Estimating Genomic Distance from DNA Sequence Location in Cell Nuclei by a Random Walk Model

Ger van den Engh,* Rainer Sachs, Barbara J. Trask*†

The folding of chromatin in interphase cell nuclei was studied by fluorescent in situ hybridization with pairs of unique DNA sequence probes. The sites of DNA sequences separated by 100 to 2000 kilobase pairs (kbp) are distributed in interphase chromatin according to a random walk model. This model provides the basis for calculating the spacing of sequences along the linear DNA molecule from interphase distance measurements. An interphase mapping strategy based on this model was tested with 13 probes from a 4-megabase pair (Mbp) region of chromosome 4 containing the Huntington disease locus. The results confirmed the locations of the probes and showed that the remaining gap in the published maps of this region is negligible in size. Interphase distance measurements should facilitate construction of chromosome maps with an average marker density of one per 100 kbp, approximately ten times greater than that achieved by hybridization to metaphase chromosomes.

Fluorescent in situ hybridization has emerged as an important tool for ordering DNA sequences along chromosomes. The positions of single-copy DNA sequences, cloned in the form of cosmids, can be localized efficiently with respect to metaphase chromosome bands, establishing the order of sequences separated by 2 to 5 Mbp (*1–5*). To order DNA markers with higher resolution, probes can be hybridized in situ to interphase nuclei (*2, 3, 5–7*), where chromatin is less condensed than in metaphase. Different colors can be used to mark the sites of different probes. In one approach, the linear order of three probes is derived from the arrangement of their hybridization sites in interphase (*3, 7*). This method establishes probe order but does not yield information about probe spacing. As a consequence, probe placement becomes more difficult as the map becomes more dense. High density maps are generated more efficiently if information about relative probe spacing is also available. We show here that the distance between hybridization sites separated by 100 kbp to 2 Mbp follows a random walk model. This model provides simple, statistical rules that relate the separation of DNA sequences along the linear DNA strand to the observed separation of DNA sequences in interphase chromatin. Therefore, pair-wise interphase measurements made among a set of cosmid probes yield both probe order and estimates of the genomic distance between the probes.

Polymer models describe the shape of long, unbranched molecules in statistical terms. The molecules are considered as flexible chains that can assume a large number of configurations. If random influences dominate systematic ones, each configuration has a probability that is predicted by a random walk model (*8–10*). Consider two points n links apart. Denote their physical distance projected onto a plane by R. In the most extensively used polymer model, the Gaussian chain (*9*), the cumulative probability P that R is less than or equal to a given value A is

$$P\,(R \leq A) = 1 - e^{-\pi A^2/(4nL^2)} \qquad (1)$$

Here the constant L is the average projected length of one link. Equation 1 implies

$$\langle R \rangle = \int_0^\infty A(dP/dA)dA = L\sqrt{n} \qquad (2)$$

Thus, the average distance $\langle R \rangle$ between the points is proportional to $\sqrt{n}$.

Cosmids hybridized to interphase nuclei mark their complementary genomic sites with well-defined spots measuring ~0.3 μm in diameter. A pair of cosmids separated by >100 kbp shows as two distinct spots in >80% of interphase chromosomes (*6*). It is reasonable to expect that the distance measured between fluorescent markers on a flexible DNA molecule may be distributed according to a polymer model. A systematic chromatin structure has not been established beyond the level of the 30-nm fiber. Even if systematic, higher levels of organization exist, aspects of the hybridization procedure (cell swelling and fixation, DNA denaturation, and incubation in high salt and formamide) are likely to randomize some levels of chromatin structure.

To determine chromatin organization in interphase, we measured the distances between 13 cosmids from a 4-Mbp region of human chromosome 4p16.3. The cosmids have been isolated by others, and the region has been extensively mapped in an effort to localize the defect in Huntington disease (*11–15*). The published genomic distances are shown in Fig. 1A. These cosmids could not be ordered by hybridization to metaphase chromosomes (*16*). Interphase nuclei were hybridized with pairs of cosmids from this 4p16.3 set. The sites of the two cosmids were labeled with green and red fluorochromes. We obtained pairwise distances by projecting photographs of interphase nuclei onto a digitizing board. Between 100 and 200 measurements were made for each tested cosmid pair.

