Behavior Genetics, Vol. 27, No. 3, 1997

DNA by Mail: An Inexpensive and Noninvasive Method for Collecting DNA Samples from Widely Dispersed Populations

Bernard Freeman,[1,2] John Powell,[1] David Ball,[1] Linzy Hill,[1] Ian Craig,[1] and Robert Plomin[1]

Received 23 Dec. 1996—Final 17 Jan. 1997

As specific genes are identified that are associated with behavior, it becomes increasingly important for behavioral geneticists to be able to incorporate these genes in their research. Rather than using blood, DNA can be extracted from cheek swabs, which makes it possible to obtain DNA inexpensively by mail from large, widely dispersed individuals. The purpose of this paper is to recommend this technique to the behavioral genetics community and to present results of our use of this technique to obtain DNA by mail for 114 2-year-olds and 116 adults.

KEY WORDS: Genes; DNA; genotyping; behavioral genetics; mail.

INTRODUCTION

Behavioral genetics is at the dawn of a new age in which quantitative genetics is beginning to merge with molecular genetics in the study of complex quantitative traits influenced by multiple genes and multiple environmental factors (Plomin *et al.*, 1994), as reflected, for example, in two recent behavioral genetics textbooks (McGuffin *et al.*, 1994; Plomin *et al.*, 1997). Behavioral genetics is not synonymous with quantitative genetic designs such as twin and adoption studies; it encompasses both quantitative genetic and molecular genetic approaches to investigate the genetic and environmental origins of individual differences in behavior. Despite the advent of molecular genetic techniques that are increasingly able to identify genes, not just for simple single-gene disorders, but also for complex behavioral dimensions and disorders, quantitative genetic research will continue to be important in pointing to the most heritable components and constellations of behavioral traits. Moreover, quantitative genetics can also go beyond the rudimentary questions of *whether* and *how much* genes influence behavior to address the question *how*, i.e. what are the mechanisms that lie in the pathway between genes and behavior. Quantitative genetics can go beyond heritability estimates by asking developmental questions about changes over time in the effects of genes; multivariate questions about heterogeneity, comorbidity, and the links between the normal and the abnormal; and questions about interactions and correlations with the environment.

Behavioral geneticists are now able to add molecular genetics to their tool kit to address these questions about the relationship between genes and behavior but with the major advantage of using identified genes, even though a particular gene is likely to account for only a small proportion of variance. Although it is difficult and expensive to find genes, especially for complex traits, it is relatively easy and inexpensive to use genes that have already been identified (Plomin and Rutter, 1997). Few behavioral geneticists will join the hunt for

[1] Social, Genetic and Developmental Psychiatry Research Centre, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London SE5 8AF, United Kingdom.

[2] To whom correspondence should be addressed. Tel.: +44 (171) 919-3011. Fax: +44 (171) 919-3866. E-mail: b.freeman@iop.bpmf.ac.uk.

0001-8244/97/0500-0251$12.50/0

genes, but it is crucial for the field that behavioral geneticists be prepared to use genes as they are found in order to investigate the relationship between identified genes and behavior.

Although initially molecular genetic techniques seem daunting, they are in fact at least as straightforward as, and no more expensive than, other physiological measures, such as cortisol assays. As with the radioimmunoassay required to assess cortisol, behavioral geneticists are unlikely to do the laboratory work themselves. In university settings, it is possible to collaborate with molecular geneticists for whom genotyping is routine. Moreover, like cortisol assays, it is possible to have the work done commercially.

There are three steps in the process: (1) obtaining biological material, (2) extracting the DNA, and (3) genotyping the DNA. Although attention is focused on the third step, the first two steps are the critical steps that inhibit behavioral geneticists from incorporating molecular genetics in their research. The purpose of this paper is to bring to the attention of the behavioral genetics community the availability of an inexpensive and noninvasive technique that makes it possible to obtain DNA from large samples though the mail. Although this technique is becoming widely used, its potential usefulness for behavioral genetic research and the relative lack of previous publications on the topic in the behavioral genetics field would seem to warrant a brief note.

