

## ORIGINAL ARTICLE

## New approaches for genotyping paraffin wax embedded breast tissue from patients with cancer: the Iowa women's health study

B Thyagarajan, K E Anderson, F Kong, F R Selk, C F Lynch, M D Gross

*J Clin Pathol* 2005;58:955-961. doi: 10.1136/jcp.2004.023374

See end of article for authors' affiliations

Correspondence to:  
Myron D Gross,  
Department of Laboratory  
Medicine and Pathology,  
University of Minnesota,  
MMC 609, 420 Delaware  
Street SE, Minneapolis,  
MN-55455, USA; gross@  
epi.umn.edu

Accepted for publication  
8 February 2005

**Background:** The use of paraffin wax embedded tissue samples as a source of DNA for genotype analysis has been limited because of difficulties in DNA extraction and single nucleotide polymorphism (SNP) analysis.

**Aims:** To test the feasibility of applying the combination of a commonly used DNA isolation procedure, PureGene, and a high throughput SNP analysis method, the polymerase chain reaction (PCR)-INVADER assay, to genotype several types of paraffin wax embedded breast tissues.

**Methods:** Twenty formalin fixed, paraffin wax blocks were obtained from five participants in the Iowa women's health study. Each participant provided several types of tissue including normal lymph node, normal nipple/areola tissue, inflammatory/fibrotic breast tissue, or normal breast tissue, and tumour tissue.

**Results:** Good quality DNA (260/280 ratio >1.6) was obtained from all tissues. Normal lymph nodes yielded the largest amount of DNA (97.1 µg). DNA obtained from the samples was tested for a germline C1183T polymorphism in the MnSOD gene by three methods—PCR-RFLP (restriction fragment length polymorphism), INVADER assay, and PCR-INVADER assay. Of the 20 samples, PCR-RFLP genotyped 16, the PCR-INVADER assay 18, and the INVADER assay two. This methodology was then used to analyse five additional genotypes and confirmed the general applicability of the method.

**Conclusions:** This study demonstrated the feasibility of (1) using several paraffin wax embedded breast tissues as a source of DNA for germline genetic analysis, with lymph nodes providing the highest yield, and (2) using the combination of a common extraction method with a high throughput SNP analysis method, the PCR-INVADER assay.

Numerous single nucleotide polymorphisms (SNPs) have been identified in susceptibility genes, which may modify the risk of cancer by influencing metabolic responses to various environmental exposures.<sup>1-3</sup> These are of interest for association studies to identify genetic determinants of complex diseases, such as cancer and cardiovascular diseases.<sup>4</sup> One of the problems in using established epidemiological studies to answer these questions is that traditional sources of DNA, such as white blood cells, are not always available. However, in many studies investigating cancer aetiology, paraffin wax embedded tissue is readily available from cases. The use of this DNA for the study of these associations within existing large cohort studies, which have collected exposure information over a long period of time, is clearly more efficient than designing new studies. To date, the use of paraffin wax embedded tissue samples as a source of DNA has been limited for a variety of reasons, including difficulties in DNA extraction and DNA amplification.<sup>5</sup> Recently, modifications in DNA extraction procedures, such as longer incubation times with proteinase K, increased quantity of proteinase K, and increased alkalinity of the buffer solution, have all been used to improve the DNA yield from paraffin wax embedded tissues.<sup>6-7</sup> In addition, it has been shown that short segments of DNA, suitable for polymerase chain reaction (PCR) amplification, can be obtained from paraffin wax embedded tissues.<sup>8-8</sup> Thus, with improved methods for DNA extraction and DNA amplification, archived paraffin wax embedded tissue blocks may be used further in epidemiological studies as a source of DNA.<sup>8-9</sup>

"Information on which tissue provides the best source of DNA would be valuable in the design of epidemiological studies and the collection of appropriate samples for genetic analysis"

Numerous tissue types such as tumour, normal breast, and lymph node are routinely obtained during mastectomy and archived in paraffin wax in pathology laboratories. However, it is still not known whether particular tissues provide higher quality DNA in larger amounts compared with others. Information on which tissue provides the best source of DNA would be valuable in the design of epidemiological studies and the collection of appropriate samples for genetic analysis.

