PAPER

Estimating properties of kinetoplast DNA by fragmentation reactions

IOP Publishing Journal of Physics A: Mathematical and Theoretical
J. Phys. A: Math. Theor. **52** (2019) 034001 (15pp) https://doi.org/10.1088/1751-8121/aaf15f

Estimating properties of kinetoplast DNA by fragmentation reactions

L Ibrahim[1,4], P Liu[2,4], M Klingbeil[3], Y Diao[2,5] and J Arsuaga[1,5]

[1] Department of Mathematics & Department of Molecular and Cellular Biology, University of California at Davis, Davis, CA 95616, United States of America
[2] Department of Mathematics and Statistics, University of North Carolina at Charlotte, Charlotte, NC 28223, United States of America
[3] Department of Microbiology, University of Massachusetts, Amherst, MA 01003, United States of America

E-mail: ydiao@uncc.edu and jarsuaga@ucdavis.edu

Received 12 July 2018, revised 8 November 2018
Accepted for publication 15 November 2018
Published 18 December 2018

Abstract

The mitochondrial DNA of trypanosomes, called kinetoplast DNA (kDNA) contains thousands of minicircles that are topologically linked into a single structure that resembles a medieval chainmail. This biological chainmail is characterized by two parameters: the link type between minicircles, and the number of minicircles linked to each minicircle (i.e. the minicircle valence). In previous works, a protocol was proposed to determine the mean value of the minicircle valence. In these experiments, minicircles were excised from the network and the products compared with those obtained from fragmenting idealized structures. These idealized structures assumed a negligible variance in the distribution of valences of the initial network. It is therefore unclear to what extent this theoretical analysis captures the true topology of the kDNA network when kDNA samples are extracted from unsynchronized cells or from cells with silenced kDNA replication genes. Subsequent studies proposed that there is a critical percolation density during network formation. We asked whether this density can be estimated using fragmentation reactions.

The goal of this work is to develop a mathematical method that can be used to estimate the mean valence of networks when the variance of the valence is non-negligible. We first show microscopy data on *Crithidia fasciculata* that, in agreement with the original experimental results, show a distribution of valences with nonzero variance. Second, we use computer simulations of network fragmentation to show that the predicted and actual mean valence are different when the valence distribution has nonzero variance. We propose a

[4] Authors contributed equally.
[5] Author to whom any correspondence should be addressed.

more general mathematical formulation and computer simulations of kDNA fragmentation to estimate this value. Last, we show that fragmentation experiments may lead to errors in the estimation of the critical percolation density since the collapsing density depends on the initial density of the network, and on the fragmentation reaction.

Keywords: statistical mechanics, polymeric systems, minicircle networks, kinetoplast DNA, network fragmentation, DNA linking

(Some figures may appear in colour only in the online journal)

1. Introduction

Trypanosomatids are single celled organisms whose mitochondrial DNA, called kinetoplast DNA (kDNA), contains thousands of small circular DNA molecules (e.g. minicicles). These minicircles are topologically (rather than chemically) linked and their centers are confined into a plane, forming a two dimensional topological network. The evolutionary origin of this network remains to be determined, however, evidence is mounting towards a model driven by confinement [4, 10, 15] together with its corresponding induced evolutionary advantage [14, 19].

A few quantitative characteristics have been identified to describe the topology of these networks. In a sequence of experimental studies, Chen and colleagues identified two experimentally measurable parameters that helped characterize kDNA networks isolated from the model organism *Crithidia fasciculata* [2, 3]. These two characterizing parameters are the link type between two minicircles and the mean minicircle valence, the latter defined as the average number of minicircles topologically linked to any chosen minicircle in the kDNA network. In their studies, they found that minicircles in *C. fasciculata* are linked through Hopf links [17] and that the mean valence was three [3]. In their approach, Chen and colleagues gradually removed minicircles from kDNA networks purified from *C. fasciculata* using limited digestion (i.e. fragmentation). As the network collapsed due to the random linearization of minicircles, clusters containing a small number of minicircles (monomers, dimers and trimers) were separated by gel electrophoresis and their topology identified by electron microscopy. The link type and the frequency of the different clusters were compared with those predicted by ideal tessellations of the plane. The mathematical models for fragmentation proposed on these tessellations all assumed that each minicircle is linked to each of its neighbors (see figure 1(A)). It is therefore unclear to what extent the same mathematical modeling can be applied to kDNA samples purified from unsynchronized cell cultures, from strains that are deficient in kDNA replication (reviewed in [10]) or even from other organisms [14].

More recent computational studies have identified properties of topological networks that can only be uncovered when the network is being formed. These are the critical percolation and mean saturation densities [4, 15]. During network formation, and as the density of minicircles grows, initial small clusters of minicircles join into larger clusters until a 'large' cluster forms. This density is called the critical percolation density [6]. Similarly, as the density of minicircles are further confined, a network in which most of the minicircles ($>90\%$) are linked into a single structure is formed. This new density is called the mean saturation density. A key question is whether these computational properties can be estimated experimentally by gradually fragmentation minicircles.

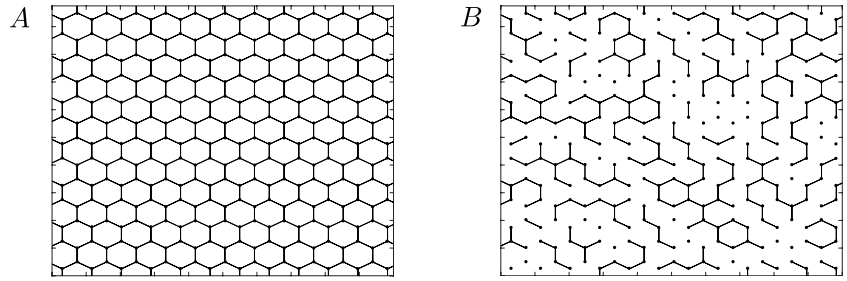

Figure 1. Two networks built on a grid in the hexagonal lattice with the same dimension. Vertices represent minicircles and edges represent topological links. (A) A non-random network in which all minicircles have valence three; (B) a random network with random linking distributions due to the random orientations of minicircles.