Blood is usually used to obtain DNA. However, it is now possible to obtain DNA much less invasively through the use of cheek cells obtained by rubbing the inside of the cheek with a cotton swab (i.e., a Q-tip) or by a mouth rinse. Samples can be collected through the mail, which greatly reduces costs. This procedure has been described previously (e.g., Meulenbelt *et al.*, 1995; Richards *et al.*, 1993); the purpose of this paper is to describe the procedures that we have used and the results we have obtained through the mail for 116 adults and for 114 2-year-olds whose cheek swabs were obtained by their parents.

METHODS

Subjects

We are currently analyzing DNA obtained by mail in two groups. The first is a group of 2-year-old twins from an ongoing study of all twins born in England and Wales in 1994, 1995, and 1996 called the Twins' Early Development Study (TEDS). The present results are based on 78 families selected for a molecular genetic analysis of developmental delays. Of the 78 families, 73% agreed to participate in the DNA study.

The second sample is from a German twin study of adult personality in which 122 individuals were selected for high or low neuroticism (Reimann and Angleitner, 1997). A response rate of 92% was obtained for cheek swab samples mailed directly to England by the subjects.

Procedures

Kits were mailed to subjects with explanatory material (available upon request) and a prepaid return envelope. DNA was extracted, quantified, and genotyped.

Cheek Swab Procedure. A kit is sent to the participants containing a prepaid return envelope, 10 cotton wool buds, and a 15-ml tube containing 2.5 ml of a storage/preservative solution [STE buffer (100 m*M* NaCl, 10 m*M* Tris–HCl, pH 8, 10 m*M* EDTA, pH 8), with 0.2 mg/ml proteinase K and 0.5% SDS] (Meulenbelt *et al.*, 1995).

The subjects are asked to swab the inside of their child's mouth or their own mouth three times during 1 day, before breakfast, lunch, and the evening meal, using three, three, and four buds, respectively. Each swab is carried out for approximately 20 s using a different area of the mouth. The swabs are then immediately placed in the tube with the storage medium and sealed. All 10 swabs are placed into a single storage tube. When all the samples are collected, they are returned to us by post.

DNA Extraction. Unlike blood, which can be frozen, DNA from cheek swabs needs to extracted within a few weeks to prevent degradation, although procedures are being developed that will allow long-term storage of samples prior to extraction (personal communication with E. Slagboom). Our samples were processed within 1 to 3 weeks after collection. On receiving the samples, they are accumulated and stored at room temperature, for up to 1 week before processing using a semiautomated large-scale cleaning method. The samples are initially incubated at 65°C for 2 h to activate the proteinase K and release the DNA. To

recover all the liquid, the buds are partially pulled out of the tube, secured in this position using a plastic pull-tie, and centrifuged at 1000 rpm (350*g*) for 5 min.

The DNA is then purified using a large-scale automated robot (Autogen 740), although DNA purification can be done manually or by commercially available kits. The samples are centrifuged to remove debris. The supernatant is then manually decanted. The robot then processes samples through a modified phenol/chloroform extraction and resuspends the DNA in water or TE (10 mM Tris-HCl, pH 8, 10 mM, EDTA, pH 8).

The DNA is quantified spectrophotometrically and an aliquot electrophoresed on a 1% agarose gel stained with ethidium bromide and compared to a DNA standard to confirm the results.

PCR and Genotyping Two genotyping procedures have been used on the DNA from the twin samples, which demonstrates the quality of the DNA. Single marker PCRs have been carried out on a small number of samples using 10 ng DNA template, 0.2 m*M* dNTP, 1× PCR Buffer, 1 U Taq polymerase, and FAM-labeled TH marker primer sets (0.5 μ*M*) (Polymeropoulos *et al.*, 1991b). The PCR was carried out using the cycling program of heating to 94°C for 3 min, then 94°C for 30 s, 55°C for 30 s, and 72°C for 1 min, for 30 cycles and a final extension of 72°C for 10 min.

