Comparison of Techniques for the Successful Detection of *BRCA1* Mutations in Fixed Paraffin-Embedded Tissue[1]

Jonine L. Bernstein,[2] W. Douglas Thompson, Graham Casey, Richard A. DiCioccio, Alice S. Whittemore, Anh T. Diep, Seema S. Thakore, Susan Vaziri, Shanyan Xue, and Robert W. Haile

Department of Community and Preventive Medicine, Mount Sinai School of Medicine, New York, New York 10029 [J. L. B., S. S. T.]; Department of Applied Medical Sciences, University of Southern Maine, Portland, Maine 04104 [W. D. T.]; Department of Cancer Biology, Cleveland Clinic Foundation, Cleveland, Ohio 44195 [G. C., S. V.]; Department of Cancer Genetics, Roswell Park Cancer Institute, Buffalo, New York 14263 [R. A. D.]; Department of Health Research and Policy, Stanford University School of Medicine, Stanford, California 94305 [A. S. W.]; Department of Preventive Medicine, Keck School of Medicine at the University of Southern California, Los Angeles, California 90033 [A. T. D., S. X., R. W. H.]

Abstract

Genomic DNA isolated from archived paraffin-embedded tissues (PETs) has important applicability in genetic epidemiological studies. To determine the accuracy of the sequence data, using DNA derived from PET among patients with known mutations characterized from blood, we conducted a blinded factorial experiment to simultaneously examine the influence of mutation type, age of the PET, PCR product type, and Taq DNA polymerase on *BRCA1* gene mutation detection. The probability of detecting sequencing artifacts was also investigated. We found that: (*a*) gene detection was most accurate for newer PET; (*b*) high fidelity Taq with shorter PCR amplicon length yielded the highest mutation detection success rate and lowest artifact rate; and (*c*) base substitutions were more often correctly identified than frameshift mutations or wild-type sequences. We concluded that DNA derived from PET that archived for less than 18 years can be used successfully for detecting *BRCA1* gene mutations if quality control is strictly maintained.

Introduction

Techniques for extracting and amplifying DNA from PET[3] blocks using PCR have been available for nearly two decades. When unfixed tissue is unavailable or not feasible to collect, archived PET is considered an important source of DNA for use in genetic epidemiological studies examining the molecular biology of both chronic and infectious diseases. For example, archived PET could greatly facilitate large-scale multicenter studies through the retrospective collection of large numbers of cases with tumors or rare diseases that would otherwise take years to accumulate. Furthermore, using PET makes possible the study of high-risk families in which affected members may be deceased and, thus, sequencing from blood or unfixed tissue is not a viable option (1). In turn, these genetic analyses can test hypotheses linking known etiological risk factors to specific molecular alterations among subtypes of tumors or for prognostic studies in which the clinical outcome has already been defined (2). Studies to date have compared fixation techniques, identified different methods for isolating DNA, examined the degree of DNA degradation after long-term storage, and suggested the ideal length of the DNA fragment to be amplified. Nevertheless, although the same factors affecting successful amplification may also affect misincorporation of nucleotides into DNA (3), surprisingly little attention has been paid to comparing the accuracy of the sequence data from the amplified PET-derived DNA with similar data ascertained from untreated tissue or blood. In addition, to our knowledge, no study has compared these factors across laboratories.

The purpose of this multilaboratory study was to determine the validity of using DNA extracted from archived PETs to detect mutations in relatively rare susceptibility genes, such as *BRCA1*, and to optimize the screening technique. To accomplish these goals, we carried out an experiment to identify sources of artifacts and errors in mutation screening by simultaneously examining mutation type, age of the PET, amplicon length, and type of polymerase in a series of cases in which *BRCA1* gene carrier status had been previously determined from lymphocytes. We also analyzed sequences derived from cloned PCR products to determine the level of sequencing artifacts under these different conditions.

Materials and Methods

Case Selection

The three laboratories participating in this study were selected because each has significant experience in screening for *BRCA1* gene mutations using high throughput technology, and each has a biorepository of specimens from patients who have already been screened for *BRCA1* using DNA extracted from WBCs. From these repositories, *BRCA1* gene mutations were characterized among breast or ovarian cancer cases for whom PET was also available, according to the following strata: "old base-substitution" mutations, with blocks from 1980 to 1985; "old frameshift" variants, with blocks from 1962 to 1983; "old wild-type" sequence, with blocks from 1984 to 1985; "new base-substitution", with blocks from 1990 to 1996; "new frameshift" with blocks from 1995 to 1997; and "new wild types" with blocks from 1997 to the present. Two cases were randomly selected from within each of the six strata ($n = 12$) and each

