Determination of microbial genome sizes by two-dimensional denaturing gradient gel electrophoresis

Saibal K. Poddar and Jack Maniloff

Department of Microbiology and Immunology, University of Rochester, Medical Center Box 672, Rochester, NY 14642, USA

Received November 4, 1988; Revised and Accepted March 14, 1989

ABSTRACT
In two-dimensional denaturing gradient gel electrophoresis, DNA is digested with a restriction endonuclease and the resulting DNA fragments are separated as a function of size by conventional agarose gel electrophoresis. Following this first dimension electrophoresis, the fragment distribution is placed at the top of a denaturing gradient slab gel and electrophoresis is carried out parallel to the gradient direction. This second dimension separation is a complex function of the base sequence of each fragment. Analysis of the DNA fragment distribution as a function of fragment size allows the DNA size to be calculated. This method has been applied to calculate three microbial genome sizes: Mycoplasma capricolum, 724 kb; Acholeplasma laidlawii, 1646 kb; and Hemophilus influenzae, 1833 kb.

INTRODUCTION
Two approaches have been described for two-dimensional separation of DNA molecules (refs. in 1): two-dimensional conventional gel electrophoresis and two-dimensional denaturing gradient gel electrophoresis. In the method using conventional gels, large DNA's are digested with a restriction endonuclease and the resulting fragments separated by gel electrophoresis. The fragment distribution is then treated with a second restriction endonuclease and the digestion products are separated by a second gel electrophoresis, with the electric field perpendicular to the original electrophoresis direction. This procedure generates a two-dimensional electrophoretic pattern of DNA fragments (refs. in 2). Two-dimensional conventional gel data can also be used to calculate microbial genome sizes (2,3). Some of these sizes (2) have recently been recalculated, based on further experience in DNA band resolution and analysis, and found to be (J. Maniloff and S.K. Poddar, unpublished data): Escherichia coli strain JM101 (a K12 derivative), 4399 kb; Acholeplasma laidlawii strain JA1, 4399 kb; Acholeplasma laidlawii strain JA1, 1719 kb; and A. laidlawii strain K2, 1680 kb.

The method using denaturing gradient gels takes advantage of the fact that the electrophoretic mobility of a double-stranded DNA molecule is markedly reduced if the duplex is partially denatured without strand separation (4—6). Melting and the consequent transition in electrophoretic mobility is related to base sequence, because melted single-stranded loops form cooperatively within a DNA molecule at regions or domains that are slightly richer in AT base pairs than neighboring domains. Therefore, electrophoretic retardation of a DNA molecule is determined by the melting of one or more domains and is a complex function of base sequence. For this procedure, a DNA restriction endonuclease digest is separated according to size by conventional agarose gel electrophoresis. The agarose gel lane containing separated DNA fragments is then placed at the top of a polyacrylamide
slab gel containing a linear denaturant gradient, such that the gradient is parallel to the direction of the second electrophoresis. Denaturant gradients are preformed using urea and formamide, and electrophoresis is at high temperature. Each DNA fragment has an initial mobility determined by its size, as in conventional electrophoresis, but then migrates into higher denaturant concentrations. At the gradient depth at which one or more critical domains melt, the fragment's mobility is sharply reduced. This produces a two-dimensional DNA pattern in which fragment positions are almost time-independent (4—6).

Both two-dimensional techniques provide a genomic fingerprint: the two-dimensional patterns contain all the DNA fragments that constitute the genome. In this report, we extend the method for genome size determination based on analysis of DNA fragment distributions, developed for two-dimensional conventional gels (2,3), to two-dimensional denaturing gradient gels and use it to determine genome sizes of three microorganisms: two mycoplasmas, *Mycoplasma capricolum* and *Acholeplasma laidlawii*, and *Hemophilus influenzae*.

**MATERIALS AND METHODS**

**DNA preparation**

*M. capricolum* strain Kid and *A. laidlawii* strain JA1 were grown overnight in tryptose broth (7). *M. capricolum* chromosomes were prepared by lysis *in situ* in agarose inserts as described by McClelland et al. (8). *A. laidlawii* high molecular weight DNA was isolated by phenol extraction as previously described (2). High molecular weight DNA from *H. influenzae* strain Rd KW35 was kindly provided by Robert Deich and Algis Anilionis (Praxis Biologics; Rochester, NY).

