JOURNAL OF BACTERIOLOGY, May 1987, p. 1861–1870
0021-9193/87/051861-10$02.00/0

Vol. 169, No. 5

Two Families of Repeated DNA Sequences in *Thiobacillus ferrooxidans*

JAMES R. YATES AND DAVID S. HOLMES*

General Electric Research and Development Center, Schenectady, New York 12301

Received 26 September 1986/Accepted 20 January 1987

The genome of *Thiobacillus ferrooxidans* ATCC 19859 is about 2.8×10^6 base pairs as determined by analysis of reassociation kinetics of sheared DNA. This is 70% of the size of the genome of *Escherichia coli*. About 6% of the genome of *T. ferrooxidans* consists of moderately repetitive DNA sequences that are repeated an average of 20 times per genome. Two distinct repeated sequences, designated family 1 and family 2, have been analyzed in more detail. Both families are approximately 1 kilobase in length and are repeated 20 to 30 times per genome. Preliminary evidence from restriction enzyme analysis, Southern blotting experiments, and thermal melting analysis indicates that members of both families are conserved and are interspersed with single-copy DNA. Six copies of one family are present on the 45-kilobase-pair plasmid of strain ATCC 19859.

Thiobacillus ferrooxidans is a gram-negative, acidophilic, chemoautotrophic bacterium capable of oxidizing ferrous iron and various reduced sulfur compounds (32). It is considered to be one of the most important organisms used for the recovery of metals in biohydrometallurgical operations, accounting for about 18% of the production of copper in the United States (30). It is also probably a major cause of acid mine drainage.

Despite its industrial and environmental significance, direct genetic analysis of *T. ferrooxidans* is only starting. Most strains of *T. ferrooxidans* have cryptic plasmids, and several of these have been cloned into *Escherichia coli* (13–15, 24, 25). Barros et al. (1) have cloned and expressed a glutamine synthetase gene (*glnA*) from *T. ferrooxidans* in *E. coli*. M. T. Roskey and C. F. Kulpa (Abstr. Annu. Meet. Am. Soc. Microbiol. 1984, H59, p. 101) have complemented histidine and proline auxotrophs of *E. coli* with DNA from *T. ferrooxidans*, and Holmes et al. (D. S. Holmes, D. Druger, and J. H. Lobos, Abstr. Annu. Meet. Am. Soc. Microbiol. 1985, Q27, p. 262) have used cloned probes to identify, in *T. ferrooxidans*, sequences homologous to the ribulose bisphosphate carboxylase gene of *Rhodospirillum rubrum*. Also, randomly cloned DNA sequences of *T. ferrooxidans* have been used as an aid in the identification and enumeration of the *thiobacilli* (35).

We report here the determination of the complexity of the genome of *T. ferrooxidans* and show that the genome contains a small percentage (about 6%) of moderately repeated DNA. Repeated DNA sequences are ubiquitous in higher organisms but are not commonly found in procaryotes. The few examples of repeated sequences in procaryotes come from widely different species. The archaebacterium *Halobacterium halobium* has at least 50 different families of repeated sequences, varying from 2 to 20 members per family (27). Repeated sequences are also known in the gram-negative bacteria *Caulobacter crescentus* (33) and *Pseudomonas cepacia* (17), as well as in several species of the gram-positive genus *Streptomyces* (8, 22, 26), including one example in which amplified repeated sequences represent 45% of the genome (10).

We have investigated and compared the reiteration frequency, size, and homology within and between two distinct families of moderately repeated DNA in the genome. These two families of repeated DNA sequences (designated family 1 and family 2) have been shown, by restriction enzyme analysis and Southern blotting experiments, to consist of conserved sequences, each about 1 kilobase (kb) in length and containing 20 to 30 copies per family. DNA-DNA reassociation kinetics was used to provide additional evidence that the reiteration frequency of family 1 repeated sequences is 20 to 30 per genome, and we confirm by thermal melting analysis that the sequence of family 1 members is conserved.

(Part of this work was conducted by J. R. Yates in partial fulfillment of the Ph.D. degree requirements at the State University of New York at Albany.)

MATERIALS AND METHODS

Bacterial strains and media. *T. ferrooxidans* ATCC 19859 was grown in a minimal-salts medium (9K salts) supplemented with ferrous iron (31). The strain was purified three times on solid media after being received from the American Type Culture Collection, Rockville, Md. *E. coli* C600 was grown in Luria broth at 37°C, supplemented with 100 μg of ampicillin per ml if recombinant plasmids were present.

Growth on solid media. Quantitative growth of *T. ferrooxidans* on solid media was carried out by modifications of published procedures (18, 20). A 35-ml portion of 10× modified 9K salts was added to 265 ml of H_2O in a 1-liter flask and adjusted to pH 2.5 with 10 N H_2SO_4 (solution A). $FeSO_4 \cdot 7H_2O$ (11g) was added to 75 ml of H_2O and adjusted to pH 2.5 with 10 N H_2SO_4 (solution B). Agarose (2 g) (type I, Low EEO; Sigma Chemical Co.) was added to 125 ml of H_2O (solution C). Solutions A and C were sterilized at 121°C and cooled to 60°C before being mixed with solution B, which had been filter sterilized. The 10× modified 9K salts was made as follows. In 3 liters of H_2O, the following were dissolved in order as listed: 150 g $(NH_4)_2SO_4$, 5 g of KCl, 2.5 g of K_2HPO_4, 25 g of $MgSO_4 \cdot 7H_2O$, and 0.72 g of $Ca(NO_3)_2 \cdot 4H_2O$; 50 ml of 10 N H_2SO_4 was added, and the mixture was stored at 4°C.

DNA preparation. High-molecular-weight genomic DNA from *T. ferrooxidans* was prepared by a modification of the Marmur procedure (19). Cells were grown to late log phase (ca. 2×10^8 cells per ml), and particulate matter was allowed to settle out. The supernatant fluid was filtered through a

* Corresponding author.

0.45-μm-pore-size nylon membrane (Zeta-probe; Bio-Rad Laboratories) backed by cheesecloth in a Buchner funnel. Cells were suspended in 30 ml of 9K salts lacking ferrous iron. This solution was allowed to stand at room temperature until the larger aggregates of ferric precipitate had settled out. Cells were collected from the partially clarified liquid by centrifugation at 1,000 × *g*. The cell pellet, together with any residual iron precipitate, was suspended in 1 ml of 9K salts, and the cells were collected by centrifugation in an Eppendorf centrifuge. The cell pellet was carefully suspended in 9K salts, leaving behind the residual iron precipitate. Cells were again collected by centrifugation. This step was repeated until no iron precipitate was apparent. Cells were suspended in 0.15 M NaCl–0.1 M EDTA (pH 8.0) and frozen at −20°C. The frozen cells were rapidly thawed at 55°C, and 1.5 mg of lysozyme (Sigma) per ml was added. This mixture was incubated at 37°C for 15 min. Sarkosyl (CIBA-GEIGY Corp.) was added to a final concentration of 1% (wt/vol), and the mixture was incubated at 55°C for 15 min. Proteinase K (Beckman Instruments, Inc.) was added (5 mg/ml), and the incubation was continued at 55°C for 1 h. The cell lysate was extracted with phenol, and total DNA was purified by CsCl-ethidium bromide buoyant density centrifugation (5).

Preparative amounts of DNA from recombinant plasmids were isolated from *E. coli* C600 by the procedure of Eisenberg and Holmes (7).

The approximately 45-kb plasmid of *T. ferrooxidans* ATCC 19859 was isolated as described previously (14).

Generation of recombinant plasmids. DNA from *T. ferrooxidans* ATCC 19859 was subjected to partial digestion with *Mbo*I, and the resulting fragments were separated on a 5 to 20% sucrose gradient. Fractions containing 5- to 10-kb DNA fragments were recovered, and the DNA was collected by precipitation with 1/2 volume of saturated ammonium acetate and 2 volumes of 100% ethanol at −18°C for 1 h. Precipitated DNA was collected by centrifugation and suspended in water. DNA fragments were ligated into *Bam*HI-cut, dephosphorylated pBR322 (New England BioLabs). Clones containing recombinant plasmids were screened for insert size by a rapid-boiling procedure (12).