Received 12/14/01; revised 4/30/02; accepted 5/14/02.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
[1] Supported by NIH Grant CA69631 and Core Grant 16056 [Roswell Park Cancer Institute (RPCI)]. The latter supported the sectioning and sequencing performed at RPCI.
[2] To whom requests for reprints should be addressed, at Mount Sinai School of Medicine, One Gustave L. Levy Place, Box 1043, New York, NY 10029. Phone: (212) 241-8495; Fax: (212) 360-6965; E-mail: jonine.bernstein@mssm.edu.
[3] The abbreviations used are: PET, paraffin-embedded tissue; HIFI, high fidelity.

case was tested in the two nonoriginating laboratories, so that each laboratory analyzed a total of 8 unique cases. The only criteria used in the selection of samples for distribution to selected laboratories for testing were that no blocks were sent for analysis to the laboratory from which it came and that each laboratory received one sample with a frameshift and one sample with a missense mutation, the remaining being wild-type samples. Because of this stratified random assignment and distribution by the coordinating laboratory, all of the laboratories did not evaluate all of the age categories of blocks. The resulting distribution of mutation type and age of the block examined by the different laboratories was as follows: Laboratory 1 was given old blocks of wild-type and frameshift mutations (frameshift: 185 AG deletion and 2080 A insertion) and new blocks of all three types (wild type; frameshift: 199 + 1 CATCTG insertion; missense: 5622 C→T substitution); Laboratory 2 evaluated old blocks of all three types (wild type; frameshift: 185 AG deletion and 2080 A insertion; missense: 300 T→G and 3508 C→G substitutions) and new blocks with frameshift (frameshift: 3053 T deletion) and missense (missense: 1806 C→T substitution) mutations; and Laboratory 3 analyzed old blocks with missense mutations (missense: 300 T→G and 3508 C→G substitutions) and new blocks of all three types (wild type; frameshift: 199 + 1 CATCTG insertion and 3053 T deletion; missense: 1806 and 5622 C→T substitutions). Furthermore, because each laboratory analyzed the eight cases they were assigned, the following two PCR conditions were also altered: amplicon lengths ("short" or 150–200 bp *versus* "long" or 250–300 bp) and type of Taq polymerase (native *versus* HIFI). Thus, each laboratory conducted a total of 32 analyses (eight blocks, each tested under four conditions). The reported mutation type of each sample was confirmed by direct sequencing, ensuring against possible sample mix-up.

Laboratory Analysis

Four 5-32 μm sections were cut from each of the identified PETs using a standard microtome with a fresh disposable blade between cases. Each section was placed on a glass microscope slide and sent to the coordinating center, where it was randomly assigned a study identification number. The tissue sections were removed from the slides, placed in microcentrifuge tubes labeled only with the study identification, and sent to the two nonoriginating laboratories for amplification and analyses. Although the mutation status of the sections was masked, the laboratories were informed of the region to amplify.

Using direct sequencing and TA cloning followed by sequencing of up to five independently isolated fragments to identify the possibility of sequencing artifacts, each laboratory assessed the same set of selected conditions, namely: (*a*) to PCR-amplify the targeted amplicon within the size range which was defined as either "short" or "long"; (*b*) to use the specific types of Taq DNA polymerase; (*c*) to restrict the maximum number of amplification cycles; and (*d*) to not perform nested PCR. The laboratories used their own protocol for DNA extraction, PCR amplification, and sequencing, but all of them used the same cloning technique (TOPO TA cloning kit, Invitrogen Corp, Carlsbad CA) and the specific methods are described below:

Laboratory Method 1. Tissue sections were extracted from the PET using the commercially available QiaAmp Tissue Extraction kit (Qiagen, Valencia, CA) in a final volume of 50 μl. Coding regions for assigned PET using this method (listed in Table 1) were PCR amplified in a final volume of 15 μl containing: 1.5 μl 10× PCR buffer (Invitrogen); 0.75 μl of dNTPs (12.5 mM stock; each); 0.45 μl of $MgCl_2$ (50 mM stock); 0.15 μM each, forward and reverse primers concentration; 0.15 μl of polymerase (5 units/μl; either native or HIFI platinum Taq); and 2 μl of DNA. The PCR product was denatured at 94°C for 5 min, followed by 35 cycles of: denaturation (94°C for 45 s); annealing (for 1 min at 53°C for exons 2, 11, and 5 and 62°C for exon 24); and annealing (72°C for 1 min). This was followed by a final extension at 72°C for 7 min. An aliquot of the PCR products was electrophoresed on an agarose gel or 6% nondenaturing polyacrylamide gel to confirm the presence and the size of a single band. The PCR products were purified using QIA quick PCR purification kit (Qiagen). An aliquot of the sample was submitted for automated sequencing in both the forward and reverse direction using the amplification primers with the ABI 377 sequencer (Perkin-Elmer: Foster City, CA) in the institutional sequencing facility. The remainder of the sample was used for cloning. Individually isolated clones were subsequently isolated, purified, and sequenced.

