Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 9052–9056, September 1994
Genetics

Genetic variation detected by quantitative analysis of end-labeled genomic DNA fragments

(two-dimensional DNA gels/detection of DNA mutations/genomic variation in humans)

JUN-ICHI ASAKAWA*†, RORK KUICK‡, JAMES V. NEEL§, MIEKO KODAIRA*, CHIYOKO SATOH*, AND SAMIR M. HANASH‡

*Department of Genetics, Radiation Effects Research Foundation, 5-2 Hijiyama Park, Minami-Ku, Hiroshima 732, Japan; and Departments of ‡Pediatrics and §Human Genetics, University of Michigan Medical School, Ann Arbor, MI 48109

Contributed by James V. Neel, May 31, 1994

ABSTRACT **The continuing efforts to evaluate specific human populations for altered germinal mutation rates would profit from more efficient and more specific approaches than those of the past. To this end, we have explored the potential usefulness of two-dimensional electrophoresis of DNA fragments obtained from restriction-enzyme-digested genomic DNA. This permits the analysis, on a single preparation, of ≈2000 DNA fragments varying in size from 1.0 to 5.0 kb in the first dimension and from 0.3 to 2.0 kb in the second dimension. To enter into a genetic analysis, these fragments must exhibit positional and quantitative stability. With respect to the latter, if spots that are the product of two homologous DNA fragments are to be distinguished with the requisite accuracy from spots that are the product of only one fragment, the coefficient of variation of spot intensity should be approximately ≤0.12. At present, 482 of the spots in our preparations meet these standards. In an examination of preparations based on three Japanese mother/father/child trios, 43 of these 482 spots were found to exhibit variation that segregated within families according to Mendelian principles. We have established the feasibility of cloning a variant fragment from such gels and establishing its nucleotide sequence. This technology should be highly efficient in monitoring for mutations resulting in loss/gain/rearrangement events in DNA fragments distributed throughout the genome.**

Research conducted in recent years has revealed a staggering wealth of genetic variation in the DNA of humans and other animals. Better techniques for the rapid identification and genetic analysis of this variation and for determining the frequency of germinal and somatic mutation, including the alterations in DNA associated with oncogenesis, are highly desirable. The advent of two-dimensional separations of genomic DNA fragments may be an important advance in this context (1–5). In this communication employing the technique of end-labeled restriction landmarks, we will describe the implementation of an approach for the quantitative analysis of the human DNA fragments visualized in autoradiographs of sheet gels prepared using two-dimensional electrophoresis. Elsewhere, we describe the qualitative variation detected with this technology (R.K., J.A., J.V.N., C.S., and S.M.H., unpublished data). Here, we emphasize the ability to distinguish between the autoradiographic intensity of a spot that is the product of two homologous DNA fragments as contrasted with the intensity of a fragment corresponding to one copy of the same DNA fragment. The ability to make this distinction with high accuracy provides the basis for employing this technique in the study of the frequency of mutation. Finally, we will consider the sources of nongenetic variation in spot density and discuss how the detection of genetic variation in gels of this type might be improved.

MATERIALS AND METHODS

DNA Samples. The DNA analyzed was obtained from three father/mother/one-child family constellations of Epstein–Barr virus-transformed human cell lines maintained by the Radiation Effects Research Foundation (6). None of the fathers or mothers had been exposed to the radiation of the atomic bombings.

Preparation of the Two-Dimensional Gels. The technique for preparing and labeling the DNA fragments and for electrophoresis has been described (ref. 7; R.K., J.A., J.V.N., C.S., and S.M.H., unpublished data). Genomic DNA was digested with *Not* I and *Eco*RV restriction enzymes and the *Not* I-derived 5′ protruding ends were α-^{32}P-labeled. These fragments were electrophoretically separated in an agarose disc gel, which was subsequently treated with *Hin*fI to further cleave the fragments *in situ*. The resulting fragments are separated perpendicularly in a 5.25% polyacrylamide gel (33 cm × 46 cm × 0.05 cm). Autoradiograms are then obtained.

