Journal of General Microbiology (1970), **64**, 377–380
Printed in Great Britain

Bacterial Genome Sizes Determined by DNA Renaturation Studies

By A. LETH BAK, C. CHRISTIANSEN AND A. STENDERUP
Institute of Medical Microbiology,
University of Aarhus, Denmark

(*Accepted for publication 12 November 1970*)

The genome size of bacterial DNA is known only for a very few species. Autoradiographic studies by Cairns (1963) have shown that the DNA in *Escherichia coli* is organized into a single circular molecule, 1100 to 1400 μm. long. *Haemophilus influenzae* was found by electron microscopy to contain a total length of DNA of about 800 μm. (MacHattie, Berns & Thomas, 1965), corresponding to the DNA content of two nucleoids per cell (Berns & Thomas, 1965). Measurements of intact DNA molecules of this length are, however, extremely difficult, and no successful experiments, other than the two mentioned, have been reported. Information on the genome size of bacterial DNA may also be obtained by chemical determinations of the DNA content per cell nucleoid or per bacterial spore (Shapiro, 1968). The main problem with these methods is that the number of nucleoids per cell or spore, and the state of replication of the chromosome, is difficult to determine (Maaløe & Kjeldgaard, 1966).

An alternative and more promising approach to the determination of genome sizes is the application of the technique of renaturation of single-stranded DNA. This process has been shown to follow second-order reaction kinetics (Britten & Kohne, 1968; Wetmur & Davidson, 1968). The genome size corresponding to a simple non-repetitive DNA can thus be calculated from the second-order reaction rate constant (K_2) and the molecular weight of the renaturating single-stranded DNA (Wetmur & Davidson, 1968), or from relative measurements of $Cot_{\frac{1}{2}}$, which is the product of the initial concentration of single-stranded DNA with a given piece length and the half time for the completion of the renaturation reaction (Britten & Kohne, 1968). The renaturation method has been successfully used in the determination of genome sizes among mycoplasmas (Bak, Black, Christiansen & Freundt, 1969) and a few other micro-organisms, including four neisseriae, two chlamydiae, two rickettsiae (Kingsbury, 1969) and four halophilic bacteria (Moore & McCarthy, 1969).

In order to obtain an estimate on the variation of genome sizes among bacteria, we have studied the DNA from 36 species representing 21 bacterial genera within 11 families.

The names and sources of the species investigated are listed in Table 1. DNA was extracted after Marmur (1961). The DNA preparations were checked by the extinction ratio E_{260}/E_{280} (1·8 to 2·0) and thermal denaturation in SSC (0·15 M-Na Cl + 0·015 M-sodium citrate, pH 7) (Marmur & Doty, 1962). The DNA was renatured following the principles of Wetmur & Davidson (1968) as described

previously (Bak *et al.* 1969). The slopes of the lines for the second-order reaction rate plots were calculated by the method of least squares. For the first 60 min., which is the time for approximately one-third of the total hyperchromicity to recover, the correlation coefficients were always better than 0·995, confirming the second-order kinetics. The DNA was denatured and degraded by heating at 100° for 30 min. in 0·01 × SSC. The concentration of DNA for renaturation was 10 to 15 μg./ml. Three to nine determinations of band sedimentation at pH 7 (Bak *et al.* 1969) and corresponding K_2 values in 1 M-Na^+ at a temperature of $Tm-25°$ were done for each DNA. The relation between *Tm* in SSC and *Tm* in 1 M-Na^+ for all the DNAs was:

$$Tm_{(1\text{M-Na}^+)} = 0{\cdot}738 \times Tm_{(1\times\text{SSC})} + 30{\cdot}3 \text{ (the correlation coefficient was 0·9965).}$$

Thus with increasing GC content, the difference between *Tm* values in 0·2 and 1·0 M-Na^+, respectively, decreases. This finding is in accordance with Gruenwedel & Hsu (1969). The average molecular weight of single-stranded DNA from about 200 determinations was 332·000 ± 36 % S.D., showing that the method for degrading the DNA gave a fairly constant molecular weight (Bak *et al.* 1969). The equation for calculation of the genome sizes

$$G_s = \frac{8{\cdot}83 \times 10^8 \, (S^{\text{pH 7}}_{20,w})^{0{\cdot}911}}{K_2}$$

was derived from Studier (1965) and Wetmur & Davidson (1968). As argued previously (Bak *et al.* 1969), we have not attempted to make any corrections corresponding to the possible small influence of the base composition of DNA on the renaturation rates (Wetmur & Davidson, 1968).

The calculated genome sizes are listed in Table 1. The range of genome sizes is seen to be from about $1{\cdot}0 \times 10^9$ to about $7{\cdot}0 \times 10^9$ daltons.

