Chromosoma (Berl.) 40, 121—126 (1973)

Genome Size and Evolution

Douglas C. Wallace and Harold J. Morowitz
Department of Microbiology and Department of Molecular Biophysics
and Biochemistry, Yale University, New Haven, Connecticut

Abstract. Examination of data on genome size for prokaryotic cells suggests an evolutionary scheme.

While it has been difficult to determine the genome size of prokaryotic cells with a high degree of precision, recent developments in DNA renaturation and electron microscopy provide substantial data on this subject. In this paper we have collected these data and have examined the frequency distribution versus genome size in the range of zero to three thousand million daltons. In analyzing these data we have excluded the obligate intracellular forms — *Rickettsiae*, *Chlamydiae*, organelles, and viruses. The distribution is shown in Fig. 1 and the data in Table 1.

The mycoplasma fall into two distinct classes as originally pointed out by Bak *et al.*, (1969). The first class falls in the range of 4×10^8 to 6×10^8 daltons, while the second class is the order of magnitude of 10^9 daltons. The genome sizes between the two classes vary by a factor of approximately 2. The lowest bacterial values are around 10^9 daltons. The striking fact is that the range between 6×10^8 and 10^9 daltons is virtually empty.

If one assumes that the mycoplasma are degenerate bacteria, the sharp break in the frequency distribution is hard to explain since deletion should lead to all values from the bacterial value of 10^9 down to the minimal value possible for a free living cell. On the other hand, if one assumes that the mycoplasma are primitive, then 5×10^9 daltons of DNA represents a primitive genome (which we shall designate a "genesistron") and subsequent evolution occurred by DNA doubling leading to the 10^9 daltons DNA of the *Acholeplasma* genus of myco-

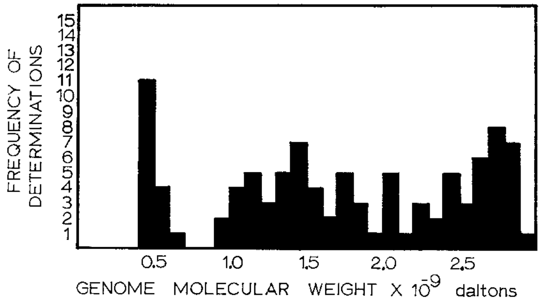

Fig. 1

plasma. The *Acholeplasma* (formerly known as *Mycoplasma laidlawii*) are saprophytic mycoplasma and are far less fastidious than the small genome forms. If successful mutation is far more probable than successful deletion, a gap would emerge between organisms having 5×10^8 daltons and 10^9 daltons DNA.

In addition to small genome and cell size, one other argument exists for the primitive nature of the mycoplasma. The tRNAs of mycoplasma show a small number of modified bases, which argues for a simpler control system (Hayashi *et al.*, 1969; Johnson *et al.*, 1970; Maniloff and Morowitz, in preparation, 1972).

Finally, the examination of the DNA of the simple eukaryotic cell *Saccharomyces cerevisiae* indicates that it is also in pieces in the size range of 4 to 6×10^8 daltons, the range of the smallest prokaryotic genomes (Blamire *et al.*, 1972).

Thus an evolutionary scheme is suggested beginning with a primitive cell with a single genesistron. Further evolution of the cell involved two possible methods of increasing genetic information: (1) increase of chromosome size by the linkage of two genesistrons or (2) increase of chromosome number with subsequent change in chromosome organization towards the eukaryotic mode. The former method allowed continuation of the original replicative and segregating systems as the genetic information is retained in one continuous piece of DNA. The

