Time for DNA Disclosure

THE LEGISLATION THAT ESTABLISHED THE U.S. NATIONAL DNA INDEX SYSTEM (NDIS) IN 1994 explicitly anticipated that database records would be available for purposes of research and quality control “if personally identifiable information is removed” [42 U.S.C. Sec 14132(b)(3)(D)]. However, the Federal Bureau of Investigation (FBI), which controls the database, has published no research derived from NDIS and has declined to disclose these records to academic scholars. The National Research Council recently noted that “methods developed in crime laboratories to aid in law enforcement” would benefit from the contributions of academic scientists (1). We believe the time has come for the FBI to release anonymized NDIS profiles to academic scientists for research that will benefit criminal justice.

Disclosure of NDIS profiles would allow independent scientists to evaluate some of the population genetic assumptions underlying DNA testing using a database large enough to allow more sensitive evaluation of population structure. The publicly available population databases used to date for statistical estimation of the frequency of DNA profiles are relatively small ($N = 1000$), consisting of convenience samples analyzed over a decade ago (2, 3). In contrast, NDIS has grown to over 7 million complete 13-locus short tandem repeat (STR) genotypes (4). Analysis of these data would allow more powerful tests of independence within and between loci, as well as assessment of the efficacy of the theta factors used to compensate for population substructure. (To the extent the data are identified by state, analysis of NDIS data could also yield important information about the most appropriate geographic scaling for allele frequency estimates.)

The large sample size also allows real-world tests of propositions that previously have been addressed only by simulation. For example, it would allow tests of the frequency with which three-person mixtures could produce profiles consistent with two contributors (5); kinship analysis could allow assessment of how match probabilities are affected by the number of close relatives in the database (6, 7); and multivariate analysis could be used to evaluate the extent to which DNA profiles cluster due to identity by descent. As studies of smaller databases have shown, researchers need not know a priori the precise number of relatives in the database, nor their ethnic/racial background, to perform these assessments (6, 8). Indeed, scholars who have examined smaller databases have called for examination of national databases (6, 8, 9). Access to the anonymized 13-locus genotypes would allow more powerful analyses of these important issues than was previously possible.

Analysis of NDIS can also yield valuable insights into the frequency and circumstances under which certain typing errors may occur. A review of a government database from Victoria, Australia, containing 15,021 9-locus STR profiles shows how important such a review can be for “quality control purposes” (10, 11). The study found an error rate of about 1 in 300 for the typing of reference samples, which raises concerns about missed opportunities to develop investigative leads.

Disclosure of NDIS profiles would not violate any meaningful privacy interests (12). (There are easier ways to determine whether an individual has a criminal record than searching such a database, and the profiles would not be useful for medical diagnoses.) The profiles in the Victoria, Australia, database have been widely circulated for years with no known harm occurring. The U.S. government regularly argues to courts that broad mandatory DNA collection statutes are not unconstitutional precisely because the 13 genetic loci are noncoding and thus have no power to reveal any sensitive information. Moreover, as most research scientists know well, the government frequently releases sensitive information under controlled conditions to verified researchers. Even within the criminal justice context, law enforcement officials have made available data about the age, race, gender, geographic residence, and a wide range of other information about criminal offenders so that researchers can conduct studies aimed at improving and enhancing effective law enforcement.

www.sciencemag.org  SCIENCE  VOL 326  18 DECEMBER 2009  PERSPECTIVES
Some have suggested that the release of NDIS profiles would be unduly burdensome (33), but the relevant fields in the SQL database could be copied in a matter of minutes.

Open access to data is a fundamental tenet of science. The need for openness was reinforced by the recent National Research Council report, which decried the insularity of forensic science and called for greater involvement of the academic community in assessment, validation, and improvement of forensic science methods (1). Law enforcement should honor the norms of science and open the NDIS and other government DNA databases to independent scientific scrutiny. Doing so poses no meaningful risk and can only strengthen the quality of forensic DNA analysis.


