NMR Diffraction and Spatial Statistics of Stationary Systems

G. A. BARRALL, L. FRYDMAN, G. C. CHINGAS*

Nuclear magnetic resonance (NMR) spatial imaging data may be acquired, processed, and interpreted in ways that provide information directly analogous to diffraction experiments, with length scales determined by gradient strengths rather than radiation wavelengths. This approach, originally considered by Mansfield nearly two decades ago, provides access to autocorrelations of sample density that statistically characterize small-scale density variations. These NMR "Patterson functions" can be acquired orders of magnitude more rapidly than comparably resolved NMR images and are suitable for spatial characterization of small features in bulk samples, such as morphology in structural materials. Unlike hindered diffusion approaches, neither mobility, penetrants, nor transport time are required for examining granularity and porosity.

PROPERTIES OF MATERIALS ARE FREQUENTLY governed by small-scale, uniformly distributed morphological features, such as domain boundaries and pores, whose size and distribution are more significant than precise location. Because material structural information resides at short length scales in small, signal-poor regions, applications of NMR imaging to materials have not had the spectacular success of biological studies.

In this report we discuss the use of NMR Patterson functions in place of images for characterizing small-scale heterogeneities. Collective properties are more succinctly described and more easily observed with this statistical approach. These NMR Patterson functions are analogous to the density autocorrelation functions of x-ray diffraction (*1*), except that nuclear rather than electronic density provides contrast and resolution is determined by magnetic field gradient strength rather than x-ray wavelength.

Two main benefits accrue. First, the number of data points required for a statistical description is much smaller than for imaging. Although resolution considerations are similar for both diffraction and imaging, the density of data points that determines the extent of the spatial "window" is quite different: the image must show the entire sample, but the Patterson function is appreciable only over several feature lengths. Hence, the required number of points is reduced by approximately the number of features in the sample, which may easily reach 10^6 to 10^{12} for small-scale substructures. Further reductions are possible for the important class of spatially isotropic systems, where full statistical information is available on two-dimensional (2-D) and 3-D isotropic structures from a 1-D scan. In addition, tighter signal filtering and more efficient signal averaging can be applied to increase sensitivity.

Second, with statistical characterization, experimental repeatability is possible even for mobile systems. For example, it is impossible to repeat imaging experiments for noise suppression or multidimensional purposes for particles undergoing Brownian motion, since the "image" is constantly changing. However, so long as system statistics remain stationary from one scan to the next, and motion during each acquisition is negligible, repeatability of statistical measurements is entirely feasible.

During early development of NMR imaging in the 1970s, Mansfield and co-workers explored analogies between scattering amplitudes and NMR signals, and even considered determination of individual nuclear positions from NMR (*2–6*), although at the time it was set aside in favor of imaging approaches (*6*). Structural studies based on mobility developed even earlier by using gradient echo methods (*7, 8*), especially the pulsed-gradient spin echo (PGSE) experiment developed by Stejskal *et al.* (*9, 10*). This method not only measures diffusion coefficients but also provides information on microstructures in the host material when boundaries hinder normal diffusive transport (*11–13*).

In 1983, Kärger and Heink directly obtained the distribution $P(\Delta\mathbf{r}, t)$ of displacements $\Delta\mathbf{r}$ occurring during a transport time t by Fourier transforming the PGSE data as a function of magnetic field gradient strength (*14*). More recently, Cory and Garroway (*15*) demonstrated that for transport times so long that initial and final particle positions become uncorrelated, the distribution $P(\Delta\mathbf{r}, t)$ approaches the fluid density autocorrelation function of the particle container (or an average of such functions when the system consists of an ensemble of such containers). They illustrated this principle for water in yeast cells, inferring the cell size of 5 μm directly from the displacement profile. They also realized that the data, regarded as a function of the gradient strength, are equivalent to that from a diffraction experiment, except that the diffraction "wave vector" is given by gradient pulse strength $\mathbf{q} = \gamma\mathbf{G}\tau$ instead of the momentum transfer $\Delta\mathbf{k}$ appearing in scattering theory (*16*). (Here γ is the magnetogyric ratio of the observed nuclear species, $\mathbf{G}$ is the pulsed magnetic field gradient vector, and τ is the duration of each of the pair of magnetic field gradient pulses.) Callaghan *et al.* have developed this analogy further for porous systems (*17*) and have presented data showing an unusual increase of signal amplitude with increasing gradient (*18*) analogous to Debye-Scherrer rings in crystallography (*16*). Invoking statistical isotropy for their system of 16-μm polystyrene beads immersed in water, they extended data taken along a single direction in q-space to a spherically symmetric 3-D function. A 3-D Fourier transformation extracted the radial density distribution characteristic of the beads in the sample. These developments [which we refer to as diffusive diffraction to avoid confusion with the existing area of x-ray dynamic diffraction (*16*)] have been recently reviewed by Cotts (*19*).

