Evolution of Genome Size by DNA Doublings

Minimum genome size in major taxonomic groups suggests an evolutionary series of DNA doublings.

Arnold H. Sparrow and Anne F. Nauman

In the course of radiobiological investigations over the past 15 years, we have determined various nuclear parameters for more than 1100 species of higher plants (*1–5*). One such parameter is DNA content per genome (DNA/G); we use the term "genome" to indicate one basic chromosome complement. It is important to note that DNA/G is the haploid (gametic) amount of DNA only for diploid species; in polyploids, it is the somatic nuclear DNA divided by ploidy level, or the gametic amount decreased proportionately (*6*). We have found that DNA/G for species within some families or genera, particularly in the gymnosperms, is lognormally distributed (*3, 5*). Such a distribution is skewed when plotted on a linear scale, but takes the shape of a normal distribution when plotted in increments of equal linear width on a log-scale abscissa. This type of distribution is not uncommon for biological parameters (*3, 7*). However, in most higher plant families for which we have adequate data, except for the gymnosperms, DNA/G is distributed neither normally nor lognormally; rather, the log-scale distributions generally form a series of several peaks (*4, 5*) (Fig. 1). Increase in DNA/G within a taxonomic group appears to be associated with an increase in chromosome size (DNA per chromosome), but not necessarily in chromosome number (*4, 5*).

In conjunction with these studies, we have compiled data, mainly from the reported work of other investigators, on DNA/G for more than 1300 additional species of lower plants, animals, and prokaryotes; we have converted these published DNA values, which are expressed in a variety of units, to the uniform unit of picograms (10^{-12} g). When the DNA/G data within each major phylogenetic grouping are similarly plotted as semilog distributions, as described above, many of these larger groups of diverse organisms also demonstrate a series of peaks, often with similar peak values, as in the three widely separated taxonomic groups shown in Fig. 2. These multimodal distributions appear to represent a series of doublings of a minimum genome size for each taxonomic group. We used the term cryptopolyploidy to define this doubling phenomenon as it was first noted in higher plants, fungi, and bacteria (*4*). Conventional polyploidy denotes a multiplication of a basic chromosome number while generally maintaining a relatively constant chromosome size. Cryptopolyploidy, on the other hand, implies an increase in genome size (DNA/G) by increase in chromosome size (DNA per chromosome). Stated simply, polyploidy results in *more* chromosomes; cryptopolyploidy results in *larger* chromosomes.

Others, notably Ohno (*8*), have observed large increases or doublings of genome size or of chromosome size in animals (*8–10*), higher plant families (*11*), bacteria (*12*), and viruses (*13*), as well as in mitochondria (*14*) and bacterial plasmids (*15*). Most investigators have assumed that such DNA increases, at least in the higher organisms, are due to polyploidy, polyteny, or extensive gene duplication. Wallace and Morowitz (*12*), however, have proposed that total genome doubling may have had an evolutionary role as a means of independent development of early prokaryotes (by genome size increase) and of eukaryotes (by genome number increase) from a "protokaryote" ancestral to the present-day *Mycoplasma*.

Dr. Sparrow is a senior cytologist and Ms. Nauman is a biology associate in the Biology Department, Brookhaven National Laboratory, Upton, New York 11973.

Exponential Periodicity

In order to determine the mean genome size of species encompassed by each of the peaks in DNA/G, we arranged the data in tabular form and correlated gaps or valleys between peaks in each histogram, as determined visually, with corresponding gaps in the tabulated data. We then calculated the mean DNA/G for species included in each peak as thus delineated. Such gaps are not always clear, particularly when large numbers of species fall into adjacent peaks, but they are often quite clearly defined, as in Figs. 1 and 2.

When we plotted these means logarithmically against a series of theoretical doublings for each taxonomic group (as done for the grasses in Fig. 3), each value fell very close to a doubling point. Numerical similarities in these means among widely different taxonomic groups (Figs. 1 and 2) suggested a series of doublings of a basic ancestral genome common to many of these groups. When the DNA/G means for all peaks within all taxa surveyed were taken together and similarly plotted against theoretical doublings, the close fit of points to the resulting line indeed gave evidence of an exponential periodicity encompassing most major life forms, and extending over nine orders of magnitude. (The large number of densely clustered points precludes clear graphic presentation here.) When we extrapolated this line downward, by halving instead of doubling, we found that RNA values for viroids and for single- and double-stranded RNA viruses fit the series at appropriate halving or doubling points, and extended the line into a tenth order of magnitude. The DNA values for mitochondria and chloroplasts, which are commonly considered to have originated as endosymbionts, also fit into the apparent pattern, although we have not presented the data here.

The periodicity becomes more clearly demonstrable when only the smallest known *individual* genome size for each phylogenetic group is considered (Table 1 and Fig. 4). These values involve neither selection nor averaging; each is simply the smallest individual value that we were able to find (from our literature search) for any species within a given taxonomic group. Presumably these values approximate the minimum amount of nucleic acid requisite for each phylogenetic group, or at least for extant members of each group. Species with lower values may exist, but these values may not yet have been determined. However,

in most taxonomic groups surveyed, the lowest peak is quite restricted in number of species and in range of DNA/G (Figs. 1 and 2). We should therefore expect that, if any lower values for a given group are found, they would either be very close to the minimum value reported here or would comprise an even lower peak in the series. These minimum DNA (or RNA) values for all taxonomic groups, when taken together, again appear to represent a continuous series of doublings.