Laboratory Method 2. PET sections were deparaffinized by washing consecutively with xylene and ethanol. After air-drying, the residue was suspended in 300 μl of 500 mM Tris-HCl (pH 9), containing 20 mM EDTA, 10 mM NaCl, 1% (w/v) SDS, and 150 μg of proteinase K. Samples were incubated at 65°C overnight, followed by 10 min at 100°C, and then were cooled to room temperature. Samples were extracted twice with 300 μl of phenol (pH 8)/chloroform (1:1). The DNA was precipitated from the upper phase by incubation with 100 μl of 10 M $NH_4C_2H_3O_2$, 3 μl of glycogen (1 mg/ml), and 900 μl of ethanol at −80°C for 30 min. DNA was collected by centrifugation at 13,000 × *g* for 15 min at 4°C, air-dried, and suspended in 50 μl of 10 mM Tris-HCl (pH 7.5) containing 1 mM EDTA. PCR amplification reaction mixtures (20 μl final volume) contained 2 μl of DNA template, 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 0.2 mM dNTPs, 0.2 mM each, forward and reverse primers (Table 1), 1.5–3.0 mM $MgCl_2$, 0.0001% (w/v) gelatin, and 0.8 units of either recombinant Taq DNA polymerase or HIFI platinum Taq DNA polymerase (Life Technologies, Inc., Rockville, MD). Reaction mixtures with template omitted served as a negative control, and DNA from an individual wild type for *BRCA1* served as a positive control (4). Reactions with recombinant Taq DNA polymerase were hot-started using AmpliWax (Perkin-Elmer) in accordance with the manufacturer's instructions. The $MgCl_2$ concentrations were selected to be optimal for each PCR assay. Cycling conditions were 1 cycle of 95°C for 5 min, followed by 35 cycles of denaturation (95°C for 45 s), annealing (at the temperatures which varied from amplicon to amplicon and depended on the melting temperature of the designed oligomers for 45 s), and extension (72°C for 1 min). Final extension was at 72°C for 7 min. The PCR products and DNA molecular size standards were electrophoresed in a 2% agarose gel containing ethidium bromide (5 mg/ml) to verify synthesis and assess the purity of each anticipated DNA fragment. Amplified products were purified from the agarose gel using the Qiaquick Gel Extraction kit (Qiagen) and used for direct DNA sequencing and cloning. The sequencing was done in both the sense and antisense directions using amplification primers with an Applied Biosystems Model 373A DNA automated sequencer (Perkin-Elmer).

Laboratory Method 3. DNA was extracted from the PET sections using a modification of the protocol of (5). Briefly, each 5 μm section was deparaffinized in a 1.5 ml polyethylene microcentrifuge tube by washing twice in xylene, twice in 100% ethanol, and air-dried. The deparaffinized specimen was digested overnight at 50°C in 50 μl buffer containing 200

Table 1 Primers used by the three laboratories to amplify selected regions of *BRCA1*

Laboratory method	Exon	Amplicon length	Annealing temperature (°C)	Sequence (5′–3′) forward	Sequence (5′–3′) reverse
1	2	Short	53	GAAGTTGTCATTTTATAAACC	AAATTAATACACTCTTGTGC
		Long	53	GAAGTTGTCATTTTATAAACC	TTTTCTTCCCTAGTATCTAAG
	5	Short	52	TTTCATGGCTATTTGCC	CCAACCTAGCATCATTAC
		Long	52	TTTCATGGCTATTTGCC	GAATGGTTTTATAGGAACG
	11	Short	53	TTGAACTAGTAGTCAGTAG	TTCATTTGGCTTGTTACTC
		Long	53	TTGAACTAGTAGTCAGTAG	CACTTAGTAAAAGAACCAGG
	24	Short	62	GCTTGTGTTCTCTGTCTCC	CTGTACCTGTGGCTGGCT
		Long	62	GCTTGTGTTCTCTGTCTCC	GTCCTTTTCAGGCTGATGT
2	2	Short	60	TATGTTTTCTAATGTG	TGTAAGGTCAATTCTGT
		Long	60	GAAGTTGTCATTTTATAAACCTTT	TGTCTTTTCTTCCTTAGTATGT
	5	Short	60	TCATGGCTATTTGCCTTTTG	TTCCTACTGTGGTTGCTTCC
		Long	62	CTCTTAAGGGCAGTTGTGAG	TTCCTACTGTGGTTGCTTCC
	11	Short	60	GGAGGAAGTCTTCTACCAGGC	CTCTTCTTGGCTCCAGTTGC
		Long	58	GCCCACCTAATTGTACTGAATT	CTTGGAAGGCTAGGATTGACA
		Short	59	TCAAAGGAGGCTCTAGGTTTTG	TTCCCATTTCTCTTTCAGGTG
		Long	59	GATAAGCCAGTTGATAATGCCA	CGGCTAATTGTGCTCACTGT
		Short	60	TAGGGGTTTTGCAACCTGAG	CACCATCATCTAACAGGTCATCA
		Long	59	CCAGTGATGAAAACATTCAAGC	TTCACCATCATCTAACAGGTCA
3	2	Short	50	GAAGTTGTCATTTTATAAACCTTT	AGGAATCCCAAATTAATACAC
		Long	54	CCTTCCAAATCTTCAAATTTAC	CTTTTCTTCCCTAGTATGTAAGGTC
	5	Short	53	AAAAGGAAGTAAATTAAATTGTTC	TGCTTCCAACCTAGCAT
		Long	55	CTCTTAAGGGCAGTTGTGAG	ATGGTTTTATAGGAACGCTATG
	11	Short	50	TCAAAGGAGGCTCTAGGT	TCAAAGTTTTCCTCTAGCAG
		Long	46	AGATAAGCCAGTTGATAATG	GGCTTCTTTAAAAACATTT
		Short	50	CTGAAATAAAAAAGCAAGAAT	TTCACCATCATCTAACAGGT
		Long	46	GTCTATAAACAAAGTCTTCCT	GTAACCCTGAGCCAAAT
		Short	53	GGGAACTAACCAAACGGA	GCTTATAGGTTCAGCTTTCG
		Long	55	GGGAACTAACCAAACGGA	ACTACTAGTTCAAGCGCATGA
	24	Short	51	AATCTCTGCTTGTGTTCTCTG	GCTGGCTGCAGTCAGTAG
		Long	52	ATGAATTGACACTAATCTCTGC	GTAGCCAGGACAGTAGAAGGA