Even though genotyping using DNA obtained from paraffin wax embedded tissues has been reported using traditional methods such as PCR followed by restriction length fragment polymorphism (RFLP), these methods are slow and labour intensive for application in large studies. Recently, many high throughput genotyping technologies have become commercially available.<sup>10</sup> Examples of these technologies include Taqman PCR, INVADER assay, pyrosequencing, and mass spectrometry. These advances have now made it possible to genotype thousands of individuals in a short period of time. Taqman PCR has been used to perform high throughput genotyping analysis from paraffin wax

**Abbreviations:** MnSOD, manganese superoxide dismutase; PCR, polymerase chain reaction; RFLP, restriction fragment length polymorphism; SNP, single nucleotide polymorphism

embedded tissues.<sup>9</sup> However, many of these technologies remain untested regarding their ability to genotype DNA obtained from paraffin wax embedded tissues.

In our study, we compared five tissue types (tumour tissue, lymph node, normal breast, diseased but non-tumorous breast tissue, and nipple/areola tissue) to determine which tissue provided the highest quality DNA in the greatest amount. The accuracy and sensitivity of the PCR-INVADER and INVADER assays were compared with PCR-RFLP in detecting a germline C1183T polymorphism in the manganese superoxide dismutase (MnSOD) gene. In addition, we assessed the reproducibility of the PCR-INVADER assay by its ability to genotype eight pairs of blinded duplicate paraffin wax embedded samples and eight blinded blank samples (paraffin wax blocks with no tissue) for seven different DNA repair gene SNPs and the feasibility of applying this methodology as a high throughput method to genotype several SNPs of DNA obtained from paraffin wax blocks.

## MATERIALS AND METHODS

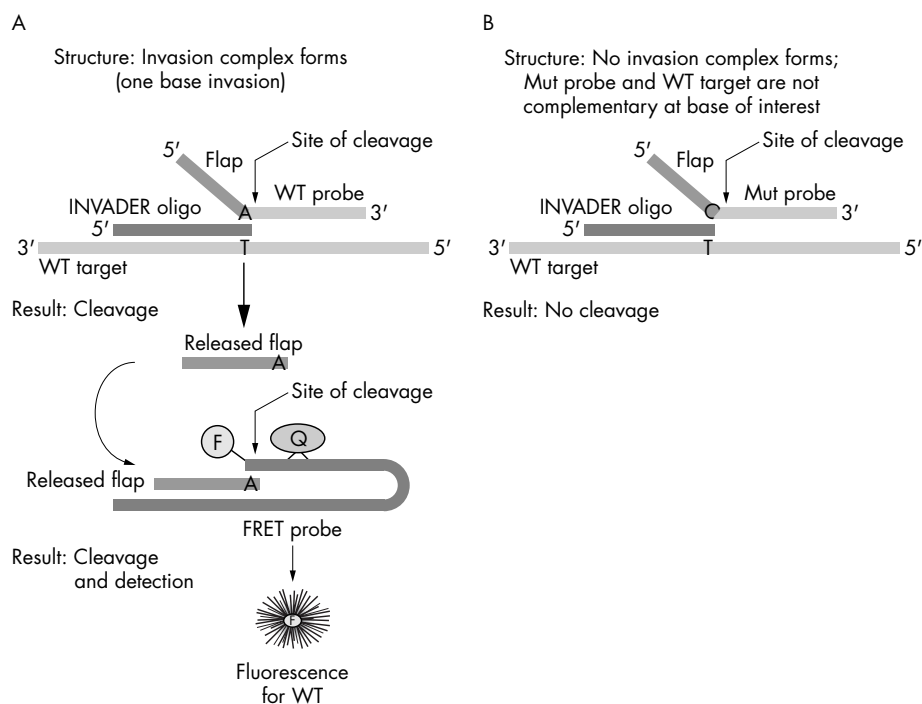
### Samples

Twenty formalin fixed, paraffin wax embedded tissue blocks were obtained from five participants in the Iowa women's health study who had undergone modified radical mastectomy for the treatment of breast cancer in 1994. The different types of tissue obtained from each participant comprised samples from tumour (infiltrating ductal carcinoma and lobular carcinoma), normal lymph node, normal nipple/areola, inflammatory/fibrotic, and normal breast. Forty 10 µm sections were cut from each block and stored at 4°C until DNA extraction. DNA was extracted from 10 of these sections. Fifteen of the 20 blocks contained tissue in at least 20% of the section surface. In addition to these 20 samples, 16 blinded duplicate samples (eight pairs) were obtained from normal lymph nodes of patients with breast cancer and eight blank samples were cut from paraffin wax blocks not

