Nucleotide Sequence Divergence Among DNA Fractions of Different Syngens of *Tetrahymena pyriformis*

Sally Lyman Allen[1] and Chi Ing Li[1]

The magnitude of the differences in base sequence of DNA fractions derived from different syngens of the ciliated protozoan Tetrahymena pyriformis *was investigated. Each DNA was fractionated into unique and repeated sequences by hydroxylapatite chromatography, and the fractions were tested by* in vitro *molecular hybridization techniques. The amount of hybrid formed and the thermal stability of the hybrid molecules were examined at different incubation temperatures (50 and 65 C) for unique sequences and at 50 C for repeated sequences. The extent of the reactions involving either unique or repeated sequences was nearly complete when the two DNAs compared were derived from the same syngen. Moreover, intrasyngenic hybrids formed at 50 C (and 65 C for unique sequences) exhibit a high degree of thermal stability. In contrast, the extent of the reactions involving sequences derived from different syngens was low, as expected from the effect of mismatching on rate of reassociation, and intersyngenic hybrids formed at 50 C have low thermal stability. The reaction of unique sequences is further reduced at 65 C and the intersyngenic hybrids formed have a higher thermal stability than those formed at 50 C. The degree to which thermal stability is lowered was then used to estimate the percentage of mispaired bases. The average divergence of unique sequences between syngens is large and of the magnitude found for rodent DNAs from different genera or for* Drosophila *DNAs from nonsibling species. The repeated sequence fraction may contain more than one component and may be more conserved than the unique sequence fraction.*

C   **213**

INTRODUCTION

Morphological criteria, along with a strong fossil record, are helpful in sorting out the relationships between higher animals and plants. Thus the greater the degree of similarity between two organisms, the closer the relationship. In microorganisms, morphological criteria become less useful in separating organisms, often vanishingly small differences being observed between groups widely separated from each other. This feature, coupled with the lack of a fossil record and barriers to genetic recombination, has made difficult the proper assessment of evolutionary distances. This problem is encountered in assessing relationships among different strains of bacteria. It is also encountered in the protozoa. For example, all strains of the ciliated protozoan *Tetrahymena pyriformis* are morphologically similar. Yet it has been possible to group many of the strains into "biological species," or "syngens," on the basis of their mating relationships (Sonneborn, 1957). Moreover, fruitful gene flow does not occur between syngens. The syngen is at least comparable to a species, but it could be comparable to an even higher taxonomic level. The difficulties of equating taxon rank in different groups of organisms were pointed out by Laird and McCarthy (1968); thus groups more closely related to man tend to be more finely divided than remote groups. We do not know how diverse are the different groups under the taxonomic umbrella *Tetrahymena pyriformis*. Comparisons of macromolecules suggest wide differences between syngens (Allen and Gibson, 1967, 1972, 1973; Allen and Weremiuk, 1971; Borden *et al.*, 1974). At a minimum, the taxon *Tetrahymena pyriformis* is equivalent to a genus in vertebrates, but it could be comparable to a higher order of taxonomic grouping if the differences between syngens are greater.

Although subtle differences in cortical architecture can be utilized in a statistical sense to distinguish strains belonging to different syngens of *T. pyriformis*, the variations are not syngen specific (Nanney, 1971; Nanney and Chow, 1974). Biochemical differences, on the other hand, have been successfully exploited to distinguish the syngens and also to determine the evolutionary distances separating syngens. A comparison of the isozymes of eight different enzymes has revealed that "vast molecular distances" separate the different syngens (Borden *et al.*, 1974). Numerical taxonomic procedures were used to assort the data and to arrange the syngens on the basis of phenotypic similarity. The most closely related syngens had a degree of similarity of 67% and the most distantly related syngens 0% similarity, using mean character differences averaged over several loci.

Nucleotide sequence relationships have been studied in a variety of different organisms. In general, these studies have supported taxonomic groupings based on morphological criteria (reviewed by Kohne, 1970). Where morphological data are lacking, they have provided information in generating taxonomic order, as, for example, in the case of bacteria (Brenner *et al.*, 1969). This approach has also been applied toward assessing the distances between the syngens of *T. pyriformis* (Allen and Gibson, 1967, 1972, 1973).

In the early studies utilizing nucleic acid reassociation to assess relationships between syngens, the DNA was unfractionated. Since then, it has been found that the DNA of *T. pyriformis*, like that of other eukaryotic organisms, contains different fractions separable on the basis of the rate of renaturation (Allen and Gibson, 1972; Wille, 1972). In other eukaryotic organisms, different fractions differ in their degree of homology between species (see review by Bostock, 1971). Thus certain specific portions of the genome evolve more rapidly than other portions. In syngens 1 and 7 of *T. pyriformis*, 15–20% of the genome behaves as repeated sequences, the remainder as unique sequences under particular growth conditions (Allen and Gibson, 1972; Allen, Douthit, and Gibson, unpublished).

In the nucleic acid reassociation experiments reported in this paper, we have fractionated the DNA of the 12 syngens into repeated and unique sequences and conducted molecular hybridization experiments on the isolated fractions. The amount of hybrid formed and the thermal stability of the hybrid molecules have been examined at different incubation temperatures in order to assess syngen relationships. Our results confirm certain of the relationships presented by Borden *et al.* (1974) and extend the thesis that the syngens in *Tetrahymena pyriformis* are indeed separated by vast molecular distances.

MATERIALS AND METHODS

Strains

The strains used and their geographic origins are shown in the first two columns of Table I.

Growth and Harvesting of Cells

The cells were grown in 1% proteose–peptone (w/v) in 5-liter Povitsky bottles without aeration or shaking for 7 days at 23 C. The yield of packed cells varied from 0.3 to 1.5 ml/liter of culture. The cells were harvested in a DeLaval cream separater, suspended in Dryl's salts solution (Dryl, 1959),

Table I. Strains of *Tetrahymena pyriformis*

Syngen/strain	Geographic origin	Percent repeated sequences	Native DNA (% G+C)	$T_{1/2}{}^a$ of fractions (C) US[b]	$T_{1/2}{}^a$ of fractions (C) RS[b]
1/D	Inbred, U.S.A. (ALP-4 × B)	25	*24–25[c]*	79.8	80.2
1/D/1	Inbred, U.S.A. (C2 × D)D^{12} × C*	22	*24*	79.5	80.0
1/B	Inbred, U.S.A. (WH-6 × WH-14)	26	*25*	79.0	79.5
2/UM-3	Massachusetts	25	25	79.2	80.0
2/UM-351	Michigan	22		79.4	80.3
3/UM-700	Mississippi	22	27	79.5	80.0
4/IN-3	India	23	*30–33*	82.0	82.5
5/UM-30	Massachusetts	33	25–27	79.5	80.5
6/UM-1091	Florida	24		79.5	80.8
6/UM-1060	Michigan	27	25	79.5	80.5
7/UC-651	California	23	25–27	79.0	79.5
8/ALP-1	Michigan	27	25	79.5	80.6
9/TC-89	Panama	17	*28–30*	80.0	81.0
10/EN-112	England	19	*31*	82.0	82.5
11/AU-50-1	Australia	22	*30*	81.8	82.3
12/AU-F$_1$-4	Australia	22	*25*	79.0	80.0

[a] $T_{1/2}$ is the temperature at which 50% of the DNA is eluted.
[b] US, unique sequences; RS, repeated sequences.
[c] The figure in italics was obtained on the particular strain used in our laboratory. The other figures were obtained by others (see Table 20 in Allen and Gibson, 1973, for a summary of all data on base composition).

and spun at 1000g in pear-shaped tubes in an oil-testing centrifuge. The pellet of cells was then used for preparing DNA.