μg/ml Proteinase K, 50 mM Tris-HCl (pH 8.5), 1 mM EDTA, and 0.5% Tween 20. The specimens were then incubated at 95°C for 8 min to inactivate the Proteinase K. Each PCR reaction used 3 μl of the crude DNA extract in a final volume of 30 μl. The reaction was amplified on an MJ Research programmable thermal controller PTC-100. PCR reactions were performed using a hot start method. The reaction mixture contained 5–10 pmol each of the forward and reverse primers (Table 1), 0.25 mM dNTPs, 1.5 units of either Taq DNA polymerase (Promega, Madison, WI) or HIFI Platinum Taq DNA polymerase (Life Technologies, Inc.), 10 mM Tris-HCL (pH 9.0), 50 mM KCl, 0.1% Triton X-100 (also provided by Promega) and with the appropriate concentration of $MgCl_2$ in the range of 2.0–3.5 mM, which is optimal for the enzyme performance in each specific PCR reaction. The thermocycling profile was set as 3 min of initial denaturation at 94°C, followed by 35 cycles of denaturation (94°C for 30 s), annealing for 45 s, and elongation (72°C for 1 min). Final extension was 72°C for 5 min. Reaction products and DNA molecular size standards were electrophoresed in a 2% agarose gel containing ethidium bromide (5 mg/ml) to confirm the quality and the size. PCR products were used for both direct sequencing and cloning. Manual direct sequencing using Thermo Sequenase kit (Amersham, Cleveland, OH) was performed in both sense and antisense directions.

Sample Cloning Technique. Products amplified by either Recombinant or HIFI platinum Taq DNA polymerase were cloned using the TOPO TA cloning kit (Invitrogen). For PCR products using platinum Taq DNA polymerase, the first step required the addition of overhanging "A"s at the 3′ end. This PCR product was inserted into the vector. The vectors were then transformed into chemically competent *Escherichia coli* and propagated in Luria Bertani medium containing ampicillin or kanamycin (50 μg/ml). Vector DNA was isolated as described by Sambrook *et al.* (6) or using Qiaprep Spin Miniprep Kit (Qiagen). The clones were linearized by digestion with *Eco*R1 restriction enzyme at 37°C for 1 h and electrophoresed in a 1% agarose gel. The molecular weight standard was loaded adjacent to the product for reference, and the final product was submitted for sequencing. Each laboratory isolated up to five independent clones per sample.

All of the laboratories used *BRCA1* nucleotide sequences in GenBank (accession nos. U14680 and L78833) as well as the standard cDNA numbering system (7, 8) and mutation nomenclature (9).

Results

Table 2 shows the number of correctly identified mutations using DNA derived from the PET samples compared with those

Table 2 Correctly identified mutations according to Taq type, amplicon length, mutation type, and block age using direct sequencing of the DNA derived from PET compared with that derived from blood