Southern blots and nick translations. DNA was digested with various restriction enzymes (New England BioLabs; Bethesda Research Laboratories, Inc.) as specified by the manufacturer. The resulting DNA fragments were separated on 0.8 or 1.5% agarose gels (Sigma) in 0.04 Tris (pH 8.0)–0.005 M sodium acetate–0.001 M EDTA. Southern blots (29) were performed under the conditions recommended by the supplier (GeneScreen; New England Nuclear Corp.). The probe was hybridized at T_m − 25°C. Nick translations were performed as previously described (34).

Reassociation kinetics. Genomic DNA was sonicated in 0.05 M phosphate buffer (pH 6.8) with a Branson Sonifier equipped with a microtip probe. Sonication was performed in ice for a total of 5 min and consisted of 30 pulses of 10 s each with 30 s between pulses for cooling. Sonicated DNA was found to have a mean average size of about 500 base pairs (bp) as determined by agarose gel electrophoresis. DNA was melted at 100°C in a microcuvette capped with paraffin oil. Hyperchromicity was monitored by increase in A_{260} and was found to be 32% for *T. ferrooxidans* and 33% for *E. coli*. Reassociation of denatured DNA was monitored at 260 nm in a Perkin-Elmer Lambda 5 spectrophotometer equipped with thermally regulated jacketed cuvettes and an automatic recorder. Reassociation was carried out at T_m − 25°C (where T_m was defined as the temperature at which half the bases of double-stranded DNA had been denatured). At the termination of the reaction, the hyperchromicity change was again determined by heating the sample to 100°C.

Hydroxyapatite chromatography. Hydroxyapatite (DNA grade; Bio-Rad Laboratories) was hydrated in 0.12 M sodium phosphate (pH 6.8), washed extensively and alternately with 0.96 and 0.01 M sodium phosphate (pH 6.8), and finally equilibrated in 0.12 M sodium phosphate (pH 6.8). Sonicated, nonradioactive *T. ferrooxidans* DNA was denatured by boiling in the presence of trace amounts of [^{32}P]DNA derived from recombinant plasmids. The [^{32}P]DNA was present at approximately 1/1,000 of the concentration of nonradioactive DNA. Reassociation of [^{32}P]DNA and nonradioactive DNA took place at T_m − 25°C in sealed capillaries. The reaction was terminated by plunging the reaction vessel into ice, and the DNA was loaded immediately onto a hydroxyapatite column equilibrated at T_m − 25°C. Single-stranded DNA was eluted with 6 column volumes of 0.12 M phosphate buffer (pH 6.8), and double-stranded DNA was eluted with 6 column volumes of 0.48 phosphate buffer (pH 6.8). Residual DNA (<5% of the total) was eluted with 0.96 M phosphate buffer (pH 6.8).

DNA melting profiles. *T. ferrooxidans* genomic DNA was sheared to a mean average size of 10 kb by repeated passage through a 26-gauge needle. The DNA was denatured in the presence of trace (1/1,000) ^{32}P-labeled plasmid fragments as described above. DNA was reassociated in sealed capillaries to a C_0T of either 1 (for repetitive DNA) or 10 (for single copy), and double-stranded DNA was collected by hydroxyapatite chromatography, where C_0T is the concentration of nucleotides (in moles per liter) multiplied by the time (in seconds) (4). DNA was removed from the column by raising the temperature in 2°C increments and holding the temperature for 5 min at each increment. Single-stranded DNA was eluted with 6 column volumes of 0.12 M phosphate buffer (pH 6.8).

RESULTS

Complexity of the *T. ferrooxidans* genome. DNA reassociation was used to determine the complexity of the *T. ferrooxidans* genome and to quantitate the proportion of the genome containing repetitive sequences. Sonicated total *T. ferrooxidans* DNA was denatured and allowed to reassociate at T_m − 25°C in phosphate buffer (pH 6.8). Reassociation was monitored by following the hyperchromicity change at 260 nm (Fig. 1). In a separate experiment, the reassociation of sonicated *E. coli* DNA was monitored in an identical manner (Fig. 1). A single second-order reaction curve was fitted to the *E. coli* data by using a computer program devised for the analysis of DNA reassociation data (23). The $C_0T_{1/2}$ of the *E. coli* was found to be 6.8 under the conditions used, which is in reasonable agreement with published values (3, 4). Since the second-order rate constant (k) is inversely proportional to the $C_0T_{1/2}$, k for the *E. coli* reaction can be calculated to be 0.15. The program was also used to find the best-fitting curve to the *T. ferrooxidans* data. The results indicate that approximately 6% of the *T. ferrooxidans* DNA has a $C_0T_{1/2}$ of 0.22 and that the remainder of the DNA, which is presumed to consist of single-copy sequences, has a $C_0T_{1/2}$ of 5. Reassociation of the *T. ferrooxidans* DNA went to 91% completion, and the two-component curve fitted the data with a root mean square of 0.003. The various computer-estimated parameters that describe the reassociation data of *T. ferrooxidans* are listed in Table 1.

A comparison between the respective DNA reassociation rate constants of *T. ferrooxidans* and *E. coli* permits the

TABLE 1. Parameters describing the reassociation kinetics of *T. ferrooxidans* DNA[a]

Component	$C_0T_{1/2}$[b]		K		% of genome[c]	Chemical complexity (10^6 bp)[d]	Kinetic complexity (10^6 bp)[e]
	Observed	Pure	Observed	Pure			
Single copy	5	4.4	0.2	0.23	94	2.6	2.6
Repetitive	0.22	0.013	4.6	75.8	6	0.16	0.008

[a] The parameters were derived from the data shown in Fig. 1.
[b] $C_0T_{1/2}$ is expressed in moles per liter multiplied by seconds.
[c] Normalized for 100% reaction.
[d] Chemical complexity of total DNA was 2.8×10^6 bp.
[e] Kinetic complexity of total DNA was 2.6×10^6 bp.

genome size of *T. ferrooxidans* to be estimated. The genome size is the total number of nucleotides (termed chemical complexity) in the haploid genome. For genomes with repetitive DNA, it is the sum of the chemical complexities of both the single-copy and repetitive DNA sequences.

The chemical complexity of the single-copy DNA can be calculated from the following equation: $N_{sc} = N_{E.\ coli}(K_{E.\ coli}/K_{sc})$, where N_{sc} is the number of base pairs of single-copy DNA of *T. ferrooxidans*; $N_{E.\ coli}$ is the number of base pairs in the *E. coli* genome (4×10^6 bp [16]); K_{sc} is the rate constant for the reassociation of *T.* ferrooxidans single-copy DNA (0.23; Table 1); and $K_{E.\ coli}$ is the rate constant (0.15) for the reassociation of *E. coli* DNA. Thus $N_{sc} = 2.6 \times 10^6$.

Six percent of the chemical complexity is represented by repetitive sequences (Table 1). Therefore, the chemical complexity of the repetitive DNA (N_R) is $0.06 \times N_{sc}$, which is 1.6×10^5. The genome size is the sum of these values, N_{sc} and N_R, which is approximately 2.8×10^6 bp.

Although the chemical complexity is a measure of the total number of base pairs in the genome, the kinetic complexity represents the total informational content. If an organism has only single-copy DNA sequences, the chemical and kinetic complexities are the same. However, in an organism with repetitive DNA sequences, the repetitive component has less information than its chemical complexity because the DNA sequences that constitute it are repeated. The kinetic complexity of the repetitive DNA can be derived from the following equation: N_R (kinetic) = $N_{E.\ coli}(K_{E.\ coli}/K_{\text{pure repetitive}})$. Therefore, N_R (kinetic) = 8×10^3 bp. This means that there are only 8×10^3 bp of information in the repetitive DNA component. However, since there are 1.6×10^5 bp of chemical complexity in the repetitive component, it can be concluded that each family of repetitive sequences is repeated an average of 20 times.

The total kinetic complexity of the *T. ferrooxidans* genome is the sum of the kinetic complexities of the single-copy (2.6×10^6 bp) and repetitive (8×10^3 bp) components, which is therefore about 2.6×10^6 bp. In summary, the chemical complexity or genome size of *T. ferrooxidans* is 2.8×10^6 bp or 70% the size of the *E. coli* genome. On the other hand, the kinetic complexity or total genetic information is 2.6×10^6 bp or 66% that of *E. coli*.