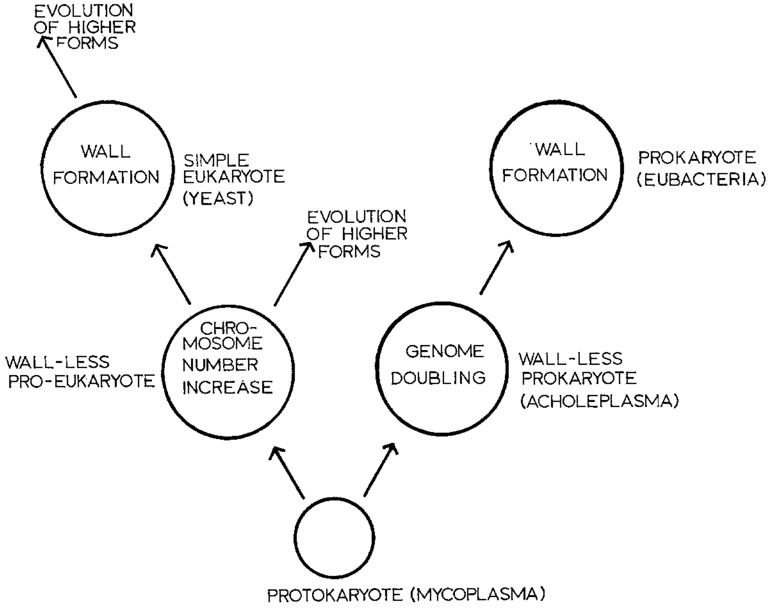

Fig. 2

Table 1. Summary of prokaryotic cell genome sizes[a]

Organism	Strains	Genome size $\times 10^{-9}$ daltons^{-1}	Reference[a]
Order: *Eubacteriales*			
Family: *Achromobacteraceae*			
Achromobacter anitratus	MMCA 19	1.44	(1)
Achromobacter sp.	M 153	2.61	(2)
Achromobacter sp.	NCIB 9650	2.18	(2)
Achromobacter turbidus	AHU 1337	2.93	(2)
Reuter's organism	NCIB 9746	2.78	(2)
Alcaligenes odorans	CCEB 554	1.89	(2)
Alcaligenes sp.	AB 717	2.65	(2)
Family: *Azotobacteraceae*			
Azotobacter agilis	K	1.71	(2)
Azotobacter agilis	SS 4	1.70	(2)
Family: *Bacillaceae*			
Bacillus anthracis	NCTC 8234	2.78	(1)
Bacillus cereus	ATCC 13061	2.60	(1)
Bacillus polymyxa	ATCC 10401	2.75	(1)
Bacillus subtilis	168^6 thy$^-$	2.31	(2)
Bacillus subtilis	BQ 2	2.47	(2)
Family: *Brevibacteriaceae*			
Brevibacterium ammoniagenes	ATCC 6871	1.98	(1)
Family: *Brucellaceae*			
Haemophilus influenzae	—	1.0	(3)
Haemophilus influenzae	MMCA 29	1.01	(1)
Haemophilus influenzae	Rd	1.66	(2)
Haemophilus influenzae	—	1.52	(4)
Haemophilus aegyptieus	NCTC 8502	1.17	(1)
Mima polymorpha	ATCC 10973	1.80	(2)
Moraxella glucidolytica var. *liquefaciens*	ATCC 17979	1.86	(2)
Moraxella osloensis	19116/51	1.49	(2)
Pasteurella multocida	MMCA 8	1.13	(1)
Family: *Corynebacteriaceae*			
Corynebacterium glutamicum	MPS 10	1.72	(2)
Family: *Enterobacteriaceae*			
Erwinia herbicola	BG 1	2.93	(2)
Escherichia coli	MMCA 38	2.75	(1)
Escherichia coli	—	2.98	(5)
Escherichia coli	B	2.2	(6)
Escherichia coli	B	2.2	(3)
Escherichia coli	B (MMCA 56)	2.84	(1)
Escherichia coli	B	2.67	(2)
Escherichia coli	C	2.42	(2)
Escherichia coli	K-12 C600	2.53	(2)
Klebsiella ozaenae	NCTC 5053	2.36	(1)
Klebsiella rubiacearum	—	2.58	(2)

---

[a] References see p. 125.