Pulsed gradient methods in general are a powerful way of characterizing systems where diffusion is affected by internal structure. Nevertheless, whether diffractive aspects are considered or not, this technique becomes inapplicable when transport is too slow, fails to reflect structure, or is entirely absent. We show here that structural information related to diffractive principles is still available for stationary objects, based directly on density variation.

In NMR imaging, spatial information is encoded by applying a magnetic field gradient $\mathbf{G}$ for a duration t so that a particle at position $\mathbf{r}$ develops a phase $\mathbf{k}\cdot\mathbf{r}$, where the spatial wave vector $\mathbf{k}$ is $\gamma\mathbf{G}t$, and where t is the time for which the gradient acts. The volume density of nuclei $\rho(\mathbf{r})$ is then extracted by Fourier transformation of the signal $S(\mathbf{k})$ considered as a function in $\mathbf{k}$-space (*20*):

$$\rho(\mathbf{r}) = \int S(\mathbf{k}) \exp(-i\mathbf{k}\cdot\mathbf{r})\, d^3\mathbf{k} \qquad (1)$$

The Wiener-Khintchin theorem, generally applied to stationary stochastic processes, relates correlation functions in one Fourier domain to the square of their transform in the other (*21, 22*). Applied here, we find that the density autocorrelation function $\Phi(\Delta\mathbf{r})$ is given by the Fourier transform of the square of the imaging data:

$$\Phi(\Delta\mathbf{r}) \equiv \langle\rho(\mathbf{r})\rho(\mathbf{r}+\Delta\mathbf{r})\rangle = \frac{1}{(2\pi)^3 V}\int |S(\mathbf{k})|^2 \exp(i\mathbf{k}\cdot\mathbf{r})\, d^3\mathbf{k} \qquad (2)$$

where $\langle\ \rangle$ denote an average over position $\mathbf{r}$ taken within the sample volume V. The

Materials Sciences Division, Lawrence Berkeley Laboratory, and Department of Chemistry, University of California, Berkeley, CA 94720.

*To whom correspondence should be addressed.

autocorrelation function can be regarded as a measure of how well the sample density comes back into registration with itself when displaced along $\Delta\mathbf{r}$ and is directly analogous to the Patterson function of x-ray scattering (*1, 16*). When statistics are isotropic it reduces to a radial distribution function. As Mansfield recognized, the nonlinear processing associated with squaring the signal modulus suppresses signal phase and makes image reconstruction impossible (*6*). Equation 2 shows, however, that the remaining amplitude-modulated component extracts statistical information from the original imaging data in a manner analogous to conventional scattering experiments (*17, 23*).

Illustrations of this concept are presented in Fig. 1. Two phantoms were prepared, consisting of 7-mm inside diameter glass tubes that were packed with nylon monofilament fibers with 0.56- and 0.33-mm measured diameters. The tubes were then filled with water, which provided a ^{1}H NMR signal from the interstitial volume.

