A Novel Method for Creating Artificial Mutant Samples for Performance Evaluation and Quality Control in Clinical Molecular Genetics

Michael Jarvis,* Ramaswamy K. Iyer,*
Laurina O. Williams,[†] Walter W. Noll,[‡]
Kirk Thomas,[§] Milhan Telatar,* and
Wayne W. Grody*^{¶||}

From the Departments of Pathology and Laboratory Medicine,*
Pediatrics,[¶] and Human Genetics,^{||} University of California at Los
Angeles School of Medicine, Los Angeles, California; the Division
of Laboratory Systems,[†] Centers for Disease Control and
Prevention, Atlanta, Georgia; the Department of Pathology,[‡]
Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire;
and the Department of Internal Medicine,[§] University of Utah
School of Medicine, Salt Lake City, Utah

The lack of readily available, patient-derived materials for molecular genetic testing of many heterozygous or rare disorders creates a major impediment for laboratory proficiency and quality control procedures. The paucity of clinically derived mutation-positive samples could be surmounted if it were possible to construct artificial samples containing mutations of interest that would sufficiently resemble natural human samples. Such samples could then function as acceptable and realistic performance evaluation challenges and quality control reagents for recipient laboratories. Using the cystic fibrosis gene (*CFTR*) as a prototype, we have devised and executed experiments designed to generate unique DNA samples that could be used for these purposes. We used site-directed mutagenesis to generate mutations of interest in plasmid DNA derived from common bacterial artificial chromosome sources containing the cystic fibrosis transmembrane conductance receptor gene. *CFTR* mutations G85E and 1078delT were chosen to represent mutations in the original American College of Medical Genetics-recommended population-screening panel of 25 mutations. DNA samples containing predetermined concentrations and ratios of wild-type and mutated plasmids, bacterial artificial chromosomes of interest, and nonhuman genomic carrier DNA were characterized and tested in-house and in a group of nine pilot testing laboratories using a variety of technical platforms. The results indicate that these constructs, containing *CFTR* mutations in heterozygous and homozygous states, can serve as valid and accessible materials for quality assurance, including performance evaluation, proficiency test-

ing, and assay quality control. (*J Mol Diagn* 2005, 7:247-251)

The rapid pace of disease gene discovery, fueled by the Human Genome Project, has in turn fueled a continuous explosion in the number of analytes tested by molecular diagnostic laboratories, especially those involved in heritable disease testing. Some have questioned whether the time interval from gene discovery to clinical test translation may be shorter than desired for full understanding of the test's clinical utility. Yet even if that were satisfied, the speed of new test development and the sheer number of genes and mutations targeted for analysis have led to another of bottleneck in laboratory practice, one of more practical daily concern: the lack of well-characterized control materials containing mutations of interest. These materials are necessary as positive controls in the assays and as resources for quality assurance programs such as the nationwide proficiency testing programs offered jointly by the College of American Pathologists (CAP) and the American College of Medical Genetics (ACMG).¹⁻³ Procurement of these materials from natural sources is hampered by the rarity of many desired mutations; the limited quantity in clinical specimens; the dependence on clinicians to recognize the need and take the trouble to deposit patient samples in existing repositories (such as the Coriell Institute); and regulatory requirements such as informed consent, sample ownership, and genetic privacy constraints. Subsequent to recommendations resulting from a Centers for Disease Control-sponsored needs-assessment project conducted in 1999,⁴ the CDC has funded studies aimed at developing more efficient methods of collecting⁵ or artificially constructing cell lines carrying defined mutations. As one of the funded centers, we have chosen the latter approach because of its practical advantages: it obviates

Supported by the Centers for Disease Control and Prevention (contract no. 200-2000-10030).

Accepted for publication December 27, 2004.

Current address of M.J.: Specialty Laboratories, Valencia, CA.

Current address of R.K.I.: Myriad Genetics Laboratories, Salt Lake City, UT.

Address reprint requests to Wayne W. Grody, M.D., Ph.D., Department of Pathology and Laboratory Medicine, UCLA School of Medicine, 10833 Le Conte Ave., Los Angeles, CA 90095-1732. E-mail: wgrody@mednet.ucla.edu.

the need to identify and approach actual patients, and potentially it puts within reach any mutation desired, no matter how rare, as long as the sequence is known.