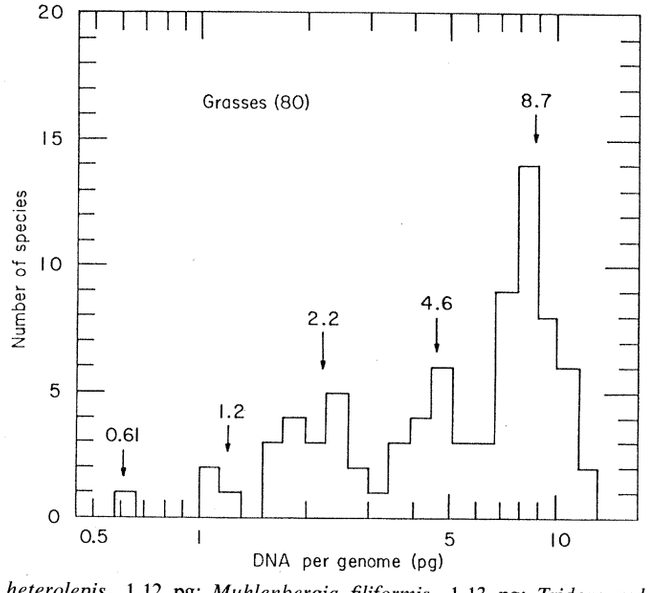
In order to present this concept graphically (Fig. 4) and to test its statistical validity, we first plotted values for organisms 1 to 23 in Table 1 against a sequence of theoretical doublings of an "elemental genome" expressed as 2^n, where n is equal to the number of genome doublings. The lowest RNA value (point 1) is taken as 1 on the abscissa, or 2^0. Each succeeding Y value (minimum DNA/G or RNA/G) was plotted against the doubling point that would best permit that Y value to fall on an estimated line of +1 slope. A line fitted to these points by the method of least squares has a slope of +1.0003 with a correlation coefficient of .9998. Next, a line with fixed +1 slope was fitted to the points by the same method (Fig. 4). The 23 points appear to fall naturally very close to the theoretical doublings. If the interval between the Y estimates for two successive doubling numbers is taken to be unity, then, under the hypothesis of a random distribution, we would expect the observed Y values to be uniformly distributed between −.5 and +.5 about a mean of zero. In fact, all of the Y values fall within ±.36 of the line. The probability of this occurring by chance alone is less than .001. Also under the random hypothesis, we would expect 50 percent of the Y values to fall between −.25 and +.25. In fact, 20 of the 23 points fall in this region closest to the line. A simple chi-square test, $\chi^2 = 12.57$ (1 d.f.; $P < .01$), thus gives evidence that the Y values correspond much more closely to doubling points that would be expected on the basis of chance.

Evolutionary Complexity

The lowest point in Fig. 4 represents the RNA viroids, which are the smallest known independent biological entities. They comprise a newly recognized class of subviral pathogenic RNA's, and include the causative agents of potato spindle tuber disease (PSTV), citrus exocortis disease (CEV), chrysanthemum stunt disease, and possibly scrapie disease of sheep (*16, 17*). These "minimal infectious molecules" are highly structured, incompletely base-paired RNA molecules, probably with a hairpin conformation, containing no DNA or protein (*17*), and with a molecular weight of about 100,000 (*16, 17*). In physical characteristics CEV and PSTV are very similar and may, in fact, be identical, since CEV has been shown to cause spindle tuber disease in potatoes (*17*). We have taken this value (equivalent to 1.65×10^{-7} pg) to represent the basic minimum genome (point 1 in Fig. 4).

Considering *all* the values surveyed, rather than just the minimum values given here (Table 1 and Fig. 4), all organisms with DNA/G below 5×10^{-3} pg are prokaryotes; all above 3.6×10^{-2} pg are eukaryotes. The range of overlap, spanning less than one order of magnitude (2^{15} through 2^{18} in Fig. 4), includes only the bacteria and blue-green algae with the largest genomes and yeasts and protozoa with the smallest genomes. The larger amounts of DNA requisite for multicellular organisms are apparently more easily maintained when organized into a multichromosomal system adapted for mitosis. The greater amounts of DNA necessary for greater diversity in form and function can then be accommodated by increases in genome and chromosome size (cryptopolyploidy) or by increases in genome and chromosome number (conventional polyploidy). However, such a system also allows for greater divergence and variability in DNA amount through such mechanisms as tandem duplications, B chromosomes, satellite DNA, viral transformation, and deletions.

The DNA/G of most species of complex life forms is in the range of 0.35 to 11 pg (2^{21} to 2^{25} in Fig. 4); the only groups of higher organisms which include appreciable numbers of species with DNA/G above this level are the lungfish, amphibi-



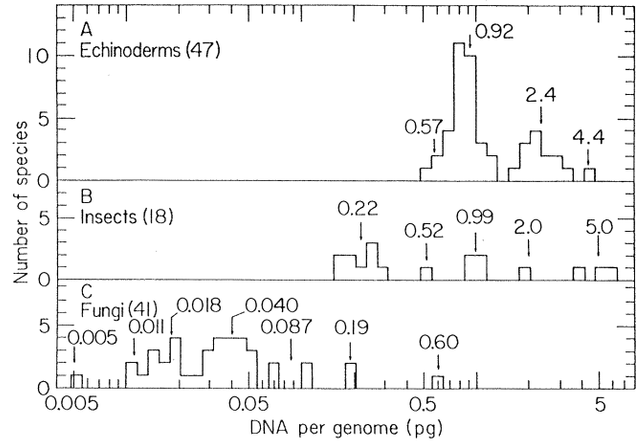


Fig. 1 (left). Logarithmic distribution of DNA per genome (DNA/G) for 80 acquisitions of grasses. Each log cycle on the abscissa is divided into 16 equal linear increments, which are the same for each cycle. Arrows indicate the mean DNA/G for species within each distributional group (see text). Modified from (*4*), with four new values added (*Sporobolus heterolepis*, 1.12 pg; *Muhlenbergia filiformis*, 1.13 pg; *Tridens pulchellus*, 1.77 pg; *Sporobolus cryptandrus*, 2.58 pg). Fig. 2 (right). Logarithmic distribution of DNA/G for 47 echinoderms (*51, 57, 59*), 18 insects (*53, 60*), and 41 fungi (*47, 61*). Increments are the same as those used in Fig. 1. Arrows indicate the mean DNA/G for species within each distributional group (see text). In (A), there is a gap in the data of nearly a full increment between the first two distributional groups; this is not apparent here because of the width of increments. The same is true in (C), between the second and third and between the third and fourth groups.