Preparation of DNA

Method 1 of Allen and Gibson (1971) was followed for preparation of DNA. DNA made by this method is highly purified, giving less than 10% RNA and/or carbohydrate and no detectable protein. It is also of high molecular weight. The yields of spoolable DNA from different syngens varied. The highest yields were usually obtained from syngen 1, with an average of 15 mg DNA being recovered from 30 ml of packed cells. Radioactive DNA was prepared from syngen 1, strain D (1/D), using thymidine-methyl-H^3, as outlined by Allen and Gibson (1971). The specific activity of this DNA was 10,000 cpm/μg.

Fractionation of DNA

DNA was sheared to segments 400–500 nucleotide pairs in length (Allen and

Gibson, 1971). It was then heat-denatured (5 min at 100 C), fast-cooled on dry ice, and then renatured to C_0t_{20} in 0.12 M sodium phosphate buffer (pH 6.8) at 50 C. This C_0t value was chosen from previous information on syngens 1 and 7 (Allen and Gibson, 1972; Allen, Douthit, and Gibson, unpublished). With this C_0t value, there may be some contamination of the repeated sequence fraction with unique sequences, but the unique sequences should be free of repeated sequences. Repeated sequences were then separated from unique sequences by salt chromatography on large hydroxylapatite columns (column volume 40 ml, maximum capacity 10 mg DNA). The hydroxylapatite was obtained from Bio Rad Laboratories (Richmond, Calif.) and was boiled before use to prevent nonspecific binding of denatured DNA. The fractions were then concentrated by lyophilization, desalted by dialysis, and dialyzed extensively against 0.12 M phosphate buffer. The samples were frozen until further use.

Hybridization of DNA Fractions

Concentrated unlabeled DNA (1–3 mg/ml) and diluted radioactive DNA fragments (0.25–2.5 μg/ml) were added to a small volume (0.25 ml final volume) in 0.12 M phosphate buffer in tightly sealed tubes. The ratio of unlabeled to radioactive unique sequence DNA was 100–250, whereas the ratio of unlabeled to radioactive repeated sequence DNA was 1000–2500. Four replicate samples were run for each experiment involving unique sequences and two replicate samples were run in the one experiment involving repeated sequences. For each set of experimental samples, a set of control samples was also set up. These samples contained only radioactive DNA and served as a measure of "self-reaction." All samples were heat-denatured (5 min at 100 C), fast-cooled on dry ice, and incubated at 50 C in one experiment involving unique and repeated sequences. In a second experiment involving only unique sequences, incubation was carried out at 50 or 65 C on replicate samples (four at each temperature). Incubation was concluded at a C_0t of 1500–2000 for unique sequences (or such that 90–95% of the unlabeled DNA had renatured at 50 C) and at a C_0t of 250 for repeated sequences (or such that 80–90% of the unlabeled DNA had renatured). At the termination of incubation, the samples were chilled on ice, and each sample was diluted with cold 0.12 M phosphate buffer to the appropriate column volume and frozen until thermal chromatography was carried out.

Thermal Chromatography

DNA samples were overlaid on small hydroxylapatite columns (1.5–2.0 ml column volumes, maximum capacity 375–500 μg DNA) held at the tem-

perature at which the fragments had been incubated and washed five times (each wash equals one column volume) with 0.14 M sodium phosphate buffer (pH 6.8) at this temperature. The fractions eluted contain *un*renatured DNA. Then, using the same buffer throughout, the temperature was raised in 2.5–5 C steps, and three washes were taken at each temperature. The total range of temperatures used was 55–95 C for a 50 C incubation and 70–95 C for a 65 C incubation. These fractions contain renatured DNA which has been denatured on the column by the increased temperature. A final step included two washes with 0.5 M sodium phosphate buffer, pH 6.8, at 90 C to recover any remaining renatured DNA. Rarely was any DNA recovered at this step. The total recovery of the DNA from the column was close to 100%. All experimental fractions were assayed for their ultraviolet absorbancy at 260 nm (unlabeled DNA) and for their radioactivity by scintillation counting (radioactive DNA). Control fractions were assayed only for their radioactivity.

RESULTS

Properties of Renatured DNA in Different Syngens

When DNA is renatured to C_0t 20, the proportion of repeated sequences shows some variation between different strains and/or syngens (third column, Table 1). The mean for all strains is 24%, which is similar to the value observed for the syngen 1 strains. A repeated sequence fraction isolated from 1/D DNA, incubated to C_0t 30, has been tested in renaturation experiments (Allen, Doughit, and Gibson, unpublished). There was considerable heterogeneity in the kinetics observed. About 15% of this fraction of DNA renatured by C_0t 0.1, and about 40% of this fraction renatured with the kinetics expected for unique sequences. The extent of the self-reaction of the radioactive repeated sequences in the experiments reported here was 24% by C_0t 0.25, suggesting that it is composed mainly of the rapidly renaturing component. Nevertheless, we probably do have contamination of the repeated sequences with unique sequences. We do not know the extent of this contamination, but, using information obtained on syngen 1 repeated sequences, it may be as high as 25% at C_0t 20.

The thermal stabilities of the renatured DNA of different syngens reflect differences in their base compositions. A comparison of the % G+C of native DNA with the $T_{1/2}$ of the fractions is shown in the last three columns of Table I. Thus syngen 4 with 30% G+C gives fractions which have $T_{1/2}$ values of 82.0 and 82.5 C compared to syngen 1 with 24–25% G+C and fractions which have $T_{1/2}$ values of 79.5 and 80.0. The repeated sequence fraction has a slightly but consistently higher $T_{1/2}$ than that of unique sequences. This appears to be true for all the strains and syngens. However,

this difference between the fractions may reflect, in part, size differences in the DNA pieces selected during fractionation on hydroxylapatite, since we have found in other experiments with syngen 1 DNA that the molecular weights of repeated sequence fractions are slightly higher than the molecular weights of unique sequence fractions (Allen, Douthit, and Gibson, unpublished).