Mutation type	Block age	Taq type						Amplicon length					
		Regular Taq			HIFI Taq			Short			Long		
		Number correctly identified	(%)	ND[a]	Number correctly identified	(%)	ND	Number correctly identified	(%)	ND	Number correctly identified	(%)	ND
Frameshift	All[b,c,d]	6/12	(50)	6	5/10	(50)	5	6/11	(55)	5	5/11	(45)	6
	New[b,c,d]	5/6	(83)	1	5/6	(83)	1	5/6	(83)	1	5/6	(83)	1
	Old[b,c,d]	1/6	(17)	5	0/4	(0)	4	1/5	(20)	4	0/5	(0)	5
Base substitution	All[b,c,d]	7/9	(78)	2	9/10	(90)	1	9/10	(90)	1	7/9	(78)	2
	New[b,c,d]	3/5	(60)	2	3/4	(75)	1	4/5	(80)	1	2/4	(50)	2
	Old[b,d]	4/4	(100)	0	6/6	(100)	0	5/5	(100)	0	5/5	(100)	0
Wild type	All[b,c,d]	7/15	(47)	8	14/16	(88)	2	11/16	(69)	5	10/15	(67)	5
	New[b,d]	4/8	(50)	4	8/8	(100)	0	6/8	(75)	2	6/8	(75)	2
	Old[c,d]	3/7	(43)	4	6/8	(75)	2	5/8	(63)	3	4/7	(57)	3
All	All[b,c,d]	20/36	(56)	16	28/36	(78)	8	26/37	(70)	11	22/35	(63)	13
	New[b,c,d]	12/19	(63)	7	16/18	(89)	2	15/19	(79)	4	13/18	(72)	5
	Old[b,c,d]	8/17	(47)	9	12/18	(67)	6	11/18	(61)	7	9/17	(53)	8

[a] ND, number not detectable because obtaining either DNA or PCR product was infeasible.
[b] Method 1 was used for the analysis (please see "Laboratory Analysis").
[c] Method 2 was used for the analysis.
[d] Method 3 was used for the analysis.

Table 3 Summary of the possible artifact using clones

Blocks	Taq type and amplicon length combinations								Taq type			
	Regular Taq/Short		HIFI Taq/Short		Regular Taq/Long		HIFI Taq/Long		Regular Taq		HIFI Taq	
	No. of additional variations identified[a]	(%)	No. of additional variations identified[a]	(%)	No. of additional variations identified[a]	(%)	No. of additional variations identified[a]	(%)	No. of additional variations identified[a]	(%)	No. of additional variations identified[a]	(%)
All[b]	22/109	(20)	5/117	(4)	12/84	(14)	11/94	(12)	34/193	(18)	16/211	(8)
New[b]	14/62	(23)	3/62	(5)	8/50	(16)	5/44	(11)	22/112	(20)	8/106	(8)
Old[b]	8/47	(17)	2/55	(4)	4/34	(12)	6/50	(12)	12/81	(15)	8/105	(8)

[a] Number of additional variations identified other than the one that matched with the gold standard, out of total.
[b] Methods 1, 2, and 3 were used for the analysis (see "Laboratory Analysis").

derived from blood, according to Taq DNA polymerase type and amplicon length by mutation type and age of the block. We observed no false negative mutations. Across all of the mutation types and different block ages, regardless of the technique, ~66% of the mutations were correctly identified in all three of the laboratories. All of the missed mutations were from specimens that did not produce DNA or PCR products. Amplification with short DNA regions using HIFI Taq provided the largest percentage of correctly identified mutations (83% overall). Across all Taq types and amplicon lengths, compared with the newer blocks, the older ones were more difficult to characterize. Base substitutions were more often correctly identified in the older samples than were frameshift mutations or wild-type sequences.

Artifacts detected from the cloned products are presented in Table 3. Regardless of the age of the PET, screening with HIFI Taq and shorter-length amplicons yielded the lowest false-positive rate compared with any other combination of Taq type and amplicon length. Compared with amplicon length, Taq type was more predictive of artifact rate. Base-substitution type mutations, specifically missense mutations, were the only type of false positives observed in the clones. Using native Taq, we found 18% additional single-base substitutions compared with the correct sequences; with HIFI Taq it was only 8%. This difference was observed only in the cloned samples.

Table 4 shows the variability in the accuracy of mutation-detection rates for the screening of the PET samples by laboratory technique. The sensitivity for detecting mutations was 29, 83, and 94% for laboratory methods 2, 1, and 3, respectively. Successful amplification of PET-derived DNA also varied by laboratory; methods 1, 2 and 3 successfully obtained PCR products for 38, 88, and 100% of the samples, respectively. Laboratory method 3 achieved the highest success rate for DNA amplification and the highest mutation detection rate. Across all laboratory methods, we found a higher success rate for amplifying products and detecting mutations for newer blocks compared with older blocks.

