JOURNAL OF BACTERIOLOGY, Dec. 1968, p. 1915-1919
Copyright © 1968 American Society for Microbiology
Vol. 96, No. 6
Printed in U.S.A.

Base Composition of Deoxyribonucleic Acid Isolated from Mycobacteria

LAWRENCE G. WAYNE AND WENDY M. GROSS
Bacteriology Research Laboratory, Veterans Administration Hospital, San Fernando, California 91342

Received for publication 11 September 1968

Guanine plus cytosine values of deoxyribonucleic acid derived from 30 cultures representing 14 mycobacterial species or varieties are presented. These data provide impressive reasons for maintaining the separation between the genera *Corynebacterium* and *Mycobacterium;* no conclusions can be arrived at from these data with respect to the *Nocardia-Mycobacterium* relationship. A bimodal clustering, in terms of guanine plus cytosine composition, is apparent within the genus *Mycobacterium.* In general, all members of any single phenetic species appear to fit into one or another of these clusters. The phenetic separation of species is, in some cases, confirmed by separation in terms of guanine plus cytosine values. The bimodal separation of guanine plus cytosine values within the genus *Mycobacterium* does not correspond to a division of the species into slow and rapid growers; it thus provides no justification for splitting *Mycobacterium* into two genera, composed of slow and rapid growers. This is not to say that such a split would not be useful, only that these data do not contribute to such a decision. Any further attempts to correlate phenetic classification with properties of mycobacterial deoxyribonucleic acid will require more specific techniques, such as molecular hybridization.

The principles of numerical taxonomy have been applied to the study of the members of the genus *Mycobacterium* by a number of investigators in recent years. It is generally recognized that conventional bacteriological tests provide little help in resolution of species within the genus *Mycobacterium,* and it has become necessary to develop specialized schemes of biochemical testing in order to provide data which may be used for taxonomic studies within this genus (11). One of the consequences of the use of specialized schemes within a genus is the difficulty encountered in attempting to relate the different levels of taxonomic subdivision within any one genus to those recognized in another genus. It is therefore desirable to attempt to correlate results of numerical taxonomic studies with determination of base composition and, eventually, of molecular homology of the deoxyribonucleic acid (DNA) derived from different members of the genus under investigation. This is a report of the base composition of DNA isolated from representatives of a number of mycobacterial species.

MATERIALS AND METHODS

Cultures from the stock culture collection in this laboratory are maintained on Lowenstein-Jensen egg medium and transferred twice a year.

The bacteria were grown in a shallow layer of liquid Sauton medium over Sauton agar (13). The thick slurry of bacterial cells so produced was aerated vigorously on a magnetic stirrer for 3 days and abruptly subjected to anaerobic conditions in the presence of sodium ethylenediaminetetraacetate (EDTA) and the proteolytic enzyme Pronase; 24 hr later, the slurry was exposed to 0.5% sodium desoxycholate at 56 C for 90 min. The extract was then deproteinized with phenol; the DNA was precipitated with ethyl alcohol, redissolved and treated with pancreatic ribonuclease, and finally deproteinized again and reprecipitated with ethyl alcohol. Details of this method of isolation of mycobacterial DNA and some characterization of the properties of the material so produced have been described previously (13).

The guanine plus cytosine (GC) values were calculated from the melting temperatures (T_m) by plotting hyperchromic melting curves. DNA was dissolved to a concentration of approximately 20 μg/ml in the 10^{-2} M sodium phosphate, 10^{-3} M EDTA buffer at *p*H 7.0 (Bohacek et al., 1). The solutions were transferred to glass-stoppered cuvettes and placed in a Beckman DU spectrophotometer equipped with thermospacers on each side of the cuvette holder. The thermospacers adjacent to the cell holder were connected to a circulating hot water-ethylene glycol bath, the temperature of which was raised at a rate of 1 C per 10 min beginning at 74 C. Temperature was measured by placing a thermometer directly into one of the cuvettes with a light shield placed over it. The per cent GC was calculated according to the equation (1): $T_m = 51 + 0.45$ (GC).

Results obtained with this method were compara-

ble to GC values of *M. tuberculosis* determined by hydrolysis and chromatography as well as by buoyant density measurement (13). Furthermore, GC values on DNA from a strain of *M. smegmatis* have been determined as 67.6% by buoyant density and 67.8% by this method of determining T_m (*unpublished data*).

RESULTS

Satisfactory melting curves were obtained for DNA isolated from 30 strains, representing 14 different species or varieties of mycobacteria. All determinations were made at least in duplicate. Assignment of an organism to a given species was based on demonstration of a similarity score of greater than 85% when compared to the hypothetical median strain pattern of that species by numerical taxonomic methods described previously (10–12). Of the species examined by this method, adequate preparations of DNA were obtained for all except *M. bovis* and *M. terrae*, which did not grow well on the glycerol rich medium employed, and *M. gastri*, which yielded a DNA preparation which coprecipitated with a substance that interfered with satisfactory determination of T_m. These species are under continued investigation.