Extent of Reaction and Thermal Stabilities of Unique Sequence Hybrids

Radioactive unique sequences from syngen 1, strain D (1/D), were incubated at 50 C in solution with unlabeled unique sequences from 16 different sources, three from strains within syngen 1 and 13 from strains in the 11 other syngens. In the homologous reaction, the percentage of radioactive fragments bound to unlabeled DNA was similar to the percentage of unlabeled DNA which

Table II. Extent of Reaction and Thermal Stabilities of Unique Sequence DNA Hybrids Formed at 50 C

Source of unlabeled DNA (syngen/strain)	Percent renaturation		Percent homology[b]	$T_{1/2}$ labeled DNA (C)	$\Delta T_{1/2}$ labeled DNA (C)	Percent mispaired bases
	Unlabeled DNA	1/D labeled DNA[a]				
1/D	89.1	83.3	100	79.0	—	—
1/D/1	83.8	82.3	104	77.0	−2.0	1
1/B	92.1	82.8	96	76.0	−3.0	2
8/ALP-1	94.7	31.9	36	75.0	−4.0	3
6/UM-1091	92.6	25.0	29	74.0	−5.0	3
6/UM-1060	90.7	24.5	29	74.0	−5.0	3
10/EN-112	90.0	15.5	18	72.0	−7.0	5
3/UM-700	91.3	18.3	21	68.0	−11.0	7
5/UM-30	94.1	15.4	18	66.0	−13.0	9
9/TC-89	90.9	14.8	17	61.0	−18.0	12
2/UM-3	90.0	12.5	15	61.0	−18.0	12
2/UM-351	90.1	13.8	16	61.0	−18.0	12
7/UC-651	86.0	10.0	12	60.0	−19.0	13
4/IN-3	90.3	9.2	11	60.0	−19.0	13
11/AU-50-1	91.9	8.4	10	60.0	−19.0	13
12/AU-F$_1$-4	92.8	7.7	9	59.0	−20.0	13

[a] Corrected for self-reaction of labeled DNA.
[b] Calculation of homology: A "value" was obtained for each mixture of unlabeled DNA with radioactive 1/D DNA by the following formula:

$$\frac{\% \text{ renatured } 1/D \text{ DNA (corrected for self-reaction)}}{\% \text{ renatured unlabeled DNA}}$$

Dividing by the % renatured unlabeled DNA adjusts for differences in the completion of the hybridization reaction between samples. The value of the homologous reaction was then set at 100% and all other values compared to it.

Table III. Comparison of Extent of Reaction and Thermal Stabilities of Unique Sequence DNA Hybrids Formed at 50 and 65 C

Incubation temperature (C)	Source of unlabeled DNA (syngen/strain)	Percent renaturation			$T_{1/2}$ labeled DNA (C)	$\Delta T_{1/2}$ labeled DNA (C)	Percent mispaired bases
		Unlabeled DNA	1/D labeled DNA[a]	Percent homology			
50	1/D	92.7	86.0	100	78.0	—	—
	1/D/1	91.0	86.6	102	77.0	−1.0	<1
	1/B	91.0	85.3	101	76.0	−2.0	1
	8/ALP-1	91.6	26.5	31	75.0	−3.0	2
	6/UM-1091	88.4	19.6	24	74.0	−4.0	3
	6/UM-1060	90.7	20.5	24	74.0	−4.0	3
	11/AU-50-1	90.6	10.4	12	60.0	−18.0	12
	12/AU-F$_1$-4	86.5	8.8	11	59.0	−19.0	13
65	1/D	87.0	84.5	100	78.0	—	—
	1/D/1	84.7	82.4	100	78.0	−0.0	0
	1/B	86.0	82.0	98	77.5	−0.5	<1
	8/ALP-4	84.8	6.1	7	77.0	−1.0	<1
	6/UM-1091	80.0	4.7	6	77.0	−1.0	<1
	6/UM-1060	80.9	4.5	6	77.0	−1.0	<1
	11/AU-50-1	80.2	4.2	5	74.0	−4.0	3
	12/AU-F$_1$-4	79.4	3.1	4	73.0	−5.0	3

[a] Corrected for self-reaction of labeled DNA.

renatured. The results of two different experiments are shown in the top line in Tables II and III. Under similar incubation conditions, the extent of the self-reaction of radioactive fragments was 5.8 and 5.6%, respectively. The thermal stability of the homologous hybrid was high ($T_{1/2}$ = 79.0 C in the first experiment and $T_{1/2}$ = 78.0 C in the second experiment). The data from these two experiments were pooled in computing the distribution of radioactive counts eluted from the hybrid with unlabeled DNA and from the hybrid resulting from self-reaction. These distributions are shown in Fig. 1. In the results to be described, the percentage of each experimental reaction was corrected by subtracting the percentage of self-reaction of that experiment. The thermal elution data were also corrected for self-reaction.

Table II includes the results of the hybridization experiments with all 16 sources of DNA. For DNA from strains D/1 and B, both in syngen 1, the extent of reaction is similar to the homologous reaction. Strain D/1 is congenic with strain D, while strains B and D are further apart genetically. That is, strain D was derived from an outcross of a wild isolate, ALP-4, and strain B, but it has been separated from strain B by selection of certain markers during the course of inbreeding by selfing for over 20 generations. The homologies to D were 104% for D/1 and 96% for B. Moreover, the thermal

Fig. 1. Thermal elution profiles of renatured 1/D unique sequence DNA from 50 C incubations. The distribution of radioactive counts between 55 C and 95 C is graphed, each point being the average of eight different hydroxylapatite columns. The curve represented by the solid line shows the thermal stability of labeled DNA which renatured with homologous unlabeled DNA at C_0t 1670. The curve represented by the dashed line shows the thermal stability of labeled DNA renatured in the absence of unlabeled DNA to C_0t 13 (self-reaction).
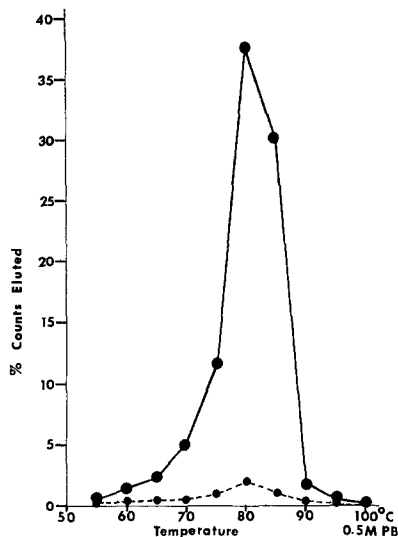
stabilities of the intrasyngenic hybrids were nearly as high as that of the homologous hybrid ($T_{1/2} = 77.0$ C for D/1; $T_{1/2} = 76.0$ C for B). In contrast, the extent of the reaction of 1/D DNA with the DNA from other syngens was significantly lower than that with syngen 1 DNA. The homology to 1/D ranges from 36% for syngen 8 to 9% for syngen 12. For two syngens (syngens 2 and 6), two different strains were included. Within the limitations of the experiment, the results from the two strains agreed. The calculations of homology, however, are most likely underestimates of relationship, since reassociation rates are reduced when complementary sequences are imperfectly matched (Bonner *et al.*, 1973; McCarthy and Farquhar, 1972). This effect is proportional to the degree of mismatching. Thus the lowest values for homology would be those which are the most underestimated.