The main goal of the work presented here is to reproduce some of the experimental results proposed by Chen and colleagues and further extend their mathematical method. First, we show electron microscopy data illustrating that the valence of minicircle networks may have a large variance. Second, we qualitatively show that the fragmentation products of these kDNA networks in which a significant variance of the valence is observed are consistent with those previously reported in [3]. Third, we test the limit of the original mathematical model by comparing its predictions with those obtained using Monte-Carlo computer simulations of fragmentation of minicircle networks with nonzero variance of the valence. We show that deviations from the predicted values rapidly accumulate, suggesting that the originally proposed method is not valid to analyze situations when the variance is non-trivial. To address this problem we propose a mathematical model that generalizes the model in the original publication by Chen and colleagues and to use computer simulations to analyze fragmentation reactions. We also investigate whether the percolation density can be estimated by the same fragmentation reactions. In particular, we ask whether the percolation density is the same as the density at which the minicircle network does not have a large component any more (i.e. the network no longer percolates). We find that in general these two densities do not agree with each other and that this difference is strongly dependent on different initial factors, including the initial minicircle density and the mechanism of fragmentation.

2. Methods

2.1. Cell growth and kDNA purification

C. fasciculata (clone HS6) promastigotes were grown at 27 °C with gentle agitation in serum-free BHI (37 g l^{-1}) containing 20 μg ml^{-1} hemin. Mid-log phase cells ($\sim 4 \times 10^7$ cells ml^{-1}) were harvested and washed once with 10 mM Tris-HCl pH 8.0, containing 100 mM NaCl and 100 mM EDTA (NET 100) and then resuspended in NET 100 at a density of 1×10^9 cells ml^{-1}. Cells were lysed in 0.5% SDS. Cell lysate was then incubated with 0.2 mg ml^{-1} proteinase K (56 °C, overnight), and then treated with 0.1 mg ml^{-1} RNase A (37 °C, 15 min). Using a modified kDNA purification method, 500 μl of cell lysate was layered on top of 700 μl of 20% sucrose in TE and centrifuged at 20 000 rpm for 60 min using a microcentrifuge. Supernatant was removed (950 μl) leaving behind approximately 250 μl. The remaining sample was resuspended with 250 μl TE and applied to a new sucrose cushion (700 μl) and centrifuged again at 20 000 rpm for 60 min. After removing the supernatant (1.1 ml), the remaining sample was resuspended with 400 μl of TE and layered onto a new sucrose cushion and centrifuged a final

time at 20 000 rpm for 60 min. The last supernatant was carefully removed leaving behind about 35 μl (containing purified kDNA). kDNA samples were dialyzed at 4 °C overnight against 1000 ml vol of TE buffer and precipitated by ethanol. If necessary the samples were ethanol precipitated before proceeding with the restriction enzyme digestion.

2.2. Restriction assay

To digest kDNA, 20 μl containing 100 mM NaCl, 50 mM Tris-HCl (pH 7.9 at 25 °C), 10 mM $MgCl_2$, 100 μg ml^{-1} BSA, 2 μg of kDNA and 2 U of the restriction enzyme XhoI were incubated at 37 °C for times ranging from 5 to 15 min. Buffer and enzyme were purchased from New England Biolabs. After heat deactivating XhoI at 65 °C for 20 min, the DNA was nicked with 10 U of Nb.BsmI by incubation at 65 °C for 1 h to compact each link type into a single band.

2.3. Gel electrophoresis

DNA links were then separated by electrophoresis (2V cm^{-1}) through a 0.8% agarose gel in containing containing 80 mM Tris-HCI (pH 7.5), 5 mM sodium acetate, 1 mM EDTA, and 0.03% SDS. After electrophoresis, gels were stained with ethidium bromide and photographed under UV light using a Amersham Imager 600. DNA was extracted from the gels using a QIAquick Gel Extraction Kit.

2.4. Electron microscopy

Electron microscopy was used to identify kDNA fragments after extraction from agarose gels. DNA-cytochrome C samples were prepared according to [16, 20]. A mixture of DNA (2.5 ng μl^{-1}) and cytochrome C (60 ng μl^{-1}) was prepared in tris-EDTA (TE) buffer. A droplet (50 μl) was placed on a piece of parafilm and allowed to sit for 2 min. Paladion-coated copper grids were touched to droplets. The grids were then rinsed with 95% ethanol for 10 s, 5% uranyl acetate for 30 s, and then again with 95% ethanol for 10 s. The grids were air dried, shadowed with Pt/Pd alloy (80%/20%) at an angle of 7°, and viewed in a JEOL-1230 electron microscope.

2.5. Assembling simulated kDNA networks

We used our previously proposed lattice models for simulating kDNA networks. These models have been detailed in our previous works [4, 5] and were based on sound biological and mathematical observations/assumptions (see [7] for the list of assumptions). In brief, we modeled minicircles by unit circles. There are two main reasons for using this model: first, minicircles in kDNA are linked by Hopf links and this is the only link type possible between two circles [17], and second their sequence is AT-rich [9, 11, 13] and therefore their folding may be different from the standard wormlike chain (e.g. [8]). The minicircle network is generated over a planar lattice grid such that the centers of the minicircles of the network are the grid points and at each grid point exactly one minicircle is placed [4, 5, 18]. Here we use the term *grid* to refer to a finite portion of the lattice on which the minicircle network is built. The orientations of the minicircles (defined as the normal vectors of the plane that contain the minicircles) are identical independent random variables that are either uniformly distributed over the unit sphere or of some other distributions [1, 7]. Thus two randomly generated minicircle networks

over the same grid will have the same number of minicircles and minicircle density, but differ in the orientations of the minicircles. Naturally, the density of minicircles in such a network is defined as the number of minicircles per unit area where the linear unit is defined by the radius of the minicircles, which is not to be confused with the distance between two adjacent vertices in the lattice. Since at each grid point exactly one minicircle is placed, the minicircle density is the same as the number of the grid points per unit area. Thus if we let a be the distance between two adjacent vertices in the lattice, then for large networks, the minicircle density can be expressed as a simple function of a. For example, it is $1/a^2$ for the square lattice and $2/(\sqrt{3}a^2)$ for the triangular lattice.