Data Collection and Analysis. Autoradiograms were digitized with a Kodak charge-coupled device camera, resulting in images of 1024 × 1024 pixels at a resolution of 0.344 mm per pixel in both dimensions, each pixel having 1 of 256 possible density values. Software to detect and quantify DNA fragments (hereafter also referred to as spots) and software for the camera were obtained from BioImage (Ann Arbor, MI). Density readings were calibrated against a wedge with steps of known optical density. Spot intensities (i.e., integrated density) were expressed as optical density units × mm^2. In the area analyzed, fragment size varied from 1.0 to 5.0 kb in the first dimension and 0.3 to 2.0 kb in the second dimension.

The analysis was based on three mother/father/child trios prepared in duplicate from the same purified DNA samples, so that two sets of nine gels were analyzed. One image was selected to serve as a master to which spots on other images were matched (Fig. 1). Of ≈2000 spots on this image, 774 were selected as potential candidates for study because they were discrete (not overlapped by other spots), were not near the margins of the gel (where spot appearance is inconsistent), and were not one of the very large spots on the image. Matching of the corresponding spots on the images was performed using software as described (8–10). However, to properly evaluate spot changes correlated with variation in these 774 spots, it was necessary to study all the spots in the preparation, whether or not the spot was among the 774 selected spots.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: CV, coefficient of variation.

†To whom reprint requests should be addressed.

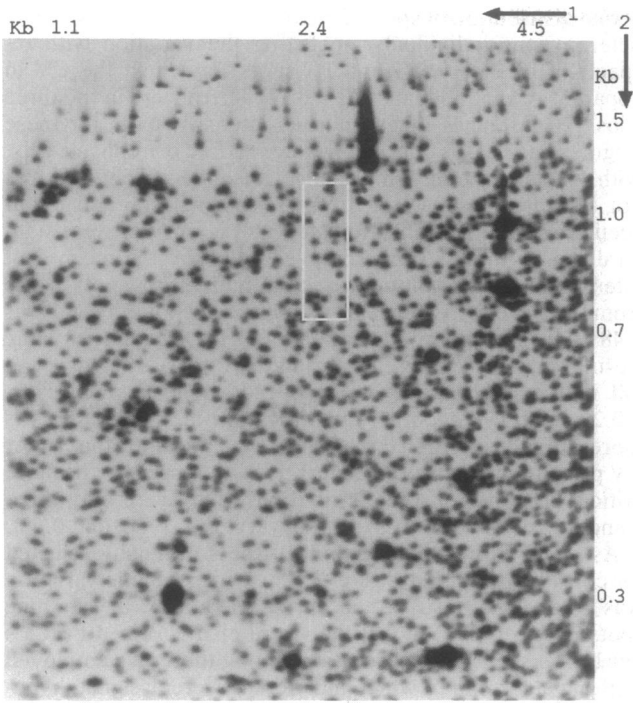

FIG. 1. Portion of the two-dimensional DNA gel analyzed in the present study. The area enclosed in the rectangle is the area included in Fig. 2.

Standardization for Gel-to-Gel Variation in the Intensity of the Autoradiograph. Of these 774 spots, 64 were singled out as "special" during preliminary matching due to one of the following reasons that suggested that the spot was a member of a genetic polymorphism: (*i*) the spot was absent on some gels or (*ii*) the spot was quantitatively diminished by ≈50% when another spot was present on the patterns. The remaining 710 spots were used to standardize the spot quantities across gels in an attempt to compensate for local or general variation in the density of the autoradiograph. For any particular spot on a study gel, the 10 closest neighboring spots (of the set of 710 spots) were used to obtain ratios of spot-integrated intensities on the study gel to those on the master. The largest and smallest of these 10 ratios were ignored, and the remaining 8 ratios were averaged by taking the antilogarithm of the mean of the logarithm of these ratios. The raw quantity for the spot of interest was divided by this local "darkness measure" to obtain the adjusted spot intensity. As a measure of the variation in the integrated density of each spot (i.e., in spot intensity), we employ a coefficient of variation (CV), obtained by dividing the square root of the unbiased estimator of the variance by the mean spot intensity for each set of nine gels; we used the average of the two intraset CVs as a single measure of spot reproducibility.