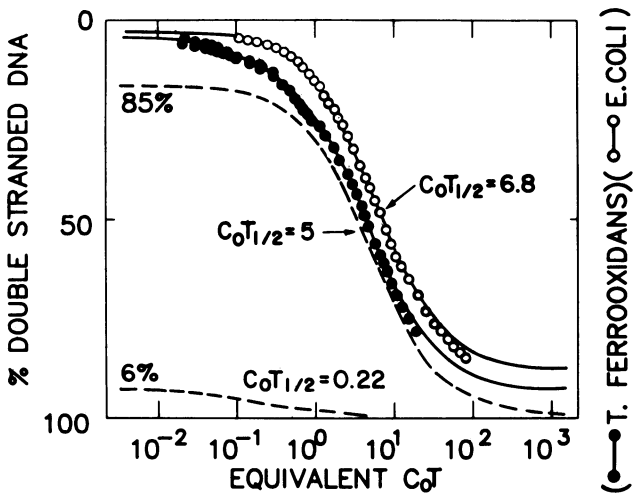

FIG. 1. C_0T curve of sheared, reassociated *T. ferrooxidans* DNA (●) or *E. coli* DNA (○). ——, Computer-fitted least-squares curves that best describe the reassociation data; – – –, the two components of the computer-fitted curve that best describe the reassociation of the *T. ferrooxidans* DNA. C_0T is measured in moles of nucleotides per liter multiplied by seconds. The data were derived from repeated experiments carried out at different phosphate buffer concentrations as described in Materials and Methods. The data have been normalized to the standard phosphate buffer concentration (0.18 M Na^+), and this gives rise to the term "equivalent C_0T" (4).

Identification of two families of repeated DNA sequences. To investigate further the properties of the repetitive DNA sequences that had been detected by C_0T analysis, a random genomic library of *T. ferrooxidans* DNA was prepared in the plasmid pBR322. A preliminary screening of 46 randomly selected recombinant plasmids revealed the presence of 6 plasmids that hybridized to multiple bands in Southern blots of *T. ferrooxidans* DNA.

Four of these plasmids were chosen for further study. Two of them, pTf8 and pTf11, gave rise to the complex pattern of hybridization shown in Fig. 2. Both probes hybridize to an almost identical set of *Ava*I, *Stu*I, or *Eco*RI restriction fragments of *T. ferrooxidans* DNA (Fig. 2A, lanes G). We suggest that this pattern of hybridization is due to the occurrence of a segment of DNA in both probes that is repeated many times per genome. We designate this repeated DNA as family 1. However, the patterns of hybridization produced by pTf8 and pTf11 were not identical. It is probable that this results from the hybridization of contiguous single-copy sequences. For example, there was a 1.14-kb band in the *Ava*I digest that hybridized effectively to pTf8 but not to pTf11. This band represents a unique *Ava*I fragment contiguous with the repeated DNA in pTf8 (Fig. 2B). This band can be uniquely identified because it corresponds to an internal fragment of the recombinant plasmid pTf8.

Both pTf8 and pTf11 hybridize to several restriction endonuclease fragments of the native *T. ferrooxidans* plasmid (Fig. 2A, lanes P). Therefore, the native plasmid also contains copies of the repeated sequences. However, it can be shown that the cloned DNA in pTf8 and pTf11 was originally derived from the chromosome and not the plasmid of *T. ferrooxidans*, because all the fragments produced by

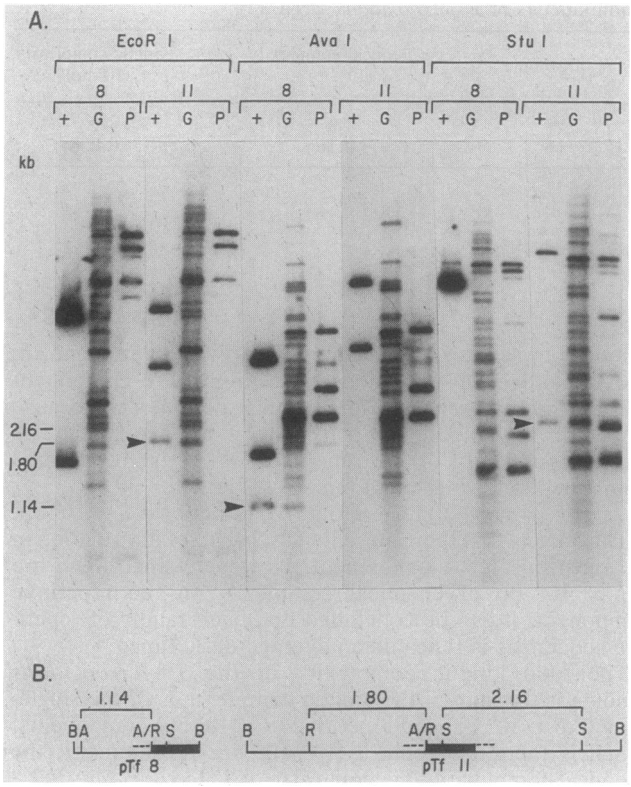

FIG. 2. (A) Autoradiogram of a Southern blot demonstrating the similarity of hybridization pattern of two probes (pTf8 and pTf11), containing the family 1 repeat. Lanes: +, positive control (pTf8 or pTf11); G, genomic DNA from *T. ferrooxidans*; P, native *T. ferrooxidans* plasmid. DNAs were cut with *Eco*RI, *Ava*I, or *Stu*I and probed with either pTf8 or pTf11 as indicated. Arrowheads indicate DNA bands corresponding to fragments produced by cleavage at sites within the inserted DNA of pTf8 or pTf11. (B) Restriction nuclease maps of the *T. ferrooxidans* DNA inserts of pTf8 and pTf11 showing the cleavage sites of the enzymes used in panel A. Abbreviations: B, *Bam*HI-*Mbo*I junction; R, *Eco*RI: S, *Stu*I; A, *Ava*I. Numbers correspond to lengths in kilobases.

digestion at sites within the inserted DNA align with hybridized bands present in the lanes containing chromosomal and not plasmid DNA (marked by arrowheads in Fig. 2A).

Figure 3A shows a Southern blot of *T. ferrooxidans* chromosomal and plasmid DNA probed with recombinant plasmids pTf32 and pTf34. We suggest that similar repeated DNA sequences are being detected by both these probes. However, because the patterns of hybridization are different from those detected with family 1 repeated sequences, we propose that they are due to another distinct group of repeated sequences, designated family 2. pTf32 and pTf34 do not effectively hybridize to the native plasmid under the hybridization conditions used (Fig. 3A). Weak hybridization to the native plasmid, just discernable in Fig. 3A, lanes P, could indicate the presence of either a sequence(s) on the native plasmid with limited homology to the probes or a relatively small region of more precise homology. The internal restriction fragments of the recombinant plasmids pTf32 and pTf34 align with hybridized genomic restriction fragments (arrowheads in Fig. 3A), indicating that the cloned sequences were derived originally from the genome and not from the plasmid of *T. ferrooxidans*.

Restriction analysis of family 1 and 2 probes. The pairs of recombinant plasmids that produced similar patterns on Southern blots contain repeated elements flanked by different, unique sequences. Restriction nuclease site maps of recombinant plasmids pTf8 and pTf11 (family 1) and pTf32 and pTf34 (family 2) are shown in Fig. 4. pTf8 and pTf11 contain 2.0 and 5.6 kb of *T. ferrooxidans* DNA, respectively. A region of about 600 bp bounded by the *Eco*RI and *Ava*I sites on one side and the*Bgl*II site on the other is present in both inserts. Restriction sites on one side of the repeated element (e.g., *Sph*I, *Pst*I, and *Bgl*II sites in pTf8) are not conserved. The conservation of restriction sites in the other direction from the repeat (downward in Fig. 4A) cannot be determined because the recombinant insert of pTf8 terminates in this region. Cross-hybridization experiments between restriction endonuclease fragments isolated from pTf8 and pTf11 showed that there was strong homology between the fragments bounded by *Bgl*II-*Eco*RI sites (Fig. 4A, solid line), as would be expected from the observed restriction site conservation (data not shown). Some cross-hybridization was also observed extending upward from the common element to another *Bgl*II site (Fig. 4A, dotted line). The extent of homology in the opposite direction could not be determined because of the truncation of the insert in pBR322. The minimal size of the homologous DNA shared by the two probes is about 600 bp (Fig. 4A, solid line).