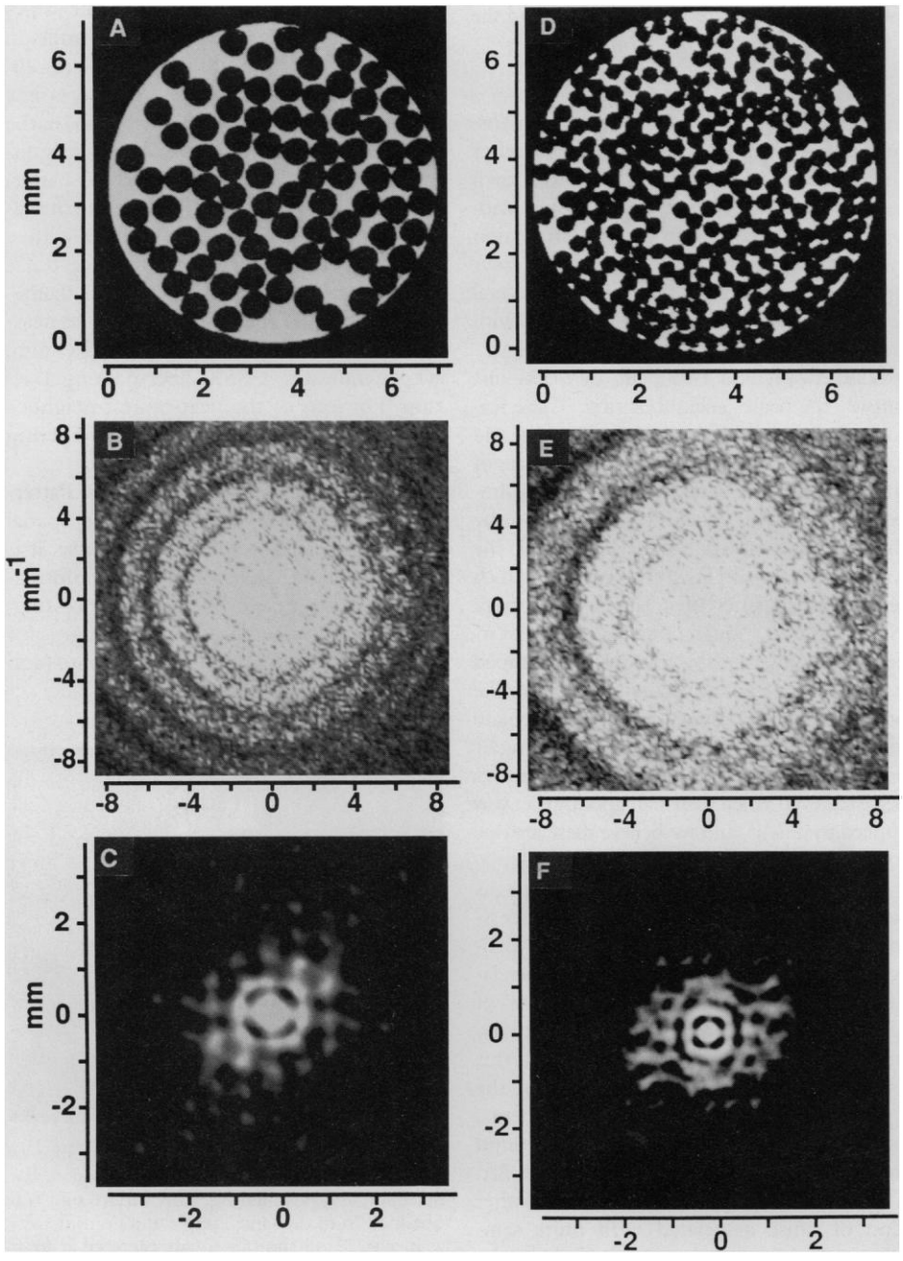

Fig. 1. (**A**) Slice image, (**B**) diffraction pattern, and (**C**) Patterson function for 0.56-mm fibers immersed in water in a 7-mm tube. The image shown in (A) is derived by Fourier transformation of NMR data, whose square, the NMR "diffraction pattern," shown in (B) reflects packing statistics. Fourier transformation of the diffraction pattern yields the Patterson function shown in (C), whose central disc reflects the fiber size and whose nearest-neighbor ring indicates packing density. (**D**) Slice image, (**E**) diffraction pattern, and (**F**) Patterson function for 0.33-mm fibers immersed in water in a 7-mm tube, to be compared directly with (A) to (C). The images and Patterson functions of the smaller fibers in (D) and (F), shown on the same scale, are smaller than those in (A) and (C), but the diffraction pattern in (E) is expanded relative to (B) since it scales according to spatial frequency.

Imaging data were taken on a Nalorac Cryogenics Quest 4300 spectrometer operating at 185 MHz by using an imaging sequence (*24*) whose excitation pulse selected a transverse slice ~0.5 mm thick. The unprocessed 2-D **k**-space data consist of complex numbers and have a rapid phase variation, but the derived density images of $\rho(\mathbf{r})$ in Fig. 1, A and D, each 128 pixels square, both confirm that the imaging signals $S(\mathbf{k})$ for the phantoms were correctly acquired and that diffusive and convective motions are negligible.