RESULTS

The Expected Manifestation of Genetic Variation in Two-Dimensional Separations of Genomic DNA Fragments. The density of any spot that appears on the gel is in the usual case expected to be determined by two homologous DNA fragments. In principle, this system will detect genetic variation of two types, namely, (*i*) that due to gain or loss of a cut site for the three restriction fragment enzymes employed in the study and (*ii*) that due to insertion/deletion/rearrangement events. On the basis of experience with Southern blot gels, we suggest the system will probably not detect alterations of fragment length of <2%. In the presence of a detectable variant in fragment length resulting from either reason above, only one DNA fragment would be at the usual position, and the autoradiographic intensity of this spot should decrease by 50%. With respect to the variant fragment, it will migrate to an altered position on the gel (a new spot), not enter the gel, or migrate off the gel. New spots may also appear on the gel as a result of variation in a fragment that does not normally appear on the gel. Some insertion/deletion/rearrangement events could eliminate a second fragment. With respect to either reason above, homozygosity for the variant should be associated with the total disappearance of the corresponding normal spot. Fig. 2 illustrates a polymorphism for which both segregating fragments are apparent on the gel.

Defining a "Normal Spot." The reliable discrimination of spots with normal intensity from the situation in which the same spot has half-normal intensity requires that the variability in normal spot intensity from gel to gel be relatively small. We have chosen a criterion of reproducibility that requires that apparently normal spots be characterized by an average CV for the two sets of gels no greater than 0.12. (The CV may reflect currently ineradicable technical problems and inherent biological variability due, for example, to inconstant methylation of a cut site.) For an idealized model in which normal spots have a mean intensity of 1.0 and spots of half-normal intensity have a mean value of 0.50, both having Gaussian distributions with CV = 0.12, ≈2 per 1000 normal spot intensities would be <0.66, and 4 per 1000 half-normal spot intensities would be >0.66. In a procedure such as this, with many complex steps, only a subset of spots is expected

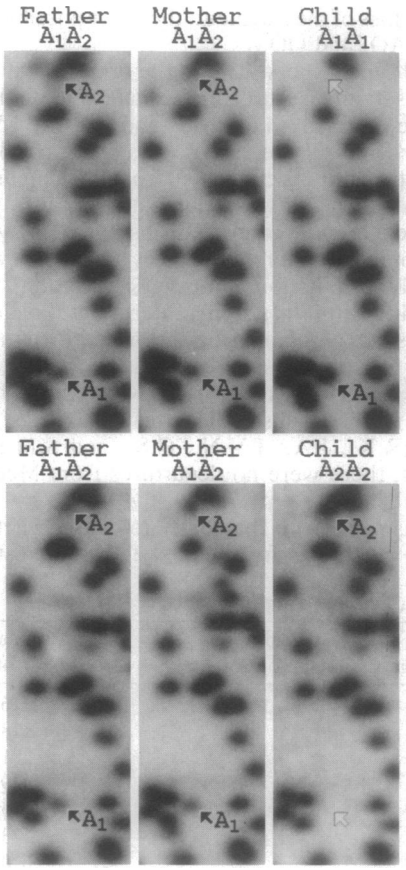

FIG. 2. Example of the genetic variation encountered in type I spots in family trios. The phenotype is indicated adjacent to the spot under consideration with the presumptive genotype given above each gel. In each case, the marriage involves two heterozygotes (A_1A_2), but in the upper examples, the child is homozygous for the A_1 fragment, and in the lower samples, the child is homozygous for the A_2 fragment. Note the greater intensity of the spot in the homozygote than in the heterozygote.

to yield measures of such quality. For the 710 selected spots that had not been singled out as likely to be associated with polymorphic systems, there were 439 for which the average CV for a given spot on the two sets of nine gels was <0.12. The mean CV for these 439 spots was 0.091. The cumulative distribution of the CVs of all 710 of the spots is shown in Table 1. (Below we will attempt to detect an additional set of spots suitable for scoring among the remaining 271 spots and among the 64 spots already identified as possibly exhibiting genetic variation.) We recognize that because of the relatively small number of subjects on whom the selection of suitable spots is based, our current spot list will almost surely undergo revision as our study progresses.

