Volume 6 Number 2 February 1979 Nucleic Acids Research

Interspersion of highly repetitive DNA with single copy DNA in the genome of the red crab, Geryon quinquedens*

Nelwyn T.Christie† and Dorothy M.Skinner

University of Tennessee-Oak Ridge Graduate School of Biomedical Sciences and the Biology Division, Oak Ridge National Laboratory, Oak Ridge, TN 37830, USA

Received 16 October 1978

ABSTRACT

Kinetic analysis of the reassociation of 420 nucleotide (NT) long fragments has shown that essentially all of the repetitive sequences of the DNA of the red crab *Geryon quinquedens* are highly repetitive. There are negligible amounts of low and intermediate repetitive DNAs. Though atypical of most eukaryotes, this pattern has been observed in all other brachyurans (true crabs) studied (1,2). The major repetitive component is subdivided into short runs of 300 NT and longer runs of greater than 1200 NT while the minor component has an average sequence length of 400 NT. Both components reassociate at rates commonly observed for satellite DNAs. Unique among eukaryotes the organization of the genome includes single copy DNA contiguous to short runs (∿300 NT) of both repetitive components. Although patent satellites are not present, subsets of the repetitive DNA have been isolated by either restriction endonuclease digestion or by centrifugation in Ag^{+} or Hg^{2+}/Cs_2SO_4 density gradients.

INTRODUCTION

Although the interspersion of single copy with repeated DNA is a general pattern of organization for eukaryotic genomes, there is considerable variability in the length of the two classes of sequences. As originally described for *Xenopus laevis* (3), most organisms have 70-80% of their single copy DNA in lengths of 800-2000 NT followed by repetitive sequences 300 NT long (short period interspersion). A distinctly different pattern (long period interspersion) is characterized by more than ten-fold greater lengths of both repetitive and single copy sequences. Until recently the long period interspersion pattern was thought to be an anomaly found in the fruit fly (4) and the honeybee (5) but not in all insects (5,6). It has now been demonstrated in the genomes of the water mold (7), the Syrian hamster (8) and the chicken (9,10).

It seemed appropriate to examine individual differences and similarities of genome organization within a particular taxon. For example, the genomes of the crabs (brachyurans, 11) *Libinia emarginata*, *Cancer borealis*, and

Gecarcinus lateralis, are all similar in their scarcity of sequences of slow and intermediate repetitive frequencies (i.e., those that reassociate 10 to 1000 times the rate of single copy DNA; 1,2). Highly repetitive satellite DNAs account for 11 to 30% of these genomes. For comparison, we have studied the organization of the genome of *Geryon*, a primitive brachyuran which lacks patent satellites but contains large amounts of highly repetitive DNA.

MATERIALS AND METHODS

Frozen testes of *Geryon quinquedens* were obtained from the Marine Biological Laboratory, Woods Hole, Massachusetts. DNA purified by adsorption to hydroxyapatite (12) was stored in 1 mM EDTA and 0.1 mM NaCl. Except where noted, all manipulations performed on the DNA including labeling with *E. coli* DNA polymerase I, shearing, reassociation, treatment with S1 nuclease and thermal dissociation of the reassociated as well as native DNA were as described (2). Heavy metal Cs_2SO_4 gradients, analytical and preparative, and CsCl gradients containing actinomycin D were prepared as described (13,14).

A. Preparation of Iodinated DNA. Sheared or unsheared DNAs were iodinated according to Commerford (15) with several modifications (16,17). Immediately after denaturation in an ethylene glycol bath (5 min, 104°) 25 µg of DNA was added to a reaction mixture of 0.25 mCi Na ^{125}I and thallium chloride (0.2 mM in 0.5 M sodium acetate buffer). Iodination was carried out at 60° for 5 min at an equivalent Cot of 1×10^{-3}. For some experiments native DNA was iodinated without denaturation. The iodinated DNA was dialyzed against 0.4 M NaCl, 15 mM sodium phosphate buffer (pH 6.0) and 0.2 mM EDTA and then placed at 90°, 15 min, to remove unstably bound iodine. Specific activities ranged from 3×10^4 to 1.5×10^5 cpm/µg for denatured DNA and $1\text{-}2 \times 10^4$ cpm/µg for native DNA. Iodinated molecules were selected from preparative alkaline sucrose gradients (18) and used within two weeks.

B. Restriction Endonuclease Analyses. Alu I was isolated according to H. Heynecker by a procedure similar to that of Roberts *et al.* (19). Hind III was isolated according to Smith (20). Restriction endonuclease digestions were carried out as described (19). Before digestion, samples of reassociated DNA were adjusted to 0.6% formaldehyde to prevent further reassociation (21). Dilute samples were concentrated by adsorption to hydroxyapatite followed by ethanol precipitation. For digestion of kinetic fractions the amount of DNA in duplex was calculated from the S1 resistance of similarly isolated fractions. Restriction fragments were sized by electrophoresis on acrylamide gels (3.5 - 5%) as described (22,23). DNA fragments less than 200 NT long were sized by co-electrophoresis with ØX174

DNA fragments of known sizes obtained by Alu I digestion of specific Hind II + III fragments (22).