The genome size of *Haemophilus influenzae* corresponds to the $0{\cdot}8 \times 10^9$ daltons, obtained by other methods (Bern's & Thomas 1965). The values for *Neisseria catarrhalis* and *N. gonorrhea* are in agreement with figures obtained by the $Cot_{\frac{1}{2}}$ method (Kingsbury, 1969). The genome size of *Escherichia coli* ($2{\cdot}8 \times 10^9$ daltons) is in agreement with Cairns ($2{\cdot}1$ to $2{\cdot}7 \times 10^9$ daltons) (Cairns, 1963). Likewise, the genome size of *Salmonella pullorum* ($2{\cdot}8 \times 10^9$ daltons) corresponds fairly well to the $2{\cdot}4$ to $3{\cdot}0 \times 10^9$ daltons determined chemically as the corrected DNA content per cell nucleoid in *S. typhimurium* (Maaløe & Kjeldgaard, 1966). The DNA content per cell nucleoid in resting cells or in uninucleated spores are particularly well investigated in several Bacillus species (Fitz-James & Young, 1959; Dennis & Wake, 1966; Eberle & Lark, 1967). The best possible value for the genome size of Bacillus species obtained from these investigations is about 3×10^9 daltons, which agrees well with the $2{\cdot}6$ to $2{\cdot}8 \times 10^9$ daltons obtained for our three species. All four species of pseudomonads have large genome sizes, 4·0 to $7{\cdot}0 \times 10^9$ daltons. These results are in disagreement with those of Park & DeLey (1967), who by chemical methods found no more than $2{\cdot}4 \times 10^9$ daltons of DNA per cell nucleoid in three Pseudomonas species.

In conclusion, the renaturation technique has proved very reliable for the determination of genome sizes of the DNA in micro-organisms and is expected to become a valuable new addition to the methods on which modern taxonomy should be founded. The information on genome sizes may create a basis for the phylogenetic differentiation of micro-organisms on a higher taxonomic level, e.g. between families,

Table 1. *Genome sizes for different bacterial species*

	Number of determinations	Genome sizes in daltons ($\times 10^9$) with standard deviation*
Achromobacteraceae		
Achromobacter anitratus (MMCA 19)†	7	1·44±0·12
Bacillaceae		
Bacillus anthracis (NCTC 8234)	4	2·78±0·35
B. cereus (ATCC 13061)	6	2·60±0·39
B. polymyxa (ATCC 10401)	4	2·75±0·37
Brevibacteriaceae		
Brevibacterium ammoniagenes (ATCC 6871)	4	1·98±0·36
Brucellaceae		
Haemophilus influenzae (MMCA 29)†	3	1·01±0·18
H. aegyptius (NCTC 8502)	4	1·17±0·11
Pasteurella multocida (MMCA 8)	4	1·13±0·16
Enterobacteriaceae		
Escherichia coli B (MMCA 56)	4	2·84±0·17
E. coli (MMCA 38)	6	2·75±0·41
Salmonella pullorum (NCTC 5776)	9	2·83±0·20
Proteus morganii (MMCA 6)	6	2·02±0·23
P. vulgaris (ATCC 13315)	6	2·09±0·49
Shigella sonnei (NCTC 8221)	4	2·09±0·25
Klebsiella ozaenae (NCTC 5053)	5	2·36±0·38
Serratia marcescens (ATCC 274)	6	5·56±0·44
S. marcescens (MMCA 55)	2	5·02±0·55
Yersinia pseudotuberculosis (ATCC 6902)	9	3·75±0·87
Lactobacillaceae		
Streptococcus pyogenes (NCTC 6175)	4	1·27±0·10
S. agalactiae (NCTC 6198 A)	4	1·20±0·15
S. faecalis (NCTC 370)	4	1·47±0·27
Diplococcus pneumoniae (MMCA 46)	3	1·45±0·23
Micrococcaceae		
Staphylococcus aureus (MMCA 1)	6	1·43±0·13
S. albus (MMCA 47)	4	1·12±0·02
Micrococcus flavus (ATCC 1024)	5	2·68±0·36
M. lysodeikticus (ATCC 4698)	4	2·82±0·10
Sarcina lutea (ATCC 9341)	6	2·78±0·35
Neisseriaceae		
Neisseria catarrhalis (NCTC 3622)	7	1·04±0·16
N. gonorrhoea (MMCA 32)	5	1·28±0·26
N. crassa (MMCA 17)	7	1·73±0·36
Pseudomonadaceae		
Pseudomonas aeruginosa (NCTC 7244)	4	6·96±1·20
Ps. fluorescens (ATCC 13525)	4	4·83±0·87
Ps. stutzeri (ATCC 17589)	4	4·21±0·62
Ps. oleovorans (ATCC 8062)	4	4·04±0·41
Rhizobiaceae		
Chromobacterium violaceum (NTCT 9371)	9	4·85±0·77
Spirillaceae		
Vibrio metschnicovii (ATCC 7708)	6	2·26±0·30