We have identified two key densities during the formation of minicircle networks: the critical percolation density and the mean saturation density. Roughly speaking, the critical percolation density is the density beyond which a 'large component spanning over the entire network of linked minicircles' forms with a probability >0, while the mean saturation density is the average density beyond which a high percentage of minicircles are linked to the same component (here the percentage is a user defined parameter). Different densities, obtained by changing the distance a between adjacent lattice points in the lattice, and topological structures of the kDNA network were selected according to the hypothesis being tested. These included: (1) hexagonal lattice in which all minicircles valence equal to three [3] (figure 1(A)), (2) triangular and square lattices with a random distribution of valences centered around valence three due to differently oriented circles or when the percolation or saturation were obtained.

For any two grid points on a grid of certain lattice, we define the *lattice distance* between the two points as the minimum number of lattice steps needed to walk from one point to the other. For two minicircles in a minicircle network of a lattice model, we define the lattice distance between the two minicircles as the lattice distance between their centers. Notice that the lattice distance between two minicircles is scale invariant. It is important to note that in our previous models linking between any two minicircles in the network, no matter how far away they are as measured by their lattice distance, is possible when the minicircle density is high enough. This is apparently an artificial effect due to our use of circles as minicircles [4, 5, 18]. From a biological point of view, it makes sense to assume that each minicircle can only link to those minicircles that are not too far from it in terms of the lattice distance. For the purpose of this paper, we shall only consider linking between a minicircle and its immediate neighbors for the triangular and hexagonal lattices. In the square lattice model, other than considering the linking of a minicircle with its four immediate neighbors, we will also consider linking between it and four minicircles closest to it along a diagonal line. Thus, if we let κ be the number of minicircles to which each minicircle can link with, then throughout this paper, $\kappa = 3$ for the hexagonal lattice, $\kappa = 6$ for the triangular lattice and $\kappa = 8$ for the square lattice.

Next, we introduce some definitions that are needed in section 3.6. Let us consider two realizations, M_1 and M_2, of a network based on the same lattice type and with the same number of minicircles; hence the connectedness of the network is uniquely determined by the individual orientations of the minicircles. We say that M_1 is *more connected* than M_2 if any two neighboring minicircles in M_1 are linked so long as their counterparts are linked in M_2. This is illustrated in figure 1 where, following the work by [3], each minicircle is represented by a vertex and topological linking between two neighboring minicircles is represented by an edge connecting the two vertices corresponding to these minicircles. Figures 1(A) and (B) show two hexagonal lattice based minicircle networks. In figure 1(A) each minicircle is linked to all of its neighbors (so there is no variance in linking) and in figure 1(B) two neighboring minicircles may or may not be topologically linked (hence there is a nonzero variance in linking) due to the random distributions of the orientations of the minicircles. In general, if M_1 and M_2 are two minicircle networks whose only difference is that M_1 has a higher density than

M_2 (i.e. the distance between any two neighboring minicircles in M_1 is smaller than that of their counterparts in M_2), then M_1 is more connected than M_2 since in our model two linked minicircles will stay linked when they are moved closer [4].

A minicircle network is said to be *completely saturated* if each minicircle in the network is linked to each of its neighboring minicircles. If M_1 is completely saturated then it is more connected than any other network realization M_2. The network in figure 1(A) is completely saturated.

Let us assume that the network has an initial density above the critical percolation density and as before p the probability that a minicircle in the network is linearized (or equivalently the percentage of minicircles linearized in the network). Let p_0 be the critical value of p such that if $p < p_0$, then the minicircle network remains percolated, but if $p > p_0$, then the minicircle network is no longer percolated (collapses). p_0 depends on the initial density D. We call the mean density D_0 of the network that is obtained by linearizing minicircles with probability p_0 the *collapsing density* of the network. Notice that D_0 is the mean number of minicircles of the networks with p_0 divided by the area of the grid, which is no longer the same as the density of the grid points.

2.6. Modeling of the fragmentation reaction by restriction enzymes

We simulated different fragmentation reactions in this study: (1) each minicircle in the entire network was linearized with a probability p [3]; (2) minicircles were uniformly selected and linearized until the network was no longer percolated; (3) minicircles were uniformly selected and all minicircles within a distance r were linearized. For example, in the case that the square lattice is used for the kDNA network, if a is the edge length between two adjacent lattice points, then if $r < a$, the result would not be any different from Case 2. But if $a \leqslant r < \sqrt{2}a$, then the four vertices closest to the chosen one would be removed, and if $\sqrt{2}a \leqslant r < 2a$, then all eight neighbors of the chosen vertex would be removed.

2.7. Estimation of network parameters

(1) We used the method described in [3] for networks in which all minicircles had equal valence. This method calculates the valence v of the network from the fragmentation probability p and its complement $1 - p = q$, the average number of monomer minicircles released after the enzyme's minicircle linearization M, and the number of minicircles remaining in the network q. The formula is given by:

$$v = \frac{\log M - \log N - \log q}{\log p}.$$

In [3], this expression was used to estimate the mean valence to be 2.96 ± 0.19.