ans, and certain groups of higher plants (gymnosperms and monocotyledonous angiosperms). These high DNA/G species may have been derived by cryptopolyploidy from related species with lower DNA/G, since they tend to have very large chromosomes and relatively low chromosome numbers (*2*, *3*, *5*). Such increases in DNA content are generally associated with corresponding increases in cell size (*18*), a parameter which is closely related to many physiological functions. Thus, increase in cell size with concomitant increase in DNA content may be of primary importance in evolutionary and adaptational change within a taxonomic group. It should be emphasized that the values used in Fig. 4 represent minimum individual values, and that within most taxonomic groups there exists a fairly wide range in DNA/G, which may encompass several intragroup doublings (as shown in Figs. 1 and 2). Thus, although the highest individual animal and plant DNA/G so far reported are in the most highly evolved groups (angiosperms and vertebrates) (*19*), there are several minimum DNA/G values for other taxa above the minima for these two groups. However, all are very ancient groups, and it seems reasonable to assume that ancestral low DNA forms within these groups may have become extinct, while actual minimum DNA per genome of extant forms may have been obscured, in some cases, by diploidization of ancient polyploids, or it may have diverged significantly through any of the mechanisms previously mentioned.

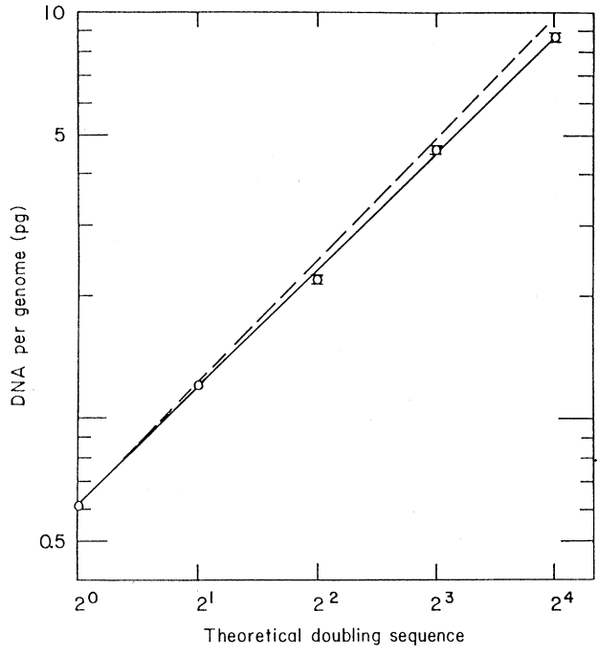


Fig. 3. Best fit of DNA/G distributional group means from Fig. 1 plotted against a theoretical doubling sequence. Dashed line indicates a slope of +1. [Modified from (*4*)]

The limited amount of data available on chromosome numbers and ploidy levels, particularly in the animals, often renders uncertain the calculation of DNA/G from total nuclear DNA. In the echinoderms, for example, most known somatic chromosome numbers are high (between 36 and 50), which strongly suggests ancient polyploidy. In these cases, if DNA/G is equated to the haploid (gametic) amounts of DNA, as is done here, such values may actually represent more than one genome, and thus be artificially high. There is a strong likelihood that the actual minimum genome sizes for echinoderms, molluscs, annelids, and crustaceans are lower than the values we have given here, because of polyploidy; that is, they actually may be similar to, or lower than, the minimum for the angiosperms and vertebrates. If we should reduce the given DNA/G values because of assumed polyploidy, the goodness of fit of data points to genome doublings demonstrated in Fig. 4 would probably not be altered; ploidy levels generally increase in even multiples, so that tetraploidy, for instance, would simply change an assumed diploid value from 2^n in Fig. 4 to 2^{n-1}.

In the ferns and some families of bryophytes, a similar situation of high chromosome number and uncertain ploidy level is common. Also, there have been virtually no DNA determinations reported for these lower plant groups. Nearly all the data available to us have been estimates of DNA content based on our measurements of nuclear volume, and the direct positive correlation of nuclear volume with DNA content demonstrated by Baetcke *et al.* (*6*) for higher plants has not yet been demonstrated to be equally valid for the lower plants. Therefore, the ferns and bryophytes have not been included in Fig. 4, even though the minimum DNA/G in each

Table 1. Smallest reported genome size for 23 major phylogenetic groups of organisms. Numbers in first column correspond to points in Fig. 4.