The phase-suppressed imaging data, $|S(\mathbf{k})|^2$, are shown in Fig. 1, B and E, for each of the two sizes. These are the NMR "diffraction patterns," defined in a **k**-space of spatial frequencies. Since the 7-mm circular images span ~118 pixels, the range is ±8.4 cycles per millimeter in each direction. This value determines the spatial resolution and depends on the size of the applied field gradients. The diffraction patterns show rings similar to those from x-ray powder patterns of granular structures, indicating an underlying regularity in the fiber separation (*16*). The dependence of this pattern on the NMR wave vector $\mathbf{k} = \gamma\mathbf{G}\tau$ here is directly analogous to the dependence of ordinary x-ray patterns on the momentum change $\Delta\mathbf{k}$ associated with scattering.

The Fourier transform of the diffraction-data yields fiber statistics even though the fibers exclude water and have zero associated signal. Indeed, if we express the spatially varying water density $\rho(\mathbf{r})$as$\rho_0 - \eta(\mathbf{r})$, where ρ_0 is the density of uniform water and $\eta(\mathbf{r})$ is the density distribution of excluded water occupied by the fibers, the water and fiber Patterson functions are related by

$$\langle \rho(\mathbf{r})\, \rho(\mathbf{r} + \Delta\mathbf{r}) \rangle$$
$$= \langle [\rho_0 - \eta(\mathbf{r})]\, [\rho_0 - \eta(\mathbf{r} + \Delta\mathbf{r})] \rangle$$
$$= \langle \eta(\mathbf{r})\eta(\mathbf{r} + \Delta\mathbf{r}) \rangle - 2\rho_0 \langle \eta(\mathbf{r}) \rangle + \rho_0^2 \quad (3)$$

Hence, the correlation functions for actual and excluded densities must be equal apart from the trivial base-line shift arising from the last two terms. Mansfield and Grannell (*5*) have also discussed this point using Babinet's principle (*25*, p. 178).

The actual Patterson functions shown in Fig. 1, C and F, obtained by Fourier transforming the diffraction data of Fig. 1, B and E, respectively, are shown on the same scale as the images in Fig. 1, A and D. The contraction of data toward the origin indicates that the fiber arrangement is ordered over a range much shorter than the image size, consistent with the random packing.

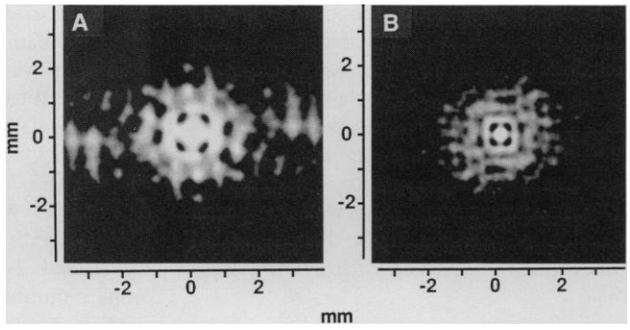

Fig. 2. Patterson functions for rotated samples. The (**A**) and (**B**) correspond, respectively, to Fig. 1, C and F, having been obtained from samples rotated 45° from the latter. Although the small fiber data have few features that follow the rotation, the large fiber data have secondary peaks which do; the latter are attributed to accidental long-range order arising from finite sample statistics. The "squareness" of the nearest-neighbor ring, the slightly diamond-shaped pattern of the central peak, and the crosslike bridges are noise artifacts.

The Patterson functions exhibit strong central peaks that are autocorrelations of each fiber with itself; as expected, these have approximately the same diameter as the fibers. A low-signal region, which indicates the excluded volume between fibers, surrounds this peak, and further out, rings appear that reflect the average separation of nearest neighbors.

The symmetry and prominence of these nearest-neighbor rings demonstrate the isotropy and short-range order of the packing statistics. These rings occur at ~0.7 and ~0.45 mm, respectively, for the 0.56- and 0.33-mm fibers, whose diameters would be the expected ring radii were the samples close packed. On this basis we predict packing fractions of $(0.56/0.7)^2 = 64\%$ and $(0.33/0.45)^2 = 54\%$ for the large and small fiber samples, respectively. Comparison of the actual fiber count (82 large and 247 small) with the number that would fit under close-packed conditions (~142 large and ~409 small) yields corresponding ratios of 58 and 60%. Agreement is better than the expected 12% uncertainty in packing based on a 6% uncertainty in ring size.