Ninety-six of the DNA samples have also been used in a seven-marker multiplex PCR reaction (Kimpton *et al.*, 1993) to test for twin zygosity. The reaction contained 10 ng DNA template, 0.2 μ*M* dNTP, 1x Buffer, 1 U Taq polymerase, and primer sets FAM-labeled HUMCD4 (0.125 μ*M*) (Edwards *et al.*, 1991b), TET-labeled HUMDHFRP2 (0.25 μ*M*) (Polymeropoulos *et al.*, 1991c), FAM-labeled HUMCYAR03 (0.25 μ*M*) (Polymeropoulos *et al.*, 1991a), FAM-labeled HUMAPOAII (0.0625 μ*M*) (Zulliani *et al.*, 1990, Hata *et al.*, 1991), TET-labeled HUMPLA2A (0.1875 μ*M*) (Polymeropoulos *et al.*, 1990b), FAM-labeled HUMFABP (0.26 μ*M*) (Polymeropoulos *et al.*, 1990a; Edwards *et al.*, 1991a), and HEX-labeled D3S1300 (0.0625 μ*M*) (Gyapay *et al.*, 1994). The PCR was carried out using the cycling program 94°C for 3 min, then 94°C for 60 s, 55°C for 60 s, and 72°C for 60 s for 30 cycles, and a final extension of 72°C for 10 min.

The PCR products were then either diluted 1-in-10 for the single-marker PCRs or used undiluted for the seven-marker multiplex PCR. Eight-tenths of a microliter were added to 1 μl formamide/EDTA mix (5:1 deionized formamide:50 m*M* EDTA with 50 mg/ml blue dextran) and 0.2 μl Genescan 500 size standard. The entire 2 μl was electrophoresed on a ABI PRISM 377 through a 4.25% polyacrylamide gel. The size fragment analysis was carried out using ABI PRISM Genescan v 2.0.0 and ABI PRISM Genotyper v 1.1 software.

RESULTS

Samples were collected and DNA extracted from the 114 2-year-olds and the 116 adults. Figure 1 shows the distribution of DNA yields for the 114 2-year-olds. The average yield was 32 μg of DNA, but with a wide range, from 3.2 to 110.8 μg. However, even the smallest amount permits 300 or more genotypings. In addition, the inexpensiveness of obtaining DNA by mail makes it possible to return to subjects with a low yield and obtain additional cheek swabs.

A small number of these DNA samples were genotyped for a tetranucleotide repeat marker in the gene for tyrosine hydroxylase. Ten nanograms of DNA template was used for the PCR reaction and analyzed on the ABI PRISM 377, with results shown in Fig. 2 for 10 individuals for this five-allele marker. Ninety-six of these DNA samples were also genotyped using the seven-marker multiplex PCR. Only one of the samples failed to amplify.

Figure 3 shows the distribution of DNA yields for the 116 adults. The average yield, 38 μg, was only slightly greater than the yield for the children and the range was just as wide, from less than 1 to 108 μg. However, the two samples with less than 1 μg of DNA contained sufficient DNA templates to yield PCR product successfully. In addition, the quality of all the DNA was assessed by genotyping with several markers, of which all 116 successfully yielded PCR products.

DISCUSSION

The use of cheek cells to obtain DNA by mail opens up many possibilities for behavioral genetics research. It facilitates the search for genes but it will have an even larger impact on research that uses identified genes to investigate developmental, multivariate, and environmental mechanisms in-