Phylogenetic group	Organism	DNA (or RNA) per genome (pg)
1. RNA viroids	Potato spindle tuber disease virus; citrus exocortis disease virus	$\sim 1.65 \times 10^{-7}$ (*16*,*17*)
2. Single-stranded RNA viruses	Satellite tobacco necrosis virus	6.60×10^{-7} (*43*)
3. Double-stranded RNA viruses	*Penicillium stoloniferum* virus S, L particles	1.53×10^{-6} (*44*)
4. Single-stranded DNA viruses	Minute virus of mice	2.48×10^{-6} (*45*)
5. Double-stranded DNA viruses	Rabbit kidney vacuolating virus	4.62×10^{-6} (*43*)
6. Bacteria	*Mycoplasma arginini; M. bovigenitalium*	6.6×10^{-4} (*46*)
7. Fungi	*Saccharomycodes ludwigii*	5×10^{-3} (*47*)
8. Blue-green algae	*Anabaena variabilis*	6×10^{-3} (*48*)
9. Protozoa	*Plasmodium berghei*	2×10^{-2} (*49*)
10. Eukaryotic algae	*Chlorella ellipsoidea*	5×10^{-2} (*50*)
11. Porifera	*Dysidea crawshagi*	5.5×10^{-2} (*51*)
12. Nematodes	*Caenorhabditis elegans*	8.8×10^{-2} (*52*)
13. Insects	*Dixa obscura*	1.56×10^{-1} (*53*)
14. Lower chordates	*Ascidia atra*	1.58×10^{-1} (*51*)
15. Slime molds	*Dictyostelium discoideum*	1.68×10^{-1} (*54*)
16. Coelenterates	*Cassiopeia* sp.	3.3×10^{-1} (*51*)
17. Angiosperms	*Bulbostylis capillaris*	3.6×10^{-1} (*5*)
18. Vertebrates	*Tetraodon fluviatilis* (fish)	3.9×10^{-1} (*55*)
19. Molluscs	*Lottia gigantea*	4.3×10^{-1} (*56*)
20. Echinoderms	*Dermasterias imbricata*	5.4×10^{-1} (*57*)
21. Annelids	*Cirratulus grandis*	7×10^{-1} (*10*)
22. Crustaceans	*Sacculina* sp.	7×10^{-1} (*58*)
23. Gymnosperms	*Podocarpus dacrydioides*	5.3 (*3*)

case, according to our best estimates (*20*), closely fits a point in the demonstrated sequence.

Extranuclear DNA such as that found in mitochondria, plastids, and ribosomal DNA of eukaryotes is generally considered to comprise a very small proportion of the total DNA, and thus would have a negligible effect on our calculations. In any case, extranuclear DNA would affect only values obtained by chemical extraction; most of the data which we have compiled for eukaryotes have been obtained by spectrophotometric methods that measure only nuclear DNA.

In view of the complicating factors involved in attempting to define the "genome" in many eukaryotes, it is possible that a more accurate correlation with evolutionary complexity might be demonstrated by determination of minimum kinetic complexity for a group (that is, the size of the unique portion of the genome, excluding repetitive DNA), rather than that of the minimum as defined here. However, as Davidson *et al.* (*21*) have demonstrated, even the size of the unique portion of the eukaryote genome is not necessarily an accurate measure of relative biological complexity; it may simply indicate potential rather than actual information content. In the prokaryotes, kinetic complexity and genome size are essentially identical, since these organisms contain little or no repetitive DNA. However, Reanny (*22*) has theorized that much repetitive DNA may once have existed in the prokaryotes, and that increases in kinetic complexity may have occurred by means of extensive base substitutions in the repetitive portion of DNA, followed by subsequent elimination of any functionally unnecessary sequences. This would explain the wide range of prokaryote genome size, and would at the same time be in accord with our proposal of evolution by genome doubling. Recent work of Zipkas and Riley (*23*) lends experimental support to our proposal, and to the essentially similar but more restricted proposal of Wallace and Morowitz (*12*). Zipkas and Riley (*23*) give evidence that the genome of *Escherichia coli* has evolved by means of two sequential duplications of an ancestral genome. Their conclusions are based on the relative positions on the chromosome of functionally related genes; 74 percent of such genes are located either 90° or 180° apart on the circular *E. coli* chromosome, which strongly suggests two successive end-to-end duplications of the total genome.

It is clear that, in the prokaryotes, increasing minimum nucleic acid content correlates with increasing evolutionary complexity. Even if we should omit many of the eukaryote values from Fig. 4 for any of the reasons outlined above, the remaining values would still demonstrate a clear series of genome doublings consistent with increasing biological complexity, at least through the Porifera (19 doublings extending over six orders of magnitude).

Chromosomal Organization

Structural elements of eukaryote chromosomes are equivalent in size to the lowest point in the sequence proposed here. The RNA viroid value (2^0 in Fig. 4), equal to about 100,000 daltons or about 300 nucleotides, is equivalent in size to the basic repetitive sequence elements reported by Davidson and coworkers (*24*). They have shown that, in a large portion of the genomes of *Xenopus* (an amphibian) and of various marine invertebrates from jellyfish to arthropods, repetitive sequences averaging 300 nucleotides in length are alternated with longer nonrepetitive sequences of up to several thousand nucleotides. Interspersed repetitive and nonrepetitive sequences of comparable length also have been noted in slime molds (*25*) and mammals (*26*), including humans (*27*), suggesting that this pattern of chromosome organization may be typical of all eukaryotes—with the single known exception of *Drosophila* (*21*). Davidson *et al.* (*28*) have further shown that single-copy structural gene sequences are contiguous to the interspersed repetitive sequences, which strongly suggests a regulatory function for these 300-nucleotide, repetitive segments. Both Diener (*16*) and Reanny (*29*) have, in fact, suggested that the 300-nucleotide RNA viroid is an aberrant form of a regulatory nuclear RNA.

Recent work by Semancik and Geelen (*30*) has shown that viroid RNA can indeed undergo hybridization with a DNA sequence in the host genome, lending support to the concept of the viroid as an aberrant regulator. If there is a functional significance to the actual size of the 300-nucleotide sequence, it has not yet been determined.