(2) When the network is random, the above approach would not apply. Here we generalize the above approach so that we can still use the quantity M to estimate the mean valence of the network. Let κ be the number of neighbors a minicircle has in the lattice. For example $\kappa = 3$ for the honeycomb lattice, $\kappa = 6$ for the triangle lattice and $\kappa = 8$ for the square lattice. We assume that τ_i $(0 \leqslant i \leqslant \kappa)$, the probability that a minicircle is linked to i of its neighboring minicircles, is the same for all minicircles in the network. This is a reasonable assumption and is not a strong condition. Thus the mean valence of a single minicircle is simply

$$\sum_{0\leqslant j\leqslant \kappa} j\tau_j = \tau_1 + 2\tau_2 + \cdots + \kappa\tau_\kappa.$$

Now let p be the probability that a minicircle is linearized and $q = 1 - p$. For a randomly chosen minicircle in the network, given that it is linked to j neighboring minicircles, the probability that it will become a monomer is qp^j. Thus the probability that a randomly chosen minicircle becomes a monomer in the network after the enzyme has been applied is

$$\sum_{0\leqslant j\leqslant \kappa} qp^j\tau_j = q\tau_0 + qp\tau_1 + qp^2\tau_2 + \cdots + qp^\kappa\tau_\kappa.$$

Let N be the total number of minicircles in the network, then $N\sum_{0\leqslant j\leqslant \kappa} qp^j\tau_j = Nq\sum_{0\leqslant j\leqslant \kappa} p^j\tau_j$ is the mean number of monomers in the network. Let M be the mean number of monomers after the enzyme is applied, then we have $M = Nq\sum_{0\leqslant j\leqslant \kappa} p^j\tau_j$ and

$$\frac{M}{Nq} = \tau_0 + p\tau_1 + p^2\tau_2 + \cdots + p^\kappa\tau_\kappa$$

is a polynomial function of p. If we write $f(p) = \frac{M}{Nq}$, then $f'(1) = \tau_1 + 2\tau_2 + \cdots + \kappa\tau_\kappa$ gives the mean valence of the network. In other word, without any prior knowledge about the network (such as how the minicircles are generated and placed in the network), the mean valence can be estimated by computing $(f(p_2) - f(p_1))/(p_2 - p_1)$ for values $p_1 \neq p_2$, p_1, $p_2 \to 1^-$. Notice that for a randomly selected minicircle, under the given condition that it is already linked to one neighboring minicircle, the probability that it is linked to j additional neighboring minicircles ($0 \leqslant j \leqslant \kappa - 1$) is given by $\tau_{j+1}/(1 - \tau_0)$, hence the probability that this minicircle is not linearized but all of its neighboring minicircles are (except the one given in the condition) is given by

$$q(\tau_1 + p\tau_2 + \cdots + p^{\kappa-1}\tau_\kappa)/(1 - \tau_0) = qh_1(p),$$

and the probability that this minicircle, as well as exactly one additional of its neighboring minicircles are not linearized, but all their neighboring minicircles are (except the minicircle given in the condition), is given by

$$q^2h_1(p)(\tau_2 + 2p\tau_3 + \cdots + (\kappa - 1)p^{\kappa-2}\tau_\kappa)/(1 - \tau_0) = q^2h_1(p)h_2(p).$$

It follows that the probabilities for a randomly selected minicircle to belong to a dimer and a trimer can be expressed as

$$(\tau_1 + 2p\tau_2 + \cdots + \kappa p^{\kappa-1}\tau_\kappa)q^2h_1(p)$$

and

$$(\tau_2 + 3p\tau_3 + \cdots + C_{\kappa,2}p^{\kappa-2}\tau_\kappa)q^3h_1^2(p) + (\tau_1 + 2p\tau_2 + \cdots + \kappa p^{\kappa-1}\tau_\kappa)q^3h_1(p)h_2(p)$$

respectively.

Notice that for a non-random network, we have $\tau_j = 1$ for some j and $\tau_i = 0$ for all other $i \neq j$. In which case $\frac{M}{Nq} = p^j$ hence j can be readily estimated by computing $\frac{\ln(M)-\ln(Nq)}{\ln(p)}$. Correspondingly, the theoretical probability for a minicircle to be in a dimer or a trimer is given by jq^2p^{2j-2} and $3j(j-1)q^3p^{3j-4}/2$. In the case that $j = 3$, we obtain $3q^2p^4$ and $9q^3p^5$, matching the results of [3].

(3) We compute the mean valence of the network directly by generating large samples of random minicircle networks, and examine how each minicircle in each sampled network is linked to its neighbors. This method is the most accurate one, but it averages the number of minicircles linked to each of the minicircles in the network and then averages the entire ensemble of networks obtained at a given density.

3. Results

All numerical results presented in this section are based on grids of size 50×50 with a sample size of 10 000.

3.1. Electron microscopy of purified networks suggest a distribution of valences with nonzero variance

To test whether purified networks had a distribution of valences with nonzero variance, we visualized them using electron microscopy. Figure 2 shows a large section of the network and different magnified subsections. Although the topology of the network cannot be appreciated from these pictures the entanglement complexity and the local aggregation of minicircles can be directly observed. Figure 2(A) shows a section of the purified network. The network is bounded by a structure that has been attributed to the more efficient spreading of DNA on the boundaries of networks [12]. Figures 2(B) and (E) show a magnification of the section near the boundary with some broken minicircles, and two regions with different minicircle densities. Figures 2(C) and (D) show regions in the interior of the network with two different minicircle densities. These figures, in particular the observations of minicircles released from the network, suggest that the distribution of valences has nonzero variance, as suggested by the experiments in [3]. These released minicircles could be from biological origin or a product of the purification process of kDNA.

3.2. Products of fragmentation reactions of kDNA networks are consistent with those previously reported

To test whether the topology of the isolated networks was consistent with previously reported results on kDNA fragmentation, we performed restriction reactions as first reported in [3]. Purified networks were treated with restriction enzyme XhoI, an enzyme that is known to cut minicircles at only one location. Figure 3 shows the product of a time course fragmentation experiment. Consistent with previous reports, minicircle networks do not enter the gel (due to their elevated molecular weight) and fractions of linear fragments, monomers, dimers, trimers can be identified. In this gel, linear fragments cannot be distinguished from circular monomers and migrate together at the bottom of the gel (near the 2.5 kb linear band). Bands associated to dimers and trimers can also be observed. Our results suggest that purified kDNA networks, with a distribution of valences with nonzero variance, produce results consistent with those previously reported [3].