Demonstration of the Genetic Basis of the Variation Observed in a Selected Battery of Spots. As noted earlier, 64 of the 774 spots in the preliminary battery appeared to exhibit genetic variation in the initial analysis. To bring the criteria for selecting these spots into conformity with the (for now) invariant spots, we required that, of the nine evaluations of each individual spot represented in the data, at least three spots be apparently the result of the presence of both homologous fragments and that the CV for such spots was also ≤0.12. These 64 spots were of two types. In type I (44 examples), a segregating fragment could be identified, so that three genotypes could be distinguished if present (AA, AA′, and A′A′). In type II (20 examples), the spot was completely absent in at least one of the nine persons examined. In this situation, the most reasonable interpretations were either a system as above (but the alternative fragment not present on the gel) or a loss of the fragment, the latter possible genotypes being AA, AO, and OO.

Of the 44 type I spots, 23 met the criteria that there were in the data set at least three individuals apparently homozygous for the fragment in question and that the CV for these spots was ≤0.12. Of the 20 type II spots, only 3 met these criteria. Of the 38 spots eliminated from further consideration at this point, 23 spots were eliminated because fewer than three apparent homozygotes were observed; i.e., with further data some of these spots should meet the criteria we have set for the study of mutation and genetic variation.

The validity of the genotypic interpretations described in the previous two paragraphs has been tested by the results of segregation analysis for spots exhibiting variation, employing an enlarged battery of fragments, of which those described in this paper were a subset. The results will be presented elsewhere (R.K., J.A., J.V.N., C.S., and S.M.H., unpublished data); there were no departures from Mendelian expectations.

Given this support to the ability to define a heterozygous state for a particular fragment, we then investigated the actual distribution of spot intensities in the normal position in heterozygous individuals, analyzing the variation with respect to the 26 (type I and II) spots for which the CV in apparent AA individuals was ≤0.12. There were in the three trios 120 instances of half-normal-intensity spots for these fragments (62 cases of heterozygotes scored on replicate gels, with 4 instances of missing data because the variant spot was not sufficiently separated from the normal spot to permit accurate quantification). Adjusted spot intensities were standardized by dividing the observed value by the mean spot intensity of apparent homozygous individuals for spots in gels from the same set. The mean standardized spot intensity was 0.48 with a standard deviation of 0.10. The latter value (resulting in a CV of 0.21) is primarily due to a heavy lower tail in the distribution of these 120 values (skewness = −0.20); because accurate quantification of smaller spots is more difficult than for larger spots, we attribute this larger CV primarily to measurement error. With respect to the more critical upper tail of the distribution, there were only three standardized values >0.65 (these were 0.66, 0.66, and 0.68).

Table 1. Variation of spot intensities

CV range	Additional number of spots	Cumulative number of spots
<0.06	16	16
<0.08	102	118
<0.10	173	291
<0.12	148	439
<0.14	76	515
<0.16	58	573
<0.18	37	610
<0.20	24	634
<0.30	54	688
>0.30	22	710

Data are the average of two CVs obtained for the two sets of gels for 710 selected spots for which probable polymorphisms had not been detected during the matching of images. Some spots later thought to be polymorphic but that were detected only by sifting through the quantitative data are included.

As noted earlier, among the 710 spots originally selected for consideration, there were 271 for which genetic variation was not originally suspected during the course of matching spots between images and that were characterized by CVs >0.12. To seek genetic variation among these, we employed a sifting algorithm in which individuals were assigned hypothetical genotypes (AA or AO) and then tested to see whether hypothetical AA individuals (at least three) yielded an average CV ≤0.12 for the spot, whereas hypothetical AO patterns all gave values <0.75 of the mean for AA individuals in the same set. This method yielded 17 potentially polymorphic spots. In no case were the assigned genotypes for these spots inconsistent with Mendelian inheritance, although adherence to Mendelian inheritance was not required by the sifting algorithm, and for each spot, all apparent AO spots yielded standardized values <0.70, increasing our confidence that the identified spots are not merely the result of random variation in the data. Further, for 3 of the 17 spots originally included in this category, a spot thought likely to represent the corresponding variant DNA fragment was identified after careful review. One of the 17 appeared to follow the pattern of sex linkage; the subject of sex-linked fragments will be treated elsewhere (R.K., J.A., J.V.N., C.S., and S.M.H., unpublished data). For the remaining 13 spots, the observed parent–offspring combinations of spots again followed Mendelian expectation, although in the presumed AA × AO matings, the observed ratio in the children was 11 AA to 5 AO. This departure from the expected 8:8 ratio was not significant ($P = 0.21$ for a two-sided test). With respect to this battery of 17 spots, we observed 66 presumed heterozygotes for the normal fragment, with a mean intensity ratio to the homozygote of 0.52 and a standard deviation of 0.11. Six of the 66 measurements exceeded 0.65, one having a value of 0.69.