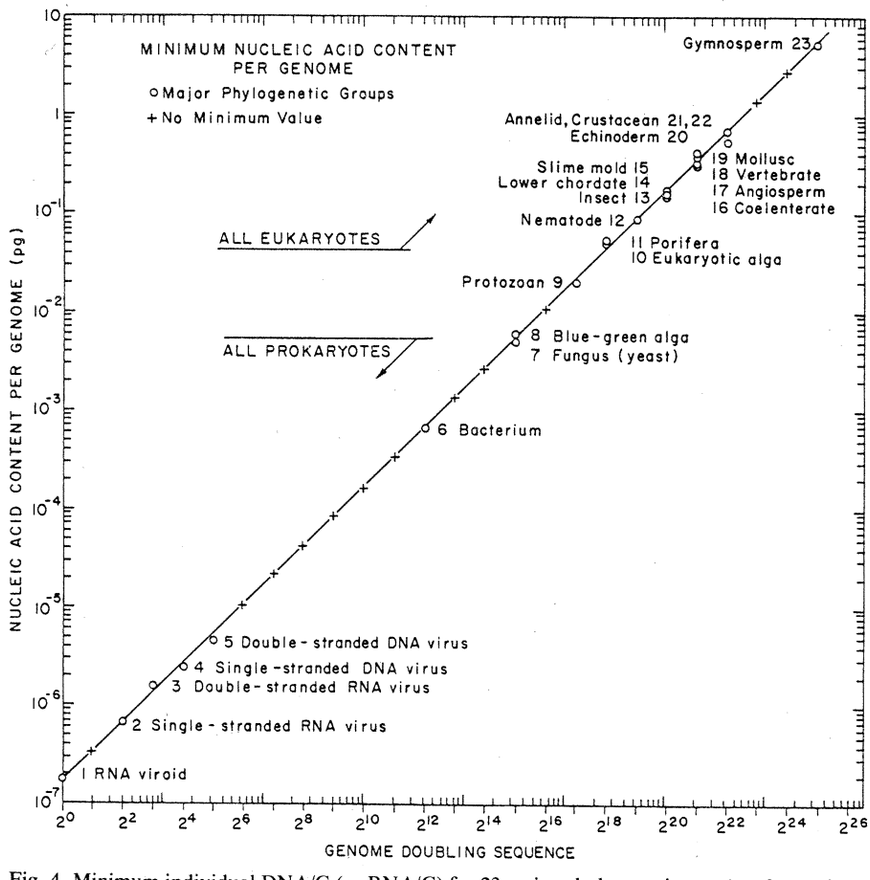


Fig. 4. Minimum individual DNA/G (or RNA/G) for 23 major phylogenetic groups of organisms as a function of a theoretical sequence of genome doublings. Each exponent of 2 indicates the number of doublings at that point on the *x*-axis. The data are taken from Table 1; the slope is equal to +1 (see text). The points indicated by + are doubling points not represented by a known minimum value for a major group. However, we do have data for organisms with these approximate values, with the exception of the lowest (2^1 on the *x*-axis). As indicated by arrows at the left, all organisms above 2^{18} are eukaryotes; all organisms below 2^{15} are prokaryotes.

Possible Doubling Mechanisms

A mechanism for saltatory total genome doubling has not been suggested previously. However, in prokaryotes, it could be accomplished by a double-stranded break in a circular DNA molecule, followed by crossing of strand ends during repair. This would essentially form a continuous single DNA strand, which after replication would become a double-length, double-stranded circular molecule. In eukaryotes, the unineme chromosome (or chromatid) has become the generally accepted model (*31*), so that polyteny is probably an unlikely mechanism for genome doubling. This single DNA molecule contains clusters of palindromic sequences (*32*) (base sequences which read the same backward and forward, that is, when read with the same polarity on complementary strands) as well as palindromic telomeres (*33–35*). A biochemical or physical aberration in meiosis (or mitosis) could thus cause all sister chromatids in a complement to join by telomeric base pairing, resulting in nondisjunction at anaphase and consequent doubling of chromosome size. This would also offer a partial explanation for the presence of internal palindromes.

A theory of telomere replication by means of self-complementary telomeric palindromes has been proposed by Cavalier-Smith (*34*). In Bateman's modification of this model (*35*), palindromic self-paired telomeres are the normal condition, implying that a eukaryotic chromosome is essentially circular, though much flattened. Scheid and Traut (*36*) have, in fact, presented visual evidence of such a conformation, demonstrating U-shaped chromosome ends in both plants and animals. Haapala and Soyer (*37*) similarly give evidence that *Euglena* chromosomes are circular uninemes. According to the Cavalier-Smith and Bateman model, after replication a sequence-specific endonuclease nicks one end of each palindrome in both old and new strands. This is followed by self-pairing of telomeres and ligation of old to new strands. If such a model is indeed correct, a doubling of chromosome length could be effected by a rejoining of nicked ends without exchange of old and new strand ends, by a twisting of old and new strands followed by rejoining, or simply by a failure of this sequence of events to occur at one set of telomeres.

An alternative mechanism for total genome doubling could be based on DuPraw's conjecture that the entire genome may consist of one enormously long circular DNA molecule, condensed into chromosomes and with attenuated interchromosomal connectives (*38*). Though admittedly highly speculative and disputed by some (*39*), this theory has generated much positive discussion, and a considerable amount of supportive evidence has accumulated (*40–42*) in addition to that cited by DuPraw (*38*). Most strikingly, Lauer and Klotz (*41*) have recently demonstrated that in the yeast *Saccharomyces cerevisiae*, which has 17 chromosomes, the largest single piece of nuclear DNA comprises at least a fourth, and possibly all, of the yeast genome. The work of Burdick *et al.* (*42*) is similarly provocative. They have demonstrated that all the pachytene chromosomes of rat and man are interconnected, but that there are always two loose ends in each interconnected configuration; although they do not offer an explanation, this could indicate that the entire genome comprises a large and much flattened loop, as in the single-chromosome model proposed above. If this conception of one DNA molecule per genome should ultimately be borne out, then a modification of the relatively simple mechanism for doubling of a circular chromosome might also apply to the doubling of an entire genome.