RESULTS

1. Reassociation Kinetics of 420 Nucleotide Fragments. Reassociation kinetics of 420 NT fragments (Fig. 1) showed that *Geryon* DNA contains two general classes of sequences, highly repetitive (60%) and presumptive single copy (35%); an additional 5% of the DNA was unreactive at Cot 4 x 10^4. The solid curve in Fig. 1 describes a least-squares analysis of the data fit by computer for three second order components which subdivides the highly repetitive DNA into two fractions (12). Table 1 lists the characteristics of these components. By these analyses very highly and highly repeated sequences comprise 22% and 34% of the DNA respectively; approximately 4% of the DNA bound to hydroxyapatite at Cot 10^{-5} ("zero-time binding DNA"). In other eukaryotic genomes, foldback sequences (palindromes) have been a major component of this fraction (24).

The stated reiteration frequencies in Table 1 are average values that may represent a range of repetition frequencies. They were calculated as the ratios of the rates of reassociation for the repetitive components and the assumed single copy component as determined in unfractionated DNA (12,25). Despite some uncertainty in the calculated frequencies, these values permit a better comparison with components in other genomes than do the rate constants alone. The Cot curve shares with those of other true crabs a most striking feature: the seemingly complete absence of middle repetitive sequences that reassociate between Cots 2 and 100 (1,2).

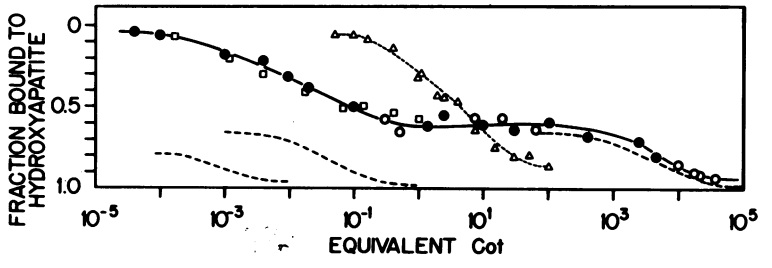

Figure 1. Reassociation kinetics of *Geryon* DNA. 420 NT long fragments of DNA were denatured, reassociated, and assayed on hydroxyapatite, unlabeled (o); Pol I labeled (●); ^{125}I-iodinated (□). The solid line represents the solution of a least-squares computer analysis which also resolved the second order components shown by the dashed lines. As a standard, the Cot 1/2 of iodinated *E. coli* DNA (Δ) 300 NT long, was measured.

Table 1

KINETIC ANALYSIS OF GERYON DNA REASSOCIATION

COMPONENT	FRACTION OF TOTAL DNA	LIMITS OF REASSOCIATION (Cot)	AVERAGE RATE CONSTANT ($l\ M^{-1}\ sec^{-1}$)	$Cot_{1/2}$	AVERAGE REPETITION FREQUENCY
ZERO-TIME BINDING DNA	0.04	10^{-5}	–	–	–
VERY HIGHLY REPETITIVE	0.22	10^{-5} - $2x10^{-3}$	984.0	$1x10^{-3}$	$3.9x10^{6}$
HIGHLY REPETITIVE	0.34	$2x10^{-3}$ - 2	20.7	$4.8x10^{-2}$	$8.3x10^{4}$
SINGLE COPY	0.35	300 - $3x10^{4}$	$2.5x10^{-4}$	$4x10^{3}$	1

2. Interspersion Studies by Hyperchromicity Measurements and S1 Digestion. If repetitive sequences are interspersed with single copy DNA, reassociation of fragments longer than the length of the repetitive sequences to Cot values at which only repetitive sequences reassociate yields duplexes with single-stranded tails (26). The fraction of paired bases in reassociated DNA can be determined from hyperchromicity measurements or by digestion of single-stranded ends by S1 nuclease (25,26). To this end, different sized fragments were incubated to Cot 2 to permit reassociation of the repetitive sequences only. The hyperchromicity of reassociated 250, 420 and 1700 NT fragments respectively was 88, 63 and 53 per cent that of native DNA (Fig. 2). The greatly reduced hyperchromicity of the longer fragments is consistent with the interspersion of short repetitive sequences and single copy DNA.

The per cent of the genome estimated as repetitive by collecting DNA reassociated to Cot 2 on hydroxyapatite includes single copy sequences attached to reassociated duplexes. The duplex content of reassociated 420

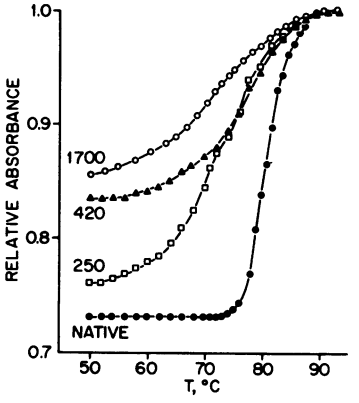

Figure 2. Thermal Dissociation of Reassociated DNA. DNA fragments of 250, 420, and 1700 NT were denatured, reassociated to Cot 2 and melted. A melt of native DNA is included for comparison.

NT fragments obtained from melting profiles allowed a correction of the estimate for the amount of repetitive DNA derived from kinetic measurements (Table 2). Although after reassociation to Cot 2, 51% of the 420 NT fragments adsorbed to hydroxyapatite only 59% of the bases in the bound DNA were paired; therefore, only 30% of the genome (0.51 x 0.59) is repetitive. Similar calculations for 250 and 1700 NT fragments gave values of approximately 39% repetitive DNA.