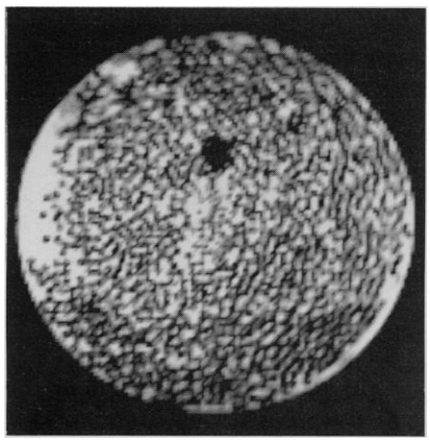

Fig. 3. Image of 0.1-mm slice of 0.12-mm fiber sample. The presence of the fibers is barely resolved in this 128 by 128 point image. The tube diameter is 7 mm as above.

In order to distinguish artifacts from correlations, we repeated the experiments after rotating each sample by ~45°; results shown in Fig. 2, A and B, may be compared with Fig. 1, C and F. (The patterns' inversion symmetry about $\Delta\mathbf{r} = 0$ follows from their being Fourier transforms of real data.) Several features of the large-fiber figures rotate with the sample, notably the large diagonal background swath and two pairs of peaks just outside the nearest-neighbor ring. These features therefore represent true anisotropy and long-range order, appearing here because of the finite statistics of the relatively small number of fibers. This conclusion is borne out by the weakness of similar structures for the small fiber data of Figs. 1F and 2B, which reflect the statistics of a larger population. However, the "squareness" of the nearest-neighbor ring, the slightly diamond-shaped pattern of the central peak, and the crosslike bridges between these two structures remain fixed, and therefore must reflect processing or noise artifacts or both. These latter features persisted even when Gaussian apodization was applied to $|S(\mathbf{k})|^2$, and we believe these are due to correlations between signal and noise.

Assessment of sensitivity and noise is complicated by the nonlinear data processing of this experiment. For example, processing pure noise would reveal the correlation function of the power spectrum of receiver noise, which can be removed by subtraction. Only the fluctuations of this estimate are a true nuisance. However, the signal itself contains random components, exemplified by the rotating background structures appearing in Figs. 1C and 2A. These anisotropic correlations are another kind of noise associated with finite sampling, which represent deviations from the ideal correlation for an infinite sample. In both cases, signal and noise input give clean statistical data, except for fluctuations due to finite sampling.

Diffraction data $|S(\mathbf{k})|^2$ are better acquired by using square law detection to suppress rapid phase oscillation characterizing image information. Since Patterson functions extend over a much smaller range than the full image, a correspondingly lower sampling rate is required, and noise can be reduced by tighter filtering before digitization. An image, obtained conventionally, of a 6-mm tube packed with smaller 0.12-mm monofilament fibers, resolved to 128 points per side, is shown in Fig. 3. The Patterson function of the same sample, derived from data acquired using square law detection prior to filtering and digitization, is shown in Fig. 4. Gradients here are four times larger than for the image, so the Patterson function has correspondingly higher resolution even though there are still 128 points on a side. In addition to reducing noise, the filters prevent aliasing and isolate the significant central 1.7-mm portion of the Patterson function. The central fiber autocorrelation function in Fig. 4 is a cone with a slightly flared base. As expected, the 0.22-mm base radius is equal to one fiber diameter within experimental accuracy. The nearest-neighbor ring radius is ~0.18 mm, which indicates a 35% fiber-packing fraction. Portions of the next-nearest-neighbor rings are also visible, because tighter filtering has greatly suppressed noise effects.

These examples show how NMR Patterson functions can characterize the spatial statistics of small structures. Of course, it is always possible to extract statistical information from an image by digital processing, but there are compelling advantages for adopting the diffraction viewpoint (and