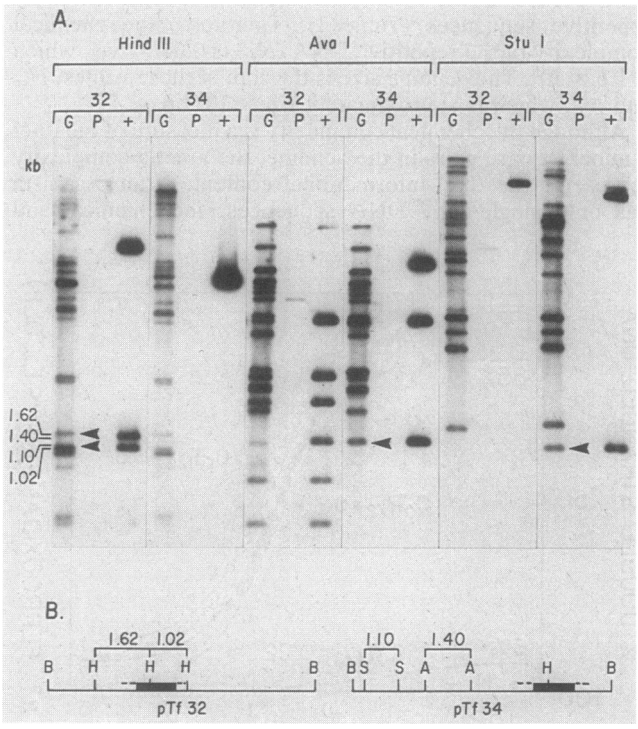

FIG. 3. (A) Autoradiogram of a Southern blot demonstrating the similarity of the hybridization patterns of two probes (pTf32 and pTf34) containing the family 2 repeat. Lanes: +, positive control (pTf32 or pTf34); G, genomic DNA from *T. ferrooxidans*; P, native *T. ferrooxidans* plasmid. DNAs were cut with *Hin*dIII, *Ava*I, or *Stu*I and probed with either pTf32 or pTf34 as indicated. Arrowheads indicate DNA bands corresponding to fragments produced by cleavage at sites within the inserted DNA of pTf32 or pTf34. (B) Restriction nuclease maps of the *T. ferrooxidans* DNA inserts of pTf32 or pTf34 showing the cleavage sites of the enzymes used in panel A. Abbreviations: B, *Bam*HI-*Mbo*I junction; H, *Hin*dIII; S, *Stu*I; A, *Ava*I. The numbers correspond to lengths in kilobases.

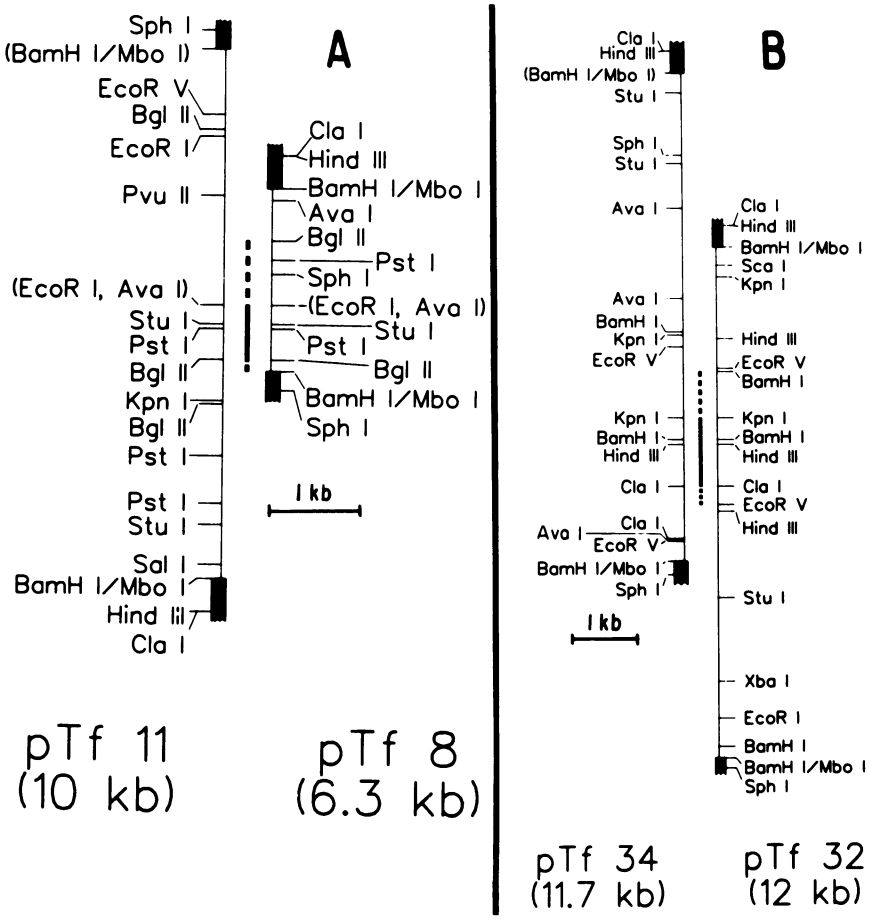

FIG. 4. Restriction enzyme site map of the *T. ferrooxidans* DNA components (—) of (A) pTf8 and pTf11 and (B) pTf32 and pTf34. A limited region of pBR322 is shown for orientation (▬). The solid line between the respective pairs of plasmids indicates the region containing conserved restriction sites and the region which cross-hybridizes strongly. The dotted line indicates a region that although lacking restriction site conservation, may contain a certain amount of homologous or related DNA according to cross-hybridization experiments.

Invoking similar arguments of restriction site conversation, we determined that the minimum size of the homologous DNA shared between pTf32 and pTf34 is 1 kb (Fig. 4B, solid line) and that the maximum size is 2 kb (Fig. 4B, dotted line). This observation has also been confirmed by comparative hybridization analysis (data not shown).

Further analysis of the family 1 repeated element. It is likely that the homologous region shared by pTf8 and pTf11 is responsible for the hybridization to the multiple bands in the genomic Southern blots shown in Fig. 1. To verify this, pTf11 was cleaved by restriction enzymes to produce the three segments indicated in Fig. 5. These three segments were isolated from an agarose gel, nick-translated, and individually used as probes against the genome of *T. ferrooxidans* in a Southern blot (Fig. 5). Segment A, corresponding to the region shown above to be conserved between pTf11 and pTf8, hybridized to multiple fragments in the genome, confirming that it contains the repeated DNA segment (Fig. 5, lanes 2 to 4). Segment A hybridized to approximately 24 *Sal*I-*Eco*RV restriction fragments of the *T. ferrooxidans* genome (lane 2). This provides a minimum estimate of the number of times this segment is repeated per genome, as discussed above. It is in good agreement with the estimate of 25 repeats derived from the data shown in Fig. 1. Segments B and C, on the other hand, hybridized predominantly to single DNA fragments in the genome (Fig. 5, lanes 5 to 10), indicating that these probes correspond to unique sequences in the genome. Since all three segments hybridized to the same genomic 4.8-kb *Sal*I-*Eco*RV fragment (lanes 2, 5, and 8) that had been cloned in pTf11 (Fig. 5, lane 1) the three segments must be located together on the genome. Therefore, the insert in pTf11 did not arise merely by random insertion of several unrelated *Mbo*I fragments. Furthermore, since two of the segments hybridized to single DNA fragments in the Southern blot, the hybridization of the third segment to multiple bands cannot be attributed merely to incomplete restriction enzyme digestion of the genomic DNA. There was no hybridization between members of family 1 and family 2 repeated sequences (data not shown).