Table 1 Continued

Organism	Strains	Genome size $\times 10^{-9}$ daltons^{-1}	Reference^{a}
Proteus morganii	MMCA 6	2.02	(1)
Proteus vulgaris	ATCC 13315	2.09	(1)
Salmonella pullorum	1	2.62	(2)
Salmonella pullorum	2	2.81	(2)
Salmonella pullorum	3	2.81	(2)
Salmonella pullorum	NCTC 5776	2.83	(1)
Salmonella typhimurium	2	2.82	(2)
Salmonella typhimurium	LT$_7$-Str^r	2.65	(2)
Shigella sonnei	NCTC 8221	2.09	(1)
Family: *Lactobacillaceae*			
Diplococcus pneumoniae	MMCA 46	1.45	(1)
Streptococcus agalactiae	NCTC 6198A	1.20	(1)
Streptococcus faecalis	NCTC 370	1.47	(1)
Streptococcus pyogenes	NCTC 6175	1.27	(1)
Family: *Micrococcaceae*			
Micrococcus flavus	ATCC 1024	2.68	(1)
Micrococcus lysodeikticus	ATCC 4698	2.82	(1)
Sarcina lutea	ATCC 9341	2.78	(1)
Staphylococcus albus	MMCA 47	1.12	(1)
Staphylococcus aureus	MMCA 1	1.43	(1)
Family: *Neisseriaceae*			
Neisseria crassa	MMCA 17	1.73	(1)
Neisseria catarrhalis	NCTC 3622	1.04	(1)
Neisseria catarrhalis	—	1.52	(5)
Neisseria flava	—	1.45	(5)
Neisseria gonorrhoeae	—	0.992	(5)
Neisseria gonorrhoeae	MMCA 32	1.28	(1)
Neisseria meningitidis	—	1.12	(5)
Neisseria sicca	—	1.45	(5)
Family: *Rhizobiaceae*			
Agarbacterium alginicum	NCMB 886	2.09	(2)
Order: *Mycoplasmatales*			
Family: *Mycoplasmataceae*			
Acholeplasma (*Mycoplasma*) *granularum*	Friend	0.95	(3)
Acholeplasma (*Mycoplasma*) *laidlawii* A	F1	1.1	(3)
Acholeplasma (*Mycoplasma*) *laidlawii* B	F8	1.0	(3)
Mycoplasma arthritidis	PG6	0.44	(3)
Mycoplasma arthritidis	—	0.444	(7)
Mycoplasma sp. (bovine strain)	Donnetta	0.585	(7)
Mycoplasma fermentans	PG18	0.48	(3)
Mycoplasma gallisepticum	PG31	0.49	(3)
Mycoplasma hominis	H39	0.51	(8)
Mycoplasma hominis	PG21	0.45	(3)
Mycoplasma hominis	H39	0.527	(6)

Table 1 Continued

Organism	Strains	Genome size $\times 10^{-9}$ daltons^{-1}	Reference[a]
Mycoplasma meleagridis	529	0.42	(9)
Mycoplasma orale I	Pratt	0.47	(3)
Mycoplasma pneumoniae	Mac	0.48	(3)
Mycoplasma salivarium	PG20	0.47	(3)
Mycoplasma sp	KID	0.684	(10)
Mycoplasma T-strain	No. 58 (Ford)	0.44	(3)
Mycoplasma T-strain	No. 27 (Ford)	0.47	(3)
Order: *Pseudomonadales*			
Family: *Nitrobacteraceae*			
Nitrosococcus sp	—	2.16	(2)
Nitrosomonas sp	—	1.40	(2)
Nitrifying organism	25	2.45	(2)
Family: *Pseudomonadaceae*			
Acetobacter rancens	23 kl+	1.71	(2)
Pseudomonas fluorescens	CCEB 488	2.72	(2)
Pseudomonas patida	CCEB 520	2.53	(2)
Xanthomonas pelargonii	ICPB P 121	2.40	(2)
Family: *Spirillaceae*			
H-I *Bdellovibrio*	100	1.410	(6)
H-I *Bdellovibrio*	109	1.340	(6)
H-I *Bdellovibrio*	114	1.390	(6)
H-I *Bdellovibrio*	118	1.340	(6)
H-I *Bdellovibrio*	120	1.300	(6)
Bdellovibrio starrii			
H-I *Bdellovibrio*	A3.12	1.670	(6)
Bdellovibrio stolpii			
H-I *Bdellovibrio*	UKi2	1.470	(6)
Vibrio metschnicovii	ATCC 7708	2.26	(1)

[a] References for data: (1) Bak *et al.* (1970); (2) Gillis *et al.* (1970); (3) Bak *et al.* (1969); (4) MacHattie *et al.* (1965); (5) Kingsbury (1969); (6) Seidler *et al.* (1972); (7) Morowitz *et al.* (1967); (8) Bode and Morowitz (1967); (9) Allen (1971); (10) Ryan and Morowitz (1969).

latter method demanded the innovation of a new ordered segregation mechanism. The scheme is diagrammed in Fig. 2.