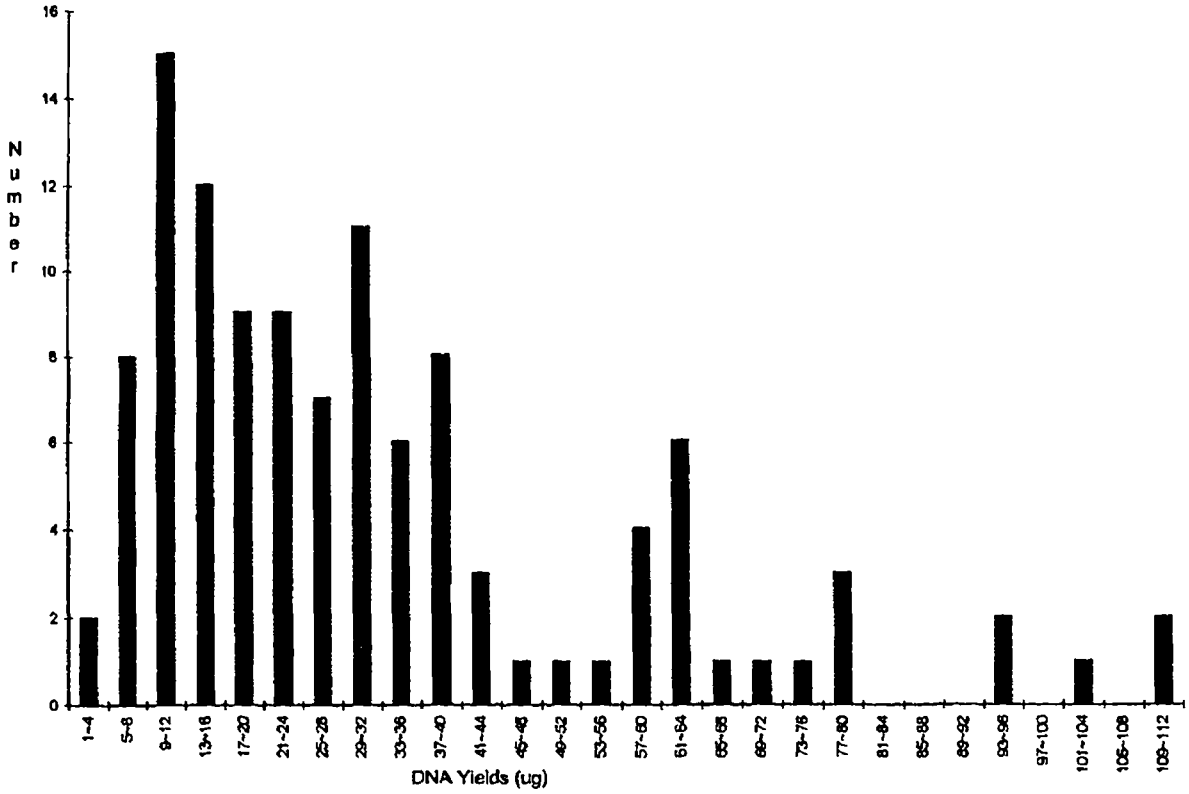

Fig. 1. Distribution of DNA yields (μg) obtained from 114 2-year old children.

volved in the interplay between genes and behavior (Plomin and Rutter, 1997).

Cheek swabs are an inexpensive and noninvasive method for obtaining DNA. For donors, it requires little time or effort and for researchers it is an inexpensive alternative to conventional blood sampling. Responses to date indicate few problems with the procedure from the subjects' perspective, and for researchers, the technique makes it possible to obtain DNA by mail rather than in person, which requires suitably trained personnel. The typical yields and quality of DNA obtained by mail are sufficient for many hundreds or even thousands of PCR reactions.

Cheek swabs are not however a replacement for venepuncture. For example, for research in hospital settings where blood can easily be obtained, blood will continue to be collected. Typical blood extractions yield in excess of 100 μg of DNA, three to four times the amount obtained from 1 day's cheek swabs, although more cheek swabs can easily be obtained.

Nonetheless, for many behavioral genetics researchers, obtaining DNA by mail using cheek swabs is an attractive alternative to blood. Moreover, parents and children are, of course, much more willing to allow a simple mouth swab compared to the trauma of a needle. The materials can be sent in a small envelope to the donor in any location, for the cost of postage. The samples can easily and effectively be taken by the donors themselves or, in the case of children, by their parents and then sent back to the researchers using a prepaid envelope. The cost of the materials is low, totaling under $1 for the tube, buffer ingredients, and cotton buds. The low cost of this procedure also means that it is feasible to obtain large samples and resample missing individuals.