An independent determination of the reiteration frequency of family 1 can be obtained from measurements of the rate of reassociation of a radioactively labeled isolated repeated sequence belonging to family 1 in the presence of a vast excess of unlabeled total *T. ferrooxidans* DNA. To investigate this point, segment A (Fig. 5) was isolated, nick-translated, and hybridized to excess unlabeled total *T. ferrooxidans* DNA. The rate of reassociation of the ^{32}P-labeled fragment was monitored by hydroxyapatite chromatography. In a separate experiment, segment B was isolated, nick-translated, and hybridized to excess *T. ferrooxidans*

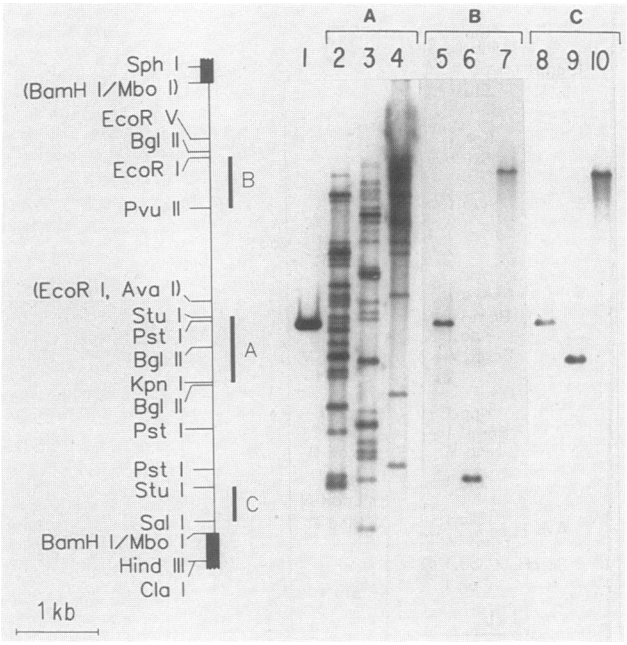

FIG. 5. Southern blot of *T. ferrooxidans* genomic DNA probed with single-copy and repeated DNA. *T. ferrooxidans* genomic DNA was cut with *Sal*I − *Eco*RV (lanes 2, 5, and 8) *Eco*RI (lanes 3, 6, and 9), or *Hin*dIII (lanes 4, 7, and 10) and probed with either segment A, B, or C derived from pTf11 (see accompanying map of pTf11) as indicated. Lane 1 is the positive control (pTf11 cut with *Sal*I − *Eco*RV). These three probes were obtained by nick translation of fragments isolated from agarose gels by the procedure of Girvitz et al. (9).

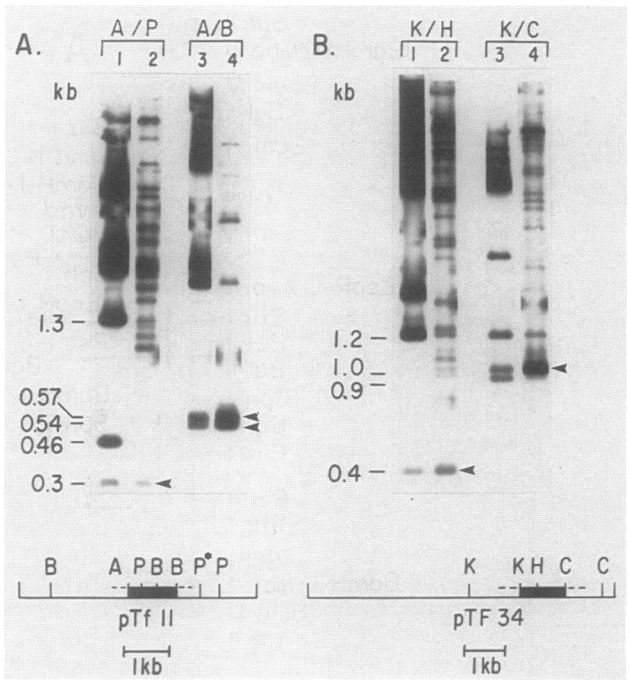

FIG. 7. Autoradiograms of Southern blots showing conservation of restriction sites in the repeated DNA of (A) family 1 or (B) family 2. (A) Lanes: 1 and 3, pTf11; 2 and 4, *T. ferrooxidans* genomic DNA. DNAs were cut with *Ava*I-*Pst*I or with *Ava*I-*Bgl*II as indicated. Abbreviations: A, *Ava*I; B, *Bgl*II; P, *Pst*I. The asterisk in the accompanying map indicates a *Pst*I site in the single-copy DNA that is discussed in the text. The probe was nick-translated pTf11. (B) Lanes: 1 and 3, pTf34; 2 and 4, *T. ferrooxidans* genomic DNA. DNAs were cut with *Kpn*I-*Hin*dIII or with *Kpn*I-*Cla*I as indicated. Abbreviations: K, *Kpn*I; H, *Hin*dIII; C, *Cla*I. The probe was nick-translated pTf34. Arrowheads indicate bands which correspond to hybridization of fragments produced by digestion at sites within the repeats.

DNA. Segment B corresponds to single-copy DNA, and its rate of reassociation provides evidence for the complexity of the single-copy DNA in addition to that derived from the C_0T analysis.

The results of the respective reassociation experiments are shown in Fig. 6. Second-order reaction curves have been

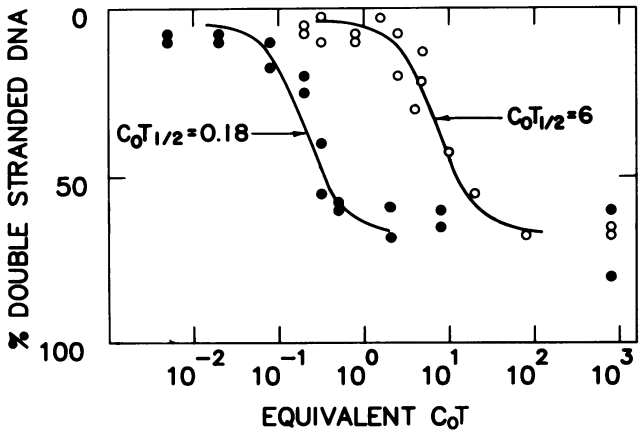

FIG. 6. C_0T curve of purified repetitive DNA (●; segment A of the recombinant plasmid pTf11) or purified single-copy DNA (○; segment B of the recombinant plasmid pTf11). Segments A and B were separately labeled with ^{32}P by nick translation, denatured, and reassociated in the presence of an excess of labeled total *T. ferrooxidans* DNA. The extent of the reassociation of the labeled DNAs was monitored by hydroxyapatite chromatography.

fitted to the data by using a least-squares computer program (23). The family 1 repeat has a $C_0T_{1/2}$ of 0.18, compared with a $C_0T_{1/2}$ of 6 for the single-copy fragment. These values have been normalized for the twofold-higher rate observed when the extent of DNA reassociation was determined by hydroxyapatite chromatography compared with optical methods. This indicates that the repetitive fragment is present about 33 times per genome. This is in good agreement with the previous estimates of a reiteration frequency of 25 times, derived from inspection of the Southern blot data (Fig. 2). It also indicates that this repetitive fragment is a representative of one of the families of repeated sequence that constitute about 6% of the genome and that exhibit an average reiteration frequency of about 20 times. The $C_0T_{1/2}$ of the single copy fragment is about 6, which is in reasonable agreement with the estimate of $C_0T_{1/2}$ = 5 for single-copy DNA derived from the DNA reassociation data shown in Fig. 1.

In control experiments, DNA segment A or B was allowed to reassociate as described above, but in the absence of excess unlabeled DNA, to determine the extent of self-association. This never exceeded 12%, and values due to self-reassociation have been subtracted from the experimental data. In duplicate experiments, neither of the two ^{32}P-labeled DNA fragments associated to more than 60 to 70% of the complete reaction. The reason for this is not known.

Conservation of restriction sites within familes 1 and 2. The mapping data presented in Fig. 4 permitted experiments to be devised that can determine the extent of restriction site conservation in all of the members of the family 1 and 2 repeated sequences. If two restriction sites are conserved in all the repeated sequences of a family, enzymatic digestion should result in a band containing more target DNA than would be produced from equivalent cutting within a single-copy sequence. Any repeated sequences in the genome that do not have one (or both) of these sites will produce a band of different molecular weight and perhaps of lower intensity.

Genomic DNA was cut with *Ava*I and *Pst*I (Fig. 7A, lane 2) or with *Ava*I and *Bgl*II (Fig. 7A, lane 4). All three of these enzymes cut within the member of family I repeated sequence that is present in pTf11, as can be seen in the accompanying map (Fig. 7A). Relatively intense bands are visible at 0.3 kb (lane 2) and at 0.54 and 0.57 kb (lane 4). These bands are marked by arrowheads in Fig. 7A and correspond to the sizes expected for internal fragments of the repeat. Since they are internal fragments, they are also found in the equivalent enzyme digestions of pTf11 (Fig. 7A, lanes 1 and 3). The lanes with the pTf11 have been deliberately overloaded to permit the low-molecular-weight bands to be visualized.