Summary

Logarithmic distributions of nucleic acid contents per genome of species within major phylogenetic groups of organisms tend to form several peaks. These peaks appear to represent intragroup doublings of DNA or RNA which, in the case of eukaryotes, are independent of polyploidy. This phenomenon has been termed cryptopolyploidy. There are numerical similarities in peak values for different taxonomic groups. A high degree of order is suggested when minimum values for the major phylogenetic groups are plotted against a series of theoretical doublings. These data demonstrate the apparent existence of an exponential periodicity over eight orders of magnitude, leading us to suggest an evolutionary continuity of doublings of a basic ancestral genome (of about 300 nucleotides), these doublings being independent of both chromosome number and ploidy level. This proposed continuity encompasses most major life forms and is generally concomitant with increasing evolutionary complexity, particularly in the prokaryotes and lower eukaryotes.

Our interpretation of the data presented here must currently be viewed as speculative, and we do not propose that genome doubling is the only mechanism for genome evolution. However, we feel that the evidence is sufficient to warrant serious scrutiny of our proposals. We hope that this approach to a synthesis of available data will provoke discussion and will stimulate further work toward either supporting, modifying, or disproving our hypothesis.

References and Notes

1. A. H. Sparrow, A. F. Rogers, S. S. Schwemmer, *Radiat. Bot.* **8**, 149 (1968); A. H. Sparrow, A. Howard, F. G. Cowie, S. S. Schwemmer, A. F. Nauman, *Int. J. Radiat. Biol.* **27**, 343 (1975).
2. A. H. Sparrow, H. J. Price, A. G. Underbrink, *Brookhaven Symp. Biol.* **23**, 451 (1972).
3. H. J. Price, A. H. Sparrow, A. F. Nauman, *ibid.* **25**, 390 (1973).
4. A. H. Sparrow and A. F. Nauman, *ibid.*, p. 367.
5. ———, in preparation: analysis of data on about 800 species of herbaceous angiosperms.
6. In higher plants, DNA content has in most cases been estimated from meristematic nuclear volume measurements, with a previously demonstrated direct correlation [(*4*); K. P. Baetcke, A. H. Sparrow, C. H. Nauman, S. S. Schwemmer, *Proc. Natl. Acad. Sci. U.S.A.* **58**, 533 (1967)]. Most of the animal data were determined from sperm or from erythrocytes.
7. K. Bachmann, O. B. Goin, C. J. Goin, *Brookhaven Symp. Biol.* **23**, 419 (1972).
8. S. Ohno, *Evolution by Gene Duplication* (Springer-Verlag, New York, 1970).
9. I. A. McLaren, S. M. Woods, J. R. Shea, Jr., *Science* **153**, 1641 (1966); E. Sexsmith, thesis, University of Toronto (1968); R. J. Britten and E. H. Davidson, *Q. Rev. Biol.* **46**, 111 (1971); K. Bachmann and E. L. Rheinsmith, *Chromosoma* **43**, 225 (1973); J. C. Vaughn and K. Bachmann, *J. Cell Biol.* **59**, 352a (1973); M. L. Beçak and W. Beçak, *Experientia* **30**, 624 (1974); S. Mizuno and H. C. Macgregor, *Chromosoma* **48**, 239 (1974).
10. W. G. Conner, R. Hinegardner, K. Bachmann, *Experientia* **28**, 1502 (1972).
11. K. Rothfels, E. Sexsmith, M. Heimburger, M. O. Krause, *Chromosoma* **20**, 54 (1966); K. Rothfels and M. Heimburger, *ibid.* **25**, 96 (1968); G. L. Stebbins, *Chromosomal Evolution in Higher Plants* (Arnold, London, 1971).
12. D. C. Wallace and H. J. Morowitz, *Chromosoma* **40**, 121 (1973).
13. M. G. Rush, A. K. Kleinschmidt, W. Hellmann, R. C. Warner, *Proc. Natl. Acad. Sci. U.S.A.* **58**, 1676 (1967); F. C. Wheeler, R. H. Benzinger, H. Bujard, *J. Virol.* **14**, 620 (1974).
14. D. A. Clayton and J. Vinograd, *Nature (London)* **216**, 652 (1967); R. Radloff, W. Bauer, J. Vinograd, *Proc. Natl. Acad. Sci. U.S.A.* **57**, 1514 (1967); J. Locker, M. Rabinowitz, G. S. Getz, *J. Mol. Biol.* **88**, 489 (1974).
15. T. F. Roth and D. R. Helinsky, *Proc. Natl. Acad. Sci. U.S.A.* **58**, 650 (1967).
16. T. O. Diener, *Annu. Rev. Microbiol.* **28**, 23 (1974).
17. J. S. Semancik, D. S. Magnuson, L. G. Weathers, *Virology* **52**, 292 (1973); J. S. Semancik, T. J. Morris, L. G. Weathers, *ibid.* **53**, 448 (1973); ———, B. F. Rodorf, D. R. Kearns, *ibid.* **63**, 160 (1975).
18. H. J. Price, A. H. Sparrow, A. F. Nauman, *Experientia* **29**, 1028 (1973).
19. The values of DNA/G are 105 and 142 pg, respectively, for the monocotyledonous angiosperm *Fritillaria assyriaca* [M. D. Bennett, *Proc. R. Soc. London Ser. B* **181**, 109 (1972)] and the lungfish *Protopterus aethiopicus* [R. A. Pedersen, *J. Exp. Zool.* **177**, 65 (1971)], equivalent to 2^{29} and 2^{30} in Fig. 4.
20. The values of DNA/G are 0.67 and 1.46 pg, respectively, for *Sphagnum* sp. [bryophyte (*2*)] and *Angiopteris hypoleuca* (fern; A. H. Sparrow, unpublished data).
21. E. H. Davidson, G. A. Galau, R. C. Angerer, R. J. Britten, *Chromosoma* **51**, 253 (1975).
22. D. C. Reanny, *J. Theor. Biol.* **48**, 243 (1974).
23. D. Zipkas and M. Riley, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1354 (1975).
24. E. H. Davidson, B. R. Hough, C. S. Amenson, R. J. Britten, *J. Mol. Biol.* **77**, 1 (1973); R. B. Goldberg, W. R. Crain, J. V. Ruderman, G. P. Moore, T. R. Barnett, R. C. Higgins, R. A.

