured more quickly than the rest, is also ruled out: in fact the obser-

(*) Centro Acidi Nucleici del C.N.R., Istituto di Fisiologia Generale, Università di Roma.
(**) Nella seduta del 10 novembre 1978.

vation of a constant fraction of repetitive DNA in different experiments and the regular slow increase of this fraction with DNA M.W. (see below) are difficult to explain by casual errors in DNA shearing. In any case a control experiment has been carried out: ^{3}H–DNA, sheared to a size of about 300 nucleotides, was reassociated at Cot 2×10^{-1} and the repetitive and unique sequences fractionated on HAP. The two fractions were then sedimented on

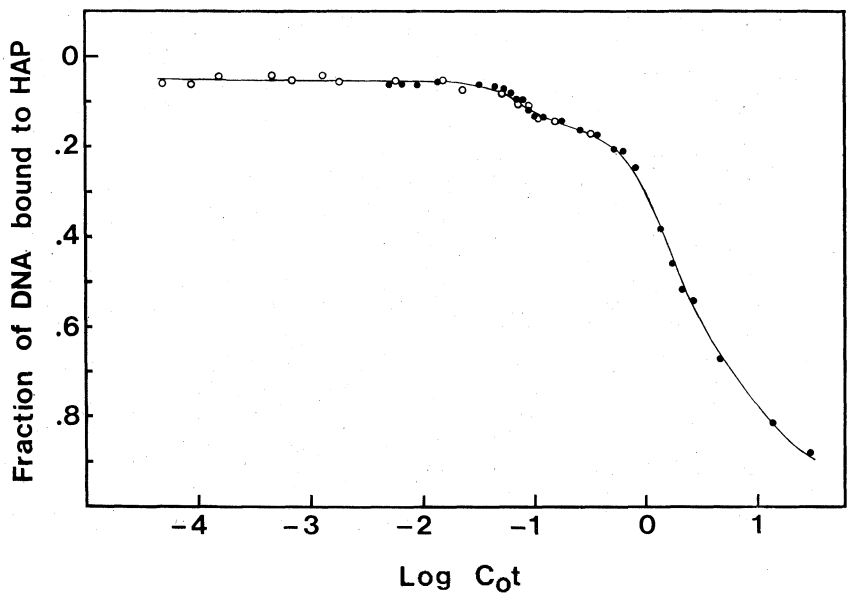

Fig. 1. – HAP reassociation kinetics of low M. W. DNA from *E.coli* B. Non radioactive DNA (●) was extracted from *E.coli* B cells grown in complete medium by a modified Marmur (1961) procedure. Further purification by preparative CsCl density gradient centrifugation, shearing and size determination of DNA fragments by analytical alkaline sedimentation were carried out as previously described (Baldari and Amaldi, 1976). 1 ml samples of DNA at a concentration of 100 μg/ml in 0.12 M sodium phosphate buffer (PB) were denatured at 100 °C for 10 min and reassociated at 65 °C for increasing times. Reassociation was evaluated by HAP fractionations as previously described (Baldari and Amaldi, Radioactive DNA (○) was extracted from *E.coli* B cells grown in minimal medium containing (methyl-^{3}H) thymine (1 μg/ml; 4.5 Ci/mmole; Radiochemical Centre, Amersham) and purified as described above. In this case reassociation was carried out at a DNA concentration of 1 μg/ml in 0.12 M PB buffer. Fractions from HAP columns were supplemented with 100 μg/ml Bovine serum albumin, TCA precipitated, filtered and counted after addition of 10 ml liquid scintillation mixture.

two alkaline sucrose gradients to check DNA M.W. No difference in M.W. was observed for the two fractions. A similar result has been obtained in an analogous experiment carried out with non-radioactive DNA; in this case the sedimentation in alkali of the two fractions was checked by analytical ultracentrifugation.