Thus, a total of 482 spots met the preliminary criteria established at the outset of the study for suitability for the simultaneous study of genetic variation and mutation. Of these 482 spots, 43 (23 + 3 + 17) were associated with segregating variation, for which there was altogether a total of 186 instances in which a spot was measured that was assumed to reflect the presence of only one of the two DNA fragments normally present in this position. The standardized intensity of all 186 spots was 0.49, with a standard deviation of 0.10.

Sources of Variation in Measuring the Intensity of the Normal (Two-Copy) Spots. Although we have not undertaken a detailed comparison of the relative contributions of various steps in spot quantification to the variability in spot intensity measures, the availability of pairs of gels for the same individual allows some preliminary conclusions. By compar-

ing the average of the spot CVs obtained within each of the two sets of gels (for the 439 nonvariable spots with average CV ≤0.12) to a CV obtained for data that was first averaged between each pair of duplicate gels, we find that on average only 22% of the variance in the adjusted spot intensity can be attributed to factors that differ between different individual's purified DNA samples. The remaining 78% of this variance is currently due to factors that contribute to differences between the replicate gels for an individual, such as variability in the results of restriction enzyme cleavage, lack of uniformity in radioactive labeling, the electrophoretic steps, autoradiography, and especially, the computer analysis steps of digitization and quantitation. This finding leads us to expect that future improvements in image-capture (higher-resolution scanning or use of the phosphor-storage plate technology) and in the quantitation algorithms will lead to decreases in the variability of adjusted spot intensity and, hence, even cleaner discrimination between spots due to one fragment and spots due to two homologous fragments. Such improvements would also increase the number of spots on a given gel that could be reliably scored for variation.

Characterization of Variants. For this approach to realize its full potential, it will be necessary to establish the sequences of the various DNA fragments visualized in the gel and, ultimately, to relate them to known sequences in the human genome. To this end, a human genomic DNA library has been prepared from DNA fragments digested with *Not* I and *Eco*RV. DNA fragments that contain a *Not* I site at one or both ends were purified from the digests, based on a specific ligation of the fragments containing *Not* I ends to Easy Anchor *Not* I (gift from Nippon Gene, Tokyo), a porous polymer solid support covalently attached to double-strand oligonucleotides whose ends are complementary with a *Not* I end. The fragments were size-fractionated and ligated to the *Not* I/*Eco*RV-digested pBluescript II vector (Stratagene). The cloned fragments were transfected into Electromax DH12S cells (GIBCO/BRL), which were grown on LB/agar plates. *Not* I/*Eco*RV-digested plasmid DNAs from white colonies were labeled with isotope and the position of each DNA clone on a standard two-dimensional gel was identified.

The usefulness of this library can be illustrated by the characterization of spot 589 (Fig. 2, spot A_1) and a common variant, spot 1996 (Fig. 2, spot A_2). The clones that correspond to the two allelic DNA fragments were identified and ≈700 bp of each was sequenced, beginning at the *Not* I site. At position ≈700, there is a *Hin*fI site sequence (_AGGAGTCGGG_) in the smaller fragment (spot 589) but the larger fragment (spot 1996) does not have this site, being characterized by the sequence (_AGGAGTTGGG_). There were no other differences in the two sequences. Thus the allelic variation exhibited by these two fragments is shown to be due to a polymorphism in a restriction site. The complete sequence data for these fragments will be presented elsewhere; no identical sequences were in the EMBL data base as of March 1994 and in the GenBank file as of February 1994.