* Genome sizes were calculated by the formula:

$$G_s = \frac{8{\cdot}83 \times 10^8 (S_{20,w}^{pH\,7})^{0{\cdot}911}}{K_2}$$

where $S_{20,w}^{pH\,7}$ is the sedimentation coefficient of single-stranded DNA measured at pH 7 (Studier, 1965) and K_2 is the second-order reaction rate constant in l./mole^{-1}/sec.$^{-1}$ (Wetmur & Davidson, 1968).

† MMCA = Medical Microbiology Culture Collection, Aarhus.

while DNA base composition determinations and the results of nucleic acid homology methods allow a differentiation at a lower level.

Finally, we wish to emphasize that the size of the genome may influence the assessments of the results of nucleic acid hybridizations (DeLey, 1969).

We wish to thank Mrs Annette Juul Thomsen for technical assistance.

This work was supported by grants from the Danish State Research Foundation, the P. Carl Petersen Foundation, the Carlsberg Foundation and the F. L. Smidth and Co. A/S Foundation.

REFERENCES

BAK, A. L., BLACK, F. T., CHRISTIANSEN, C. & FREUNDT, E. A. (1969). Genome size of mycoplasmal DNA. *Nature, London* **224**, 1209–1210.

BERNS, K. I. & THOMAS, C. A. JUN. (1965). Isolation of high molecular weight DNA from *Hemophilus influenzae*. *Journal of Molecular Biology* **11**, 476–490.

BRITTEN, R. J. & KOHNE, D. E. (1968). Repeated sequences in DNA. *Science, New York* **161**, 529–540.

CAIRNS, J. (1963). The chromosome of *Escherichia coli*. *Cold Spring Harbor Symposia on Quantitative Biology* **28**, 43–46.

DELEY, J. (1969). Compositional nucleotide distribution and theoretical prediction of homology in bacterial DNA. *Journal of Theoretical Biology* **22**, 89–116.

DENNIS, E. S. & WAKE, R. G. (1966). Autoradiography of the *Bacillus subtilis* chromosome. *Journal of Molecular Biology* **15**, 435–439.

EBERLE, H. & LARK, K. G. (1967). Chromosome replication in *Bacillus subtilis* cultures growing at different rates. *Proceedings of the U.S. National Academy of Sciences* **57**, 95–101.

FITZ-JAMES, P. C. & YOUNG, I. E. (1959). Comparison of species and varieties of the genus *Bacillus*. *Journal of Bacteriology* **78**, 743–754.

GRUENWEDEL, D. W. & HSU, C. H. (1969). Salt effects on the denaturation of DNA. *Biopolymers* **7**, 557–570.

KINGSBURY, D. T. (1969). Estimate of the genome size of various micro-organisms. *Journal of Bacteriology* **98**, 1400–1401.

MACHATTIE, L. A., BERNS, K. I. & THOMAS, C. A. JUN. (1965). Electron microscopy of DNA from *Hemophilus influenzae*. *Journal of Molecular Biology* **11**, 648–649.

MAALØE, O. & KJELDGAARD, N. O. (1966). *Control of Macromolecular Synthesis*, p. 64. New York and Amsterdam: W. A. Benjamin, Inc.

MARMUR, J. (1961). A procedure for the isolation of deoxyribonucleic acid from micro-organisms. *Journal of Molecular Biology* **3**, 208–218.

MARMUR, J. & DOTY, P. (1962). Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *Journal of Molecular Biology* **5**, 109–118.

MOORE, R. L. & MCCARTHY, B. J. (1969). Base sequence homology and renaturation studies of the deoxyribonucleic acid of extremely halophilic bacteria. *Journal of Bacteriology* **99**, 255–262.

PARK, I. W. & DELEY, J. (1967). Ancestral remnants in deoxyribonucleic acid from Pseudomonas and Xanthomonas. *Antonie van Leeuwenhoek* **33**, 1–16.

SHAPIRO, H. S. (1968). In *Handbook of Biochemistry*, H-52. Edited by H. A. Sober. Cleveland, Ohio: Chemical Rubber Co.

STUDIER, F. W. (1965). Sedimentation studies of the size and shape of DNA. *Journal of Molecular Biology* **11**, 373–390.

WETMUR, J. G. & DAVIDSON, N. (1968). Kinetics of renaturation of DNA. *Journal of Molecular Biology* **31**, 349–370.