3.3. Our simulations can reproduce previously published analytical results

To test whether our simulations produced outcomes consistent with previous studies [3], we generated tessellations of the plane in which all minicircles had equal valence (such as the one shown in figure 1(A) where every minicircle had equal valence 3). Simulations of kDNA

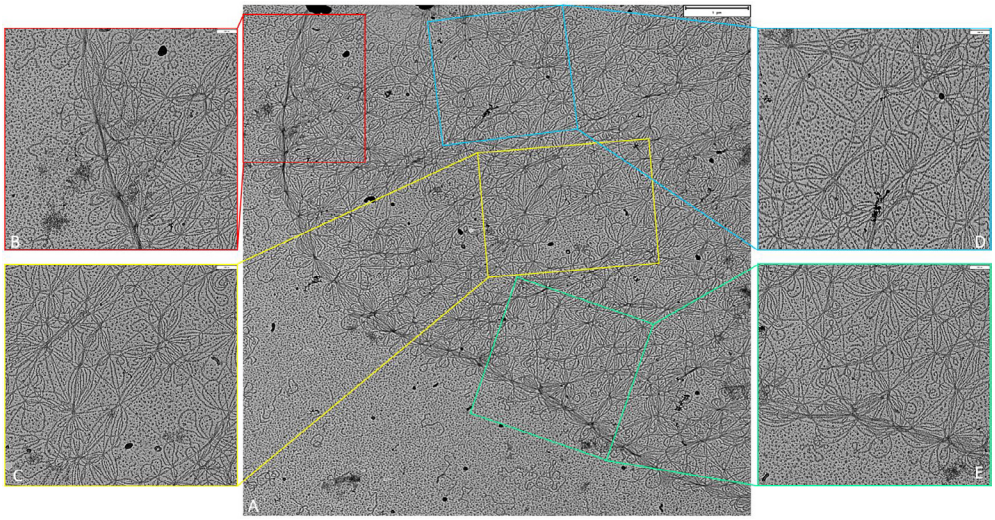

Figure 2. Electron microscopy images of kNDA networks purified from *C. fasciculata*. (A) The kDNA network visualized by electron microscopy. Insets (B) and (E) show regions of high density of minicircles with some minicircles detached from the network. The minicircles and maxicircles at the periphery of the network. Insets (C) and (D) show the interior of the network that is less dense.

fragmentation were performed as described before and results were compared with those obtained using the mathematical model presented in [3]. Figure 4(A) shows the proportion of minicircles present in the monomer, dimer and trimer populations as a function of the linearization probability p. Solid lines are the theoretically derived proportions presented in [3]: qp^3 for monomers, $3q^2p^4$ for dimers and $9q^3p^5$ for trimers. Figure 4(B) shows the estimated mean valence around 3. For a valence three network with zero variance of the valence distribution, the estimated value of the mean minicircle valence is accurately estimated for all values of p.

These calculations suggest that our simulations can reproduce the results presented in [3].

3.4. Initial configurations with nonzero variance produce results that may significantly deviate from those of networks with zero variance

Next, we hypothesized that the proportion of minicircles in monomers, dimers and trimers as well as the estimated mean valence should change when the ensemble of initial network configurations consisted of randomly generated networks, with a fixed mean valence but nonzero variances. As shown in figure 1(B), an underlying hexagonal lattice cannot produce mean valence three upon removing linkage between neighboring minicircles, we thus tested this hypothesis using the triangular and square lattices as the underlying lattices. We have previously analyzed networks built on these lattices and estimated that the critical percolation densities for both lattices are $D^{\text{perc}}_{\text{triang}} \sim D^{\text{perc}}_{\text{sqr}} \sim 0.637$ and mean valence three is achieved when $D^{\text{val=3}}_{\text{triang}} \sim 0.925$ and $D^{\text{val=3}}_{\text{sqr}} \sim 0.923$ [5]. We generated ensembles of simulated networks with density 0.91 for the triangular lattice and 1.02 for the square lattice which are both near the densities corresponding to a mean valence 3; since both are larger than their corresponding percolation densities, we expect that the generated networks will have a 'large component' and a distribution of small clusters with variable amounts of minicircles in them. When purifying kDNA networks, one would expect to lose most of these small clusters. We therefore only

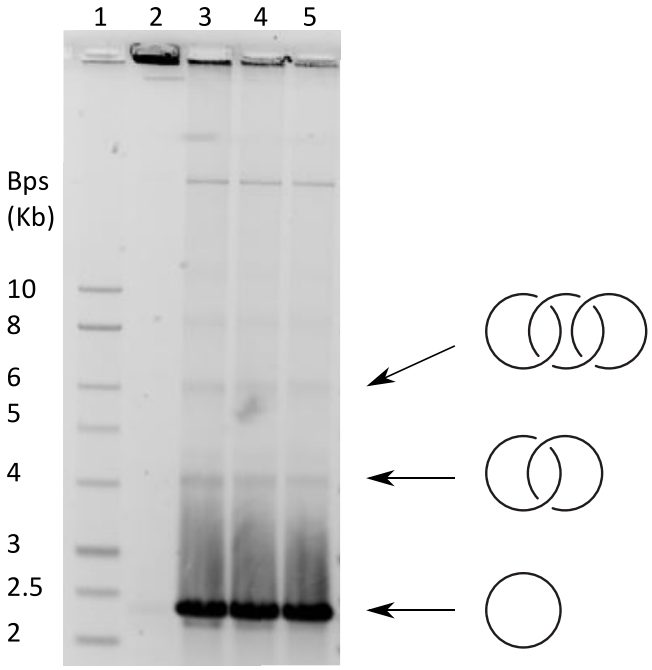

Figure 3. Electrophoretic analysis of kDNA decatenation products obtained by XhoI restriction enzyme. After partial digestion of 2 μg of kDNA, the products were resolved by electrophoresis through a 0.8% agarose gel. Lane 1 contains 1 kb linear DNA ladder (Promega). Lane 2 contains the undigested kDNA network. Lane 3–5 contains the products after digestion with 2 U of XhoI enzyme for 5, 10, and 15 min. A trimer (link with three components) is illustrated but other possibilities for the links type are possible.