The thermal stabilities of the intersyngenic hybrids show a wide range but are lower than for any of the intrasyngenic hybrids. The $T_{1/2}$ is lowered by 4 C in syngen 8 but by 20 C in syngen 12. The difference in $T_{1/2}$ ranges in between for the other syngens. Often bimodal curves are observed for the thermal elution profiles of the intersyngenic hybrids, illustrated for the hybrid formed between syngens 1 and 8 (Fig. 2a). Such curves indicate that a portion of the heterologous hybrid is made up of mismatched sequences which melt at low temperatures and a portion is made up of well-matched sequences which melt at temperatures characteristic of homologous hybrids of syngen 1 DNA. In the case of the syngen 1–12 hybrid, most of the sequences appear to be mismatched (Fig. 2a).

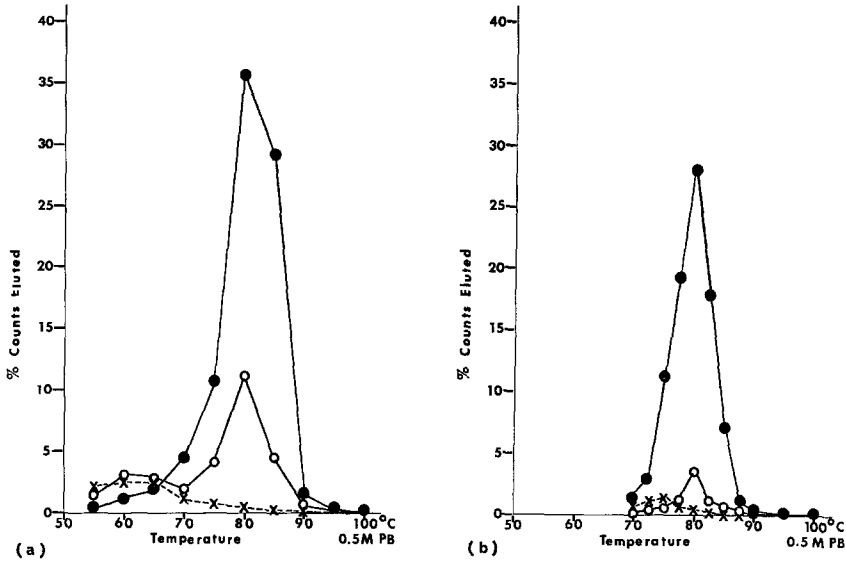The choice of 50 C as the incubation temperature for the hybridization

Fig. 2. Comparison of thermal elution profiles of homologous and heterologous unique sequence DNA hybrids formed at (a) 50 C and (b) 65 C. Labeled 1/D DNA in homologous hybrid (●) and in heterologous hybrids, syngen 8 (○) and syngen 12 (×). The distribution of radioactive counts between 55 (or 70) and 95 C is graphed, each point being the average of four different columns and corrected for self-reaction.

reaction was dictated by the consideration that the optimal temperature for renaturation is roughly 30 C below $T_{1/2}$ (Marmur and Doty, 1961). It also happens to be closest to the optimum temperatures for comparing relative rates of reassociation between well-matched and imperfectly matched *Tetrahymena* DNAs (see Fig. 8 of Bonner *et al.*, 1973). One can, however, make the conditions for hybridization more stringent by using a higher temperature for incubation. In this case, only well-matched sequences will be able to react and the rate of reassociation involving imperfectly matched sequences is far from the optimum. In a second experiment, then, we compared the extent of reaction and the thermal stabilities of the hybrids formed at 50 C and at 65 C. We used unique sequences from eight different sources, three from strains within syngen 1 and five from strains in syngens 8, 6, 11, and 12. The average percentage renaturation of the unlabeled DNAs was 90.3% for reactions carried out at 50 C. At 65 C, the average percentage renaturation was 82.9%, or a 7.4% reduction. A similar reduction in the amount of renaturation with higher temperature was found for bacterial DNA (Brenner *et al.*, 1969). At 65 C, the radioactive fragments bound to a similar extent to homologous unlabeled DNA, but the self-reaction was reduced from 5.6% (at 50 C) to 2.4%.

The data obtained with a 50 C incubation shown in Table III (top) are, in general, in good agreement with the data obtained in the first experiment, shown in Table II, except that some of the reactions are slightly lower. The range of homologies for the five strains in syngens 8, 6, 11, and 12 runs from 31% for syngen 8 to 11% for syngen 12. The difference in $T_{1/2}$ ranges from 3 to 19 C. In contrast, the data obtained from the 65 C incubation show a marked difference in the extent of reaction between strains within syngen 1 and those from other syngens. The homologies of the syngen 1 strains are close to 100%, while those of other syngens are reduced to 4–7%. The difference in $T_{1/2}$ is also reduced to between 1 and 5 C, indicating that those sequences capable of reacting are better matched than is the case at 50 C. This is revealed in comparing the thermal elution profiles. The contrast in the melting of hybrids formed at 50 C and at 65 C is illustrated in Fig. 2(a,b) for the homologous hybrid and two intersyngenic hybrids, 1–8 and 1–12. After an incubation at 65 C, the curve for the homologous hybrid is sharpened, compared to the 50 C incubation, with a symmetrical peak being observed over 80 C. The intersyngenic 1–8 hybrid formed at 50 C shows a bimodal melting curve and 20% of the radioactivity is eluted at temperatures between 75 and 85 C. The hybrid formed at 65 C, however, is restricted in its melting to a small peak over 80 C. The intersyngenic 1–12 hybrid formed at 50 C shows only low-melting sequences. When formed at 65 C, most of the 1–12 hybrid melts at temperatures lower than 77.5 C. Thus there appear to be few well-matched sequences present. This is also true for the 1–11 hybrid, which is not shown.

In summary, the extent of the reactions involving radioactive unique sequences from syngen 1 is nearly complete when the unlabeled DNA is also derived from syngen 1. Moreover, intrasyngenic hybrids formed at 50 C and at 65 C exhibit a high degree of thermal stability. In contrast, the extent of the reaction is low when the source of unlabeled DNA is from other syngens, due to the effect of mismatching on the rate of reassociation, and the hybrids formed at 50 C have lowered thermal stability compared to intrasyngenic hybrids. The reaction is further reduced at 65 C and the intersyngenic hybrids formed have higher thermal stability.