Gelfand, G. A. Galau, R. J. Britten, E. H. Davidson, *Chromosoma* **51**, 225 (1975); R. C. Angerer, E. H. Davidson, R. J. Britten, *Cell* **6**, 29 (1975).
25. R. A. Firtel, A. Jacobson, J. Tuchman, H. F. Lodish, *Genetics* **78**, 355 (1974); R. A. Firtel and K. Kindle, *Cell* **5**, 401 (1975).
26. E. Pays and A. Ronsse, *Biochem. Biophys. Res. Commun.* **62**, 862 (1975).
27. C. W. Schmid and P. L. Deininger, *Cell* **6**, 345 (1975).
28. E. H. Davidson, B. R. Hough, W. H. Klein, R. J. Britten, *ibid.* **4**, 217 (1975).
29. D. C. Reanny, *J. Theor. Biol.* **49**, 461 (1975).
30. J. S. Semancik and J. L. M. C. Geelen, *Nature (London)* **256**, 753 (1975).
31. F. Crick, *ibid.* **234**, 25 (1971); C. D. Laird, *Chromosoma* **32**, 378 (19➡ H. G. Callan, *Proc. R. Soc. London Ser. B* **181**, 19 (1972); R. Kavenoff and B. H. Zimm, *Chromosoma* **41**, 1 (1973); K. Brasch and G. Setterfield, *Exp. Cell Res.* **83**, 175 (1974); H. G. Callan, *Cold Spring Harbor Symp. Quant. Biol.* **38**, 195 (1974); D. E. Comings, in *Chromosomes and Cancer*, J. German, Ed. (Wiley, New York, 1974), pp. 95–133; D. E. Comings and T. A Okada, *Cold Spring Harbor Symp. Quant. Biol.* **38**, 145 (1974); J. H. Taylor, M. Wu, L. C. Erickson, *ibid.*, p. 225.
32. D. A. Wilson and C. A. Thomas, Jr., *J. Mol. Biol.* **84**, 115 (1974); R. D. Wesley, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 678 (1975).
33. C. F. Garon, K. W. Berry, J. A. Rose, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 2391 (19➡ J. Wolfson and D. Dressler, *ibid.*, p. 30➡ R. J. Roberts, J. R. Arrand, W. Keller, *ibid.* **71**, 3829 (1974); P. H. Steenbergh, J. S. Sussenbach, R. J. Roberts, H. S. Jansz, *J. Virol.* **15**, 268 (1975).
34. T. Cavalier-Smith, *Nature (London)* **250**, 467 (1974).
35. A. J. Bateman, *ibid.* **253**, 379 (1975).
36. H. Traut and W. Scheid, *Mutat. Res.* **17**, 335 (1973); W. Scheid and H. Traut, *ibid.* **25**, 321 (1974); *Can. J. Genet. Cytol.* **17**, 129 (1975).
37. O. K. Haapala and M.-O. Soyer, *Hereditas* **80**, 185 (1975).
38. E. J. DuPraw, *DNA and Chromosomes* (Holt, Rinehart and Winston, New York, 1970).
39. M. J. D. White, *Animal Cytology and Evolution* (Cambridge Univ. Press, Cambridge, 1973); D. E. Comings and T. A. Okada, *Cold Spring Harbor Symp. Quant. Biol.* **38**, 145 (1974).
40. E. G. Diacumakos, S. Holland, P. Pecora, *Nature (London)* **232**, 33 (1971); P. M. Kraemer, D. F. Petersen, M. A. Van Dilla, *Science* **174**, 714 (1971); L. K. Schneider, *J. Cell Biol.* **56**, 608 (1973); T. Ashley and E. B. Wagenaar, *Can. J. Genet. Cytol.* **16**, 61 (1974); H. M. Golomb and G. F. Bahr, *Exp. Cell Res.* **84**, 79 (1974); J. Y. Jaffray and A. Geneix, *Humangenetik* **25**, 119 (1974); S. Takayama, *Exp. Cell Res.* **91**, 408 (1975).
41. G. D. Lauer and L. C. Klotz, *J. Mol. Biol.* **95**, 309 (1975).
42. A. B. Burdick, B. L. Libbus, V. Athanassiou, A. Falek, *Can. J. Genet. Cytol.* **16**, 481 (1974).
43. H. Fraenkel-Conrat and R. R. Wagner, Eds., *Comprehensive Virology* (Plenum, New York, 1974), vol. 1, pp. 39–40.
44. K. W. Buck and G. F. Kempson-Jones, *J. Gen. Virol.* **18**, 223 (1973).
45. L. V. Crawford *et al.*, *ibid.* **4**, 37 (1969).
46. G. Askaa, C. Christiansen, H. Erno, *J. Gen. Microbiol.* **75**, 283 (1973).
47. E. N. Sokurova, *Mikrobiologiya* **42**, 1020 (1973).
48. B. A. Whitton, N. G. Carr, I. W. Craig, *Protoplasma* **72**, 325 (1971). DNA/G varies with growth rate; value given is the lowest extreme.
49. P. R. Whitfield, *Aust. J. Biol. Sci.* **6**, 234 (1953). Lowest of four determinations; mean = 2.7×10^{-2} pg.
50. T. Iwamura and S. Kuwashima, *Biochim. Biophys. Acta* **174**, 330 (1969).
➡ A. E. Mirsky and H. Ris, *J. Gen. Physiol.* **34**, 451 (1951).
52. J. E. Sulston and S. Brenner, *Genetics* **77**, 95 (1974).
53. E. Jost and M. Mameli, *Chromosoma* **37**, 201 (1972).
54. C. K. Leach and J. M. Ashworth, *J. Mol. Biol.* **68**, 35 (1972).
55. R. Hinegardner and D. E. Rosen, *Am. Nat.* **106**, 621 (1972).
56. R. Hinegardner, *Comp. Biochem. Physiol. A* **47**, 447 (1974).
57. ———, *Comp. Biochem. Physiol. B* **49**, 219 (1974).
58. E. L. Rheinsmith, R. Hinegardner, K. Bachmann, *ibid.* **48**, 343 (1974).
59. H. S. Shapiro, in *Handbook of Biochemistry*, H. A. Sober, Ed. (Chemical Rubber Company, Cleveland, Ohio, 1970), p. H-107.
60. H. Swift and R. Kleinfeld, *Physiol. Zool.* **26**, 301 (1953); B. Daneholt and J.-E. Edström, *Cytogenetics* **6**, 350 (1967); E. M. Rasch, in *Introduction to Quantitative Cytochemistry*, G. L. Wied and G. F. Bahr, Eds. (Academic Press, New York, 1970), vol. 2, pp. 357–397; J. M. Lagowski, M.-Y. W. Yu, H. S. Forrest, C. D. Laird, *Chromosoma* **43**, 349 (1973); A. Lima-de-Faria, T. Gustaffson, H. Jaworska, *Hereditas* **73**, 119 (1973); E. M. Rasch, *Chromosoma* **45**, 1 (1974).
61. A. von Hofsten, *Physiol. Plant.* **16**, 709 (1963); K. Esser and R. Kuenen, *Genetics of Fungi* (Springer-Verlag, New York, 1967); J. N. Bicknell and H. C. Douglas, *J. Bacteriol.* **101**, 505 (1970); C. Christiansen, A. L. Bak, A. Stenderup, G. Christiansen, *Nature (London) New Biol.* **231**, 176 (1971); R. A. Firtel and J. Bonner, *J. Mol. Biol.* **66**, 339 (1972); J. Mohberg, K. L. Babcock, F. B. Haugli, H. P. Rusch, *Dev. Biol.* **34**, 228 (1973); S. K. Dutta, *Nucleic Acids Res.* **1**, 1411 (1974); R. L. Dusenbery, *Biochim. Biophys. Acta* **378**, 363 (1975).
62. We thank K. H. Thompson for statistical analysis of data and R. B. Setlow and H. H. Smith for helpful comments on the manuscript. Supported by the Energy Research and Development Administration.