Associations between DNA markers and behavior are appearing at an increasing rate and we

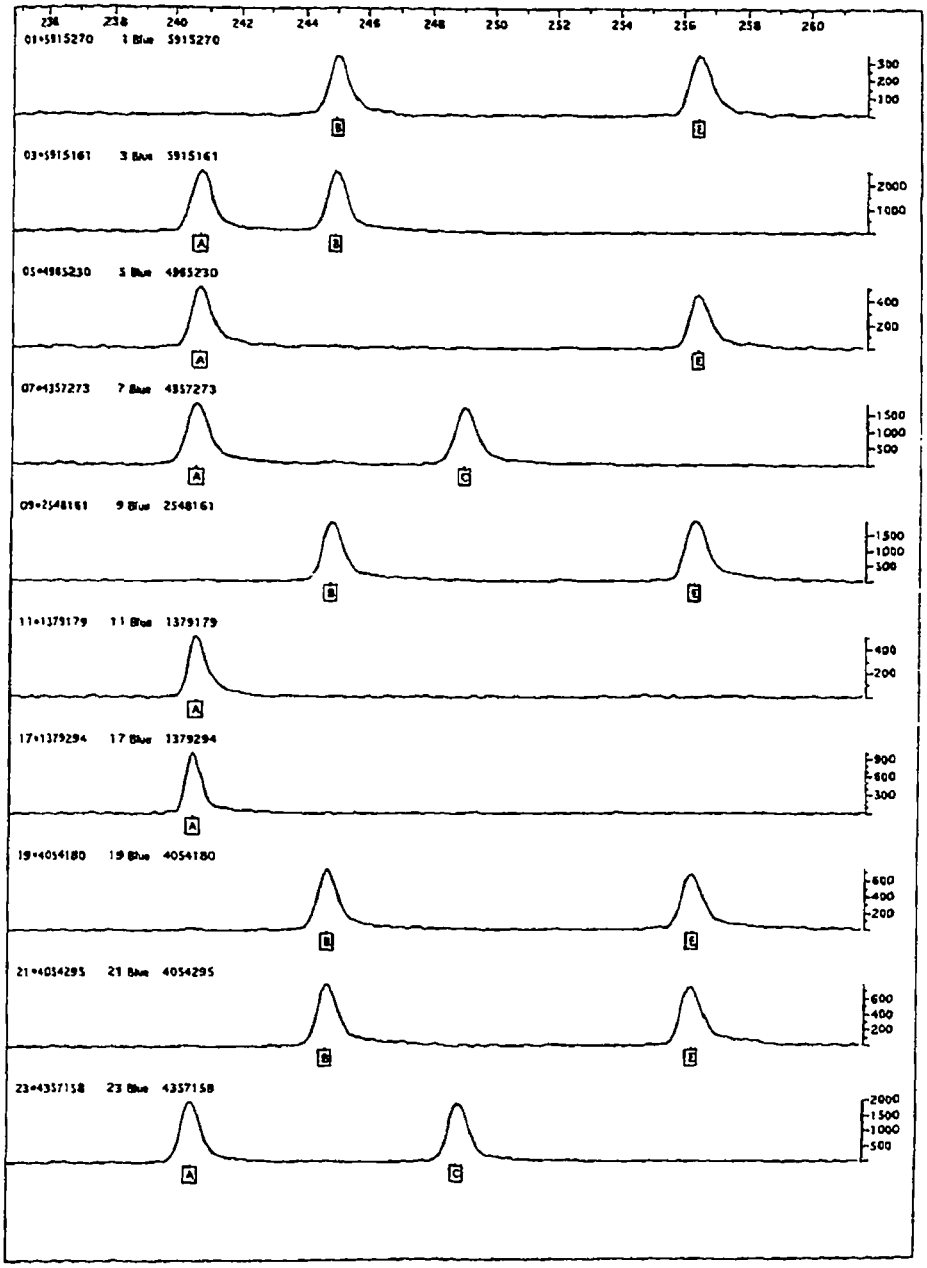

Fig. 2. Representative output obtained from fluorescent genotyping. Data from 10 individuals genotyped using the tyrosine hydroxylase tetranucleotide repeat polymorphism (Polymeropoulos *et al.*, 1991b). Horizontal scale is in base pairs; vertical scale, in fluorescent intensity (individually scaled). Alleles are assigned according to positions. For example, individual 1 (ID Number 5915270) is heterozygous for allele *B* (245 base pairs) and allele *E* (257 base pairs), while individual 3 (ID Number 5915161) is also heterozygous and shares allele *B* with individual 1 but has a different second allele, *A* (241 base pairs). Individuals 11 and 17 are homozygous for allele A. The rare allele *D* is not represented in these 10 individuals.