The measured distributions of three typical probe pairs are shown in Fig. 2. The distributions are fit with a function of the form of Eq. 1. There is agreement between the expected and observed distributions. After determining a mean interphase distance for many probe pairs, it becomes possible to check Eq. 2 as well. At genomic distances <2 Mbp, the square of the mean interphase distance ($\langle R \rangle^2$) is linearly related to known genomic distance (n) as predicted by Eq. 2 (Fig. 3). From the observation that interphase distances are distributed accord-

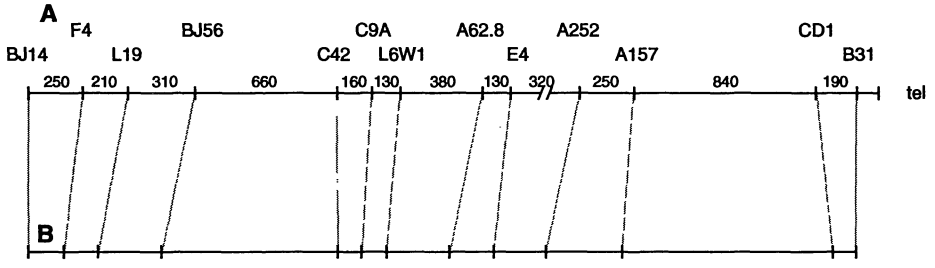

Fig. 1. Comparison of published map with map constructed from interphase distance measurements. (**A**) The published map of the terminal 4 Mbp of 4p16.3 showing the most likely locations of 13 cosmid probes used in this study (*11–15*). The numbers are the published estimates of genomic distance between the probes in kilobase pairs (*11–15*). Literature values vary by 10 to 30%. There is a gap in the map between E4 and A62.8. Cosmid probes used were BJ14 (locus D4S126), BJ56 (D4S127), C42 (42RB1.8, at D4S43), C9A (at D4S43), L6W1 (at J107, D4S166), A62.8 (D4S113), E4 (D4S168), A252 (D4S115), A157 (D4S111), CD1 (D5, D4S90), and B31 (D4S182) (*11–15*). (**B**) Map derived from relative physical distance measured among the cosmids in interphase chromatin. Distance estimates are the squares of the mean interphase distances between nearest neighbors.

G. van den Engh and B. J. Trask, Human Genome Center, Biomedical Sciences Division, Lawrence Livermore National Laboratory, Livermore, CA 94550. R. Sachs, Department of Mathematics, University of California, Berkeley, CA 94720.

*Present address: Department of Molecular Biotechnology GJ-10, University of Washington School of Medicine, Seattle, WA 98195.
†To whom correspondence should be addressed.

ing to the random walk model for sites separated by 0.1 to 2 Mbp, it can be inferred that interphase chromatin behaves over this range as a flexible chain without major constraints. Failure of the model at genomic distances >2 Mbp suggests that some constraining higher order structure exists at these distances. Relations similar to Fig. 3 have been observed in other chromosome regions. Deviation from linearity occurs at ≈1 Mbp in these regions [Xq28 (*3*), 6p21 (*16*)].

The observed relation between physical interphase distance and genomic distance leads to a rapid method for generating dense maps of genomic markers. By determining the interphase distance among cosmids, we can calculate their relative position along the linear DNA molecule. The data in Fig. 3 reveal that markers can be positioned if they fall within 1 to 2 Mbp (~1.5 μm) of other markers. Although each pair-wise determination is relatively inaccurate (relative variation of 0.3), measurements among a set of markers will reveal trends that predict the relative positions with reasonable confidence. Table 1 illustrates this ordering process. We established the order of the probes by arranging the matrix of pair-wise interphase distance measurements so that the values in the rows and columns show an increasing trend away from the diagonal. The squared distance measurements between neighboring probes provide an estimate of their relative separation. The map obtained in this manner (Fig. 1B) agrees with the published map (Fig. 1A), obtained by genetic linkage, somatic cell hybrid, and pulsed-field gel electrophoresis (PFGE) analyses (*11–15*). The map at the bottom (Fig. 1B) was produced by 60 hybridizations. The interphase distance map confirms the existing map around the Huntington locus and indicates that little if any DNA remains to be accounted for in the unmapped gap between E4 and A62.8.