3. Estimate of Amounts and Lengths of Repetitive Components. A more accurate measure of the amount of DNA in each repetitive component was obtained by digesting single-stranded ends of reassociated fragments with S1 nuclease. Following isolation on hydroxyapatite, the S1 resistant duplexes were sized by chromatography on agarose $A_{50}m$. The size distribution for zero-time binding DNA in 2000 NT fragments was determined after reassociation to Cot 10^{-5} (Fig. 3A). Although these sequences ranged in size from less than 300 to 1200 NT, the majority were approximately 300 NT. The DNA reassociating between Cot 10^{-5} and 2 x 10^{-3} (9% of the genome) is composed of zero-time and very highly repetitive sequences. The size distribution for the latter was corrected for the contribution of zero-time binding DNA (5% of the genome) to obtain the distribution of only the very highly repeated sequences (lower curve, Fig. 3B). Approximately 80% of the very highly repeated sequences ranged in size from 300 to 500 NT. By correcting for the amount of DNA in duplex at Cot 2 x 10^{-3} recovered in the Cot 2 fraction, it was found that 29% of the genome was comprised of highly repetitive DNA (lower curve, Fig. 3C). Approximately 60% of that component (17% of the genome) was arranged in repetitive clusters longer than 1200 NT. The remainder of the fraction was estimated as stretches of 300 NT interspersed

Table 2

DUPLEX CONTENT OF DNA FRAGMENTS REASSOCIATED TO Cot 2.0

FRAGMENT LENGTH (NUCLEOTIDES)	250	420	1700
FRACTION BOUND TO HYDROXYAPATITE (F)	0.45	0.51	0.81
HYPERCHROMICITY (H)[a]	0.241	0.172	0.145
T_m (°C)	76	75.5	76
AVERAGE DUPLEX CONTENT FROM HYPERCHROMICITY OF BOUND FRAGMENTS (D)[b]	0.86	0.59	0.48
PERCENT OF GENOME, (D X F)	0.39	0.30	0.39

[a] Calculated as (A_{260} at 100°C - A_{260} at 55°C)/A_{260} at 100°C.

[b] Calculated as (H - 0.025)/(H_{native} - 0.025). A correction of 2.5% is made for the hyperchromicity of single strands. H_{native} = 0.275.

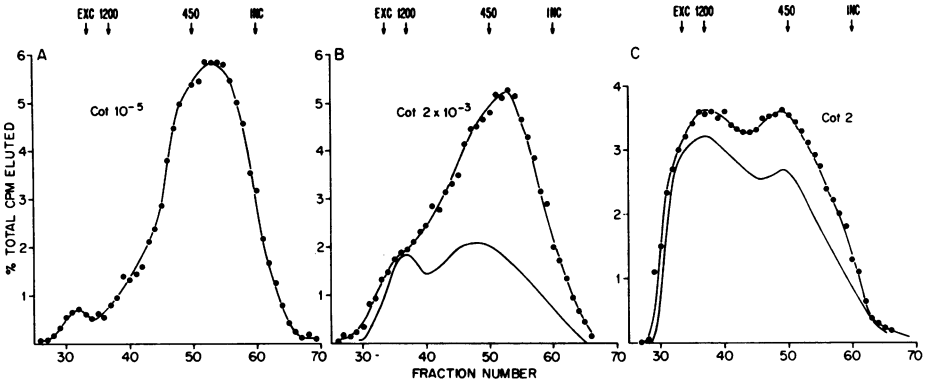

Figure 3. Size Distribution of S1 Resistant Sequences of DNA. ^{125}I-iodinated 2000 NT fragments were reassociated to (A) Cot 10^{-5}, (B) Cot 2 x 10^{-3} or (C) Cot 2 and digested with S1 nuclease. Positions are indicated for an included marker (Inc; ^{3}H-uridine), DNA fragments 450 and 1200 NT long, and an excluded marker (Exc; 4000 NT). The lower curve of (B) shows the size distribution of repetitive sequences which reassociate between Cot 10^{-5} and Cot 2 x 10^{-3}, i.e., the very highly repetitive component. The lower curve of (C) represents the size distribution for sequences which reassociate between Cot 2 x 10^{-3} and Cot 2, i.e., the highly repetitive component.

with single copy DNA.

Although the data on ^{125}I-iodinated DNA included in the Cot curve (Fig. 1) were superimposable on those from either Pol I labeled or unlabeled DNA, a report of destabilization of DNA duplexes by the presence of 5-iodocytosine (27) prompted us to determine the susceptibility to S1 nuclease of reassociated repeated regions in iodinated *Geryon* DNA. As a control unlabeled 2000 NT fragments were reassociated to Cot 2 and treated with S1 nuclease. The size distribution of the unlabeled duplexes (data not shown) was nearly identical to that of iodinated duplexes isolated in the same manner (Fig. 3C). We conclude that the reassociation kinetics of iodinated DNA did not yield an underestimate of the lengths of the repetitive sequences or an overestimate of the fraction of interspersed repetitive sequences.