Since these internal fragments are relatively small, it is necessary for comparisons of hybridization intensities to have an internal control that is a single-copy sequence of approximately the same size as the fragment(s) containing the repeated sequence. Such a control is provided by the 0.46-kb *Pst*I fragment that resides outside the repeated sequence of pTf11. The position of this 0.46-kb fragment can be seen in the map presented in Fig. 7A. Although the 0.46-kb band is visible in the pTf11 control (lane 1), the corresponding band cannot be seen in the genomic digest (lane 2) (this band is detectable in longer exposures). This contrasts with the visibility of the smaller (0.3-kb) band that corresponds to an internal fragment of the repeat. Therefore, the 0.3-kb band is present in many more copies than the 0.46-kb unique band, and we conclude that the 0.3-kb band is highly conserved with family 1. The *Pst*I site marked by an asterisk in the map (Fig. 7A) lies outside the repeat, and, therefore, multiple bands are expected in the genomic blot that correspond to fragments generated by cleavage with *Pst*I once within the repeated sequence and once at a random site in the single-copy DNA outside the repeat. This presumably accounts for the multiple bands observed in lane 2. In contrast, there are very few hybridized bands in the genomic blot after cleavage with *Ava*I and *Bgl*II. Our interpretation is that one *Bgl*II site resides within the repeat and the other *Bgl*II site and the *Ava*I site virtually mark the ends of the repeat (map in Fig. 7A). Therefore, an *Ava*I-*Bgl*II double digest produces two bands at 0.54 and 0.57 kb corresponding to the conserved sites of the repeat (arrowheads in lane 4). Two of the three remaining bands presumably correspond to flanking fragments, one of which is delineated by the *Ava*I site at the end of the repeat and the *Bgl*II site to the left of the repeat and the other of which is delineated by the *Bgl*II site at the end of the repeat and the next rightward *Bgl*II or *Ava*I site which is not present in the cloned insert of pTfll. A fourth band visible in lane 4 may result from a restriction site polymorphism either in the repeat or in the adjacent single-copy DNA. Thus, we conclude that all the repeats of family 1 (with one possible exception) contain conserved *Ava*I and *Bgl*II sites.

A similar experiment involving the use of pTf34 as a probe to examine the conservation of restriction sites in the family

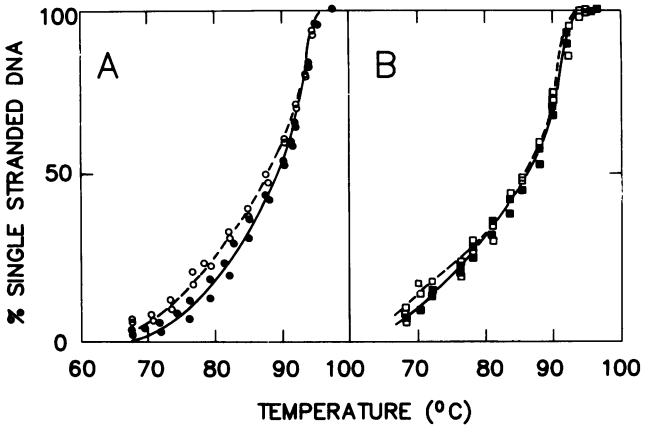

FIG. 8. Thermal melting profiles (A) repetitive DNA (segment A of pTf11) reassociated either to itself (●) or to total *T. ferrooxidans* DNA (○) or (B) single-copy DNA (segment B of pTf11) reassociated either to itself (■) or to total *T. ferrooxidans* DNA (□).

2 repeated sequences was conducted (Fig. 7B). A low-molecular-weight band (0.4 kb) of moderate intensity, produced by a *Kpn*I-*Hin*dIII double digest, is present (Fig. 7B, arrowhead in lane 2) and corresponds to a band in the positive control (lane 1). The positions of these restriction sites can be seen by inspection of the accompanying map in Fig. 7B. We interpret the relatively strong signal from the 0.4-kb band to be due to more target DNA than would be expected from a unique sequence. These data are consistent with the hypothesis that the *Kpn*I and *Hin*dIII sites are conserved in the family 2 repeated sequences (resulting in the band of 0.4 kb) but that the adjacent *Kpn*I and *Hin*dIII sites are not. In lane 4, the *Kpn*I-*Cla*I double digest of genomic DNA resulted in two very strong bands (a 0.98-kb band [arrowhead] and a high-molecular-weight band) and several weaker bands. The intense 0.98-kb band in lane 4 corresponds to a band of similar size in the positive control (lane 3), indicating that these *Kpn*I and *Cla*I sites are conserved in the family 2 repeated sequences. The presence of additional weaker bands is either due to a slight extension of family 2 repeated sequences beyond the conserved *Kpn*I or *Cla*I sites (or both) or because this family exhibits more sequence divergence than family 1.

Extent of homology between the family 1 repeated sequences. An estimate of the degree of homology between the members of family 1 can be determined from analyses of the melting profiles of heteroduplex DNA. Hetroduplexes were formed between nick-translated segment A of pTf11 and excess unlabeled total genomic DNA. Hybridization was carried out at a criterion of T_m − 25°C and to C_0T = 1. Resulting double-stranded DNA was bound to a hydroxyapatite column and eluted by incremental temperature increases. The amount of radioactively labeled single-stranded DNA released at each temperature was determined. The results are shown as a melting profile (Fig. 8A). Single-copy DNA segment B (Fig. 5) was similarly reassociated to total DNA at a criterion of T_m − 25°C and to $C_0T = 10^2$. The resulting double-stranded DNA melting profile is shown in Fig. 8B. The repetitive and single-copy segments were also reassociated separately to the plasmid pTf11 from which they were derived. Since pTf11 contains only one copy of the repeated segment, the melting profile of segment A should be indicative of a perfectly matched duplex, provided that there is no significant amount of internally reiterated

DNA. Melting profiles of DNA segment A reassociated either to itself on plasmid pTf11 or to total genomic DNA are shown in Fig. 8A. Melting profiles of the single-copy DNA segment B reassociated either to itself on plasmid pTf11 or to total genomic DNA are shown in Fig. 8B. Both the repetitive DNA and the single-copy DNA segments when hybridized to their homologous sequences on the plasmid pTf11 melted over a 30°C range. The reason for this rather broad melt, which is not characteristic of homologous DNA, is not known but could result from the use of DNA labeled via nick translation. Nick-translated DNA exhibits a broad range of sizes, including fragments sufficiently small to cause a significant reduction in melting temperature (11). For this reason, the thermal melting profiles are not sufficiently accurate to be used for base composition determinations of segments A and B. However, both melting profiles of the single-copy DNA are virtually identical with regard to both the shape of the melting curves and the temperatures of the midpoint of strand dissociation. Thus the single-copy DNA exhibits little if any sequence divergence as expected. The shape of the melting profiles and the midpoint temperature of the repetitive DNA are also quite similar to each other, with a reduction of only 2°C from 88 to 86°C between the melting curves and a slight increase in the breadth of the melting profiles.

DISCUSSION

The size of the *T. ferrooxidans* genome is estimated to be about 2.8×10^6 bp. This value is derived from two lines of evidence. First, a comparison of the optically determined reassociation kinetics of sheared *T. ferrooxidans* DNA and *E. coli* DNA was used to deduce the complexity of the *T. ferrooxidans* genome. Reassociation kinetics were monitored with a spectrophotometer, and the resulting data were analyzed by using a computer-fitted least-squares programs, which has been shown to be the best approach to the interpretation of reassociation kinetics (23). The computer-generated curve fitted the data with a goodness of fit and a root mean square indicative that the data had been accurately described statistically. Further, the inclusion of *E. coli* DNA as a standard in the same experiment reduces the likelihood that fluctuations in temperature or buffer concentration could contribute to errors in kinetic measurements.

A second approach to estimating the genome size comes from hydroxyapatite measurements of the reassociation of a cloned single-copy fragment to excess DNA. This can provide one of the most reliable determinations of single-copy rate constants (23). The $C_0T_{1/2}$ of the single-copy component derived from the hydroxyapatite experiments was 6. This value has been corrected for the difference between the hydroxyapatite and optical determinations. This difference arises from the fact that the hydroxyapatite determination measures the total amount of DNA that is double stranded (both partial and complete duplexes), whereas the optical determination measures the extent of formation of DNA base pairs (4). The hydroxyapatite determination of the $C_0T_{1/2} = 6$ and the optical determination of the $C_0T_{1/2} = 5$ for the single-copy DNA are in good agreement.