Figure 1 presents the distribution of GC within the genus *Mycobacterium* as well as a comparison with data derived from members of other genera which may or may not be closely related to this genus (5).

Data from this study are presented in the lower half of Fig. 1. As might be expected, when all the work is done in a single laboratory, the range of values is narrower than that seen in a compilation from a number of investigators. Thus, in his compilation, Hill noted the reported range for mycobacterial DNA derived from 10 strains, representing 3 species, to be between 62 and 73% GC; 6 of these cultures were identified as *M. tuberculosis* and covered a range of 62 to 70% GC. We have found, with a larger number of strains and species, that mycobacterial DNA covers a range of GC composition from 64 to 70%.

A striking feature of our data is at least a bimodal distribution of values within the genus *Mycobacterium*. Members of the genus may be characterized as possessing DNA of GC composition of less than 66.5% or greater than 66.5%. As indicated in Table 1, all tested members of a given species or subgroup fit into one or the other of the two clusters in terms of GC composition. Thus, the four strains of *M. kansasii* tested exhibited base compositions in the range of 64.1 to 65.8% GC, whereas three strains of *M. avium* fell into the range of 67.4 to 69.0%.

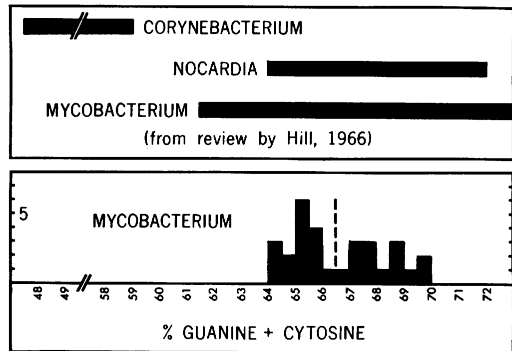

FIG. 1. *Distribution of GC composition of mycobacteria, nocardias, and corynebacteria. The upper half is an expression of the range of reported values for the three genera, collected by Hill (5). The lower half is a histogram of distribution of cultures with different GC values as determined in the present study.*

DISCUSSION

Investigators have long been concerned with the relationship between *Nocardia* and *Mycobacterium*, and there is complete overlap of reported GC values of DNA for members of these two genera (Fig. 1). Thus, these two genera cannot be distinguished on the basis of DNA base composition, but neither can it be concluded that these two genera have very similar DNA. Although the proportions of the bases are essentially the same, the similarity of sequences of bases cannot be known until studies on molecular hybridization have been carried out.

On the other hand, Harrington (4), after completing a numerical taxonomic study of the genus *Corynebacterium* and having included a few mycobacteria, proposed that the corynebacteria, the mycobacteria, and the nocardias could all be lumped in a single genus. We have some doubts about this, initially, on the basis of the methodology used in his numerical taxonomy, because in our opinion the selection of tests is not truly unbiased. Mycobacteria which exhibit striking differences in our hands show no significant differences when schemes of numerical taxonomy, which are most appropriate for non-acid-fast organisms (9), are applied to them. Even by using the broad range of unselected GC values which were included in Hill's review, it is quite evident that the upper limit of the range of corynebacterial DNA does not impinge on the lowest reported limit of members of the genus *Mycobacterium;* that is, they clearly differ in DNA base composition.

On further examination of the data compiled by Hill (5) for many genera, it is seen that a phenetically defined genus usually covers a range

TABLE 1. *GC values of DNA derived from representative mycobacteria*

Runyon group	GC range 64.0–66.4%				GC range 66.5–70.0%			
	Species	Culture no.[a]	% GC	AD[b] (±)	Species	Culture no.[a]	% GC	AD (±)
I	*M. tuberculosis*	H37Rv	65.0	0.15				
	M. tuberculosis	731	64.3	0.1				
	M. kansasii (high catalase)	P1	65.7	0.35				
	M. kansasii (high catalase)	A12478	65.8	0				
	M. kansasii (low catalase)	427	64.1	0.1				
	M. kansasii (low catalase)	448	65.3	0				
	M. marinum	467	64.2	0.2				
	M. marinum	468	65.0	0.1				
II	Tap-water scotochromogen	905	65.9	0.15	*M. scrofulaceum*	952	68.0	0.25
	Tap-water scotochromogen	A19277	65.4	0.1	*M. scrofulaceum*	A19981	67.9	0.15
					M. marianum	899	68.5	0.1
III	*M. xenopei*	A19276	65.5	0.2	*M. avium*	226	68.5	0.1
	M. xenopei	1183	64.5	0.1	*M. avium*	P14	69.0	0.2
					M. avium	A15769	67.4	0.1
					M. intracellulare (Battey)	294	67.3	0.2
					M. intracellulare (Battey)	708	68.6	0.45
IV	*M. fortuitum*	201	65.1	0	*M. smegmatis*	763	66.7	0.3
	M. fortuitum	A6841	66.1	0.1	*M. smegmatis*	A14468	67.8	0.25
	M. flavescens	272	65.4	0.1	*M. phlei*	762	67.4	0.1
	M. flavescens	A14474	64.9	0.25	*M. phlei*	A19249	69.9	0.35
					M. rhodochrous	305	67.9	0.15
					M. rhodochrous	355	70.0	0.25