In order to distinguish whether the fast reassociating DNA fraction represents self-reassociating or highly repetitive sequences, it was necessary to extend the analysis to lower Cot values. For this purpose radioactively labeled

E.coli DNA was prepared, which made it possible to operate at low DNA concentrations; the left part of the curve in Fig. 1 (open circles) shows the results obtained in a reassociation experiment carried out with ^{3}H-DNA sheared to a fragment size of about 300 nucleotides. It is evident that the fast reassociating DNA has already reannealed at a Cot value as low as 3×10^{-5}. This non dependence of the reaction on DNA concentration indicates a self-

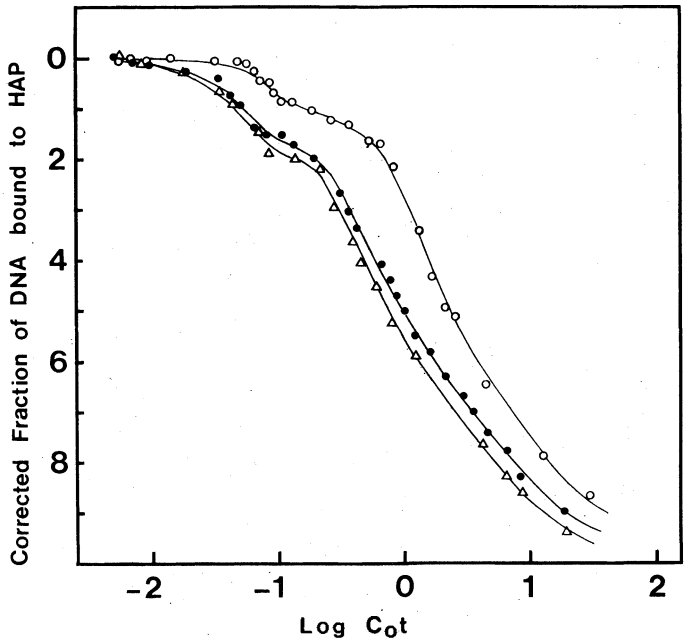

Fig. 2. – HAP reassociation kinetics of *E.coli* B DNA sheared to various fragment lengths (○), 300 nucleotides; (●), 4,000 nucleotides; (△), 8,500 nucleotides. Data are corrected for zero-time binding (Davidson *et al.*, 1973).

reassociation of these DNA fragments which would thus represent palindromic sequences, the zero time binding sequences described by Davidson *et al.* (1973).

To study the organization of the about ten fold repeated and the unique sequences along the genome of *E.coli*, DNA was sheared to different fragment sizes (Davidson *et al.*, 1973). Reassociation kinetics at DNA fragment sizes of about 300, 4,000 and 8,500 nucleotides are shown as examples in Fig. 2 after correction for zero time binding sequences. There is an evident dependence of the kinetics on the size of the DNA fragments as a progressive lowering of the curves. The percentage of DNA fragments behaving as repeated sequences on HAP, in these and in other curves not shown, are plotted against fragment length in Fig. 3. The slow increase of the amount of DNA bound to HAP indicates that the repetitive sequences are very heterogeneous in size and in any case quite long.

In order to obtain an estimate of the faithfulness of duplex reassociation, melting profiles have been studied for reassociated unique and repetitive DNA

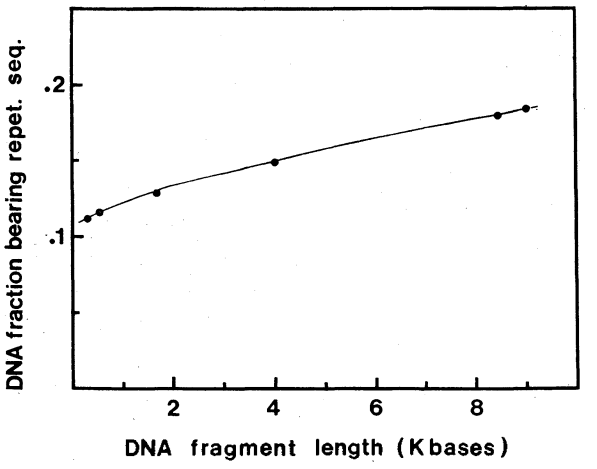

Fig. 3. – Fraction of DNA fragments bearing repetitive sequences, as a function of fragment length. Data are obtained from the curves shown in Fig. 2 and others not shown.