The last column in Tables II and III gives an estimate of the percentage of mispaired bases in the hybrids. This percentage is estimated from the relationship described by Ullman and McCarthy (1973) in which the overall effect of mismatched base pairs on thermal stability is roughly a lowering of $T_{1/2}$ by 1.5 C for 1% altered base pairs. Intrasyngenic hybrids formed at 50 C have at most 1–2% mispaired bases, while those formed at 65 C have less than 1% mispaired bases. In contrast, intersyngenic hybrids formed at 50 C have between 2–3 and 13% of the bases mispaired, and those formed at 65 C have between 1 and 3% mispaired bases. In Table II, the syngens are ordered

in terms of their relationship to syngen 1 using the criterion of mispaired bases. On this basis, syngen 8 is most closely related to syngen 1 while syngen 12 is the least closely related. The other syngens would fall in between in terms of their relationship.

Extent of Reaction and Thermal Stabilities of Repeated Sequence Hybrids

Radioactive repeated sequences from 1/D were incubated at 50 C in solution with unlabeled repeated sequences from the same 16 sources as in the experiment with unique sequences. In the homologous reactions, 87.2% of the radioactive fragments bound to unlabeled DNA, and 87.3% of the unlabeled DNA renatured. The self-reaction was high: 24% of the radioactive fragments renatured (C_0t 0.25). The thermal stability of the radioactive hybrid with unlabeled DNA was slightly lower than that found for unique sequences in terms of $T_{1/2}$ ($= 77.5$ C). The thermal elution profile of this hybrid and that of the hybrid resulting from self-reaction are shown in Fig. 3. Note the wider spread in these curves compared to the curves for unique sequences in Fig. 1. In particular, there appear to be more sequences with low thermal stability, and a "shoulder" is observed in the region of 60–70 C for the hybrid with unlabeled DNA. In the curve corrected for self-reaction (Fig. 4) this shoulder is less apparent, and the $T_{1/2}$ is 78.5 C.

In the results to be described, all experimental reactions were corrected for self-reaction by subtracting the average percentage of self-reaction from
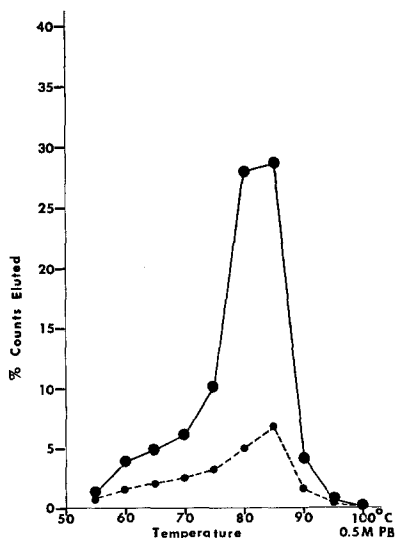

Fig. 3. Thermal elution profiles of renatured 1/D repeated sequence DNA from 50 C incubations. The distribution of radioactive counts between 55 and 95 C is graphed, each point being the average of two different columns. The curve represented by the solid line shows the thermal stability of labeled DNA which renatured with homologous unlabeled DNA at C_0t 250. The curve represented by the dashed line shows the thermal stability of labeled DNA renatured in the absence of unlabeled DNA to C_0t 0.25 (self-reaction).
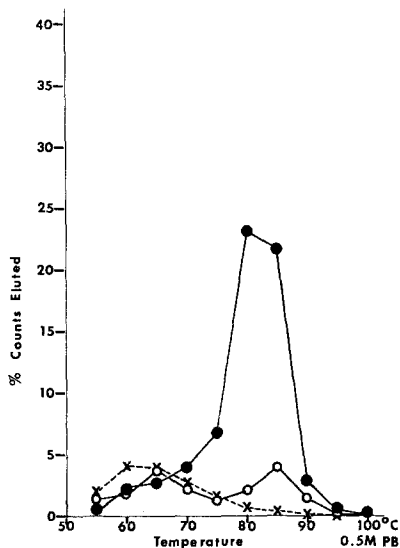
Fig. 4. Thermal elution profiles of homologous and heterologous repeated sequence DNA hybrids formed at 50 C. Labeled 1/D DNA in homologous hybrid (●) and in heterologous hybrids, syngen 8 (○) and syngen 12 (×). The distribution of radioactive counts between 55 and 95 C is graphed, each point being the average of two different columns and corrected for self-reaction.

all experiments (24%). The thermal elution data were also corrected for self-reaction.

Before describing the results, we should point out that the experiments with repeated sequences suffer due to two technical considerations: the high background "noise" of the self-reaction and contamination by unique sequences. The large effect of self-reaction is felt, in particular, in correcting the intersyngenic reactions by subtracting out the percentage of the reaction due to self-reaction. This procedure results in low values for the intersyngenic reactions; moreover, these percentages vary by less than 5% for different syngens. The second effect, that of contamination by unique sequences, needs also to be considered. At most, we estimate 25% contamination by unique sequences. Thus the majority of sequences which react would fall into the "repeated sequence" class; nevertheless, we might expect heterogeneity in this fraction and variable participation of sequences depending on their degree of reiteration, including participation of sequences present in one copy only. The results then, to be described, need to be evaluated with these considerations in mind, and any interpretations made are made with reservations.

Table IV summarizes the data of the hybridization experiments. The extent of the reaction for strains in syngen 1, D/1 and B, is similar to the homologous reaction, and the homologies to strain D are 97% and 95%, respectively. The thermal stabilities of the intrasyngenic hybrids are close to that of the homologous hybrid ($T_{1/2}$ = 77.5 C for D/1, $T_{1/2}$ = 77.0 C for B).

The extent of the intersyngenic reactions is low (12.7–17.3%) and less variable than those of unique sequences (7.7–31.9%). The average homology

Table IV. Extent of Reaction and Thermal Stabilities of Repeated Sequence DNA Hybrids Formed at 50 C

Source of unlabeled DNA (syngen/strain)	Percent renaturation		Percent homology	$T_{1/2}$ labeled DNA (C)	$\Delta T_{1/2}$ labeled DNA)C)	Percent mispaired bases
	Unlabeled DNA	1/D labeled DNA[a]				
1/D	87.3	63.2	100	78.5	—	—
1/D/1	79.6	56.0	97	77.5	−1.0	<1
1/B	83.5	57.7	95	77.0	−1.5	1
8/ALP-1	85.2	16.0	26	74.0	−4.5	3
6/UM-1091	84.0	16.3	27	74.0	−4.5	3
6/UM-1060	86.9	16.5	26	73.5	−5.0	3
10/EN-112	81.0	16.2	28	72.5	−6.0	4
3/UM-700	82.4	17.3	29	75.0	−3.5	2
5/UM-30	86.5	15.1	24	67.0	−11.5	8
9/TC-89	80.6	14.4	25	66.0	−12.5	8
2/UM-3	84.6	14.2	23	67.0	−11.5	8
2/UM-351	80.9	13.6	23	67.5	−11.0	7
7/UC-651	83.6	15.0	25	66.0	−12.5	8
4/IN-3	85.0	14.5	23	65.0	−13.5	9
11/AU-50-1	81.0	12.8	22	63.5	−15.0	10
12/AU-F_1-4	82.0	12.7	21	63.0	−15.5	10

[a] Corrected for average self-reaction of labeled DNA.