4. Organization of Repetitive DNA Classes and Single Copy DNA. A measure of the spacing of repetitive sequences is obtained by reassociating different sized fragments to the same Cot in the presence of an excess of short fragments (250-400 NT) as driver DNA (3,26). For the *Geryon* genome the spacing of the sequences of both repetitive components was measured simultaneously by reassociation to Cot 2 (Fig. 4, Table 3). Corrections were made for zero-time binding DNA and adjacent sequences as for other

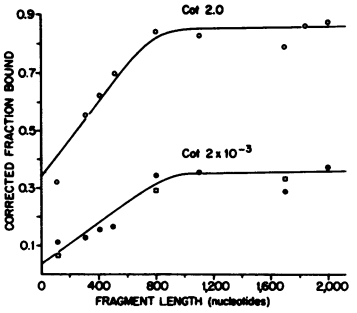

Figure 4. Fractions of *Geryon* DNA Containing Repetitive Elements as a Function of Fragment Length. ^{125}I-iodinated DNA fragments ranging in size from 110 to 2000 NT as determined by alkaline sucrose gradients were combined with a 100-fold excess of unlabeled 300 NT driver DNA. Reassociation, monitored by hydroxyapatite chromatography, was performed to driver Cot 2 with tracer DNA iodinated after denaturation (o) and to driver Cot 2 x 10^{-3} for tracer DNA iodinated with (●) and without (□) denaturation. See Table 3 for calculation of the corrected fraction bound.

genomes in which such sequences are interspersed with both single copy and repetitive sequences (3,25). When 2000 NT long fragments were reassociated to Cot 2 in the presence of a 100-fold excess of short unlabeled fragments, 89% bound to hydroxyapatite. This result indicates that most of the genome is organized in a short period interspersion pattern with repetitive sequences interspersed every 800-1000 NT as in many other eukaryotes (2,6,28). No more than 11% of the genome contains stretches of single copy DNA longer than 800-1000 NT. Interspersion measurements also provided an independent measure of the amount of repetitive DNA. The Y-intercept of the line for the amount of DNA bound when the fragments are less than 1000 NT is an

Table 3

FRACTION OF FRAGMENTS CONTAINING REPETITIVE SEQUENCES

FRAGMENT LENGTH (NUCLEOTIDES)	FRACTION BOUND TO HYDROXYAPATITE (F)			CORRECTED FRACTION BOUND (F_C)[a]	
	Cot 10^{-5}	Cot 2 X 10^{-3}	Cot 2	Cot 2 X 10^{-3}	Cot 2
110	0.09	0.2	0.38	0.12	0.31
300	0.08	0.2	0.58	0.14	0.56
400	0.09	0.23	0.65	0.16	0.63
500	0.19[b]	0.33[b]	0.78[b]	0.18	0.72
800	0.21[b]	0.48[b]	0.87[b]	0.35	0.84
1100	0.17[b]	0.46[b]	0.85[b]	0.36	0.82
1700	0.24	0.46	0.84	0.29	0.79
2000	0.21[b]	0.5[b]	0.89[b]	0.36	0.86

[a] The fraction bound to hydroxyapatite (F) is corrected for zero-time sequences (Z) according to the formula, $F_c = \frac{F - Z}{1 - Z}$, where F_c is the corrected fraction bound.

[b] The fraction bound to hydroxyapatite is an average of two determinations that differ by 3% or less.

estimate of the amount of duplex DNA which reassociates by the specified Cot (26). Since essentially all the repetitive DNA reassociates by Cot 2, the amount of repetitive DNA is 35% excluding zero-time binding DNA.

The interspersion of the very highly repetitive component alone was measured with 300 NT fragments of driver DNA at Cot 2 x 10^{-3} by which less than 5% of the highly repeated component would have reassociated. A significant amount of the very highly repeated fraction was interspersed with other sequences in the genome. Approximately 30% of the genome is arranged as very highly repeated sequences interspersed with either highly repeated or single copy sequences averaging 1000 NT (Fig. 4; Table 3). The negligible increase in hydroxyapatite binding of fragments longer than 1000 NT indicates that longer spacings between these sequences are uncommon. This interspersion pattern for the very highly repeated sequences was confirmed by determining the fraction of DNA bound to hydroxyapatite at Cot 2 x 10^{-3} for 100, 800, and 1700 NT fragments that had been iodinated without denaturation. The data points fell on the same curve derived from similarly sized fragments denatured before iodination, indicating that reassociation during the iodination reaction did not cause significant underiodination of very highly repeated sequences. The sequences adjacent to the very highly repeated component could include repeated sequences that reassociated after Cot 2 x 10^{-3}, single copy sequences, or both. The Y-intercept of the line for the amount of DNA fragments bound at Cot 2 x 10^{-3} for lengths less than 1000 NT corroborates the S1 digestion data in Section 3. Taken together, they indicate that only 4% of the genome is comprised of sequences of the very highly repeated component rather than 22% as calculated from kinetic analyses.