There are several possible experimental approaches for creating artificially constructed mutation samples, including sample spiking, transient transfection, permanent transfection, and genetic engineering through such techniques as homologous recombination. The aims of the project required us to develop and compare two different approaches, and we settled on site-directed mutagenesis and homologous recombination. During the course of our experiments, we have evaluated intermediate by-products of these methods and have used them to create DNA samples that could be used more readily for these purposes. Here we show that these artificial DNA samples can effectively substitute for conventional patient-derived mutant samples. Our protocol is relatively simple, robust, and reproducible, and should be applicable to the production of a wide variety of mutation samples for many genes and diseases. The DNA samples have been evaluated both in-house and at several pilot testing facilities using a variety of mutation detection methods.

For these initial experiments, we have chosen cystic fibrosis as an especially appealing model target for a number of reasons. With a carrier frequency of ~1 in 25 to 30 among Caucasians (and present in other ethnic groups as well), it is the most common lethal autosomal recessive disorder in North America. More than 1000 mutations have been reported in the *CFTR* gene, most of which are extremely rare. No Food and Drug Administration-licensed commercial test kit complete with a comprehensive set of mutation controls is available. Large-scale population carrier screening for cystic fibrosis (CF) mutation carriers has recently been launched as recommended by a National Institutes of Health consensus panel, the ACMG and the American College of Obstetricians and Gynecologists.⁶ This is by far the largest molecular genetic screening program ever conducted, yet in the absence of standards and proficiency testing challenges for the full panel of 25 (now 23) recommended mutations and associated polymorphisms, the testing community has been at some disadvantage in meeting routine quality assurance standards. Although the CAP/ACMG proficiency-testing program has offered CF surveys for several years,⁷ it has challenged the participating laboratories with only a small subset of the recommended core mutation panel, mostly because of lack of availability of samples containing the other mutations. For these reasons, we have chosen CF, and the *CFTR* gene, as our first model target.

Methods and Results

Sample Construction and Preparation

The *CFTR* gene is located on chromosome 7 (7q31.2). It is ~250 kb in size and contains 27 exons.⁸ Several mutations were chosen for preliminary targeting experiments, including G85E (exon 3), N1303K (exon 21), and

Figure 1. Diagram of general strategy for amplification of *CFTR* target arms from genomic DNA followed by plasmid cloning and site-directed mutagenesis.

1078delT (exon 7). The first two were chosen for transgene construction purposes because they represent the most 5' and 3' mutations in the original ACMG panel, respectively, whereas the third was chosen because it is one of the mutations not currently available from the Coriell repository or other accessible sources. The study described here focused on the G85E and 1078delT target sequences because they are rarer and harder to obtain from natural sources than N1303K; the latter was used as a marker primarily to ensure that our *CFTR* constructs potentially encompassed all possible mutations in the original ACMG panel. It may be noted that the 1078delT was recently removed from the core mutation screening panel because of its rarity.⁹

Our original homologous recombination strategy, modeled on the methodology used to create gene knockouts in mice, called for the design of two targeting arms for each mutation site of interest: a long arm, 4 to 6 kb in size, and a short arm, 2 to 4 kb in size. These targeting arms, contiguous in sequence, were polymerase chain reaction (PCR) amplified from human bacterial artificial chromosome (BAC) DNA (ResGen, Birmingham, AL) containing the *CFTR* gene. BAC clones CIT-B 068P20 (AC000111) and CIT-B 133K23 (AC000061), which together contain the complete *CFTR* gene and flanking sequences, served as our source for PCR amplification. These two BACs were completely sequenced as part of the Human Genome Project. While the homologous recombination sequence is still in progress, we fortuitously observed that these target arms could themselves be used to create artificial DNA testing samples. Figure 1 illustrates the experimental strategy for this portion of our project. The short DNA targeting arm for each desired *CFTR* mutation was PCR amplified with the mutation site for each segment generally centered in the amplified gene product. Primer sequences 5'-tgg gga ggg aaa tag atg gga aaa ggt aat-3' and 5'-tta caa gcc aag cag agc ata gaa agg-3' generated a 3-kb amplicon that contained the G85 mutation site, whereas primer sequences 5'-aaa tgc cag gta ccc aca tgc act atg cca-3' and 5'-tct tca ttt tct tct ctg ctc tct acc-3' generated a 2.4-kb amplicon that contained the 1078 mutation site. These short target arms were subsequently ligated to standard cloning vectors and subjected to one round of site-directed mutagenesis (Promega, Madison, WI) to introduce the desired *CFTR*