NEWS AND COMMENT

Cancer Institute: Expert Charges Neglect of Carcinogenesis Studies

A prominent scientist has resigned from directorship of a key program in the National Cancer Institute for reasons which, if well founded, could provoke a serious perturbation in the agency's affairs. The scientist, Umberto Saffiotti, heads the NCI's program in chemical carcinogenesis, a subject whose importance has been increasingly acknowledged by cancer epidemiologists, by regulatory agencies, and in Congress. A threatened fragmentation of the program which could prolong the public's exposure to carcinogens is one of the reasons behind his decision to quit.

NCI director Frank Rauscher pays tribute to Saffiotti's scientific expertise but regards the issue of his resignation as the result of a difference in approach to the management of certain programs under his control, which Rauscher believes could have been pushed ahead faster.

Saffiotti believes that the carcinogenesis program has long been denied the manpower necessary to keep pace with its growing responsibilities. It has only 20 percent more staff than in 1968, but almost 8 times the amount of money to administer. This year the program received only 3 of the 79 new staff positions assigned to the NCI, although Rauscher told the House Appropriations Committee that he was giving Saffiotti the highest priority possible. With a current budget of $47 million, the program conducts basic research on chemical carcinogenesis as well as developing bioassay tests for carcinogens.

Saffiotti also feels that he and other colleagues with relevant expertise have been excluded from a series of decisions on chemical carcinogenesis, the most recent being the announcement of a National Clearinghouse on Environmental Carcinogenesis, on which he says he was not consulted until a late stage. The final straw for Saffiotti was a recent decision to split away from his program the responsibility for developing bioassay tests for chemical carcinogens. The move will, in his view, compromise the scientific credibility of the tests, delay their being put into action, and increase the time that people will be exposed to the chemicals the tests may show to be carcinogenic.

Although he has been asked to remain as director for the research part of the carcinogenesis program, Saffiotti has chosen to resign altogether from the program management, lest he seem by staying to concur with the decision on the bioassay tests. He plans to take up full time research in his laboratory at the NCI. "I am glad to call it quits rather than endorse a mode of operation I disagree with," he told *Science* in an interview last week before announcing his resignation.

Saffiotti adds that a fundamental reason for resigning is his belief that active scientists have very little voice in setting policy or priorities in his division of the NCI, and that the division is being run by managers with the help of scientists rather than the other way around. Because of the growth of successive layers of bureaucracy, whose actions are not accountable to detailed peer review by scientists, Saffiotti says, "There seems to be a growing gap between the top policy-making decisions of the institute and the