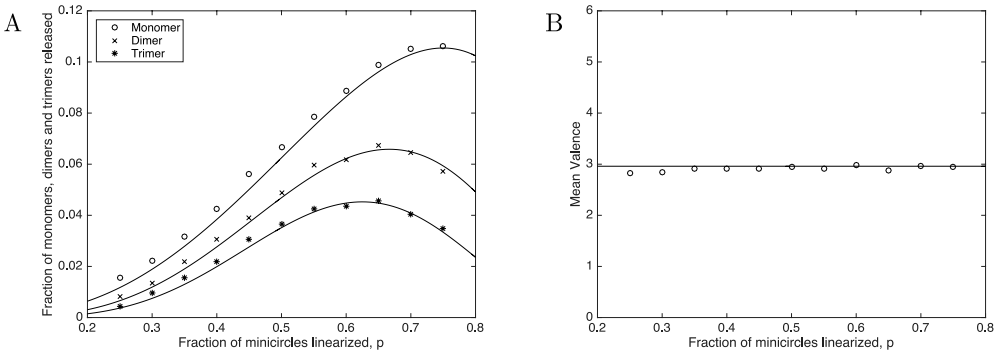

Figure 4. (A) Estimation of the frequency of monomers, dimers and trimers and (B) estimation of the mean valence for minicircle networks with zero variance as a function of fragmentation probability p. The data points are obtained from our simulations.

show the expected distributions of minicircles belonging to monomers, dimers and trimers excised from the large component. Results are shown in figure 5(A). Both lattices give results clearly distinct from the predictions in [3] (solid lines) but, interestingly very similar to each other. This similarity suggests that the distributions of monomers, dimers and trimers are

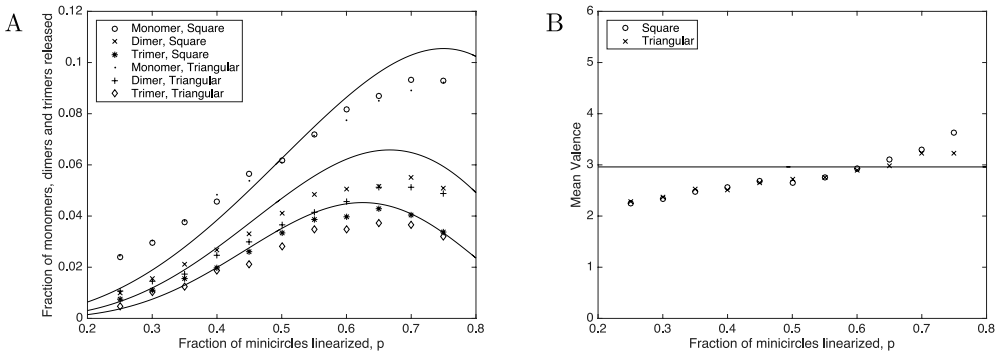

Figure 5. Fragmentations of a triangular and a square lattice at densities at which minicircles have mean valence near or slightly over 3. (A) Fragmentation of randomly generated networks. Solid lines correspond to the predictions presented in [3]; (B) mean valence of the starting network determined theoretically from the simulation output.

sensitive to the valence distribution of the network but not to the different lattices used in the simulations. Interestingly the estimated mean valence, figure 5(B), is clearly dependent on the value of p and reflects the variance of the valence distribution.

3.5. Computer simulations can be used to estimate the distribution of valences in kDNA

In order to use the computer simulations proposed above to estimate the distribution of valences in kDNA experiments, we tested whether the results were independent of the chosen model (i.e. *triangular and square lattices*) and whether the distributions of monomers, dimers and trimers would be different for different mean valences. To test this hypothesis we selected the percolation and 95% saturation densities for the triangular and square lattices which for the latter are $D_{\text{sqr}}^{95\%} \sim D_{\text{triang}}^{95\%} \sim 0.85$ and correspond to mean valences $\sim$2.68 and $\sim$2.69 [5]. Figure 6 shows the proportion of minicircles in the monomer, dimer and trimer populations for percolation ((A) and (C)) and for 95% saturation ((B) and (D)) for both lattices. As expected, the graphs are very similar for both lattices but different for different densities, even within the same underlying lattice.

3.6. The collapsing density of a minicircle network is dependent on the initial minicircle density and on the fragmentation reaction

Next we wanted to address the question of whether the percolation density can be estimated by the collapsing density, that is by the number of minicircles that need to be removed from the network to lose the 'large component'. To address this question we first present a new analytical result that states that the collapsing density depends only on the number of minicircles linked on the initial network; second we present computer simulations results that estimate the collapsing density. We illustrate these results in the square lattice although they hold for any of the other regular lattices we have previously studied [5].

Theorem 1. *If the minicircle network is modeled by unit circles with random normal vectors and centers placed on the points of some planar lattice with an initial density* D*, then the collapsing density* $p_0(D)$ *is a non-decreasing function of* D *with* $p_0(D)/D$ *having a horizontal asymptote, where the value of the asymptote can be obtained from the collapsing density of a completely saturated minicircle network.*