Routine mapping based on interphase measurements could proceed as follows. Randomly selected cosmids are first mapped to 2 to 5 Mbp subregions of a chromosome by fluorescent hybridization to metaphase chromosomes. This step may be omitted if the starting materials are cosmids derived from large yeast artificial chromosome (YAC) clones (*17*). Interphase distances among probes mapping to the same metaphase region are determined until a set of reference markers separated by 0.5 Mbp on average is obtained. Additional probes mapping within this region are then positioned by interphase distance measurements with respect to the reference set. As more probes are placed in a region, the reference set can be updated to obtain an evenly spaced grid. The procedure for ordering the matrix of interphase distance measurements is mathematically straightforward and can be automated. Algorithms for solving similar ordering problems in genetic linkage analysis can be applied with little modification (*18, 19*). Suitable computer programs will calculate the most likely probe arrangement that fits a given set of interphase distance measurements.

The speed of mapping by interphase distance measurements should compare favorably to other methods for obtaining a large set of ordered markers along a chromosome. Because distances of <1 Mbp are informative, a small chromosome such as chromosome 19 (60 Mbp) need be covered by only ~120 reference markers. The order of these reference markers can be established with ~540 hybridizations (*20*). An additional 500 probes can be positioned relative to the reference grid in fewer than 5000 hybridizations. Thus, it is technically feasible for a small team to produce in 1 to 2 years a map with a ~100-kbp average marker density for an entire human chromosome.

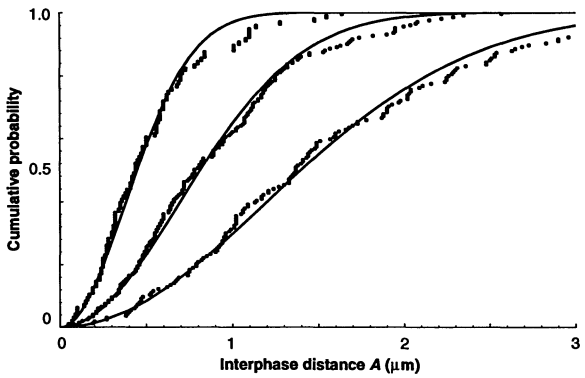

Fig. 2. Three representative cumulative frequency distributions of physical distance measurements made for cosmid pairs hybridized to interphase nuclei (*21*) (dotted line). The distributions have been fitted with a function of the form of Eq. 1 (solid line) by finding the best value for nL^2. The mean measured distances were 0.51, 0.90, and 1.6 μm for probes separated by 190, 500, and 1030 kbp, respectively.

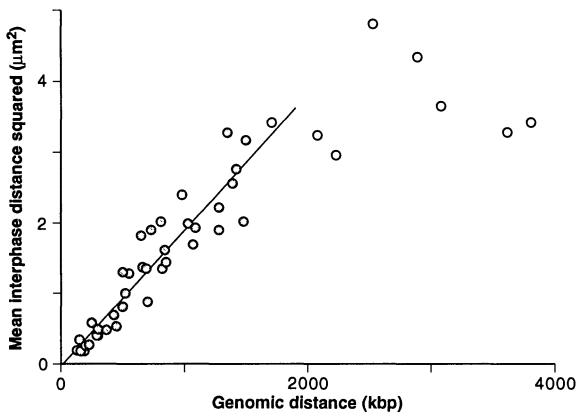

Fig. 3. Linear relation between the square of mean interphase distance ($\langle R \rangle^2$) and known genomic distance (n). Interphase distance measurements for 45 probe pairs with known genomic separations were made (*21*). Results from replicate hybridizations were averaged for 15 of the pairs. The line is the result of linear regression analysis of data points for genomic distances <1.8 Mbp (gray symbols).

Table 1. Mean interphase distance between cosmid hybridization sites in interphase chromatin.