(25%) is, however, higher than the average homology (18%) for unique sequences. The range in homologies is also more limited for repeated sequences (21–29%) than for unique sequences (9–36%).

The thermal stabilities of the intersyngenic hybrids are lower than those of any of the intrasyngenic hybrids. The difference in $T_{1/2}$ ranges between 3.5 and 15.5 C. This is an average difference for all "repeated" sequences in each of the hybrids. How meaningful an average figure is could be disputed since there appear to be two components with different melting properties in most of the intersyngenic hybrids. Thus bimodal curves are observed for the thermal elution profiles of the intersyngenic hybrids, with the exception of those involving syngen 11 or 12. Two peaks occurred: one at 65 C and a second one at 85 C. The percentage of counts eluting in the region of 65 C was relatively constant in all hybrids, but the peak at 85 C decreased, with minor fluctuations, in the series from syngen 8 to syngen 12. The hybrid formed between syngens 1 and 8 gives roughly similar percentages of counts eluting in each peak (see Fig. 4). A similar, or somewhat smaller, percentage of counts eluting at 85 C is found for the other intersyngenic hybrids (except for the 1–3 hybrid). Only low-melting sequences are observed for the hybrid formed between syngens 1 and 11 and between 1 and 12 (see Fig. 4 where the

1–12 hybrid is illustrated). The 65 C peak appears to correspond to the "shoulder" observed in the homologous hybrid. A peak at 85 C is prominent in the hybrid formed by self-reaction (see Fig. 3).

The last column of Table IV gives a rough estimate of the percentage of mispaired bases in the hybrids. This estimate ignores the possibility of heterogeneity within the repeated sequence fraction or some contamination by unique sequences. The intrasyngenic hybrids have at most 1% mispaired bases. In contrast, the intersyngenic hybrids have between 3 and 10% of the bases mispaired. In Table IV the same ordering of the syngens in relation to syngen 1 is followed as for Table II. With minor departures, the two orders appear to be in agreement. The average percentage of mispaired bases is 6% for repeated sequences compared to 9% for unique sequences, suggesting that the repeated sequence fraction may be slightly more conserved in the different syngens with respect to syngen 1. However, what is being measured appears to be differences in the amount of high-melting sequences. If the 65 C peak represents sequences present in syngen 1 which are homologous to sequences in the other syngens, then we may be overestimating the degree to which bases are mispaired. Also, there is the problem of contamination by unique sequences. Both of these considerations, the possibility of a conserved low-melting fraction and contamination by unique sequences—depending on their effects—might imply a greater degree of conservatism than what is actually observed for some or all of the repeated sequences within different syngens.

DISCUSSION

The syngens of *Tetrahymena pyriformis* are morphologically similar. Even a detailed analysis of cortical organelles shows that the variations which occur, often in a complex way, are not syngen specific and thus cannot be used fruitfully to distinguish between syngens (Nanney and Chow, 1974). The syngens share many properties, including the same micronuclear chromosome number ($2N = 10$). They vary in the details of their breeding systems, in temperature tolerance, in nutritional requirements, and in their geographic distribution (Elliott, 1973). Syngens 1 to 8 are found in North America; syngen 2 is also found in Central America, syngen 3 in Europe, syngen 4 in England, continental Europe, and Asia, and syngen 6 in England, continental Europe, and Africa. Syngen 9 has been collected in Central and South America and certain Pacific islands, while the syngen 10 collections have been restricted to three southern counties in England. Syngens 11 and 12 are found in Australia. No mating or gene flow occurs between any of the syngens, except for syngens 6 and 8, which were shown to mate and yield some viable offspring (Orias, 1959). Some of the strains found in syngens 2 to 12 are now

senile and do not produce viable progeny even when crosses are made within the syngen; thus it is possible that more than 12 syngens are represented in this collection. However, strains can now be assigned to syngens on the basis of their isozyme patterns (Borden *et al.*, 1974), and most of the strains used in this study were also included in the isozyme study.

The base composition of DNA varies between 24 and 33% G+C for strains in different syngens (see Table I). This variability in base composition encompasses a range wider than that reported for vertebrates and is more similar to that observed in bacteria. Heterogeneity within the genome may occur in the form of satellites observed by density gradient centrifugation (Flavell and Jones, 1970). We have observed a satellite in syngen 1 DNA, but this represents a minor fraction of the DNA (Allen, Douthit, and Gibson, unpublished).

Renaturation experiments also reveal heterogeneity within the genome. Under optimal conditions for renaturation, about 15–20% of the genome behaves as repeated sequences. This was found earlier for syngens 1 and 7 (Allen and Gibson, 1972). It also appears to be true for the other syngens. The $T_{1/2}$ of renatured unique and repeated sequences is similar, but there are syngenic differences which reflect the G+C content. Additional heterogeneity within the repeated sequence fraction is suggested from the wider distributions of thermal stabilities compared to unique sequences, not only in syngen 1, but also in the other syngens.

Nucleotide sequences from different strains were compared by measuring the extent of the hybridization reaction and the thermal stabilities of the molecular hybrids. The effect of sequence divergence on the rate of reassociation of DNA is about a factor of 2 in the rate per 10 C reduction in $T_{1/2}$ of the reassociation strand pair after correcting for suboptimal renaturation temperature (Bonner *et al.*, 1973). Thus the extent of the reaction is reduced in proportion to the degree of mismatching. This means that calculations of homology lead to values which are underestimated, particularly for those DNAs which are most divergent. Greater stress then is placed in this discussion on the differences in thermal stabilities of the reassociated DNAs as indicators of sequence divergence.

Comparison of the nucleotide sequences of different strains within syngen 1 indicates minimal differences as measured both by the extent of the hybridization reaction and by the thermal stabilities of the molecular hybrids. This applies to unique as well as to repeated sequences. At most, the percentage of mismatched bases is 1–2%.