By contrast the interspersion of the very highly repeated fraction with single copy DNA was demonstrated by recovering from hydroxyapatite DNA that had been reassociated to Cot 2 x 10^{-3}. This fraction which contains zero-time binding and very highly repetitive DNAs and their adjacent sequences, was further reassociated to Cot 2 and digested with S1 nuclease. If the very highly repeated component is interspersed with single copy DNA, the fraction recovered at Cot 2 will contain S1 sensitive DNA. Subtracting from the total S1 sensitive DNA (33%; Fig. 5B) in the isolated fraction, the 15% S1 sensitive DNA adjacent to zero-time sequences (Fig. 5A), we obtained an estimate for the amount of single copy DNA adjacent to the very highly repeated component. 18% is a minimum estimate since some of the DNA which abuts zero-time binding DNA may be highly repetitive. Verification of this value was obtained by comparing estimates of the amount of repetitive DNA from the two

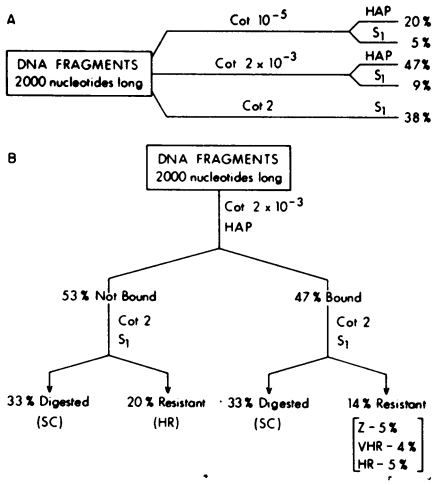

Figure 5. Analysis of Single Copy DNA Adjacent to the Very Highly Repetitive Component.

A. *Direct Measurements:* The percent of ^{125}I-iodinated DNA fragments containing a duplex was monitored by hydroxyapatite (HAP) chromatography. The percent of the genome in duplex at the different Cots was measured as the S1 resistance of DNA fragments.

B. *Fractionation Measurements:* ^{125}I-Iodinated DNA fragments with and without very highly repetitive sequences were isolated by reassociation to Cot 2 x 10^{-3}, denatured and reassociated to Cot 2 in the presence of driver DNA of the same size. The S1 resistance of the DNA fragments was determined after the second reassociation. SC refers to single copy DNA, HR to highly repetitive DNA, VHR to very highly repetitive DNA and Z to zero-time binding DNA.

types of experiments (*Direct and Fractionation*, Fig. 5). After incubation of 2000 NT fragments directly to Cot 2, 38% of the DNA was in duplex (Fig. 5A); after the two-step reassociation experiment (Fig. 5B), 34% (20% + 14%) had formed duplexes. Despite the 4% difference in the two estimates of the amount of repetitive DNA, the data clearly show interspersion of the very highly repetitive component with single copy DNA.

5. Further Characterization of the Repetitive DNA: a. Density gradient centrifugation in the presence of ligands. Although 17% of the *Geryon* genome is arranged as long repetitive sequences (Fig. 3C), the total DNA formed a single band (1.700 g/cm^3) in neutral CsCl gradients (Fig. 6A). Three minor satellites were observed on centrifugation of the DNA in the presence of either Hg^{2+} or Ag^+/Cs_2SO_4 gradients (Fig. 6B,C). Two satellites comprising 3-5% of the genome were detected in the presence of Hg^{2+} while a

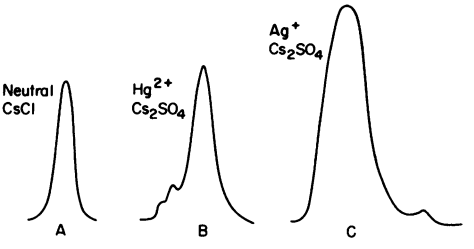

Figure 6. Model E Analytical Ultracentrifuge Tracings of Total DNA. (A) 3 µg of DNA in neutral CsCl; (B) 5 µg of DNA in neutral Cs_2SO_4 with Hg^{2+} (rf = 0.26). (C) 15 µg of DNA in neutral Cs_2SO_4 with Ag^+ (rf = 0.19).

single satellite comprising 1-2% of the genome was seen in the presence of Ag^+. No additional satellites were detected with actinomycin D at a ratio as high as 1:1. In sum, the cryptic satellites could account for less than one third of the long repetitive sequences detected by the kinetic analyses.

While these results do not exclude the presence of long tandem arrays of short repeating sequences, they do indicate a greater probability that the majority of the sequences in the long repetitive class have a longer basic repeat. The probability of a highly repetitive DNA being separated from the bulk of the DNA in density gradients is greater for a short repeating sequence in which an overall bias in base composition is more likely. The longer and more complex the repeat, the greater the probability that the density will be the same as the major component, even in the presence of base-specific ligands.

b. Digestion with restriction endonucleases. Another means of excising and characterising the repetitive sequences in *Geryon* DNA, all of which are present at frequencies characteristic of satellites, is by digestion with restriction enzymes. Such enzymes have been used to excise highly repetitive DNAs from total DNA of numerous eukaryotic genomes (29-34).