DISCUSSION

In this initial study, we have demonstrated the feasibility of detecting with high accuracy in a single preparation quantitative variation in a selected set of some 482 DNA fragments, averaging ≈2 kb in length. Forty-three of these fragments exhibited genetic variation. The present estimate is based on the variation encountered in only three mother/father/child trios. Furthermore, for a DNA fragment to be included in this study, we required that in at least three of nine gels under consideration, the spot under consideration be the product of two homologous DNA fragments, with a CV for spot intensity of ≤0.12. We have encountered a number of spots for which the requisite three homozygotes with respect to this fragment have not yet been observed; some of these will certainly be eligible for the battery of variable spots under consideration as the gel series is expanded.

Of the many possible applications of this technology, we will discuss just one: the potentialities of the system for the study of mutation, using the 482 spots for which there were at least three AA homozygotes that had an average CV for the two sets of nine gels ≤0.12. For this subset, there was a total of 8595 spots on the 18 gels with nonmissing data, of which 186 were half-normal-intensity spots. For each spot in each set, the spot intensities were standardized by dividing them by the mean value for "obvious" AA individuals in the set. To obtain this mean, first the largest two values of the nine for the set were averaged, and spots with values >75% of this mean were considered clearly AA. The mean for AA spots was recomputed, and again spots with values >75% of this mean were also considered to be AA, and the mean for AA individuals was recomputed a final time.

Table 2 illustrates the distribution of standardized spot intensities for fragments from both homozygous and heterozygous individuals. If we select a standardized value of 0.66 as the cut point for discriminating between normal and half-normal-intensity spots, Table 2 demonstrates that none of 8409 normal-intensity spots would be misclassified whereas 5 of the 186 spots thought to represent the product of a single fragment would be misclassified as normal (two copy) spots. In the simplest case, a mutation consists of a 50% intensity spot in a child, both of whose parents exhibit 100% intensity for that spot, with or without the appearance of a "new" fragment in the child. In a study of germ-line mutation, searching for mutations involving loss of a fragment in a child's gel, the misclassification of a normal-intensity spot as half-normal would present as a false-positive error (i.e., spurious mutation), whereas the latter type of misclassification results in a false negative (a missed mutation). Our method (i.e., selecting the 482 spots by demanding small CVs for normal-intensity spots and then using the same data for estimating the rate of false positives) leads to an underestimate of the number of false positives to be expected for such spots on future gels, but we note that the false-negative-rate estimate of 5 per 186 observations is not similarly flawed.

Since all apparent mutations must be verified, the cost of false-positive observations is the work of preparing additional gels (which we assume would correct the mistake). If we assume actual false-positive rates of 1 per 5000 spots scored, a study monitoring 500 spots per family would require that additional gels be prepared for ≈10% of the children. For false-positive rates of 1 per 2000, additional preparations

Table 2. Distribution of standardized spot sizes for one-DNA-fragment and two-DNA-fragment spots

Standardized spot size	No. spots due to two DNA fragments	No. spots due to one DNA fragment
<0.61	0	156
0.61–0.62	0	8
0.62–0.63	0	7
0.63–0.64	0	2
0.64–0.65	0	4
0.65–0.66	0	4
0.66–0.67	2	0
0.67–0.68	0	2
0.68–0.69	4	2
0.69–0.70	1	1
0.70–0.71	4	0
>0.71	8398	0
Total	8409	186

Decision as to whether a spot is due to one or two homologous fragments is reached by independent criteria.

would be needed 22% of the time. With respect to false-negative rates, with the present technology, and with single gels per individual, ≈3% of the mutations might be missed if attention were restricted to the quantitative aspects of the spot in question, but some fraction of these missed mutations might be detected by the appearance of a "new" spot not present in either parent that ultimately could be traced to its spot of origin, both by the accompanying correlated quantitative change in the latter and, ultimately, by the DNA identities. We note that in studies involving both study and control populations, this missed 3% would not introduce bias into estimates of population differences in the mutation rates. The frequency of both false negatives and positives in mutation studies will in the future undoubtedly be decreased by a reduction in the errors in quantification currently embedded in the various steps in the analysis.

A potential complication in this approach to detecting mutation is the occurrence among the battery of spots under consideration of one spot or several spots for which the corresponding *Not* I sites are subject to occasional methylation, as is now thought to be the explanation of gene imprinting. Hayashizaki *et al.* (11) have suggested that in the restriction landmark genomic scanning of mouse DNA, conducted in a manner similar to these studies, 8 out of 3100 DNA fragments were identified as imprinted through the use of reciprocal F_1 hybrids. If this phenomenon were to occur in these preparations and masquerade as a mutation, the nature of the event would become clear in the types of studies of genomic DNA that should be performed on all putative mutations.