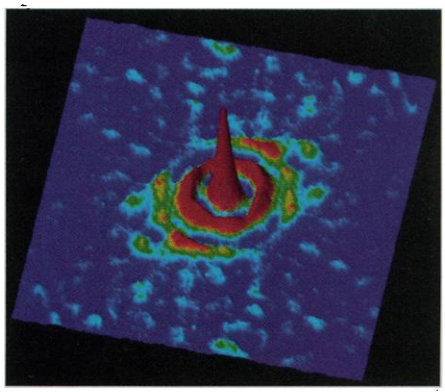

Fig. 4. Patterson function of 0.1-mm slice of 0.12-mm fiber sample obtained using square law detection and postfiltering. This surface plot was obtained from the same sample and point density as in Fig. 3, but shows a square centered at $\Delta\mathbf{r} = 0$, approximately 1.7 mm on a side (that is, expanded four times larger than the image scale of Fig. 3). This expansion was achieved by using larger gradients in conjunction with square law detection and tighter anti-alias filtering. The central fiber autocorrelation function appears as a conical structure with a base twice the fiber diameter. Both nearest- and next-nearest-neighbor rings are visible.

square law detection), which fall under the headings of acquisition economy and, most generally, statistical regularity.

Acquisition economy means that fewer points are required to specify the Patterson function of a sample compared to its image. Although range in **k**-space (maximum gradient level and duration) must be sufficient for good spatial resolution for both diffraction and imaging, density of data points in **k**-space determines the spatial range, and this is much smaller for diffraction.

Statistical regularity refers to the aforementioned fact that statistical descriptions of systems often possess higher symmetry than the system itself. We have indicated that statistical characterization permits repeated signal acquisition for rearranging systems when statistical data are time-dependent but imaging data are not. Likewise, the rotational symmetry of Fig. 1, B and E, demonstrates angular symmetry, showing how 2-D features can be characterized by 1-D radial information. The lower density of required points allows a correspondingly reduced sample rate. The fact that diffraction information resides in an intrinsically narrower bandwidth than data for comparably resolved images is encouraging, and preliminary results indicate that Patterson functions are indeed "cleaner" than corresponding images. Signal averaging is quicker because the data are simpler, and extensions that would impractically lengthen imaging experiments become possible. For example, contrast-enhancing preparation sequences (*26*) may be applied to selectively weighted regions according to spectroscopic or mobility differences (*27*) between regions with different morphologies or composition. The statistical approach described here should eventually make NMR studies of small-scale inhomogeneities in plastics, ceramics, and structural materials a practical possibility.

REFERENCES AND NOTES

1. A. L. Patterson, *Phys. Rev.* **46**, 372 (1934).
2. P. Mansfield and P. K. Grannell, *J. Phys. C* **6**, L422 (1973).
3. ———, A. N. Garroway, D. C. Stalker, *Proc. Coll. Ampere (Krakow)* **1**, 16 (1973).
4. P. Mansfield, P. K. Grannell, A. A. Maudsley, *Proc. Coll. Ampere (Nottingham)* **2**, 431 (1974).
5. P. Mansfield and P. K. Grannell, *Phys. Rev. B* **12**, 3618 (1975).
6. P. Mansfield, *Contemp. Phys.* **17**, 553 (1976).
7. E. L. Hahn, *Phys. Rev.* **80**, 580 (1950).
8. H. Y. Carr and E. M. Purcell, *ibid.* **94**, 630 (1954).
9. E. O. Stejskal, *J. Chem. Phys.* **43**, 3597 (1965).
10. J. E. Tanner and E. O. Stejskal, *J. Phys. Chem.* **49**, 1768 (1968).
11. P. T. Callaghan, *Aust. J. Phys.* **37**, 359 (1984).
12. P. Stilbs, *Prog. Nucl. Magn. Reson. Spectrosc.* **19**, 1 (1987).
13. J. Kärger, H. Pfeiffer, W. Heink, *Adv. Magn. Reson.* **12**, 1 (1988).
14. J. Kärger and W. Heink, *J. Magn. Reson.* **51**, 1 (1983).
15. D. G. Cory and A. N. Garroway, *Magn. Res. Med.* **14**, 435 (1990).
16. L. H. Schwartz and J. B. Cohen, *Diffraction from Materials* (Springer-Verlag, Berlin, ed. 2, 1987).
17. P. T. Callaghan, D. McGowan, K. J. Packer, F. O. Zelaya, *J. Magn. Reson.* **90**, 177 (1990).
18. P. T. Callaghan, A. Coy, D. McGowan, K. J. Packer, F. O. Zelaya, *Nature* **351**, 467 (1991).
19. R. M. Cotts, *ibid.*, p. 443.
20. D. Twieg, *Med. Phys.* **10**, 610 (1983).
21. N. Wiener, *The Fourier Integral and Certain of Its Applications* (Dover, New York, 1958).
22. A. Khintchin, *Math. Ann.* **109**, 604 (1934).
23. S.-H. Chen, B. Chu, R. Nossal, Eds., *Scattering Techniques Applied to Supramolecular and Nonequilibrium Systems* (Plenum, New York, 1980).
24. W. A. Edelstein, J. M. S. Hutchison, G. Johnson, T. Redpath, *Phys. Med. Biol.* **25**, 751 (1980).
25. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, M. Hammermesh, Transl. (Addison-Wesley, New York, 1962).
26. R. R. Ernst, G. Bodenhausen, A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions* (Oxford Univ. Press, Oxford, 1990).
27. D. B. Twieg, J. Katz, R. M. Peshock, *Magn. Reson. Med.* **5**, 32 (1987); G. C. Chingas, J. B. Miller, A. N. Garroway, *J. Magn. Reson.* **66**, 530 (1986).