Following a series of digestions with numerous restriction enzymes, Alu I and Hind III were selected since they yielded fragments of convenient length for subsequent sequencing and their recognition sites were similar. The Alu I site is comprised of the four central nucleotides of the Hind III site. Accordingly, the Hind III fragments must include many but not necessarily all of the Alu I fragments. Treatment of toal *Geryon* DNA with Hind III (Fig. 7A, Lane 4) and Alu I (Fig. 7B, Lane 4) produced a background of heterogeneously sized fragments with large amounts of fragments of discrete sizes superimposed on the background. Treatment of each of the three kinetic fractions with Alu I yielded different amounts of the discrete fragments (Fig. 7B). Zero-time binding DNA contained none (Lane 1); the very highly repeated (Lane 2) and highly repeated (Lane 3) fractions contained large amounts of 75 and 150 NT fragments. Digests of the highly repeated, but not the very highly repeated, fraction also contained fragments 85, 95, 130, 140 and 200 NT long. Because of the frequency class from which the latter fragment are derived, it is clear that there are fewer of them than of those sequences that contain the 75 and 150 NT fragments. Because of the lower specificity of the Alu I site (i.e., the shorter recognition sequence may be expected to occur at higher frequency) there was a higher background of heterogeneously sized fragments in that digest than in the Hind III digest.

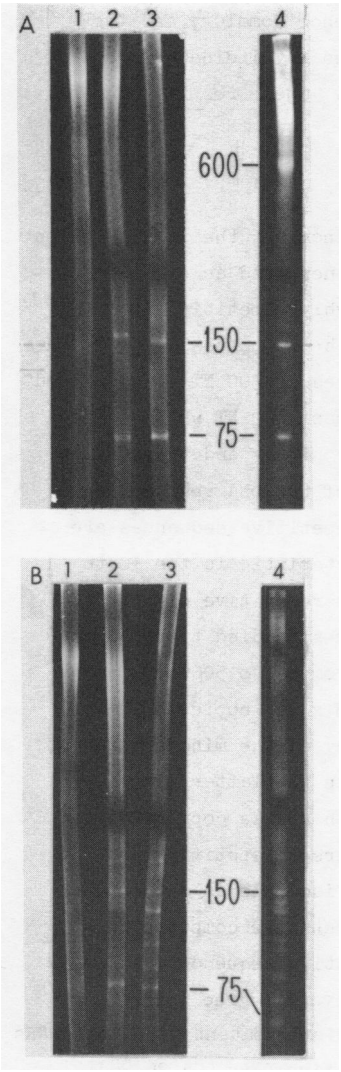

Figure 7. Restriction Endonuclease Digestion of Three Kinetic Fractions and Native DNA. (A,B) (1) Zero-time binding DNA was isolated after rapid addition of formaldehyde to heat-denatured 1500 NT fragments. (2) Sequences of the very highly repetitive component were isolated by reassociation of 1500 NT fragments to Cot 2 x 10^{-3}. (3) Highly repetitive sequences were collected by further reassociation to Cot 2 of the single-stranded DNA from a Cot 2 x 10^{-3} sample. Each kinetic fraction (5 µg duplex DNA) and (4) native DNA (6 µg) were digested with either Hind III (A) or Alu I (B). Electrophoresis on acrylamide gels was for 1.5 hr for A, Lanes 1-4 and B, Lanes 1-3 and 2 hr for B, Lane 4. (C) Densitometer tracings of the Alu I digests of very highly repetitive (1; B, Lane 2) and highly repetitive (2; B, Lane 3) sequences.

We have begun sequence analyses (35) of the Hind III fragments because of the lower background. As in highly reiterated human (33) and simian DNAs (34), within each size class there are several populations of molecules since there is more than one residue at certain positions. Nevertheless, there is extensive similarity among the molecules in any one size class. Furthermore, evidence to date indicates considerable homology among the fragments of different sizes. Even from these initial results it is clear that the sequences are not composed of short, simple repeats as found in several crus-

tacean satellite DNAs (23,36-38). The partial sequence homology and the variation in copy number of closely related sequences may indicate extensive alterations in a previously homogeneous reiteration. These results will be discussed in detail elsewhere.

DISCUSSION

One distinctive characteristic of the *Geryon* genome is the interspersion of single copy DNA with both a highly repeated component and an even more highly repeated component; interspersion of such highly repetitive DNA with single copy DNA has not been observed previously. The sequences of the highly repeated component are subdivided into short stretches of 300 NT interspersed with single copy DNA and longer stretches greater than 1200 NT whose interspersion pattern has not been defined. Although the highly and the very highly repetitive components account for 29 and 4% of the DNA respectively (Table 4), as in other brachyuran genomes, middle repetitive sequences are few in number or absent. Another distinctive characteristic is the 10 to 1000 times greater repetition frequency of the major repetitive component than has been found in other genomes. In most genomes studied to date, the frequencies of the major repetitive classes range from 10 to 5000 copies (3,6,25,26) while in *Geryon* there are as many as 8.3×10^4 copies of the major repetitive component. The repetition frequency of the minor component is similar to that observed for *Aplysia* DNA (25). In the latter genome, however, the minor component is not interspersed with single copy DNA. The rate constants for reassociation of the two interspersed repetitive components of *Geryon* are more similar to those of DNAs isolable as satellites (39,40) than to those for the major interspersed repetitive components in other genomes. In addition, the reassociated repetitive sequences of *Geryon* exhibit only 4% mismatch (assuming 1% mismatch for every degree lowering of the T_m, 12), a value also characteristic of reassociated patent satellite DNAs (40) but less than that of other reassociated repetitive DNAs (25,26,28). Furthermore, the pairing of the reassociated repetitive sequences in *Geryon* DNA is sufficiently faithful to allow recognition and cutting by restriction enzymes under standard conditions. Although the restriction fragments could be obtained from long repetitive sequences, there is no *a priori* reason to assume that they are excluded from the short repetitive sequences.