[a] Culture numbers preceded by a code letter are those which were received from special collections and have received wide distribution. The letter "P" indicates a culture from a set distributed to many investigators under the indicated code number by E. H. Runyon. The letter "A" indicates the ATCC catalog number of the culture.
[b] Average deviation.

of base ratios of about 10%. There do appear to be much broader ranges for some genera. However, on closer inspection of the data, these very broad ranges can be explained either because a single value in the literature fell far outside the range of the great majority of cultures and thus may well have represented an error in identification, or the genera exhibited well-separated bi- or polymodal distribution of GC values, and, as in the cases of the genera *Bacillus*, *Pseudomonas*, *Vibrio*, and others, each of these might be separated into two or more different genera. On the other hand, it is not uncommon to see further finer subdivisions into bi- or trimodal distribution within a fairly narrow range of GC values for a genus; this is what is seen among the mycobacteria as well. There is at least a bimodal distribution of frequencies of GC values within the range 64 to 70%. In an attempt to relate this to phenetic characterizations, we chose a point of lowest frequency at approximately 66.5% GC and separated our organisms into those which had a GC value somewhere between 64 and 66.5% GC and those that had a GC value in excess of 66.5%.

After this division has been made (Table 1), organisms which cluster together on phenetic grounds also appear in the same cluster in terms of the GC composition of the bacterial DNA. Thus, all four strains of *M. kansasii* examined fell in the 64 to 66.5% cluster, as did both strains of *M. tuberculosis*. On the other hand, both

strains of *M. smegmatis*, both strains of *M. phlei*, and all three strains of the *M. scrofulaceum-M. marianum* group (it is generally accepted that these two species are synonymous on phenetic grounds; 12) fell in the cluster exhibiting GC values in excess of 66.5%. The two strains of *M. intracellulare* or Battey bacillus included in these data also fell in the same range as *M. avium*.

We had initially expected that a bimodal distribution of per cent GC within the genus might reflect a difference between members of the the slow-growing mycobacteria and the rapid growers, which are different in a number of physiological properties. This did not prove to be the case; the *M. fortuitum* and *M. flavescens* cultures, which belong with the rapid growers of Runyon's group IV, have a low GC value, whereas the strains of *M. smegmatis* and *M. phlei* are in the high GC range. Similarly, many slow growers are in the low range; other slow growers have high GC values.

Clearly, these data do not provide very fine species distinctions. Allowing for experimental deviation alone, the best that could be done was to place the organisms in one or the other of these two clusters of GC ratios; however, this has proven useful because it has helped to confirm some of the phenetic distinctions made on the basis of numerical taxonomic procedures. Thus, on rather limited evidence, it had been decided to separate the tap-water scotochromogens of group II from the scrofula scotochromogens of this same Runyon group. The distinction initially was based primarily on differences in ability to hydrolyze Tween 80, but was also supported by differences in amidase content and some serological evidence (12). We feel more confident in this distinction now because the scotochromogens are clearly separated into different GC ranges, and there is no overlap on the basis of this prior phenetic distinction.

On a phenetic basis, *M. xenopei* appears closely related to both *M. scrofulaceum* and *M. avium*, but the *M. xenopei* DNA preparations we examined fell into the low GC range.

We had hoped to be able to clarify the position of *M. rhodochrous* which has been described by Gordon as a species in search of a genus (3). The DNA from the cultures we examined had GC values of 67.9 and 70%, excluding them from the genus *Corynebacterium*. From these data, this organism could be a *Mycobacterium* or a *Nocardia*, if one limits them to *Actinomycetales*. Recent observations of Tewfik and Bradley (8) suggest a close relationship of *Nocardia corallina* and *N. opaca* to *Nocardia* and *Streptomyces*, and the question of the synonymy of these organisms with *M. rhodochrous* is discussed.