Discussion

In etiological and clinical prognostic studies of selected diseases, tissue blocks may be retrieved from several hospitals and stored for different time intervals. Thus, factors pertaining to hospital practices, such as the preservative used in tissue processing, the duration of fixation, and length of time between surgical removal of tissue and fixation are out of the investigator's control. Nevertheless, any of these can affect the yield and quality of the amplification product from PET (10–14). In this study, we focused on factors that can be controlled by the investigator, including the effect of the age of the PET on the

Table 4 Successful mutation detection screening rate using PET by different laboratory (Lab) methodologies and age of the PET[a]

Age of blocks	Lab method 1				Lab method 2				Lab method 3			
	No. agreed/ total[b]	(%)	No. detectable/ total[c]	(%)	No. agreed/ total[b]	(%)	No. detectable/ total[c]	(%)	No. agreed/ total[b]	(%)	No. detectable/ total[c]	(%)
New	7/7	(100)	7/16	(44)	0/8	(0)	8/8	(100)	22/24	(92)	24/24	(100)
Old	3/5	(60)	5/16	(31)	8/20	(40)	20/24	(83)	8/8	(100)	8/8	(100)
All	10/12	(83)	12/32	(38)	8/28	(29)	28/32	(88)	30/32	(94)	32/32	(100)

[a] Please see "Laboratory Analysis."
[b] Number in agreement between pair and gold standard, proportion of correctly identified mutations out of total.
[c] Number detectable, proportion of feasibility out of total.

accuracy of mutation detection. Several studies have shown that DNA derived from archived PET stored over long periods of time, even up to 40 years, can be successfully amplified (2, 10, 15–18) yet others report anomalies related to the misincorporation of nucleotides during the amplification of DNA from PET stored for 5–36 years (3, 17, 18). We found, regardless of laboratory method, PET that was archived for less than 18 years provided accurate mutation detection, whereas those stored longer did not. Obviously, we did not test enough samples across a full range of calendar years to identify precisely a cutpoint for the age of the blocks in terms of accurate mutation detection with the methods used in this study.

Because DNA is extracted from PET (2, 12, 19), its degradation from the aging of the sample (2, 10, 12) precludes analysis with techniques requiring high-molecular weight DNA (2). Our success using the shorter-length primer is consistent with published reports suggesting that the size of the damaged genomic DNA fragment prepared from older archival samples may be smaller than that derived from newer samples (11, 12, 15). Although suboptimal for PCR amplification (13), using shorter fragments can enhance the sensitivity (10, 20). One possibility is that the pH of the original preservative rather than the length of storage *per se*, impedes successful amplification (10) and that the fixation-induced DNA degradation may be due to the extensive cross-linking of proteins to DNA and acid depurinization of the DNA (13). As a result, the DNA template can be altered or destroyed, preventing accurate amplification (2, 18), which, in turn, yields the relatively short PCR fragments with low recovery (13). Nevertheless, when the expected length of the PCR product is between 200 and 300 bp, the fixation time used by the diagnostic histology laboratory only slightly decreases the sensitivity of the PCR reaction (21). The results of amplifying from PET are comparable with that using unfixed tissue (12, 22).

The type of Taq or length of amplicon might lead to more false negatives in that certain mutation types might be missed as a result of the physical damage to DNA that can result in misincorporated bases, gaps, and cross-over products, and misincorporation from different Taq polymerases with reduced fidelity. Although we found that all four combinations of techniques (*e.g.,* HIFI and native Taq types and long and short amplicon lengths) detected 100% of the wild type, the rates for detecting frameshift mutations appeared to be dependent on the age of the block; DNA from a 38-year-old block that was known to contain a frameshift mutation was not detected by either of the two testing laboratories, whereas an 18-year-old block with a frameshift mutation amplified as well as the similarly aged PET sample with a missense mutation.

Although this study focused on detecting *BRCA1* mutations in PET, our findings should not be mutation specific, because the factors that most impeded our ability to successfully detect mutations (*i.e.,* false negatives) and amplify the PCR product likely reflect the different laboratory methods. For example, differences in the type of sequencing (manual *versus* automated) and the sequencing enzymes used by each of the laboratories may have had a significant effect on the processivity and nucleotide incorporation capabilities in the sequencing reaction. Laboratory method 3 may have been advantaged by using ThermoSequenase DNA polymerase for manually sequencing because the impure DNA templates from the PET samples can result in uneven peak patterns making it difficult to call single-base substitutions. Although convenient for high throughput work, automated sequencing is prone to misclassifying the small peaks of a real base as background, while identifying high background signals as real bases (23). In contrast, manual sequencing provides more uniform band intensities, allowing heterogeneity in the DNA template to be reproducibly identified without ambiguous interpretation of sequences using software that discerns signal strength patterns indicative of mixed-base positions. In our study, missed frameshift mutations accounted for a majority of the discrepant laboratory results reported by the laboratories (data not shown), possibly caused by the compression arising from the template secondary structure. An important caveat to this discussion is that every laboratory did not evaluate every sample included in the study because the originating laboratory could not test its own block. Laboratory 3 donated the oldest sample and, therefore, did not test it. Thus, in the absence of data, we can conjecture that this methodology may account for Laboratory 3's 100% success in amplifying samples, but only using younger blocks. Furthermore, primer position, length, and composition are known to affect the ability to amplify a DNA fragment as also do the method of fixation, the length of time between tumor excision and fixing, and the condition (*e.g.,* humidity etc.) of storage of the tissue blocks. Therefore, differences among these in the various laboratories may have contributed to differences in the frequency of successful amplification and mutation detection, in addition to the age of the block.