**Restriction endonuclease digestion**

About 30 µg DNA was digested with 30 units of restriction endonuclease in a total volume of 60 µl of reaction mixture for at least 3 hr, using buffer and temperature conditions as specified by the supplier (BRL, Gaithersburg, MD). Following digestion, 5 µl of gel loading solution (25% glycerol, 60 mM EDTA, 0.1% bromphenol blue) was added to the mixture.

To digest chromosomal DNA in an agarose insert, the insert was put into a microcentrifuge tube with 1 ml of reaction buffer containing 100 µg BSA/ml. The insert was dialyzed at least 3 times (30 min each) in this buffer, and 900 µl of buffer was then removed and replaced by a restriction endonuclease solution (10 units enzyme/µg insert DNA). The mixture was incubated in ice for 1.5—2.0 hr, to allow enzyme diffusion into the insert, and digestion was then carried out at an appropriate temperature for the enzyme. After 2 hr, an additional aliquot of enzyme was added and digestion was continued overnight.

**First dimension electrophoresis**

A 0.7% agarose (BRL, Gaithersburg, MD) gel in TAE buffer (40 mM Tris, 20 mM sodium acetate, 1 mM EDTA; pH 8.0) was cast on a glass plate. Gels were 15 cm wide, 19 cm long, and 2.0—2.5 mm thick; with sample wells 8.5 mm wide and 3 mm long. Each gel was placed in a horizontal electrophoresis tank and the wells were loaded with either a chromosome digestion mixture or a digested agarose insert. Inserts were sealed in wells with molten agarose. In addition, one well was loaded with a 1 kb DNA ladder (BRL, Gaithersburg, MD) as size markers.

Following electrophoresis at 1 V/cm for 16 hr, two longitudinal cuts were made in each gel lane, one just inside each side of the lane. This allowed a strip of agarose containing
the central 5 mm of the lane to be removed. The remaining gel was stained with ethidium bromide (2) and DNA fragments in the trimmed edges were visualized with UV illumination to examine the quality of the first dimension separation. From the band positions in the 1 kb DNA ladder, the lower portion of the unstained gel strip corresponding to DNA fragments <1 kb in size was removed. These small fragments were not needed for the subsequent analysis. The agarose strip was then used for second dimension denaturing gradient gel electrophoresis.

Preparation of denaturing gradient slab gels
Denaturing gradient polyacrylamide gels were formed in a vertical slab gel apparatus, following the protocol of Lerman et al. (6). Glass plates (17 cm × 15 cm) separated by 2.5—3.0 mm spacers were secured in a vertical frame with stainless steel clamps and the sides were sealed with 1% agarose. A bottom gel was formed by pouring a 12 ml solution of 8% polyacrylamide and 0.1% bisacrylamide in 0.5 × TAE buffer, containing 60 μl ammonium persulfate (200 mg/ml) and 12 μl TEMED. About 1 hr was allowed for completion of the polymerization reaction.

Stock solutions of 33% polyacrylamide, 1% bisacrylamide, and deionized formamide (Sigma Chemical Co., St. Louis, MO) were used for making denaturing gradient gels. The buffer was 0.5 × TAE. Using these solutions, 20 ml mixtures were prepared containing 4% polyacrylamide, 0.1% bisacrylamide, and appropriate concentrations of formamide and urea (added by weight) to form the denaturant. For each gel, two mixtures, one containing a high denaturant concentration and the other a low denaturant concentration, were made and extensively degassed. Just before forming the gel, 50 μl ammonium persulfate (200 mg/ml) and 10 μl TEMED was added to each 20 ml polyacrylamide-denaturant mixture. An equal weight of each polyacrylamide-denaturant mixture was put into each chamber of a gradient former (Bio-Rad Model 230; Bio-Rad Laboratories, Richmond, CA) and, using thin tubing, the gradient was poured on top of the bottom gel, such that the denaturant concentration decreased from the bottom to the top of the gel. At least 2 hr were allowed for polymerization of the denaturing gradient gel.

Second dimension electrophoresis
The bottom spacer was removed and the slab gel was placed in a vertical gel apparatus. The bottom buffer chamber was filled with 1% agarose (in 0.5 × TAE buffer), to serve as a plug and contact between the slab gel and the anode reservoir. The agarose lane containing the DNA fragment distribution from the first dimension electrophoresis was set on top of the denaturing gradient slab gel and sealed into place with molten 1% agarose.