Our GC values of DNA from *M. tuberculosis* are in close agreement with others determined by different methods and reported in recent publications (7, 8). On the other hand, marked differences are seen between reported values for *M. kansasii*. Thus, Tewfik and Bradley (8), employing buoyant density determinations on DNA obtained by mechanical grinding of cells, found their strain of this species to have a GC value of 68.5%, as compared to the lower range of 64.1 to 65.8% obtained with the four strains we examined. Tarnok et al. (7), on the other hand, reported a low value of 60.8% as the average of seven strains examined. They did not report the range of values observed when a number of strains of a given species were examined, but only the average value, so it is difficult to evaluate the discrepancies between their data and ours for the 10 species or subgroups that we have both examined.

Several differences in techniques employed may help to explain the discrepancies observed between the data presented here and those published by Tarnok et al. (7). We employed an autolytic system for release of the DNA, whereas Tarnok et al. employed a process of mechanical grinding for 30 min, a process which causes shearing of DNA. Although shearing per se should not affect base ratios of purified DNA, one may speculate that shearing effects might yield a disproportionate distribution of smaller fragments, derived from the ends rather than the center of the molecule. This in turn could lead to selection of the larger fragments during the deproteinizations and ethyl alcohol precipitations. If the distribution of bases was not uniform throughout the original molecule, this enrichment could yield a product either richer or poorer in a given base pair than was the original molecule. Furthermore, the isolation procedure employed by Tarnok et al. was carried out in saline without addition of citrate or other chelating agents which inactivate deoxyribonuclease. This also could have contributed effects similar to those caused by mechanical shearing. An additional technical difference was the failure of Tarnok et al. to employ a digestion with ribonuclease during purification of their product. Although they state that no ribonucleic acid (RNA) was detectable in their final product, no details of sensitivity of their orcinol assay procedure were given. Their GC values were calculated from the ratio of absorbance at 260 and 280 nm, and, as pointed out by Fredericq et al. (2), this method yields significant errors when traces of RNA are present. Savitsky and Stand (6) have presented evidence that orcinol methods for determination of traces of RNA in DNA preparations would not dis-

tinguish levels of RNA between 5 and 10% of the product because of some masking reactivity of DNA itself.

ACKNOWLEDGMENT

This investigation was supported by a grant from the American Thoracic Society, the Medical Section of the National Tuberculosis and Respiratory Disease Association.

LITERATURE CITED

1. Bohacek, J., M. Kocur, and T. Martinec. 1967. DNA base composition and taxonomy of some micrococci. J. Gen. Microbiol. **46**:369–376.
2. Fredericq, E., A. Oth, and F. Fontaine. 1961. The ultraviolet spectrum of deoxyribonucleic acids and their constituents. J. Mol. Biol. **3**: 11–17.
3. Gordon, R. E. 1966. Some strains in search of a genus—*Corynebacterium*, *Mycobacterium*, *Nocardia* or what? J. Gen. Microbiol. **43**:329–343.
4. Harrington, B. J. 1966. A numerical taxonomical study of some corynebacteria and related organisms. J. Gen. Microbiol. **45**:31–40.
5. Hill, L. R. 1966. An index to DNA base compositions of bacterial species. J. Gen. Microbiol. **44**:419–437.
6. Savitsky, J. P., and F. Stand. 1965. A simple method for measuring ribonucleic acid content of preparations of deoxyribonucleic acid. Nature **207**:758–759.
7. Tarnok, I., E. Röhrscheidt, and R. Bönicke. 1965. Basenzusammensetzung der Deoxyribonukleinsäure (DNS) von Mykobakterien ver wandten Mikroorganismen. Rassegna Patol Apparato Respiratorio **17**:3–16.
8. Tewfik, E. M., and S. G. Bradley. 1967. Characterization of deoxyribonucleic acids from streptomycetes and nocardiae. J. Bacteriol. **94**: 1994–2000.
9. Wayne, L. G. 1964. The mycobacterial mystique: deterrent to taxonomy. Am. Rev. Respirat. Diseases **90**:255–257.
10. Wayne, L. G. 1966. Classification and identification of mycobacteria III. Species within group III. Am. Rev. Respirat. Diseases **93**:919–928.
11. Wayne, L. G. 1967. Selection of characters for an Adansonian analysis of mycobacterial taxonomy. J. Bacteriol. **93**:1382–1391.
12. Wayne, L. G., J. R. Doubek, and G. A. Diaz. 1967. Classification and identification of mycobacteria. IV. Some important scotochromogens. Am. Rev. Respirat. Diseases **96**:88–95.
13. Wayne, L. G., and W. M. Gross. 1968. Isolation of deoxyribonucleic acid from mycobacteria. J. Bacteriol. **95**:1481–1482.