A prior study by Wong *et al.* (3) indicated that the vast majority of artifacts detected in fixed tissues were base substitutions, not insertion/deletion. So, in this study, we considered frameshift as well as base substitutions as two different types of mutations, even though sequencing is equally sensitive in detecting both types of mutations. The clone data indicate that the errors may be attributable to the specific nature of the mispairs (A:C or T:G) generated by the Taq polymerase because all of the additional variants were single-base substitutions, and almost all of those observed were transitions, with AT→GC being the most predominant (34 of 50; data not shown). Katagiri *et al.* (24) and Tindall *et al*. (25) found that the PCR error rate when using native Taq was estimable from the template product ratio of the PCR assay. However for PET, this may not be feasible because the amplification often involves the use of

crude DNA extracts to avoid loss of DNA during the purification process. Additionally, repeating all of the base-substitute variants for quality control is costly. An alternative would be to reanalyze the samples with the mutation using a technique that can differentiate true positives from false positives. For example, using retroviral or nonretroviral primers, which have a built-in signature sequence, would assure the detection of extremely low amounts of carryover as a form of false positives (26). Regardless, as these artifacts were observed only in the cloned samples, amplification using native Taq might be sufficient for use in large-scale epidemiological studies if the screening method is carefully selected and quality control maintained.

In summary, the primary goal of this study was to compare the accuracy of different techniques used for mutation detection in *BRCA1* to identify optimal conditions for analyzing PETs. We also examined whether artifacts would result from the routine mutation detection methods as performed in this study and, if so, whether the methods could be easily modified to minimize the number of artifacts without requiring a labor-intensive cloning technique to be used as quality control. Although our study is limited by the relatively small sample size within each analytic stratum, our factorial study design enabled us to simultaneously examine the effects of block age, mutation type, amplicon size, and polymerase type, on the accuracy of *BRCA1* mutation screening, and our results provide insight as to techniques that can be used most successfully. Specifically, we observed that mutation detection by direct sequencing of DNA isolated from PET was more accurate using younger blocks. For any aged block, the shorter amplicon length, regardless of the type of Taq, was more predictive of success, although HIFI Taq was optimal. Successful amplification greatly depended on the laboratory method used and on the age of the block. These results demonstrate that DNA derived from PET archived for up to 18 years can be used successfully to detect *BRCA1* gene mutations in a multicenter laboratory study if quality control is strictly maintained.

Acknowledgments

We are indebted to Dr. Andrew Futreal for his invaluable advice and guidance during the study design phase of this project. We are also grateful to Drs. Joellen Schildkraut and Marilie Gammon for their help with a related project that led to our developing this one and to Monica Katyal for her administrative help with the manuscript.

References

1. Gross, E., Arnold, N., Goette, J., Schwarz-Boeger, U., and Kiechle, M. A comparison of BRCA1 mutation analysis by direct sequencing, SSCP, and DHPLC. Hum. Genet., *105:* 72–78, 1999.

2. Shibata, D., Martin, W. J., and Arnheim, N. Analysis of DNA sequences in forty-year-old paraffin-embedded thin-tissue sections: a bridge between molecular biology and classical histology. Cancer Res., *48:* 4564–4566, 1988.

3. Wong, C., DiCioccio, R., Allen, H., Werness, B., and Piver, M. Mutations in BRCA1 from fixed, paraffin-embedded tissue can be artifacts of preservation. Cancer Genet. Cytogenet., *107:* 21–27, 1998.

4. Tworek, H., Peng, R., Fetzer, S., Werness, B., Piver, M., Allen, H., and DiCioccio, R. A. Mutation analysis of BRCA1, TP53, and Kras2 in ovarian and related pelvic tumors. Cancer Genet. Cytogenet., *112:* 105–118, 1999.

5. Wright, J., Teraoka, S., Onengut, S., Tolun, A., Gatti, R., Ochs, H., and Concannon, P. A high frequency of distinct *ATM* gene mutations in ataxia-telangiectasia. Am. J. Hum. Genet., *59:* 839–846, 1996.

6. Sambrook, J., Fritsch, E. F., and Maniatis, T. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1989.

7. Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L., Ding, W., Bell, R., Rosental, J., Hussey, C., Tran, T., McClure, M., Frye, C., Hattier, T., Phelps, R., Haugen-Strano, A., Katcher, H., Yakumo, K., Gholami, Z., Shaffer, D., Stone, S., Bayer, S., Wray, C., Bogden, R., Dayananth, P., Ward, J., Tonin, P., Narod, S., Bristow, P., Norris, F., Helvering, L., Morrison, P., Rosteck, P., Lai, M., Barrett, J., Lewis, C., Neuhausen, S., Cannon-Albright, L., Goldgar, D., Wiseman, R., Kamb, A., and Skolnick, M. A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. Science (Wash. DC), *266:* 66–71, 1994.