The size of the *T. ferrooxidans* genome (2.8×10^6 bp) is about two-thirds that of *E. coli* (4×10^6 bp) but is larger than those of two members of the archaebacterial group, *Thermoplasma acidophilum* (1.2×10^6 bp) and *Methanobacterium thermoautotrophicum* (1.7×10^6 bp) (16, 21, 28). It is not clear whether the reduction in genome size of *T. ferrooxidans* compared with *E. coli* is correlated with the obligate chemolithotrophy of *T. ferrooxidans*, because the genome sizes of other obligate chemolithotrophs have not been determined.

The C_0T curve illustrated in Fig. 1 indicates that about 6% of the genome of *T. ferrooxidans* consists of moderately repeated DNA sequences. A comparison of the $C_0T_{1/2} = 0.22$ for this repetitive component with the $C_0T_{1/2} = 5.0$ for the single-copy component suggests that the average reiteration frequency of the DNA sequences in the repetitive component is about 23. However, it should be emphasized that this value may represent only the averaged value of a number of familes of repetitive DNA sequences, where the reiteration frequency of each family is actually different from 23. In addition, the accuracy of the estimate of a total of 6% moderately repeated DNA is suspect. Despite the statistical accuracy of the computer-fitted curve, the reliability of the data in this part of the C_0T curve is questionable. This stems from the inherent inaccuracy in measuring the relatively small percentage of double-stranded DNA present in the early C_0T points. In our opinion, the C_0T curve of *T. ferrooxidans* demonstrates the existence of a small percentage of moderately repeated DNA sequences, almost certainly representing at least 2% and probably less than 10% of the genome.

The reiteration frequency of the ribosomal genes of *T. ferrooxidans* has not been determined; therefore, we do not know whether they contribute to the fraction of moderately repeated DNA. The strain of *T. ferrooxidans* (ATCC 19859) used in this study contains a 45-kb plasmid. This plasmid is represented only once or twice per genome and is unlikely to contribute significantly to the moderately repeated DNA. Furthermore, the chromosomal DNA used in the C_0T analysis was derived from the upper band of a CsCl-ethidium bromide gradient. This will minimize contamination from supercoiled plasmids, further reducing the representation of the plasmid in the moderately repeated DNA.

We have demonstrated the existence of two distinct families of repetitive DNA sequences in the genome of *T. ferrooxidans*, designated family 1 and family 2. The number of repeats per family, their location on the chromosome and plasmid, their size, their pattern of interspersion, and the conservation of restriction sites within and between the two families have been investigated.

Families 1 and 2 each contain about 20 to 30 repeats. This estimate is derived from several lines of evidence. Analysis of Southern blots permits only a semiquantitative determination of the number of members in families 1 and 2. A given band in the autoradiogram could be produced by two similarly sized but different fragments, or, alternatively, one fragment might contain more than one repeated sequence. Therefore, quantitation of the bands in the autoradiograms will almost certainly underestimate the number of repeated sequences. Hybridization of pTf8 and pTf11 to *Ava*I and *Eco*RI digests of chromosomal DNA resulted in approximately 25 bands per digest (Fig. 2A). These two probes hybridized to six bands in the corresponding plasmid lanes. There are more bands in the *Stu*I digests of chromosomal (ca. 33) and plasmid (ca. 12) DNA than in the previous digests. This is to be expected, since *Stu*I cuts within the repeated sequence and each copy of the repeated sequence generates two bands. Approximately 50 bands are predicted from the *Stu*I digest. Since fewer bands are observed, it is most likely that several of the bands comigrate on the gel.

A comparison of the number of family 1 members on the chromosome and plasmid revealed that a higher density of

repeated sequences is present on the plasmid. Assuming that there are 25 repeated sequences on the chromosome and that the size of the chromosome is 2.8×10^6 bp and there are 6 repeated sequences on the 45-kb plasmid, then the relative density of family 1 repeats is approximately 15 times higher on the plasmid than on the chromosome.

The Southern blot shown in Fig. 5 can be used to confirm our estimate of the number of members in family 1. In Fig. 5, segment A of the cloned insert, containing part of the repeated sequence, hybridized to 24 bands in the *Sal*I-*Eco*RV and *Eco*RI digests. This is in good agreement with the estimate of 25 derived from Fig. 2.

In a similar analysis of the number of repeats in family 2, it was observed that probes pTf32 and pTf34 hybridize to about 16 bands in the *Ava*I and *Stu*I digests of chromosomal DNA and to about 23 bands in the *Hin*dIII digest. Since *Hin*dIII cuts within the repeated sequence, it is expected to produce more bands than do the other enzymes.

The four recombinant plasmids we have used as probes were chosen essentially at random from a population of 46. A preliminary screening of this population revealed the presence of six probes that produced multiple bands in Southern blots of genomic DNA. Further analysis demonstrated that three probes contained family 1 repeated sequences and three contained family 2 sequences. Since the genome size of *T. ferrooxidans* is 2.8×10^6 bp, and assuming that the repeated sequences are dispersed throughout the genome and that the average insert size in our recombinant library is 6 kb, we would expect to find three examples of family 1 repeated sequences (assuming 25 per genome) and two examples of family 2 sequences (assuming 16/genome) in a sample of 46 random clones. These numbers are in agreement with our experimental findings.

The reiteration frequency of family 1 has also been determined by kinetic analysis. This was done by determining the $C_0T_{1/2}$ of a radioactively labeled fragment of a repeated DNA sequence derived from family 1 reassociated with an excess of unlabeled total DNA (Fig. 6). A comparison of the $C_0T_{1/2}$ value of the repeated DNA fragment ($C_0T_{1/2} = 0.18$) with that of a single-copy fragment in the same experiment ($C_0T_{1/2} = 6$) provides an estimate of 33 for its reiteration frequency. This is in good agreement with the estimate of 26 derived from inspection of Southern blots.

Three pieces of circumstantial evidence indicate that family 1 and 2 repeated sequences are dispersed throughout the genome and are not organized in a block of tandemly repeated sequences. First, a wide variety of sizes exist in the bands produced by probes specific for family 1 or 2 in Southern blots. For the family 1 repeated sequences, the *Stu*I digest in Fig. 2A and the *Ava*I-*Pst*I double digest in Fig. 7A produced many different-sized fragments that hybridize to the probe even though these enzymes cut within the repeated sequence. If these sequences were tandemly repeated, one would expect to find one very intense band in the autoradiogram. Similarly, the *Hin*dIII digest in Fig. 3A and the *Kpn*I-*Cla*I double digest in Fig. 7B produce fragments of various sizes that bind probes specific for family 2, a result that argues against tandem repetition of these sequences. Finally, the autoradiogram shown in Fig. 5 clearly demonstrates that one member of family 1 is flanked on both sides by single-copy DNA sequences.

Although the majority of repetitive elements are interspersed with single-copy sequences, it is not known whether they are localized to a portion of the chromosome or are dispersed around it, nor is there evidence concerning the relative locations of the two families with respect to each other.

An estimate of the minimal size of the family 1 and 2 members can be derived from the data we have presented. Inspection of the restriction maps in Fig. 4A and B (with the associated cross-hybridization blots [data not shown]) and the sizes of the conserved fragments in Fig. 7 indicates that the minimum size for the members of family 1 and 2 is about 1,000 bp.

There was no hybridization between family 1 and family 2 repeated sequences (data not shown), and the two families are presumably not related at the nucleotide level.

The restriction maps and Southern blots can be used to address the question of sequence conservation within the members of family 1 or 2. The restriction maps in Fig. 4A and B and the results of experiments shown in Fig. 7 clearly show that several restriction sites are conserved in both members of family 1 and 2 repeats. However, since only a limited number of restriction enzyme sites were analyzed, the possibility remains that there are regions of more marked sequence divergence between the restriction enzyme sites that were analyzed. To address this point, a comparison was made between the thermal melting profiles of a repeated sequence belonging to family 1 reassociated with other members of family 1 or, as a control, reassociated with itself. Sequence divergence is characterized by a reduction in the temperature of the midpoint of base-pair dissociation and in an increase in the breadth of the thermal melting transition of base pairs (11). A mismatch of 1% is generally thought to result in a depression of the T_m by about 1°C (11). Our melting curves were determined by measuring the midpoint of strand dissociation (T_d) and by measuring the breadth of the transition of strand disassociation. The effect of mismatch of T_d is not well characterized, so that it is not possible to derive a quantitative estimate of mismatch from our data. However, qualitatively, both the T_ds and the breadth of transition of the repetitive sequence reassociated either with other members of family 1 or with itself are very similar. This is strong circumstantial evidence for sequence homogeneity among members of family 1 repeated DNA sequences. Confirmation of this would require a comparison of the nucleotide sequence of several members of family 1.