Figure 2. Results of in-house pilot testing of constructed heterozygous and homozygous products for CFTR mutations G85E and 1078delT using a commercial reverse hybridization strip system (Roche Linear Array CF Gold 1.0). Test results from an actual patient sample are also shown for comparison (**lane 6**). The observed genotypes are: **lane 1**, negative for tested mutations; **lane 2**, G85E homozygote; **lane 3**, 1078delT homozygote; **lane 4**, G85E heterozygote; **lane 5**, 1078delT heterozygote; **lane 6**, negative for tested mutations.

gene mutation. Plasmids containing target arms with the desired mutated gene sequences were identified by both sequence analysis of the exon of interest and restriction digest pattern analysis.

These plasmids, containing a single *CFTR* mutation and flanking gene regions, served as the basis for the creation of artificial mutation samples. Large preparations of each plasmid type (with and without the mutation of interest) as well as BACs 068P20 and 133K23 were purified, characterized, and quantitated. Each artificial sample was first formulated by combining equimolar ratios of these separate DNA sources. Calculations to determine the approximate concentration of each DNA source for each sample were based primarily on the relative molar concentration of each allele that would be expected in a typical CAP/ACMG proficiency sample (50 μ g of DNA in 20 μ l of buffer). [50 μ g of genomic DNA is approximately equivalent to 2×10^{-17} mol of one set of human chromosomes (3×10^9 bp)]. The addition of flanking DNA to the mutation site allows for more flexibility

in the type of processes that can be used in detection and identification. The two BAC clones make available the entire *CFTR* gene sequence, increasing the apparent authenticity of the sample and also presenting the above-mentioned platform flexibility. To enhance the realism of these samples, salmon sperm DNA was also added for bulk and background genomic carrier. Because it was the major DNA component, we determined that when salmon sperm DNA was used as template for PCR analysis using any of our usual oligonucleotide primer sets, no amplified products of the anticipated sizes were observed by either an in-house amplification restriction digestion method or a commercial *CFTR* hybridization assay (Roche Diagnostics, Indianapolis, IN).¹⁰ Samples were formulated and analyzed for each of the following five genotypes: wild-type (homozygous normal), homozygous G85E, homozygous 1078delT, heterozygous G85E, and heterozygous 1078delT. Each of these samples was tested and found to behave indistinguishably from natural samples when assayed with both our in-house PCR/re-

Figure 3. Exon 7 sequencing of wild-type and 1078delT mutation-containing plasmids. **A:** A segment of the plasmid containing the wild-type exon 7 sequence. **B:** The corresponding segment of the plasmid containing the 1078delT mutation. The **arrows** indicate the position of the T that is present in the wild-type but deleted in the mutation-containing sequence.

Table 1. Pilot Testing Results

Sample	Laboratory analysis (correct results/total results)	Unable to analyze
Normal (w.t.)	7/8	1
G85E, heterozygous	8/9	
G85E, homozygous	8/9	
1078delT, heterozygous	8/9	
1078delT, homozygous	7/9	

striction digest assay (data not shown) and standard reverse line blot technology (Figure 2).

Confirmatory DNA Sequencing

To ensure fidelity of the constructed fragments and detect the introduction of unexpected variants due to PCR misincorporation errors or other sources, plasmids containing the *CFTR* exon 7 with or without the 1078delT mutation were sequenced using standard technology. Both the wild-type and mutant sequences were found to be present in the respective constructs with no alterations (Figure 3).

External Pilot Testing

The five artificial DNA samples along with a methods questionnaire were sent to nine pilot testing facilities that together encompass a wide range of testing platforms for *CFTR* mutation screening. All nine facilities have reported their testing results, which are summarized in Table 1. The results indicate that the five DNA testing samples were remarkably reproducible in mimicking the desired human genotypes across multiple testing platforms. Although some facilities noted slight anomalies in sample behavior, only one mutation sample was incorrectly identified (as a heterozygote 1078delT instead of a homozygote). Because the five test samples were optimized using reverse dot blot technology, facilities using similar testing methods (by two different manufacturers) had little difficulty discerning the designed genotype of each sample. Other testing platforms included the ABI Genotyper technology (Applied Biosystems, Foster City, CA), a proprietary multiplex hybridization assay, amplification refractory mutation system, and two different proprietary DNA microarray systems. Each of these platforms obtained generally correct results (Table 2), although one microarray facility reported slightly more