	BJ14	F4	L19	BJ56	C42	C9A	L6W1	A62.8	E4	P252	P157	CD1	B31
BJ14	0	0.52	0.83	1.38	1.66		1.85	1.80	1.72	2.19		1.81	1.85
F4		0	0.51	0.90			1.42						
L19			0	0.70			1.38						
BJ56				0	1.16	1.20	1.55	1.81	1.78			2.08	1.91
C42					0	0.43	0.63	1.17	1.42				
C9A						0	0.44	1.14	1.35				
L6W1							0	0.69	1.00	1.16	1.30		
A62.8								0	0.58	0.73	0.94		
E4									0	0.63	1.13	1.60	
P252										0	0.76	1.39	1.49
P157											0	1.27	1.41
CD1												0	0.43
B31													0

REFERENCES AND NOTES

1. P. Lichter *et al.*, *Science* **247**, 64 (1990).
2. J. B. Lawrence, R. H. Singer, J. A. McNeil, *ibid.* **249**, 928 (1990).
3. B. J. Trask, H. F. Massa, S. Kenwrick, J. Gitschier, *Am. J. Hum. Genet.* **48**, 1 (1991).
4. V. Montanaro *et al.*, *ibid.*, p. 183.
5. B. J. Trask *et al.*, *Genomics*, in press.
6. B. J. Trask, D. Pinkel, G. J. van den Engh, *ibid.* **5**, 710 (1989).
7. B. J. Trask, H. F. Maşsa, M. Burmeister, *ibid.* **13**, 455 (1992).
8. C. R Cantor and P. R. Schimmel, *Biophysical*

Chemistry (Freeman, New York, 1980), chaps. 18 and 19.
9. M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics* (Oxford Univ. Press, Oxford, 1988).
10. R. P. Feynman, R. B. Leighton, M. Sands, *The Feynman Lectures on Physics* (Addison-Wesley, Reading, MA, 1963), chap. 6.
11. M. Bućan *et al.*, *Genomics* **6**, 1 (1990).
12. G. P. Bates *et al.*, *Am. J. Hum. Genet.* **46**, 762 (1990).
13. G. P. Bates *et al.*, *ibid.* **49**, 7 (1991).
14. L. Whaley *et al.*, *Somatic Cell Mol. Genet.* **17**, 83 (1991).
15. C. S. Lin *et al.*, *ibid.*, p. 481.
16. B. J. Trask, unpublished results.
17. I. M. Chumakov *et al.*, *Nat. Genet.* **1**, 222 (1992).
18. K. Buetow and A. Chakravarti, *Am. J. Hum. Genet.* **41**, 180 (1987).
19. ———, *ibid.*, p. 189.
20. About 600 randomly selected markers are required to establish a 100-kbp average density map of a 60-Mbp chromosome. All probes are mapped first to metaphase chromosomes (600 hybridizations). When ten probes have been mapped to the same ~5 Mbp region of the chromosome, they can be ordered by measuring the distance between all possible pairs (45 hybridizations). These probes form a reference grid into which additional probes assigned to the same region can be mapped. About 540 (12 × 45) hybridizations are required to establish a 120-probe reference grid for a 60-Mbp (12 × 5 Mbp) chromosome. Additional probes can be added to the map by measuring their distance to ten reference markers (480 × 10 = 4800). Thus, the maximum number of hybridizations required is ~6000.
21. Cosmid DNAs were labeled and hybridized in pairs to G1 interphase cell nuclei from fibroblast cell cultures as described (*3, 6, 22*). One site was labeled with Texas red (biotinylated probe detected with avidin–Texas red), and one site was labeled with fluorescein isothiocyanate (FITC) [digoxigenin-labeled probe detected with sheep antidigoxigenin and FITC-conjugated donkey anti-sheep immunoglobulin G (IgG) antibodies]. The two fluorochromes were viewed simultaneously through a double bandpass filter (Chromatechnology). Photographic slides of randomly selected nuclei were made for each tested pair (15-s exposures, ≥10 nuclei per field, Scotch 3M 640T film). Slides were projected at ~10^4× magnification onto a digitizing board with a nominal resolution of 40 lines per millimeter (Summagraphics, Seymour, CT). The coordinates of paired red and green fluorescent spots were identified and were entered through the digitizing board into a computer for further analysis. With this approach, ~5000 measurements can be accumulated per person·day.
22. B. J. Trask, *Methods Cell Biol.* **35**, 3 (1991).
23. We are grateful to M. MacDonald and G. Bates for supplying the probes used in this report, L. Hlatky and P. Hahnfeldt for discussions on the applications of the polymer model, and H. F. Massa, R. Esposito, S. Warren, and S. Allen for technical assistance. Work performed under the auspices of the U.S. Department of Energy at Lawrence Livermore National Laboratory, under contract W-7405-ENG-48 with support from NIH HG00256 (B.J.T.) and NSF DMS9025103 (R.S.).