The nucleotide sequences of strains in other syngens appear to have diverged considerably from those in syngen 1. Measurements of the extent of the reaction, but, more particularly, of the thermal stabilities of the intersyngenic hybrid molecules, indicate large differences in unique sequences

when renaturation occurs under optimal renaturation conditions at 50 C. At 65 C, only those sequences which are well matched react and these appear to include only a small percentage of the molecules. Earlier work employing unfractionated DNA and the agar or filter techniques of hybridization also indicated extensive sequence divergence between the syngens (Allen and Gibson, 1967, 1973). The degree of base sequence divergence is correlated to some extent with differences in gross base composition. Syngen 1 has 25% G+C. Syngens with detectably higher G+C contents (4, 9, 10, and 11) tend to show the most extensive sequence divergence with respect to syngen 1, except for syngen 10. On the other hand, some syngens which have the same G+C content as syngen 1 also show large differences, such as syngens 2, 7, and 12.

The degree of base sequence divergence of the other syngens with respect to syngen 1 shows the following order: 8 and 6, followed by 10, 3, and 5, then 9, 2, 7, 4, 11, and 12. Earlier work with unfractionated DNA from syngens 1, 8, 9, 10, and 12 showed a similar ordering (Allen and Gibson, 1967, 1973). This ordering also shows some similarities to, and some differences from, the ordering obtained by comparisons of isozymes (Borden *et al.*, 1974). It is interesting that syngens 6 and 8, the only syngens which have shown intersyngenic mating, are ranked together with respect to syngen 1, yet they appeared less related to one another on the basis of the isozymes studied by Borden.

The difference in $T_{1/2}$ from the homologous hybrid can be used to calculate the percentage of mispaired bases, using data obtained under optimal renaturation conditions—in this study at 50 C. The average percentage of mispaired bases is 9% for the intersyngenic hybrids, with a range between 3 and 13%. This degree of mispairing is of the order of magnitude found by McConaughy and McCarthy (1970) for molecular hybrids formed between rodent genera. The thermal stabilities of unique sequence hybrids of mouse–rat, mouse–hamster, and mouse–guinea pig indicated 9, 11, and 12% mispaired bases. Studies on *Drosophila* species by Laird and McCarthy (1968) gave $\Delta T_{1/2}$ values of 3 C for the sibling species *melanogaster–simulans* and 13 C for the nonsibling species *melanogaster–funebris*. These values would lead to estimates of 2% mispairing for sibling species and 9% for nonsibling species using the relationship between $\Delta T_{1/2}$ and % mispairing suggested by Ullman and McCarthy (1973). Thus the average magnitude of the difference between the syngens of *T. pyriformis* is roughly similar to that observed between rodent genera and *Drosophila* nonsibling species. It is also similar to the difference observed between the enterobacteria *Escherichia coli* and *Salmonella typhimurium* or *E. coli* and *Aerobacter aerogenes* for DNA–DNA (Brenner *et al.*, 1969) or *trpA* mRNA–DNA (Li *et al.*, 1973).

The repeated sequence fraction appears to be more heterogeneous and,

D

on average, may be more conserved than the unique sequence fraction with respect to syngen 1. In other organisms, heterogeneity has also been found in the repeated sequence fraction (see Kohne, 1970; McCarthy and Farquhar, 1972). Some of the sequences are conserved, such as ribosomal DNA, and this also is the case for intersyngenic comparisons in *T. pyriformis* (Allen and Gibson, 1973). Other sequences are rapidly evolving such as the highly reiterated satellites found in mammalian DNA. In *T. pyriformis* DNA, the reality of satellites remains, as yet, controversial (see Allen and Gibson, 1973). If present, they are certainly minor contributions to the total DNA molecules. The majority of sequences in other organisms appear to show rates of divergence similar to those of unique sequences. Before any solid conclusions can be reached with respect to the situation in *T. pyriformis*, further characterization of the repeated sequence fraction will be necessary, particularly in terms of the cellular source of the component fractions. For example, one of the subfractions could be derived from mitochondrial DNA, which is known to be present in many copies (Flavell and Jones, 1970).

The DNA used in these experiments was isolated from whole cells. There are three known cellular sources for ciliate DNA: macronucleus, micronucleus, and mitochondria. Estimates of the percentage contributions from each of these sources have been made for syngen 1 of *T. pyriformis* (Allen and Gibson, 1972). The vast majority of the DNA molecules are derived from the macronucleus (89%), while the micronucleus would contribute about 7.5% and the mitochondria 3.5%. The macronucleus is rather specialized in the ciliated protozoa. Well-developed chromosomes are lacking, and in some species of hypotrichs—such as *Stylonychia*—genes may be lost from the macronucleus during its development (Prescott *et al.*, 1973). The situation with respect to gene loss and the macronucleus of *T. pyriformis* is, at present, not completely established, but it appears unlikely (Allen and Gibson, 1972; Gorovsky, 1973). However, since most of the sequences which are being compared between syngens are derived from the macronucleus, it becomes important to have this information in interpreting the biological meaning of the large differences in nucleotide sequences which we are seeing between syngens. If gene loss does occur in the macronucleus, then the differences we see may derive, in part, not from evolutionary processes, but perhaps from developmental events. In early studies using the agar technique, we could distinguish small differences between clones in the extent of the hybridization reaction (Allen and Gibson, 1967). The question of whether we are measuring differences in the nucleotide sequences which have arisen during the evolution of syngens may only be settled by examining micronuclear DNA. The micronucleus serves as the germline nucleus. It is possible that the rate of evolutionary change for micronuclear DNA molecules may be different from that for macronuclear DNA molecules.

The nucleus which functions to control the phenotypic expression of genes is the macronucleus. It is thus of interest to compare measures of evolutionary distance between syngens based on the divergence of nucleotide sequences, mainly of macronuclear origin, and those based on comparisons of isozymes. Borden *et al.* (1974) utilize Nei's method (1971) to translate their measures of isozyme differences into estimates of absolute time divergence. According to their calculations, the most closely related syngens would have diverged 3×10^5 years ago, and widely separated syngens, such as 1 and 11, would have diverged 1.7×10^6 years ago. The average time of divergence for all the syngens can be estimated as 1.23×10^6 years (Adams and Allen, 1974). The interesting feature of these values is that they can be compared to the average divergence times estimated by Nei (1971) for *Drosophila* sibling and nonsibling species. These are 5.7×10^5 and 1.4×10^6 years, respectively. The data of Laird and McCarthy (1968) on nucleotide sequence divergence for *Drosophila* sibling and nonsibling species in terms of percentage mispairing are 2 and 9%, respectively. The average percentage mispairing for unique sequences between the syngens of *T. pyriformis* is 9%. Thus there appears to be remarkable agreement between estimates based on isozyme comparisons and nucleotide sequence comparisons, and the average evolutionary distance between the syngens is similar to the distance between *Drosophila* nonsibling species. It is also interesting that in *Paramecium aurelia* the syngens do not appear to have diverged as extensively as those in *T. pyriformis*. The proportion of nucleotide sequences which react between syngens is greater and the thermal stabilities of the intersyngenic hybrid molecules are higher than those observed for *T. pyriformis* (Gibson, personal communication). The isozyme data give an average divergence time for the syngens of *P. aurelia* of 6.15×10^5 years (Adams and Allen, 1974). This is similar to the average time estimated for the divergence of sibling species of *Drosophila*.