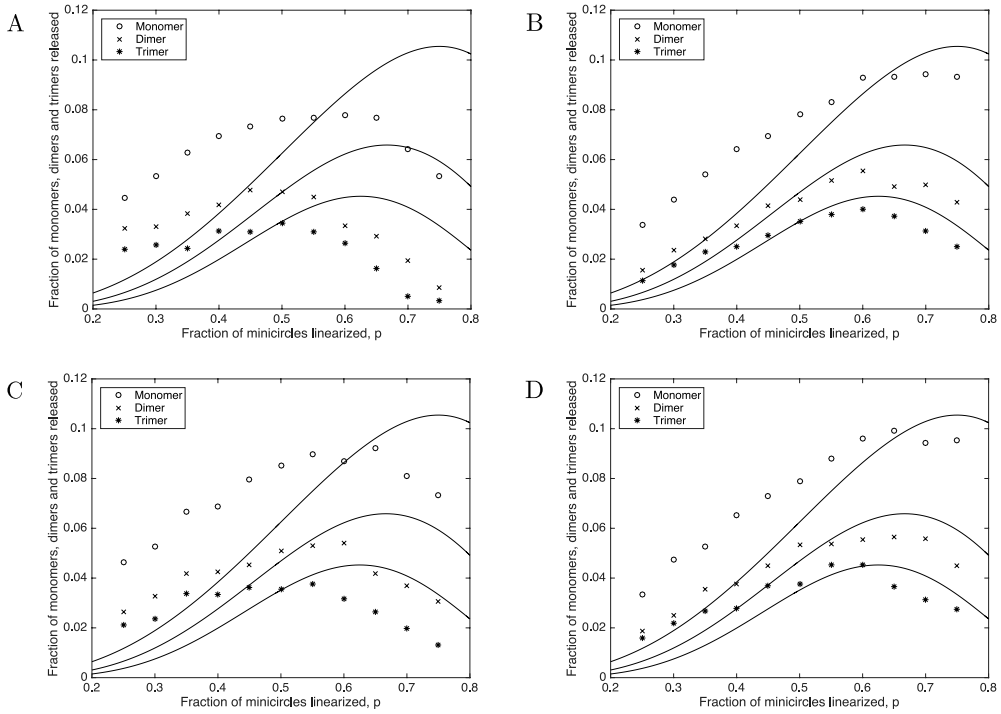

Figure 6. Fragmentation results on square and triangular lattices for key network densities: (A) percolation density in a square lattice; (B) 95% saturation density in a square lattice; (C) percolation density in a triangular lattice and (D) 95% saturation density in a square lattice.

Proof. Let M_1 and M_2 be the two minicircle networks generated under the same model but corresponding to two different densities D_1 and D_2 such that $D_1 > D_2$. That is, M_1 and M_2 have the same number of minicircles that are placed on two lattice grids that differ only by scale. By the remark above, M_1 is more connected than M_2. Thus if we consider a fixed minicircle network in M_2 (meaning each minicircle has a fixed orientation), then two neighboring minicircles in M_1 remain linked if they are linked in M_2. This means that the probability of two neighboring minicircles in M_1 are linked after the enzyme is applied is the same as its counterpart in M_2 if the two minicircles are linked in M_2 (hence also in M_1), or are not linked in M_1 (hence not linked in M_2 either), but is higher in the case that they are linked in M_1 yet not in M_2 (which is possible). Thus under the same enzyme action (namely the same value of probability p), the minicircle network M_1 is more connected than M_2 after the enzyme action is applied. Let D_0 be the critical percolation density of the minicircle network. If $D_2 \leqslant D_0$ then $p_0(D_2) = 0$ since the network is below the critical percolation level for any $p > 0$. So in this case we apparently have $p_0(D_1) \geqslant p_0(D_2)$. If $D_2 > D_0$, then for $p < p_0(D_2)$, the network under M_2 remains percolated after the enzyme is applied by the definition of $p_0(D_2)$. However M_1 is more connected than M_2, hence the network under M_1 will also remain percolated if the same enzyme level (namely probability p) is applied. That is, $p_0(D_1) \geqslant p_0(D_2)$ by definition. This argument applies to any M_2 if M_1 is completely saturated. However, if M_1 is completely saturated, then increasing its density will not change that. Thus at the collapsing density the same number of minicircles should remain regardless

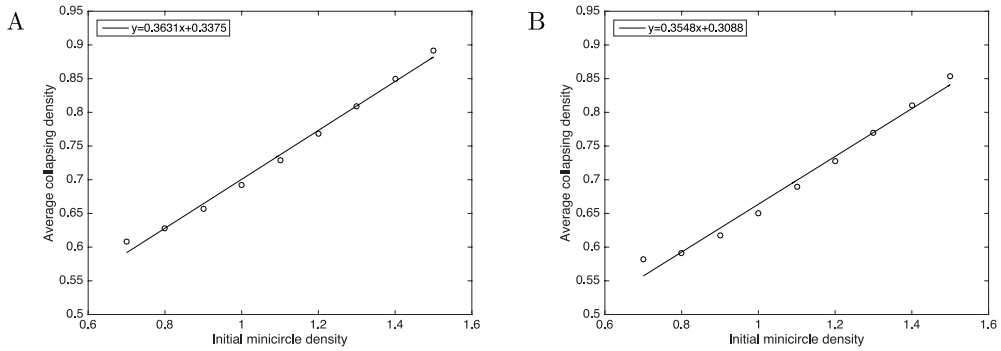

Figure 7. Average collapsing density as a function of the initial density. Minicircles were removed (A) one at a time in a uniformly random fashion and (B) in groups of five.

of its initial density. Let N be the number of minicircles initially in M_1 and N_0 be the number of minicircles remaining when M_1 reaches its collapsing density, then $D_1 = N/A_1$, where A_1 is the area of the grid upon which M_1 is built. It follows that $p_0(D_1) = N_0/A_1 = (N_0/N)D_1$. Since N_0 and N are constants, this proves that the collapsing density of M_1 is proportional to its initial density if M_1 is completely saturated. □

Remark. Notice that in the above proof, the argument does not depend on the value of κ. So the statement of the theorem in fact holds for models with different choices of the κ values.

We performed the corresponding simulations based on square lattice with $\kappa = 8$. Here we varied the initial density values between 0.7 and 1.5 (in increments of 0.1). The values of the network collapsing density were estimated and results averaged over the whole ensemble of networks. Figure 7(A) shows the average collapsing density as a function of the initial density in the lattice. Not surprisingly we observed that the average collapsing density is different from the critical percolation density ($D_c \approx 0.64$) since it depends on the initial density D_0 in a strong linear fashion as predicted by the theorem. Therefore the collapsing density can be lower than, equal to or larger than the critical percolation density.

Next we wanted to test whether the way in which minicircles were removed from the network could affect our estimation of the percolating density. To test this hypothesis we implemented a program that removed minicircles in clusters (see methods) rather than individually. Results for the removal of minicircles in groups of five are presented in figure 7(B). We observe that when minicircles were removed in clusters, the larger the number of minicircles in the cluster, the smaller the average collapsing density. However the overall behavior of $p_0(D)$ with respect to D is very similar in all cases as shown in figure 7. We therefore conclude that without information on the initial density of minicircles and the fragmentation reaction one cannot estimate the percolation density by fragmentation reactions.