The function of these two families of repeated sequences is not known. If our estimate of the minimum size is correct (i.e., about 1,000 bp), then they would be similar in size to insertion sequences (6). Alternatively, this size is similar to that of the terminal modules of compound transposable elements (2). Experiments are now under way to determine the entire nucleotide sequence of a family 2 member so that it can be compared with other known repeated elements.

ACKNOWLEDGMENTS

This work was supported in part by grants CPE-8420881 and MSM-8215220 from the National Science Foundation.

We thank W. R. Pearson, Department of Biochemistry, University of Virginia, Charlottesville, for kindly providing a Fortran version of the least-squares fitting program.

LITERATURE CITED

1. **Barros, M. E. C., D. E. Rawlings, and D. R. Woods.** 1985. Cloning and expression of the *Thiobacillus ferrooxidans* glutamine synthetase gene in *Escherichia coli*. J. Bacteriol. **164**:1386–1389.
2. **Bennett, P.** 1985. Bacterial transposons, p. 97–115. *In* J. Scaife, D. Leach, and A. Galizzi (ed.), Genetics of bacteria. Academic Press, Inc., London.
3. **Britten, R. J.** 1968. Reassociation of non-repeated DNA. Carnegie Inst. Wash. Year Book **65**:330–332.

4. **Britten, R. J., and D. E. Kohne.** 1968. Repeated sequences in DNA. Science **161**:529–540.
5. **Clewell, D. B., and D. R. Helinski.** 1972. Effect of growth conditions on the formation of the relaxation complex of supercoiled ColE1 deoxyribonucleic acid and protein in *Escherichia coli*. J. Bacteriol. **110**:1135–1141.
6. **Cullum, J.** 1985. Insertion sequences, p. 85–95. *In* J. Scaife, D. Leach, and A. Galizzi (ed.), Genetics of bacteria. Academic Press, Inc., London.
7. **Eisenberg, A. J., and D. S. Holmes.** 1982. A note on the use of CsCl centrifugation to purify bacterial plasmids prepared by the rapid boiling method. Anal. Biochem. **127**:434.
8. **Fishman, S. E., and C. L. Hershberger.** 1983. Amplified DNA in *Streptomyces fradiae*. J. Bacteriol. **155**:459–466.
9. **Girvitz, S. C., S. Bacchetti, A. J. Rainbow, and F. L. Graham.** 1980. A rapid and efficient procedure for the purification of DNA from agarose gels. Anal. Biochem. **106**:492–496.
10. **Hasegawa, M., G. Hintermann, J.-M. Simonet, R. Crameri, J. Piret, and R. Hutter.** 1985. Certain chromosomal regions in *Streptomyces glaucescens* tend to carry amplifications and deletions. Mol. Gen. Genet. **200**:375–384.
11. **Hayes, F. N., E. H. Lilly, R. L. Ratliff, D. A. Smith, and D. L. Williams.** 1970. Thermal transitions in mixtures of polydeoxyribonucleotides. Biopolymers **9**:1105–1117.
12. **Holmes, D. S.** 1984. Improved rapid heating techniques for screening recombinant DNA plasmids in *E. coli*. Biotechniques **2**:68–69.
13. **Holmes, D. S., J. H. Lobos, L. H. Bopp, and G. C. Welch.** 1983. Setting up a genetic system *de novo* for studying the acidophilic thiobacillus *T. ferrooxidans*, p. 541–544. *In* G. Rossi and A. E. Torma (ed.), Recent progress in biohydrometallurgy. Associazone Minearia Sarda, Sardinia, Italy.
14. **Holmes, D. S., J. H. Lobos, L. H. Bopp, and G. C. Welch.** 1984. Cloning of a *Thiobacillus ferrooxidans* plasmid in *Escherichia coli*. J. Bacteriol. **157**:324–326.
15. **Holmes, D. S., J. R. Yates, J. H. Lobos, and M. V. Doyle.** 1984. Genetic engineering of biomining organisms, p. A64–A81. *In* The world biotech report 1984. Online Publishers, Ltd. Pinner, United Kingdom.
16. **Klotz, L. C., and B. H. Zimm.** 1972. Size of DNA determined by viscoelastic measurements: results on bacteriophages, *Bacillus subtilis* and *Escherichia coli* J. Mol. Biol. **72**:779–800.
17. **Lessie, T. G., and T. Gaffey.** 1986. Catabolic potential of *Pseudomonas cepacia*, p. 439–481. *In* J. R. Sokatch and L. N. Ornston (ed.), The bacteria. Academic Press, Inc., Orlando, Fla.
18. **Manning, H. L.** 1975. New medium for isolating iron-oxidizing and heterotrophic acidophilic bacteria from acid mine drainage. Appl. Microbiol. **30**:1010–1016.
19. **Marmur, J.** 1961. A procedure for the isolation of deoxyribonucleic acid from micro-organisms. J. Mol. Biol. **3**:208–218.
20. **Mishra, A. K., and R. Pradosh.** 1979. A note on the growth of *Thiobacillus ferrooxidans* on solid medium. J. Appl. Bacteriol. **47**:289–292.
21. **Mitchell, R. M., L. A. Loeblich, L. C. Klotz, and A. R. Loeblich.** 1979. DNA organization of *Methanobacterium thermoautotrophicum*. Science **204**:1082–1084.
22. **Ono, H., G. Hintermann, R. Crameri, G. Wallis, and R. Hutter.** 1983. Reiterated DNA sequences in a mutant strain of *Streptomyces glaucescens* and cloning of the sequence in *Escherichia coli*. Mol. Gen. Genet. **186**:106–110.
23. **Pearson, W. R., E. H. Davidson, and R. J. Britten.** 1977. A program for the least squares analysis of reassociation and hybridization data. Nucleic Acids Res. **4**:1727–1737.
24. **Rawlings, D. E., C. Ganith, A. Petersen, and D. R. Woods.** 1983. Characterization of plasmids and potential genetic markers in *Thiobacillus ferrooxidans*, p. 555–570. *In* G. Rossi and A. E. Torma (ed.), Recent progress in biohydrometallurgy. Associazone Mineraria Sarda, Sardinia, Italy.
25. **Rawlings, D. E., and D. R. Woods.** 1985. Mobilization of *Thiobacillus ferrooxidans* plasmids among *Escherichia coli* strains. Appl. Environ. Microbiol. **49**:1323–1325.
26. **Robinson, M., E. Lewis, and E. Napier.** 1981. Occurrence of reiterated DNA sequences in strains of *Streptomyces* produced by an interspecific protoplast fusion. Mol. Gen. Genet. **182**:336–340.
27. **Sapienza, C., and W. F. Doolittle.** 1982. Unusual physical organization of the *Halobacterium* genome. Nature (London) **295**:384–389.
28. **Searcy, D. G., and E. K. Doyle.** 1975. Characterization of *Thermoplasma acidophilium* deoxyribonucleic acid. Int. J. Syst. Bacteriol. **25**:286–289.
29. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
30. **Torma, A. E.** 1986. Biohydrometallurgy as an emerging technology. Biotech. Bioeng. Symp. **16**:49–63.
31. **Tuovinen, O. H., and D. Kelly.** 1973. Studies on the growth of *Thiobacillus ferrooxidans*. Arch. Microbiol. **88**:285–298.
32. **Vishniac, W. V.** 1974. *Thiobacillus*, p. 456–461. *In* R. E. Buchanan and N. E. Gibbons (ed.), Bergey's manual of determinative bacteriology, 8th ed. The Williams & Wilkins Co., Baltimore.
33. **Wood, N., A. Rake, and L. Shapiro.** 1976. Structure of *Caulobacter* deoxyribonucleic acid. J. Bacteriol. **126**:1305–1315.
34. **Yates, J. R., and D. S. Holmes.** 1986. Molecular probes for the identification and quantitation of microorganisms found in mines and mine tailings. Biotech. Bioeng. Symp. **16**:301–309.
35. **Yates, J. R., and D. S. Holmes.** 1986. The use of genetic probes to detect microorganisms in biomining operations. J. Ind. Microbiol. **1**:129–135.