skewing of the mutant:wild-type ratios in the heterozygote samples and spurious detection of wild-type sequence in the 1078delT homozygous sample as noted above. These inaccuracies should not be construed to mean that there are technical problems with any of the platform systems involved in our pilot testing. The composition of our artificial DNA samples, which had initially been optimized on other platforms, in all likelihood, led to the perceived discrepancies. In fact, our method of construction dictated that the homozygous samples were constituted with a trace amount of wild-type sequence, and the microarray facility that detected this in the 1078delT sample stated that their platform is extremely sensitive in its allelic detection of heterozygote sequences. The only platform that produced a more significant range of spurious results was the amplification refractory mutation system assay; the laboratory using this technology detected several extra *CFTR* mutations for all five samples in addition to the ones we had intended. On further investigation it was revealed that the concentration of template DNA in the artificial samples was 10- to 100-fold higher than that used from patient samples in this assay. Some of the extra bands (but not all), which were rather faint to begin with, disappeared when the laboratory diluted our samples before analysis.

Discussion

The lack of easily accessible, validated mutant controls has proven to be a major obstacle to the advancement of clinical molecular genetic testing. To attempt to address this need, two pilot projects have been initiated with governmental support. One of them addressed methods for the efficient collection and cell culture transformation of residual blood samples from patients with genetic disorders, and the first report was recently published.^{5,11} The other project, from our group, has focused on the development and construction of artificial mutant controls, using pure *in vitro* genetic engineering methods with no patient-derived material at all. The protocol described here was initially developed to provide intermediate preliminary samples as we move toward our ultimate goal of creating human cell lines with locus-specific *CFTR* gene mutations. However, it appears to be a useful method in itself for designing and developing control samples for quality control and proficiency testing purposes. It requires some facility with recombinant DNA techniques, BACs, plasmid vectors, and site-directed mutagenesis. Once established, the method could be amenable to any

Table 2. Pilot Performance by Method

Method	Number of laboratories	Results	Comments
Reverse ASO line blots	3	Correct	Exon 7 partial PCR failure*
Genotyper	1	Correct	Heterozygous ratios not exactly 50:50
Microarray	2	Mostly correct	Trace w.t. in 1078delT homozygous
Proprietary multiplex assay	1	Correct	Some exon signals uneven
Sequencing	1	Correct	M470V polymorphism also detected
Amplification refractory mutation system	1	Equivocal	Several extraneous mutations detected

*One laboratory; did not affect interpretation.

known mutation in any characterized gene. Scale-up to large preparations, providing enough control material to last for years, should be easy and relatively inexpensive. Although the approach is clearly more involved and cumbersome than simple spiking of samples with artificially synthesized mutant oligonucleotides, as some laboratories are using now, we feel that it provides a more realistic-looking product that behaves much like natural human samples on most of the commonly used technical platforms in molecular genetic testing.

Having said that, however, our pilot testing results did reveal some important caveats. The artificial samples did not perform equally well with all platforms. They performed most reliably on allele-specific oligonucleotide (ASO) hybridization platforms, whether in reverse line blots or more sophisticated microarray instrumentation. They should be used with caution, if at all, in amplification refractory mutation system assays, because of the appearance of spurious mutation results, at least at certain dilutions. These signals were specific for mutations that were never constituents of our samples, and the source is unclear. Even with the ASO methods, there were a few scattered anomalies reported in the respondent questionnaires (Table 2). Although all intended mutations were detected correctly, two labs using a reverse line blot from a particular manufacturer reported that one of the expected exon control products did not amplify. As described above, extreme sensitivity in microarray hybridization could pick up trace wild-type sequence in at least one of our intended homozygote samples, and several laboratories using other methods reported variably unequal mutant versus wild-type signal in the intended heterozygote samples; with further study, it is possible that customized products could be developed to eliminate this problem. Because most current assays for *CFTR* mutations target multiple sites in the gene and not just a single target mutation, we felt that inclusion of the BAC DNA was necessary so that negative (wild-type) signal would be detected for all those mutations targeted by the assay but not present in the sample. It is true that this adds, in effect, a third allele to certain of the samples, although the amount is slight enough that it does not affect the biallelic signal balance in most assays, as we have shown. We feel this owes, at least in part, to the more efficient PCR amplification of the plasmid DNA compared to the BAC. Parenthetically, it can be noted that even CAP/ACMG proficiency testing samples, derived from cell lines, have produced variable results throughout the years, some of which are undoubtedly platform-dependent. Overall, however, the error rate has been low.¹² The one laboratory using DNA sequencing methodology was able to analyze the entire *CFTR* gene in our product, and even detected an unexpected (although common) M470V polymorphism in all of the samples—apparently carried by whatever person had donated DNA for construction of these BACs in the early years of the Human Genome Project.