4. Conclusion

In recent years the three dimensional organization of genomes has taken center stage due to its connections with functional aspects of the cell. However, determination of the topological properties of any genome remains elusive. This is clearly illustrated by the numerous and diverse, and in many cases contradictory, set of 3D organization models that have been

proposed for chromatin. There are a few examples where evolution has selected topologically complex genomes. One of them is the mitochondrial DNA of trypanosomatids that is partly organized into several thousand minicircles that are topologically linked into a single network. Several parameters determine the topology of the network. In this study we have investigated the distribution of minicircle valences and the critical percolation density.

At any given time point the network can be described by two parameters: the distribution of valences and the link type between every pair of minicircles. Chen and colleagues [3] proposed a combination of experiments and mathematical modeling to determine the mean of the distribution of valences. Our experimental results, in agreement with those in [3] suggest that kDNA purified from *C. fasciculata* have nonzero variance. In [3], a mathematical method to analyze these experiments was proposed. In their method several tessellations of the plane were considered, and the estimated fragmentation probabilities derived. Since these were regular tessellations of the plane the variance of the distribution of valences was assumed to be equal zero. Whether these theoretical results could be generalized to cases where the distribution of valences are affected by the distribution of cells in the different phases of the cell cycle or whether they could be applied to situations in which it is firmly known that the topology of the network deviates from the wild-type remained open. In this work, we have presented a more general mathematical formulation of the problem (section 2.7) and performed Monte-Carlo simulations of fragmentation reactions of networks with nonzero variance. We therefore suggest simulations provide a method to estimate the parameters of the network such as the variance of the distribution of valences and the number of holes.

The percolation density, first proposed in [4, 6], is an important property of the minicircle networks that may provide information on the evolution of the topological structure of the kDNA network. Our results, however show that the estimated value depends on the initial minicircle density and of the fragmentation reaction mechanism. It is our goal to use the methods presented here to estimate the topological properties of kDNA networks purified from replication deficient mutants and other organisms.

Acknowledgment

This work was supported by Internal funds of the University of California Davis (JA and LI) and National Institutes of Health GM58016, CA92276 (LI) and a grant from the National Science Foundation to Prof O Jenda (1343651) and by DMS 1817156 (JA). We want to thank Dr Fei Guo and Dr Samantha Lewis for their assistance with microscopy, and Prof W Heyer for providing laboratory space and advice. Finally, we sincerely thank the referees for reading our paper carefully and for providing some very important suggestions that made the paper more readable and more accurate.

ORCID iDs

J Arsuaga https://orcid.org/0000-0002-4979-8215

References

[1] Arsuaga J, Diao Y and Hinson K 2012 The effect of angle restriction on the topological characteristics of minicircle networks *J. Stat. Phys.* **146** 434–45

[2] Chen J, Englund P T and Cozzarelli N R 1995 Changes in network topology during the replication of kinetoplast DNA *EMBO J.* **14** 6339–47
[3] Chen J, Rauch C A, White J H, Englund P T and Cozzarelli N R 1995 The topology of the kinetoplast DNA network *Cell* **80** 61–9
[4] Diao Y, Hinson K, Kaplan R, Vazquez M and Arsuaga J 2012 The effects of minicircle density on the topological structure of the mitochondrial DNA from trypanosomes *J. Math. Biol.* **64** 1087–108
[5] Diao Y, Hinson K and Arsuaga J 2012 The growth of minicircle networks on regular lattices *J. Phys. A: Math. Theor.* **45** 035004
[6] Diao Y and van Rensburg J 1998 The percolation of linked circles *Topology and Geometry in Polymer Science* (*The IMA Volumes in Mathematics and its Applications* vol 103) ed S Whittington *et al* (New York: Springer) pp 79–88
[7] Diao Y, Rodriguez V, Klingbeil M and Arsuaga J 2015 Orientation of DNA minicircles balances density and topological complexity in kinetoplast DNA *PLoS One* **10** e0130998
[8] Geggier S and Vologodskii A 2010 Sequence dependence of DNA bending rigidity *Proc. Natl Acad. Sci. USA* **107** 15421–6
[9] Griffith J, Bleyman M, Rauch C A, Kitchin P A and Englund P T 1986 Visualization of the bent helix in kinetoplast DNA by electron microscopy *Cell* **48** 717–24
[10] Jensen R E, Simpson L and Englund P T 2012 Network news: the replication of kinetoplast DNA *Annu. Rev. Microbiol.* **66** 473–91
[11] Lankas F, Lavery R and Maddocks J 2006 Kinking occurs during molecular dynamics simulations of small DNA minicircles *Structure* **14** 1527–34
[12] Laurent M and Steinert M 1970 Electron microscopy of kinetoplastic DNA from trypanosoma mega *Proc. Natl Acad. Sci. USA* **66** 419–24
[13] Levene S D, Wu H M and Crothers D M 1986 Bending and flexibility of kinetoplast DNA *Biochemistry* **25** 3988–95
[14] Lukes J, Guilbride D L, Votýpka J, Zíková A, Benne R and Englund P T 2002 Kinetoplast DNA network: evolution of an improbable structure *Eukaryotic Cell* **1** 495–502
[15] Michieletto D, Marenduzzo D and Orlandini E 2015 Is the kinetoplast DNA a percolating network of linked rings at its critical point? *Phys. Biol.* **12** 036001
[16] Perez-Morga D L and Englund P T 1993 Microtechnique for electron microscopy of DNA *Nucleic Acids Res.* **21** 1327
[17] Rauch C A, Perez-Morga D, Cozzarelli N R and Englund P T 1993 The absence of supercoiling in kinetoplast DNA minicircles *EMBO J.* **12** 403–11
[18] Rodriguez V, Diao Y and Arsuaga J 2013 Percolation phenomena in disordered topological networks *J. Phys.: Conf. Ser.* **454** 012070
[19] Savill N J and Higgs P G 1999 A theoretical study of random segregation of minicircles in trypanosomatids *Proc. R. Soc.* B **266** 611–20
[20] Thresher R and Griffith J 1992 Electron microscopic visualization of DNA and DNA-protein complexes as adjunct to biochemical studies *Methods in Enzymology* (vol 211) (New York: Academic) pp 481–90