8. Smith, T. M., Lee, M. K., Szabo, C. I., Jerome, N., McEuen, M., Taylor, M., Hood, L., and King, M. C. Complete genomic sequence and analysis of 117 kb of human DNA containing the gene *BRCA1*. Genome Res., *6:* 1029–1049, 1996.

9. Ad hoc committee on mutation nomenclature update on nomenclature for human gene mutations. Hum Mutat, *8:* 197–202, 1996.

10. Frank, T., Svoboda-Newman, S., and Hsi, E. Comparison of methods for extracting DNA from formalin-fixed paraffin sections for nonisotopic PCR. Diagn. Mol. Pathol., *5:* 220–224, 1996.

11. Greer, C., Peterson, S., Kiviat, N., and Manos, M. PCR amplification from paraffin-embedded tissues. Am. J. Clin. Pathol., *95:* 117–124, 1991.

12. Goelz, S. E., Hamilton, S. R., and Vogelstein, B. Purification of DNA from formaldehyde fixed and paraffin embedded human tissue. Biochem. Biophys. Res. Commun., *130:* 118–126, 1985.

13. Akalu, A., and Reichardt, J. A reliable PCR amplification method for microdissected tumor cells obtained from paraffin-embedded tissue. Genet. Anal. Biomol. Eng., *15:* 229–233, 1999.

14. Jackson, D., Lewis, F., Taylor, G., Boylston, A., and Quirke, P. Tissue extraction of DNA and RNA and analysis by the polymerase chain reaction. J. Clin. Pathol., *43:* 499–504, 1990.

15. Pavelic, J., Gall-Troselj, K., Bosnar, M. H., Kardum, M. M., and Pavelic, K. PCR amplification of DNA from archival specimens. A methodological approach. Neoplasma (Bratisl.), *43:* 75–81, 1996.

16. Soong, R., and Iacopetta, B. A rapid and nonisotopic method for the screening and sequencing of *p53* gene mutations in formalin-fixed, paraffin-embedded tumors. Mod. Pathol., *10:* 252–258, 1997.

17. Nadji, M., Meng, L., Lin, L., Nassiri, M., and Morales, A. Detection of *p53* gene abnormality by sequence analysis of archival paraffin tissue: a comparison with fresh-frozen specimens. Diagn. Mol. Pathol., *5:* 279–283, 1996.

18. De Giorgi, C., Sialer, M. F., and Lamberti, F. Formalin-induced infidelity in PCR-amplified DNA fragments. Mol. Cell Probes, *8:* 459–462, 1994.

19. Dubeau, L., Chandler, L. A., Gralow, J. R., Nichols, P. W., and Jones, P. A. Southern blot analysis of DNA extracted from formalin-fixed pathology specimens. Cancer Res., *46:* 2964–2969, 1986.

20. Ben Ezra, J., Johnson, D., Rossi, J., Cook, N., and Wu, A. Effect of fixation on the amplification of nucleic acids from paraffin-embedded material by the polymerase chain reaction. J. Histochem. Cytochem., *39:* 351–354, 1991.

21. Marchetti, A., Merlo, G., Buttitta, F., Pellegrini, S., Callahan, R., Bistocchi, M., and Squartini, F. Detection of DNA mutations in acid formalin-fixed paraffin-embedded archival tumor specimens by polymerase chain reaction-single strand conformation polymorphism analysis. Cancer Detect. Prev., *19:* 278–281, 1995.

22. Alaibac, M., Filotico, R., Giannella, C., Paradiso, A., Labriola, A., and Marzullo, F. The effect of fixation type on DNA extracted from paraffin-embedded tissue for PCR studies in dermatopathology. Dermatology, *195:* 105–107, 1997.

23. Kwok, P. Y., Carlson, C., Yager, T. D., Ankener, W., and Nickerson, D. A. Comparative analysis of human DNA variations by fluorescence-based sequencing of PCR products. Genomics, *23:* 138–144, 1994.

24. Katagiri, T., Emi, M., Ito, I., Kobayashi, K., Yoshimoto, M., Iwase, T., Kasumi, F., Miki, Y., Skolnick, M., and Nakamura, Y. Mutations in the *BRCA1* gene in Japanese breast cancer patients. Hum. Mutat., *7:* 334–339, 1996.

25. Tindall, K. R., and Kunkel, T. A. Fidelity of DNA synthesis by the Thermus aquaticus DNA polymerase. Biochemistry, *27:* 6008–6013, 1988.

26. Abbott, L. Z., Spicer, T., Bryz-Gornia, V., Kwok, S., Sninsky, J., and Poiesz, B. Design and use of signature primers to detect carry-over of amplified material. J. Virol. Methods, *46:* 51–59, 1994.