Another caution to keep in mind is that, because this method is entirely synthetic, with no starting material derived from actual patients having the disease in ques-

tion, it is critically important to verify that the mutation sequence reported in the literature or database on which it is based is actually correct. We suggest that this approach could be applied to many other genetic and nongenetic diseases, including cancer markers and even pathogen and host markers in infectious diseases. It can be used to generate large amounts of sample containing any mutation of interest in relatively short order and at far less expense than clinical or laboratory approaches involving patient sample collection and/or human cell culture. We expect that further development of this approach, particularly by adjusting the ratios of mutant plasmid to BAC DNA, could be used to produce a wide array of positive control samples that can meet a critical need of the genetic testing community.

Acknowledgments

We thank BioArray Solutions, Foundation for Blood Research, Genzyme Genetics, Kimball Genetics, Nanogen, Quest Diagnostics, Specialty Laboratories, Ambry Genetics, and the University of Colorado for generously agreeing to participate in the pilot testing of our artificial samples; Ellen Carpenter, Bradley Popovich, and Dalal Shamoun of our expert consultant panel for their helpful advice; and Linda McCabe for her help in designing the pilot questionnaire.

References

1. Grody WW: Proficiency testing in diagnostic molecular pathology. *Diagn Mol Pathol* 1994, 3:221–223
2. Dequeker E, Ramsden S, Grody WW, Stenzel TT, Barton DE: Quality in molecular genetic testing. *Nat Rev Genet* 2001, 2:717–723
3. Richards CS, Grody WW: Alternative approaches to proficiency testing in molecular genetics. *Clin Chem* 2003, 49:717–718
4. Williams LO, Cole EC, Lubin IM, Iglesias NI, Jordan RL, Elliott LE: Quality assurance in human molecular genetics testing: status and recommendations. *Arch Pathol Lab Med* 2003, 127:1353–1358
5. Bernacki SH, Stankovic AK, Williams LO, Beck JC, Herndon JE, Snow-Bailey K, Prior TW, Matteson KJ, Wasserman LM, Cole EC, Stenzel TT: Establishment of stably EBV-transformed cell lines from residual clinical blood samples for use in performance evaluation and quality assurance in molecular genetic testing. *J Mol Diagn* 2003, 5:227–230
6. Grody WW, Cutting G, Klinger K, Richards CS, Watson MS, Desnick RJ: Laboratory standards and guidelines for population-based cystic fibrosis carrier screening. *Genet Med* 2001, 3:149–154
7. Grody WW, Desnick RJ, Carpenter NJ, Noll WW: Diversity of cystic fibrosis mutation screening practices. *Am J Hum Genet* 1998, 62:1252–1254
8. Zielenski J, Tsui L-C: Cystic fibrosis: genotypic and phenotypic variations. *Ann Rev Genet* 1995, 29:777–807
9. Watson MS, Cutting GR, Desnick RJ, Driscoll DA, Klinger K, Mennuti M, Palomaki GE, Popovich BW, Pratt VM, Rohlfis EM, Strom CM, Richards CS, Witt DR, Grody WW: Cystic fibrosis population carrier screening: 2004 revision of American College of Medical Genetics mutation panel. *Genet Med* 2004, 6:387–391
10. Makowski GS, Nadeau FL, Hopfer SM: Single tube multiplex PCR detection of 27 cystic fibrosis mutations and 4 polymorphisms using neonatal blood samples collected on Guthrie cards. *Ann Clin Lab Sci* 2003, 33:243–250
11. Grody WW: Quest for controls in molecular genetics. *J Mol Diagn* 2003, 5:209–211
12. Palomaki GE, Bradley LA, Richards CS, Haddow JE: Analytic validity of cystic fibrosis testing: a preliminary estimate. *Genet Med* 2003, 5:15–20