The anode reservoir was a 14 liter plexiglas tank, with a horizontal platinum electrode. After the agarose plug solidified, the vertical gel apparatus was placed in the 14 liter tank and the reservoir was filled with 0.5 × TAE buffer, up to the top of the denaturing gradient gel. The temperature of the buffer in the anode reservoir was thermostatically controlled and maintained at either 50°C or 55°C, and the buffer was circulated between the reservoir and the cathode chamber of the vertical gel apparatus at about 5—10 ml/min by a peristaltic pump. Electrophoresis was carried out perpendicular to the first dimension gel at 100 V for 22 hr.

RESULTS AND DISCUSSION
Two-dimensional denaturing gradient gels
Mycoplasma DNA's have a low G+C content: M. capricolum DNA is 24% G+C and A. laidlawii is 32% G+C. These DNA's were digested with restriction endonuclease HindIII
Figure 1: Two-dimensional denaturing gradient gel electrophoresis pattern of *A. laidlawii* genomic DNA. First dimension electrophoresis was from left to right, and second dimension electrophoresis was from top to bottom. The top (horizontal) lane is a first dimension agarose gel lane containing a 1 kb DNA ladder as size markers. Some of the marker sizes are noted at the top. The lane below the marker lane is a first dimension gel lane containing a *HindIII* digest of *A. laidlawii* genomic DNA, as it would have been laid across the top of a denaturing gradient polyacrylamide gel. Below that is a second dimension denaturing gradient gel after electrophoresis.

(recognition sequence AAGCTT). *H. influenzae* DNA, with a G+C content of 40%, was digested with *PstI* (recognition sequence CTGCAG).

Each genomic DNA digest was analyzed by conventional agarose gel electrophoresis.

Figure 2: Semilogarithmic plot of the number of DNA fragments $\geq L$, $S(\geq L)$, as a function of fragment length, $L$. The number of fragments $\geq L$ was counted from the negatives of two-dimensional denaturing gradient gels (as in Fig. 1) for a number of values of $L$ and plotted as a function of $L$. Data are shown for *M. capricolum* (Mca), *A. laidlawii* (Ala), and *H. influenzae* (Hin) genomic DNA's.
These gels also had lanes containing a 1 kb DNA ladder, for use as size markers. Following the first dimension electrophoresis, lanes of genomic DNA were cut out and used for second dimension denaturing gradient gel electrophoresis. The second dimension electrophoretic conditions for the three genomic DNA’s were as follows: HindIII-digested M. capricolum DNA was analyzed by second dimension electrophoresis in denaturing gels containing a gradient of 0—3.5 M urea and 0—20% (v/v) formamide; HindIII-digested A. laidlawii DNA and PstI-digested H. influenzae DNA were both analyzed by second dimension electrophoresis in gels containing a denaturant gradient of 0.87 M urea and 5% (v/v) formamide to 5.25 M urea and 30% (v/v) formamide.

Fig. 1 shows the typical pattern observed for HindIII-digested A. laidlawii DNA. Similar types of patterns were found for HindIII-digested M. capricolum DNA and PstI digested H. influenzae DNA.

Determination of genome sizes
Restiction endonuclease digestion of a DNA of M base pairs containing n cleavage sites will produce fragments with a number average molecular size m = M/n. Assuming the cleavage sites are randomly distributed, the number distribution of fragments as a function of fragment length L, N(L), is given by (2):

\[ N(L) = \frac{M}{m^2} \exp(-L/m) \]

The number of fragments of size ≥ L, S(≥ L), is obtained by summing N(L) from L to M and is (2):

\[ S(\geq L) = \frac{M}{m} \exp(-L/m) \]

A more useful form of this equation is:

\[ \ln[S(\geq L)] = \ln(M/m) -(L/m) \]

Hence, S(≥ L), the number of fragments ≥ L, can be counted (relative to DNA size markers) from a two-dimensional denaturing gradient gel for a number of values of L. A plot of \( \ln[S(\geq L)] \) as a function of L will have a slope of \(-1/m\) and an intercept of \( \ln(M/m) \), and the genome size can then be calculated from these values.