Because of the extensive interspersion of single copy DNA with repetitive DNA in the *Geryon* genome the amount of single copy sequences was underesti-

mated by the usual kinetic analyses. From either interspersion measurements or S1 digestion (Table 4) the more accurate value of 57% was obtained. Although the very highly repeated component accounts for only 4% of the genome, it is interspersed with approximately 25% of the genome, including single copy DNA and probably some of the highly repeated sequences. Several possible arrangements of both repetitive components with single copy DNA are shown in Fig. 8A and B. A fraction of the highly repeated component is also interspersed only with single copy DNA. Two other proposed arrangements show clustering of repetitive sequences as might occur in cryptic satellites (Fig. 8C,D). Short runs (300-500 NT) of the very highly repeated component could be adjacent to sequences of the highly repeated component which, in turn, could contain internal repeats related but not identical to those of the very highly repeated sequences. From the partial sequence homologies observed for three different Hind III fragments (Christie and Skinner, unpublished data) this postulated organization might involve the two classes of sequences containing Alu I sites, the abundant sequences juxtaposed with those of less abundance. Alternatively, a cryptic satellite might consist primarily of very highly repeated sequences, i.e. only the abundant sequences containing the Alu I sites.

Many but not all brachyuran genomes have relatively high percentages of patent satellite DNAs (36,41-43). Of the crustacean genomes whose organization has been examined, all have at least 25% highly repeated DNA. In *Libinia emarginata* satellites constitute 11% of the genome which has a total

Table 4

ESTIMATES OF THE AMOUNT OF EACH COMPONENT IN THE GENOME

COMPONENT	KINETIC[a] (420 nt)	INTERSPERSION[b] (110-2000 nt)	S_1[c] (2000 nt)
ZERO-TIME BINDING DNA	4.0%	6.0%	5.0%
VERY HIGHLY REPETITIVE	22.0	4.0	4.0
HIGHLY REPETITIVE	34.0	28.0	29.0
SINGLE COPY	35.0	57.0	57.0
UNREACTIVE	5.0	5.0	5.0

[a]Derived from the least squares fit to the data for the reassociation of 420 nucleotide fragments assuming three second order components.

[b]Determined from reassociation of labeled fragments of increasing length with 300 nucleotide driver DNA to Cots 10^{-5}, $2x10^{-3}$, and 2.

[c]S_1 resistance at Cot 10^{-5}, $2x10^{-3}$, and 2 corrected for each repetitive component. Undigested DNA at Cot 2 is corrected for unreactive sequences to estimate single copy DNA.

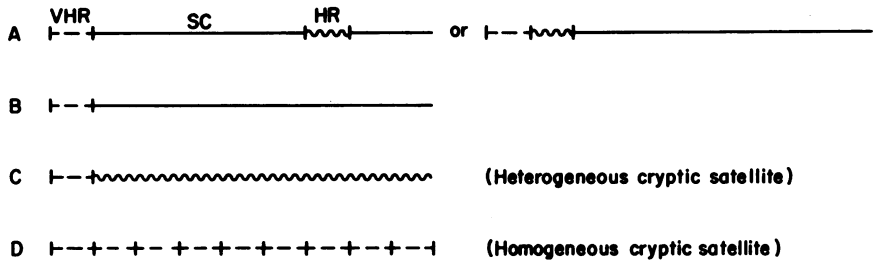

Figure 8. Possible Arrangements for Sequences of the Very Highly Repetitive Component. Symbols are as follows: (---) VHR, very highly repetitive sequences; (∿∿) HR, highly repetitive sequences; and (___) SC, single copy sequences.

of 25% highly repetitive DNA; in *Cancer borealis* satellites account for 30% of the DNA with an additional 10% kinetically defined as highly repetitive sequences (1). Approximately 55% of the DNA of *Gecarcinus lateralis* is in long repetitive sequences, half of which are isolable as satellites (2,36). In the *Geryon* genome the highly repeated DNA is nearly equally divided between long and short repetitive sequences. If, as suggested (44), short repetitive sequences are derived from long repetitive sequences, the interspersion of the very highly repeated sequences with single copy DNA in the *Geryon* genome may indicate a more rapid transfer of long repetitive sequences to single copy regions as compared to other eukaryotes. While it is likely that the same general mechanisms of DNA evolution occur in all eukaryotes, it is this abundance of both satellite DNAs and other highly repetitive fractions in crustacean genomes which may aid the study of possible relationships among classes of repeated DNAs.

Studies of the DNAs of distantly related crustaceans, such as shrimps and lobsters, will establish whether the organization of crab genomes is typical of the entire Class *Crustacea*.

ACKNOWLEDGMENTS

We thank W. L. Carrier, W. E. Masker and S. Mitra for materials; and C. A. Holland and D. E. Graham for discussion.

†N. T. Christie is a predoctoral fellow supported by Grant GM 1974 from the National Institute of General Medical Sciences, National Institutes of Health. *This research was sponsored by the Division of Biomedical and Environmental Research, U.S. Department of Energy, under contract W-7405-

eng-26 with the Union Carbide Corporation.