We consider that it is desirable in any genetic monitoring at the DNA level to employ DNA fragments primarily derived from functional genes, rather than fragments that are of unknown or dubious function. This is because of the better possibility of projecting the phenotypic impact of mutation in DNA of known function, a necessary objective in interpreting the significance of an increased mutation rate to a concerned public. *Not* I sites are disproportionately frequent in the unmethylated "CpG islands" so common at the 5′ end of genes (12–17). Thus, the use of *Not* I as one of the restriction fragment enzymes in the first dimension of these gels may result in the derivation of a very high proportion of the visualized fragments from active genes.

We are grateful to Mr. Satomichi Kaneoka for his excellent technical assistance. This publication is based on Research Protocol 7-85 and is MS 20-94 of the Radiation Effects Research Foundation (RERF), Hiroshima and Nagasaki, Japan. RERF is a private foundation funded equally by the Japanese Ministry of Health and Welfare and the U.S. Department of Energy through the National Academy of Sciences. J.V.N. and S.M.H. acknowledge the support of Department of Energy Grant ER60533 and Public Health Service Grant CA26803.

1. de Wachter, R. & Fiers, W. (1982) in *Gel Electrophoresis of Nucleic Acids: A Practical Approach*, eds. Rickwood, C. & Hames, B. D. (IRL, Oxford), pp. 77–116.
2. Uitterlinden, A. G., Slagboom, P. E., Knook, D. L. & Vijg, J. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 2742–2746.
3. Yi, M., Au, L., Ichikawa, N. & Ts'o, P. O. P. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 3919–3923.
4. Hatada, I., Hayashizaki, Y., Hirotsune, S., Komatsubara, H. & Mukai, T. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9523–9527.
5. te Meerman, G. J., Muulaart, E., van der Meulen, M. A., den Daas, J. H. G., Morolli, B., Uitterlinden, A. G. & Vijg, J. (1993) *Am. J. Hum. Genet.* **53**, 1289–1297.
6. Neel, J. V., Satoh, C. & Myers, R. (1993) *Mutat. Res.* **291**, 1–20.
7. Hayashizaki, Y., Hirotsune, S., Okazaki, Y., Muramatsu, M. & Asakawa, J. (1994) in *The Encyclopedia of Molecular Biology, Fundamentals and Applications*, ed. Myers, R. A. (VCH, Cambridge, U.K.), in press.
8. Skolnick, M. M., Sternberg, S. R. & Neel, J. V. (1982) *Clin. Chem.* **28**, 969–978.
9. Skolnick, M. M. & Neel, J. V. (1986) in *Advances in Human Genetics*, eds. Harris, H. & Hirschhorn, K. (Plenum, New York), Vol. 15, pp. 55–160.
10. Kuick, R. D., Skolnick, M. M., Hanash, S. M. & Neel, J. V. (1991) *Electrophoresis* **12**, 736–746.
11. Hayashizaki, Y., Shibata, H., Hirotsune, S., Sugino, H., Okazaki, Y., Sasaki, N., Hirose, K., Imoto, H., Okuizumi, H., Muramatsu, M., Komatsubara, H., Shiroishi, T., Moriwaki, K., Katsuki, M., Hatano, N., Sasaki, H., Ueda, T., Mise, N., Takagi, N., Plass, C. & Chapman, V. M. (1994) *Nat. Genet.* **6**, 33–40.
12. Bird, A., Taggart, M., Frommer, M., Miller, O. J. & Macleod, D. (1985) *Cell* **49**, 91–99.
13. Bird, A. P. (1987) *Trends Genet.* **3**, 342–347.
14. Gardiner-Garden, M. & Frommer, M. (1987) *J. Mol. Biol.* **196**, 261–282.
15. Bickmore, W. & Sumner, A. (1989) *Trends Genet.* **5**, 144–148.
16. Larsen, F., Gundersen, G., Lopez, R. & Prydz, H. (1992) *Genomics* **13**, 1095–1107.
17. Antequera, F. & Bird, A. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 11995–11999.