Genome size measurements
For each genomic DNA the number of fragments ≥ L was counted from two-dimensional denaturing gradient gels, relative to size marker bands, and plotted as a function of L (Fig. 2). At low fragment sizes, it was apparent that many small DNA fragments in the two dimensional gels could not be visualized; for M. capricolum and H. influenzae DNA’s, this occurred for fragments <2 kb, and for A. laidlawii DNA, to fragments <5 kb. The problem arises because ethidium bromide binding and fluorescence is much greater for double-stranded than single-stranded DNA (9) and, therefore, in denaturing gradients as DNA fragment size gets smaller an increasing fraction of the fragment base sequence will be in single-stranded denatured regions, leading to decreasing ethidium bromide binding and fluorescence.

The linear relationship between \( \ln[S(\geq L)] \) and L confirms the validity of the assumption, that cleavage sites are randomly distributed, used to derive the equation relating S(≥ L), L, M, and m. The intercepts of the curves with the ordinate show M. capricolum contains 346 HindIII sites, A. laidlawii contains 598 HindIII sites, and H. influenzae contains 245 PstI sites.

Applying the above equation to the curves in Fig. 2, genome sizes were calculated as follows: M. capricolum, 724 kb; A. laidlawii, 1646 kb and H. influenzae, 1833 kb. Based on least square analyses of these data, genome sizes determined by this method have about a 5% standard deviation.
The analyses also show the number average molecular size of the *M. capricolum* DNA fragments after *Hind*III digestion is 2094 bp, of the *A. laidlawii* DNA fragments after *Hind*III digestion is 2754 bp, and of the *H. influenzae* DNA fragments after *Pst*I digestion is 7493 bp.

The 724 kb genome size for *M. capricolum* is close to the value of 758 kb, based on renaturation kinetics measurements, for *M. mycoides* (10), a species phylogenetically close to *M. capricolum* (11). The genome sizes of other *Mycoplasma* species are in the range of 606–803 kb, as determined by renaturation kinetics, electron microscopy, construction of an ordered clone library, and two-dimensional field-inversion gel electrophoresis (10,12–15). The 346 *Hind*III sites found in the *M. capricolum* genome is greater than the 218 sites expected based on a random distribution of bases, showing that (although the sites are randomly distributed in the genome) the number of sites cannot be approximated by a random distribution of bases in a genome with such a low G+C content.

Recent pulse-field electrophoresis studies, using a CHEF apparatus, have given a genome size range of 900–1300 kb for a number of *Mycoplasma* species and a range of 1182–1330 kb for *M. mycoides* (16,17). However, the pulse-field data exhibit a G+C content bias (18), indicating anomalous mobility of the mycoplasma DNA’s in this study. Hence, these higher values cannot be compared to other measurements of *Mycoplasma* genome sizes.

The 1646 kb genome size for *A. laidlawii* is within the range of 1483–1703 kb measured for a number of *A. laidlawii* strains, by renaturation kinetics (10,13), and close to the value of 1719 determined for this strain by two-dimensional conventional gel electrophoresis (2). The 598 *Hind*III sites found in the *A. laidlawii* genome is close to the 520 sites expected assuming a random distribution of bases.

The 1833 kb genome size for *H. influenzae* is experimentally indistinguishable from the 1810–1878 kb values recently determined by pulse-field gel electrophoresis (19), for a strain that should have the same genome size as the one used in these studies (20). The 245 *Pst*I sites found in the *H. influenzae* genome is close to the 265 sites expected based on a random distribution of bases.

Conclusions

Analysis of genomic DNA fragment distributions by two-dimensional denaturing gradient gel electrophoresis provides an alternate physical method for microbial genome size determination. Genome sizes of *M. capricolum*, *A. laidlawii*, and *H. influenzae* measured this way are in agreement with published values using other methods. The accuracy of sizes measured by two-dimensional denaturing gradient gel electrophoresis appears comparable to those from pulse-field gel electrophoresis, and both of these methods appear to provide better accuracy than sizes from renaturation kinetics studies.

ACKNOWLEDGMENTS

These studies were supported by a USPHS—University of Rochester Biomedical Research Support Grant. We thank George Kampo for his technical assistance, and Robert Deich and Algis Anilionis of Praxis Biologies for supplying the *H. influenzae* DNA.

*To whom correspondence should be addressed

REFERENCES