28. We acknowledge the advice, encouragement, and support of A. Pines, in whose laboratory this work was carried out, and also acknowledge helpful discussions with M. Denn. Supported by the U.S. Department of Energy through the Director, Office of Energy Research, Office of Basic Energy Sciences, Materials Sciences Division, under contract no. DE-AC03-76SF00098.

2 October 1991; accepted 19 December 1991

On the Probability of Matching DNA Fingerprints

NEIL J. RISCH* AND B. DEVLIN

Forensic scientists commonly assume that DNA fingerprint patterns are infrequent in the general population and that genotypes are independent across loci. To test these assumptions, the number of matching DNA patterns in two large databases from the Federal Bureau of Investigation (FBI) and from Lifecodes was determined. No deviation from independence across loci in either database was apparent. For the Lifecodes database, the probability of a three-locus match ranges from 1 in 6,233 in Caucasians to 1 in 119,889 in Blacks. When considering all trios of five loci in the FBI database, there was only a single match observed out of more than 7.6 million comparisons. If independence is assumed, the probability of a five-locus match ranged from 1.32×10^{-12} in Southeast Hispanics to 5.59×10^{-14} in Blacks, implying that the minimum number of possible patterns for each ethnic group is several orders of magnitude greater than their corresponding population sizes in the United States. The most common five-locus pattern can have a frequency no greater than about 10^{-6}. Hence, individual five-locus DNA profiles are extremely uncommon, if not unique.

VNTR (VARIABLE NUMBER TANDEM repeat) loci are used to generate the "DNA fingerprints" that have been presented as evidence in criminal and paternity cases. These loci are extremely polymorphic, having potentially hundreds of alleles at a single locus (*1*). Any particular genotype at a collection of such loci is deemed to be so rare that many forensic scientists believe the probability two unrelated individuals have matching genotypes across a set of loci to be extremely small. When many VNTR loci are tested (for example, up to five), the probability of a matching pattern occurring by chance has been reported in criminal cases to be extremely small, often on the order of 10^{-7} to 10^{-8} or even less, and sometimes the probability suggests less than one matching pattern in the total population of North America. Yet it is often argued that these probabilities are calculated in a conservative fashion, that is, the true probabilities are even smaller (*2, 3*). Others have argued, however, that the probabilities are invalid and are unrealistically small [(*4*) but see (*5*)].

In forensic cases, probability estimates are obtained by the multiplication rule. For multiplication to be valid, the events must be statistically independent. Statistical independence allows one to multiply allele frequencies within a locus to derive a single-locus genotype probability and to multiply genotype probabilities across loci to obtain a multilocus genotype probability. Statistical independence within a locus is referred to as Hardy-Weinberg equilibrium (HW), while statistical independence across loci is called linkage equilibrium (LE).

The assumption of independence, both within and across loci, has been challenged (*6*). For forensics, the reference population is divided into major ethnic components: for instance, Caucasians, Blacks, and Hispanics. Sometimes Hispanics are further subdivided by geography. The argument is put forth that none of these ethnic components is genetically homogeneous and that mating patterns are

N. J. Risch, Division of Biostatistics, Department of Epidemiology and Public Health, and Department of Genetics, Yale University, Post Office Box 3333, 60 College Street, New Haven, CT 06510.
B. Devlin, Division of Biostatistics, Department of Epidemiology and Public Health, Yale University, Post Office Box 3333, 60 College Street, New Haven, CT 06510.

*To whom correspondence should be addressed.