1Department of Biological Sciences, Wright State University, Dayton, OH 45435, USA. 2Institute of Genetics, University College London, London WC1E 6BT, UK. 3Office of the Public Defender, San Francisco, CA 94103, USA. 4Gene Codes Corporation, Ann Arbor, MI 48108, USA. 5Benjamin and DesPortes, PC, Richmond, VA 23218–2464, USA. 6Office of Biochemistry and Medicine, New York University School of Medicine, New York, NY 10016, USA. 7STAR Institute, Stanford University, CA 94305, USA. 8Department of Computer Science and Engineering, Wright State University, Dayton, OH 45435–0001, USA. 9Institute of Cognitive Neuroscience, University College London, London WC1E 6BT, UK. 10Lexigen Science and Law Consultants, San Francisco, CA 94107, USA. 11Office of the Public Defender, Trial Team, Hastings, NJ 07533, USA. 12Forensic Bioinformatics Inc., Fairborn, OH 45324, USA. 13Department of Biology and Criminal Justice, Boise State University, Boise, ID 83725–1515, USA. 14Department of Criminal Justice Administration, California State University East Bay, Hayward, CA 94542, USA. 15The Forensic Institute, Glasgow G2 6HJ, UK. 16Forensics Division, Maryland Office of the Public Defender, Baltimore, MD 21202, USA. 17Institute for Forensic Science Administration, Silberman College of Criminal Justice, Fairleigh Dickinson University, Madison, NJ 07940, USA. 18Molecular Forensics Laboratory, University of Maine, Orono, ME 04469, USA. 19Department of Urban and Environmental Policy and Planning, Tufts University, Medford, MA 02155, USA. 20University of California, Los Angeles, School of Law, Los Angeles, CA 90095–1476, USA. 21Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697–2525, USA. 22University of California School of Law, Berkeley, CA 94720–7200, USA. 23Department of Engineering, Pennsylvania State University Beaver, Monaca, PA 15061, USA. 24Department of Biology, Stanford University, Stanford, CA 94305–0520, USA. 25Seton Hall University School of Law, South Orange, NJ 07079, USA. 26Stanford University School of Law, Stanford, CA 94305–8610, USA. 27Forensic DNA Consultant, Mountain View, CA 94041, USA. 28Department of Environmental and Forest Biology, State University of New York, Syracuse, NY 13210, USA. 29Department of Chemistry and Chemical Biology and the Forensic and Investigative Sciences Program, Indiana University Purdue University Indianapolis, Indianapolis, IN 46202, USA. 30Department of Integrative Biology, University of California, Berkeley, CA 94720–3140, USA. 31Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720–1776, USA. 32Department of Statistics and Program in Biostatistics, University of California, Berkeley, CA 94720–3860, USA. 33Department of Statistics, Texas A&M University, College Station, TX 77843–3143, USA. 34Office of the Public Defender, Hennequin County–Fourth Judicial District, Minneapolis, MN 55445–1600, USA. 35SciLawForensics Ltd., Germantown, MD 20874, USA. 36Department of Mathematics and Statistics, Wright State University, Dayton, OH 45435–0001, USA. 37Department of Criminology, Law and Society, University of California, Irvine, CA 92697–7080, USA. 38Northern Virginia Capital Defender Office, Arlington, VA 22201, USA. 39Departments of Mathematics and Statistics, Northwestern University, Evanston, IL 60208, USA.

*To whom correspondence should be addressed. E-mail: Dan.Krane@wright.edu

References

Letters to the Editor
Letters (~300 words) discuss material published in Science in the previous 3 months or issues of general interest. They can be submitted through the Web (www.submit2science.org) or by regular mail (1200 New York Ave., NW, Washington, DC 20005, USA). Letters are not acknowledged upon receipt, nor are authors generally consulted before publication. Whether published in full or in part, letters are subject to editing for clarity and space.