predict that behavioral geneticists will routinely use DNA markers in their research. The time has come for behavioral geneticists to consider routinely obtaining DNA. With DNA in the freezer, it is possible to genotype subjects for relevant genes when they are found. In our experience,

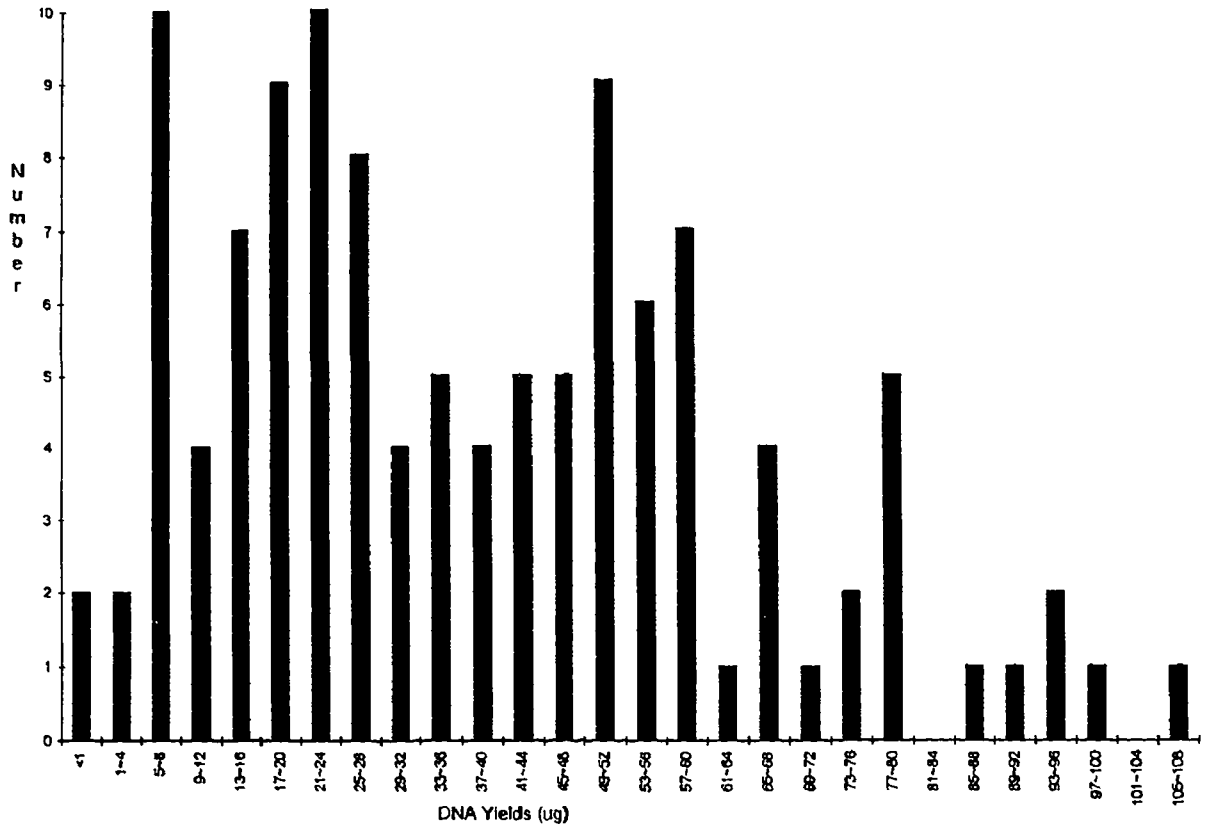

Fig. 3. Distribution of DNA yields (μg) obtained from 116 adults.

cheek swabs make it possible to obtain reasonable amounts of DNA in a simple and inexpensive manner by mail.