Borden *et al.* (1974) make the point that the degree of divergence in the syngens of *T. pyriformis* is of an order of magnitude higher than that found previously for sibling species. Our data strongly support their contention that "vast molecular distances" separate the syngens. Either molecular changes are more rapid in these protozoa or the syngens of *T. pyriformis* are of ancient evolutionary origin. Rates of change in nucleotide sequences appear to be generation dependent (Kohne, 1970; Laird *et al.*, 1969). The number of sexual generations per year for *T. pyriformis* is roughly similar to the number for *Drosophila* in the laboratory, but how these numbers would compare in nature is totally unknown. If rates of evolution are generation dependent, and the numbers of generations are greater for *T. pyriformis* compared to *Drosophila*, this difference could account for the rapid divergence of the syngens.

Members of the taxonomic "basket" *T. pyriformis* are diverse in their evolutionary relationship. Certainly the syngens should at least be accorded the distinction of a status change to *species*.

REFERENCES

Adams, J. P., and Allen, S. L. (1974). Genetic polymorphism and differentiation in *Paramecium*. In *Proceedings of the Third International Isozyme Conference*, Academic Press, N.Y. (in press).

Allen, S. L., and Gibson, I. (1967). Genetic homologies and drift within populations of DNA molecules. *Science* **158**:523.

Allen, S. L., and Gibson, I. (1971). The purification of DNA from the genomes of *Paramecium aurelia* and *Tetrahymena pyriformis*. *J. Protozool.* **18**:518.

Allen, S. L., and Gibson, I. (1972). Genome amplification and gene expression in the ciliate macronucleus. *Biochem. Genet.* **6**:293.

Allen, S. L., and Gibson, I. (1973). Genetics of *Tetrahymena*. In Elliott, A. M. (ed.), *Biology of Tetrahymena*, Dowden, Hutchinson & Ross, Stroudsburg, Pa., pp. 307–373.

Allen, S. L., and Weremiuk, S. L. (1971). Intersyngenic variations in the esterases and acid phosphatases of *Tetrahymena pyriformis*. *Biochem. Genet.* **5**:119.

Bonner, T. I., Brenner, D. J., Neufeld, B. R., and Britten, R. J. (1973). Reduction in the rate of DNA reassociation by sequence divergence. *J. Mol. Biol.* **81**:123.

Borden, D., Miller, E. T., Whitt, G. S., and Nanney, D. L. (1974). Electrophoretic analysis of evolutionary relationships in *Tetrahymena*. *Evolution* (in press).

Bostock, C. (1971). Repetitious DNA. *Advan. Cell Biol.* **2**:153.

Brenner, D. J., Fanning, G. R., Johnson, K. E., Citarella, R. V., and Falkow, S. (1969). Polynucleotide sequence relationships among members of Enterobacteriaceae. *J. Bacteriol.* **98**:637.

Dryl, S. (1959). Antigenic transformation in *Paramecium aurelia* after homologous antiserum treatment during autogamy and conjugation. *J. Protozool. Suppl.* **6**:25.

Elliott, A. M. (1973). Life cycle and distribution of *Tetrahymena*. In Elliott, A. M. (ed.), *Biology of Tetrahymena*, Dowden, Hutchinson & Ross, Stroudsburg, Pa., pp. 259–286.

Flavell, R. A., and Jones, I. G. (1970). Mitochondrial deoxyribonucleic acid from *Tetrahymena pyriformis* and its kinetic complexity. *Biochem. J.* **116**:811.

Gorovsky, M. A. (1973). Macro- and micronuclei of *Tetrahymena pyriformis*: A model system for studying the structure and function of eukaryotic nuclei. *J. Protozool.* **20**:19.

Kohne, D. E. (1970). Evolution of higher-organism DNA. *Quart. Rev. Biophys.* **3**:327.

Laird, C. D., and McCarthy, B. J. (1968). Magnitude of interspecific nucleotide sequence variability in *Drosophila*. *Genetics* **60**:303.

Laird, C. D., McConaughy, B. L., and McCarthy, B. J. (1969). Rate of fixation of nucleotide substitutions in evolution. *Nature* **224**:149.

Li, S. L., Denney, R. M., and Yanofsky, C. (1973). Nucleotide sequence divergence in the α-chain-structural genes of tryptophan synthetase from *Escherichia coli*, *Salmonella typhimurium*, and *Aerobacter aerogenes*. *Proc. Natl. Acad. Sci.* **70**:1112.

Marmur, J., and Doty, P. (1961). Thermal renaturation of DNA. *J. Mol. Biol.* **3**:585.

McCarthy, B. J., and Farquhar, M. N. (1972). The rate of change of DNA in evolution. *Brookhaven Symp. Biol.* **23**:1.

McConaughy, B. L., and McCarthy, B. J. (1970). Related base sequences in the DNA of simple and complex organisms. VI. The extent of base sequence divergence among the DNA's of various rodents. *Biochem. Genet.* **4**:425.

Nanney, D. L. (1971). Cortical characteristics of strains of syngens 10, 11 and 12 of *Tetrahymena pyriformis*. *J. Protozool.* **18**:33.

Nanney, D. L., and Chow, M. (1974). Basal body homeostasis in *Tetrahymena*. *Am. Naturalist* **108**:125 (abst.).

Nei, M. (1971). Interspecific gene differences and evolutionary time estimated from electro-
    phoretic data on protein identity. *Am. Naturalist* **105**:385.
Orias, E. (1959). Mating interaction between varieties 6 and 8, *Tetrahymena pyriformis*.
    J. Protozool. Suppl. **6**:19.
Prescott, D. M., Murti, K. G., and Bostock, C. J. (1973). Genetic apparatus of *Stylonychia*
    sp. *Nature* **242**:576.
Sonneborn, T. M. (1957). Breeding systems, reproductive methods and species problems
    in protozoa. In Mayr, E. (ed.), *The Species Problem*, American Association for the
    Advancement of Science Symposium, Washington, D.C., pp. 155–324.
Ullman, J. S., and McCarthy, B. J. (1973). The relationship between mismatched base
    pairs and the thermal stability of DNA duplexes. II. Effects of deamination of cytosine.
    Biochim Biophys. Acta **294**:416.
Wille, J. (1972). Physiological control of DNA nucleotide sequence redundancy in the
    eukaryote, *Tetrahymena pyriformis. Biochem. Biophys. Res. Commun.* **46**:677.