According to this proposal, the genus *Mycoplasma* represents the descendants of the organisms that preceded the prokaryotic-eukaryotic cell split. We can designate such an organism a "protokaryote". The *Acholeplasma* can then be regarded as intermediates in the evolution from protokaryotic to prokaryotic cells. In this scheme it is important to note that the eukaryotic cell is no longer regarded as being derived from prokaryotic cells. Rather, both are regarded as having arisen from the protokaryotes which now survive as the *Mycoplasma*.

Acknowledgements. The authors are grateful for the financial assistance contributed by the U.S. Public Health Service (Traineeship No. GM00275-13) and the National Aeronautics and Space Administration during the preparation of this manuscript.

Added in the Proof. Another recent study of the sedimentation pattern of yeast DNA has come to our attention [Petes, T. D., Fangman, W. L.: Sedimentation properties of yeast chromosomal DNA. Proc. nat. Acad. Sci. (Wash.) **69**, 1188–1191 (1972)]. In this paper a number average molecular weight of 6.2×10^8 daltons was observed, but the range of molecular weights was distributed from 5×10^7 to 1.4×10^9 daltons. Although strains used, DNA standards, and methods of spheroplast formation differ slightly; the source of the discrepancy between this paper and that of Blamire *et al.*, 1972, is not clear.

References

Allen, T. C.: Base composition and genome size of Mycoplasma meleagridis deoxyribonucleic acid. J. gen. Microbiol. **69**, 285–286 (1971).

Bak, A. L., Black, F. T., Christiansen, C., Freundt, E. A.: Genome size of mycoplasma DNA. Nature (Lond.) **224**, 1209–1210 (1969).

Bak, A. L., Christiansen, C., Stenderup, A.: Bacterial genome sizes determined by DNA renaturation studies. J. gen. Microbiol. **64**, 377–380 (1970).

Blamire, J., Cryer, D. R., Finkelstein, D. B., Marmur, J.: Sedimentation properties of yeast nuclear and mitochondrial DNA. J. molec. Biol. **67**, 11–24 (1972).

Bode, H. R., Morowitz, H. J.: Size and structure of the Mycoplasma hominis H39 chromosome. J. molec. Biol. **23**, 191–199 (1967).

Gillis, M. de Ley, J., de Cleene, M.: The determination of molecular weight of bacterial genome DNA from renaturation rates. Europ. J. Biochem. **12**, 143–153 (1970).

Hayashi, H., Fisher, H., Söll, D.: Transfer ribonucleic acid from Mycoplasma. Biochemistry 8, 3680–3686 (1969).

Johnson, L., Hayashi, H., Söll, D.: Isolation and properties of a transfer ribonucleic acid deficient in ribothymidine. Biochemistry **9**, 2823–2831 (1970).

Kingsbury, D. T.: Estimate of the genome size of various microorganisms. J. Bacteriol **98**, 1400–1401 (1969).

MacHattie, L. A., Berns, K. I., Thomas, C. A., Jr.: Electron microscopy of DNA from Hemophilus influenzae. J. molec. Biol. **11**, 648–649 (1965).

Morowitz, H. J., Bode, H. R., Kirk, R. G.: The nucleic acids of Mycoplasma. Ann. N.Y. Acad. Sci. **143**, 110–114 (1967).

Ryan, J. L., Morowitz, H. J.: Partial purification of native rRNA and tRNA cistrons from Mycoplasma sp. (Kid). Proc. nat. Acad. Sci. (Wash.) **63**, 1282–1289 (1969).

Seidler, R. J., Mandel, M., Baptist, J. N.: Molecular heterogeneity of the Bdellovibrios: evidence of two new species. J. Bacteriol. **109**, 209–217 (1972).

Received August 14, 1972/Accepted by J. G. Gall
Ready for press August 23, 1972

Mr. D. C. Wallace
Department of Microbiology
Yale University
New Haven, Conn. 06520
U.S.A.

Dr. H. J. Morowitz
Department of Molecular Biophysics
and Biochemistry
Yale University
New Haven, Conn. 06520
U.S.A.