REFERENCES

Edwards, A., Civitello, A., Hammond, H. A., and Caskey, C. T. (1991a). DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am. J. Hum. Genet.* **49**:746–756.

Edwards, M. C., Clemens, P. R., Tristan, M., Pizzuti, A., and Gibbs, R. A. (1991b). Pentanucleotide repeat length polymorphism at the human CD4 locus. *Nucleic Acids Res.* **19**:4791.

Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lathrop, M., and Weissenbach, J. (1994). The 1993–1994 Genethon Human Genetic Linkage Map. *Nature Genet.* **7**:246–339.

Hata, A., and Lalouel, J.-M. (1991). Micro VNTR (CTTT)n at the apo C-III locus. *Nucleic Acids Res.* **19**:5098.

Kimpton, C. P., Gill, P., Walton, A., Urquhart, A., Millican, E. S., and Adams, M. (1993). Automated DNA profiling employing multiplex amplification of short tandem repeat loci. *PCR Methods Appl.* **3**:13–22.

McGuffin, P., Owen, M. J., O'Donovan, M. C., Thapar, A., and Gottesman, I. I. (1994). *Seminars in Psychiatric Genetics*, Gaskell Press, London.

Meulenbelt, I., Droog, S., Trommelen, G. J., Boomsma, D. I., and Slagboom, P. E. (1995). High yield noninvasive human genomic DNA isolation method for genetic studies in geographically dispersed families and populations. *Am. J. Hum. Genet.* **57**:1252–1254.

Plomin, R., DeFries, J. C., McClearn, G. E., and Rutter, M. (1997). *Behavioral Genetics*, 3rd ed., W. H. Freeman, New York.

Plomin, R., and Rutter, M. (1997). Child development and molecular genetics: What do we do with genes once they are found? *Child Dev.* (in press).

Plomin, R., Owen, M. J., and McGuffin, P. (1994). The genetic basis of complex human behaviors. *Science* **264**:1733–1739.

Polymeropoulos, M. H., Rath, D. S., Xiao, H., and Merril, C. R. (1990a). Trinucleotide repeat polymorphism at the human intestinal fatty acid binding protein gene (FABP2). *Nucleic Acids Res.* **18**:7198.

Polymeropoulos, M. H., Rath, D. S., Xiao, H., and Merril, C. R. (1990b). Trinucleotide repeat polymorphism at the hu-

man pancreatic phospholipase A-2 gene (PLA2). *Nucleic Acids Res.* **18**:7468.

Polymeropoulos, M. H., Xiao, H., Rath, D. S., and Merril, C. R. (1991a). Tetranucleotide repeat polymorphism at the human aromatase cytochrome P-450 gene (CYP19). *Nucleic Acids Res.* **19**:195.

Polymeropoulos, M. H., Xiao, H., Rath, D. S., and Merril, C. R. (1991b). Tetranucleotide repeat polymorphism at the human tyrosine hydroxylase gene (TH). *Nucleic Acids Res.* **19**:3753.

Polymeropoulos, M. H., Xiao, H., Rath, D. S., and Merril, C. R. (1991c). Tetranucleotide repeat polymorphism at the human dihydrofolate reductase psi-2 pseudogene (DHFRP2). *Nucleic Acids Res.* **19**:4792.

Richards, B., Skoletsky, J., Shuber, A. P., Balfour, R., Stern, R. C., Dorkin, H. L., Parad, R. B., Witt, D., and Klinger, K. W. (1993). Multiplex PCR amplification from the CFTR gene using DNA prepared from buccal brushes/swabs. *Hum. Mol. Genet.* **2**:159–163.

Riemann, R., and Angleitner, A. (1997). Genetic and environmental influences on personality: A study of twins reared together using the self- and peer report NEO-FFI scales. *J. Pers.* (in press).

Zulliano, G., and Hobbs, H. (1990). Tetranucleotide repeat polymorphism in the op-lipoprotein C-III gene. *Nucleic Acids Res.* **18**:4299.

Edited by David W. Fulker