27 April 1992; accepted 10 July 1992

Dependence of Cortical Plasticity on Correlated Activity of Single Neurons and on Behavioral Context

Ehud Ahissar,*† Eilon Vaadia, Merav Ahissar,‡ Hagai Bergman, Amos Arieli,† Moshe Abeles

It has not been possible to analyze the cellular mechanisms underlying learning in behaving mammals because of the difficulties in recording intracellularly from awake animals. Therefore, in the present study of neuronal plasticity in behaving monkeys, the net effect of a single neuron on another neuron (the "functional connection") was evaluated by cross-correlating the times of firing of the two neurons. When two neurons were induced to fire together within a short time window, the functional connection between them was potentiated, and when simultaneous firing was prevented, the connection was depressed. These modifications were strongly dependent on the behavioral context of the stimuli that induced them. The results indicate that changes in the temporal contingency between neurons are often necessary, but not sufficient, for cortical plasticity in the adult monkey: behavioral relevance is required.

Learning processes are usually assumed to be mediated by lasting changes in synaptic efficacies, a phenomenon known as "synaptic plasticity." However, the underlying mechanisms are only partially understood (*1*). Many computational models of associative learning have adopted Hebb's postulate (*2*) and suggested that experience-dependent synaptic changes depend mainly on the contingency [that is, the correlation (*3*)] between the firing times of two interacting neurons (*4, 5*). Although much evidence supports the necessity of correlated activity for synaptic plasticity in vivo in the brain of mammals (*5, 6*), that assumption has not been tested at the level of two single cells. Furthermore, little is known about other factors that may affect this requirement for synaptic plasticity (*7*). Evidence from psychological studies suggests a crucial role for behavioral factors in neural plasticity: Thorndike argued that a connection is significantly modified only if its activation is associated with outcomes important to the animal's behavior (*8*). Because of methodological and technical difficulties, the role of behavioral context has been often overlooked in the study of the synaptic mechanisms underlying learning in mammals (*9*).

To overcome such difficulties, we used a different approach: instead of studying the efficacy of single synapses, we examined the correlation between the firing times of pairs of neurons. The correlation manifests the net effect of the whole synaptic substrate through which the two neurons interact, including both direct and indirect connections; it represents the "functional connection" (*10*) between the two neurons. Lasting changes of functional connections ("functional plasticity") represent lasting changes of cortical functioning as mediated by synaptic plasticity. Our methodological approach was to examine whether and how "general behavioral factors," such as attention, motivation, and reinforcement, affect functional plasticity. These factors were not differentiated in this study and henceforth are included in the term "behavior."

Neuronal activity was recorded from the auditory cortices of two adult monkeys (*11*). In each session, the extracellular activities of two to ten single neurons were recorded simultaneously. One hundred and fifty-eight pairs of neurons that exhibited positive or negative correlation in their activity were selected for this study. The dependence of functional plasticity on the contingency between the activities of the two neurons (henceforth referred to as "contingency") and on behavior was tested by combined cellular conditioning and behavioral paradigms (Fig. 1). The activity of one neuron in each pair (the "CS neuron") was regarded as the conditioned stimulus (CS), and the activity of the other neuron (the "CR neuron") as the conditioned response (CR) (*12*). An auditory stimulus capable of eliciting or suppressing activity in the CR neuron was used as the unconditioned stimulus (US). The US served both for pairing the activities of the two neurons and for guiding the monkey's behavior during the performance of an auditory discrimination task.

The combined paradigm yielded three combinations. (i) "Conditioning associated with behavior" occurred when the monkey performed the task, and the connections (direct, indirect, or both) between the neurons were conditioned. Conditioning was applied by pairing the US with the CS neuron activity (the auditory stimulus was

Department of Physiology, Hebrew University, Hadassah Medical School, Jerusalem 91010, Israel.

*To whom correspondence should be addressed.
†Present address: Department of Neurobiology, Brain Research Building, Weizmann Institute of Science, Rehovot 76100, Israel.
‡Present address: Department of Neurobiology, Institute of Life Sciences, Hebrew University, Jerusalem 91904, Israel.