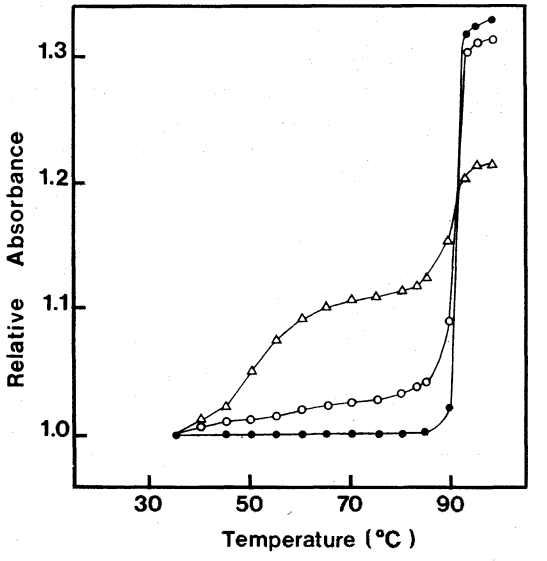

Fig. 4. – Melting profiles of native *E.coli* DNA (●), reassociated unique DNA (○) and reassociated repeated DNA sequences (△). DNA sheared to a fragment size of about 1000 nucleotides was denatured and passed through HAP to remove zero-time binding sequences. The unbound fraction was reassociated to a Cot value of 1.5×10^{-1} and passed again on HAP to separate reassociated repetitive fraction (A) from single stranded unique DNA. This was further reassociated to a Cot value of 2.5×10^{1} and again passed through HAP from which reassociated unique sequences were finally obtained (B). Reassociated repetitive (A) and reassociated unique (B) fractions were dialyzed against SSC (0.015 M Na citrate, 0.15 M NaCl, pH 7.0). DNA samples were melted in SSC in water-jacketed cuvettes and the hyperchromicity monitored at 260 mμ with a Zeiss spectrophotometer.

fractions. The two melting profiles obtained are compared in Fig. 4 with that of native *E.coli* DNA. It is evident that reassociated unique DNA has a Tm (91 °C) very similar to that of native DNA (92 °C). On the contrary, the melting profile of reassociated repeated sequences appears to be biphasic, denaturing in part with a Tm of 89 °C and in part with a Tm as low as 50 °C. Note that this temperature is lower than the one at which reassociation had taken place, that is 65 °C. This result can be interpreted as due to the single-stranded tails adjacent to the repeated sequences reannealed at 65 °C. Such tails, which cannot reassociate at this temperature (and are thus included in the unique fraction), would do so at lower temperatures during concentration and dialysis of the sample, indicating a low sequence homology. The plateau between the two transitions would be due to the particular reannealing conditions utilized (65 °C etc.; criterion of Davidson *et al.* (1973) by which the fragments with at least one region of fairly good pairing are selected).

In conclusion, the results here presented suggest that the genome of *E. coli* is made up of interspersed repeated sequences with different degrees of divergence, some so similar to each other as to be able to reanneal quite well, others completely different (unique sequences). Such a structural organization of the *E.coli* genome can be interpreted according to the hypothesis put forward recently (Zipkas and Riley, 1975; Masters, 1975) on the basis of studies of the genetic map of *E.coli*. These authors have proposed that the *E.coli* genome has evolved by two or three subsequent duplications of a smaller primitive genome, occurring during evolutive times, followed by mutations and divergence of functions of the replicated genes. Different selective pressure on the various genome regions probably leads to a different degree of divergence, and is responsible for the about ten times repeated sequences (10 % of the genome) we observed and for the peculiar melting profile these have after reassociation.

We thank Mr. Di Francesco for excellent technical assistance in preparative and analytical ultracentrifugations.

References

ATTARDI G. and AMALDI F. (1970) - « Ann. Rev. Biochem. », *39*, 183-226.
BALDARI C. T. and AMALDI F. (1976) - « Chromosoma », *59*, 13-22.
BRITTEN R. J. and KOHNE D. E. (1968) - « Science », *161*, 529-540.
DAVIDSON E. H. and BRITTEN R. J. (1973) - « Quant. Rev. Biol. », *48*, 565-613.
DAVIDSON E. H., HOUGH B. R., AMENSON C. S. and BRITTEN R. J. (1973) - « J. Mol. Biol. », *77*, 1-23.
MARMUR J. (1961) - « J. Mol. Biol. », *3*, 208-218.
MASTERS M. (1975) - « Nature », *256*. 616-617.
MIZOBUCHI K. and SAITO H. (1975) - « Proc. of the Mol. Biol. Meeting of Japan », p. 60.
TALFORD J., BOSELEY P., SCHAFFNER W. and BIRNSTIEL, M. (1977) - « Science », *195*, 391-393.
ZIPKAS D. And RILEY M. (1975) - « Proc. Nat. Acad. Sci. U.S.A. », *72*, 1354-1358.