REFERENCES

1 Vaughn, J. C. (1975) Chromosoma (Berl.) 50, 243-257
2 Holland, C. A. and Skinner, D. M. (1977) Chromosoma (Berl.) 63, 223-240
3 Davidson, E. H., Hough, B. R., Amenson, C. S. and Britten, R. J. (1973) J. Mol. Biol. 77, 1-23
4 Manning, J. E., Schmid, C. W. and Davidson, N. (1975) Cell 4, 141-155
5 Crain, W. R., Davidson, E. H. and Britten, R. J. (1976) Chromosoma (Berl.) 59, 1-12
6 Efstratiadis, A., Crain, W. R., Britten, R. J., Davidson, E. H. and Kafatos, F. C. (1976) Proc. Natl. Acad. Sci. (USA) 73, 2289-2293
7 Hudspeth, M. E. S., Timberlake, W. E. and Goldberg, R. B. (1977) Proc. Natl. Acad. Sci. (USA) 74, 4332-4336
8 Moyzis, R., Bonnet, J. and Ts'o, P. O. P. (1977) J. Cell Biol. 75, 130a
9 Arthur, R. R. and Straus, N. A. (1978) Can. J. Biochem. 56, 257-263
10 Eden, F. C. (1978) Biochemistry. In press.
11 *Brachyura* is an infraorder of the decopod crustaceans that includes the true crabs.
12 Britten, R. J., Graham, D. E. and Neufeld, B. R. (1974) in Methods in Enzymology, Grossman, L. and Moldave, K., Eds., Vol. 29E, pp. 363-418 Academic Press, New York
13 Skinner, D. M. and Beattie, W. G. (1973) Proc. Natl. Acad. Sci. (USA) 70, 3108-3110
14 Kersten, W., Kersten, H. and Szybalski, W. (1966) Biochemistry 5, 236-244
15 Commerford, S. L. (1971) Biochemistry 10, 1993-1999
16 Getz, M. J., Altenburg, L. C. and Saunders, G. F. (1972) Biochim. Biophys. Acta 287, 485-494
17 Orosz, J. M. and Wetmur, J. G. (1974) Biochemistry 13, 5467-5473
18 Noll, H. (1969) in Techniques in Protein Biosynthesis, Campbell, P. N. and Sargent, J. R., Eds., Vol. 2, pp. 101-179 Academic Press, New York
19 Roberts, R. J., Myers, P. A., Morrison, A. and Murray, K. (1976) J. Mol. Biol. 102, 147-165
20 Smith, H. O. (1974) in Methods in Molecular Biology, Wickner, R. B., Ed., Vol. 7, pp. 71-85 Academic Press, New York and London
21 Feldman, M. Ya. (1973) in Progress in Nucleic Acid Research and Molecular Biology, Davidson, J. N. and Cohn, W. E., Eds. Vol. 13, pp. 1-49 Academic Press, New York and London
22 Maniatis, T., Jeffrey, A. and van deSande, H. (1975) Biochemistry 14, 3787-3794
23 Chambers, C. A., Schell, M. P. and Skinner, D. M. (1978) Cell 13, 97-110
24 Wilson, D. A. and Thomas, C. A., Jr. (1974) J. Mol. Biol. 84, 115-144
25 Angerer, R. C., Davidson, E. H. and Britten, R. J. (1975) Cell 6, 29-39
26 Graham, D. E., Neufeld, B. R., Davidson, E. H. and Britten, R. J. (1974) Cell 1, 127-137
27 Anderson, D. M. and Folk, W. R. (1976) Biochemistry 15, 1022-1030
28 Walbot, V. and Dure, L. S., III (1976) J. Mol. Biol. 101, 503-536
29 Botchan, M. R. (1974) Nature (Lond.) 251, 288-292
30 Cooke, H. J. (1975) J. Mol. Biol. 95, 87-99
31 Gruss, P. and Sauer, G. (1975) FEBS Letters 60, 85-88
32 Mowbray, S. L., Gerbi, S. A. and Landy, A. (1975) Nature (Lond.) 253, 367-370

33 Manuelidis, L. and Wu, J. C. (1978) Nature (Lond.) 276, 92-94
34 Rosenberg, H., Singer, M. and Rosenberg, M. (1978) Science 200, 394-402
35 Maxam, A. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. (USA) 74, 560-564
36 Skinner, D. M. (1967) Proc. Natl. Acad. Sci. (USA) 58, 103-110
37 Skinner, D. M., Beattie, W. G., Blattner, F. R., Stark, B. P. and Dahlberg, J. E. (1974) Biochemistry 13, 3930-3937
38 Skinner, D. M. (1977) BioScience 27, 790-796
39 Waring, M. and Britten, R. J. (1966) Science 154, 791-794
40 Cordeiro-Stone, M. and Lee, C. S. (1976) J. Mol. Biol. 104, 1-24
41 Sueoka, N. (1961) J. Mol. Biol. 3, 31-40
42 Skinner, D. M. and Kerr, M. S. (1971) Biochemistry 10, 1864-1872
43 Beattie, W. G. and Skinner, D. M. (1972) Biochim. Biophys. Acta 281, 169-178
44 Britten, R. J. and Davidson, E. H. (1971) Quart. Rev. Biol. 46, 111-138