containing tissue. These blinded samples were collected as a part of an ongoing study that collected paraffin wax blocks from cancer cases within the Iowa women's health study cohort. These samples were processed using the same techniques as described for the 20 paraffin wax embedded tissue samples. The University of Iowa human subjects office approved our study protocol.

### DNA extraction

Precautions to prevent contamination were taken during the cutting of the paraffin blocks and during DNA extraction. These included wearing gloves, changing blades for sectioning, and extensive cleaning of utensils. These methods have been used previously and were shown to prevent technician and cross contamination of samples. DNA was extracted using a PureGene DNA extraction kit (Gentra Systems, Minneapolis, Minnesota, USA). This method uses a salting out procedure to extract DNA from biological samples.<sup>11</sup> DNA extraction was performed using procedures described by the manufacturer with minor modifications. Briefly, the paraffin wax was dissolved in 300 µl of Citrisolv (xylene substitute) and washed with ethanol to remove the Citrisolv. Then, 300 µl of the cell lysis solution was added, followed by the addition of 3 µl of proteinase K (20 mg/ml) and incubation overnight at 55°C. After cooling to room temperature, 100 µl of the protein precipitation solution was added and the cell lysate was centrifuged. The supernatant was isolated, 300 µl of 100% isopropanol was added, and the solution was centrifuged to precipitate the DNA. The DNA was washed with 300 µl of 70% ethanol, dried, and 100 µl of DNA hydration solution was added. The amount of DNA was measured by absorbance at 260 nm using an ultraviolet spectrometer, and the purity of the DNA was evaluated by measuring the 260/280 absorbance ratios. The minor modifications in this procedure were the use of a longer incubation time (overnight, approximately 16 hours) and doubling the concentration of protease K used.



**Figure 1** Schematic representation of the monoplexed INVADER<sup>®</sup> assay. The Cleavase<sup>®</sup> enzyme recognises the invasive structure formed by the probe and INVADER oligonucleotides in the presence of the appropriate target. ©Third Wave Technologies, Inc 2003; figure and legend reproduced with permission from Third Wave Technologies Inc. Mut, mutant; WT, wild-type.

**Table 1** PCR reaction mixtures for 5 DNA repair gene SNPs

Reagents	XPA (Ala23Gly) (12.5 µl)	XPD-23 (Lys751Gln) (12.5 µl)	XPG-15 (Asp1104His) (12.5 µl)	XRCC3 -7 (Thr241Met) (12.5 µl)	XRCC1-6 (Arg194Trp) (6.25 µl)
DNTP	1	1	1	1	1
MgCl <sub>2</sub>	1.25	1.25	1.25	1.25	1.25
10× buffer	1.25	1.25	1.25	1.25	1.25
Primer (F)	0.5	0.5	0.5	0.5	0.5
Primer (R)	0.5	0.5	0.5	0.5	0.5
Taq	0.2	0.125	0.1	0.1	0.1
H <sub>2</sub> O	1.55	1.625	1.65	1.65	3.4
DNA	6.25	6.25	6.25	6.25	5

F, forward; R, reverse; PCR, polymerase chain reaction; SNP, single nucleotide polymorphism.

### Quality of the slides

The cellularity of the tissue and the percentage of the block occupied by tissue were determined by a single pathologist (CFL) by reviewing a haematoxylin and eosin stained thin section from the tissue block under a light microscope. The cellularity was judged by categorising the number of nuclei in the tissue section into the following subgroups: 1, low; 2, low to medium; 3, medium; 4, medium to high; and 5, high. Lymph nodes typically had high cellularity. The percentage of the block occupied by tissue was assessed by categorising the tissue section into the following subgroups: 1, < 20%; 2, 20–39%; 3, 40–59%; 4, 60–79%; and 5, > 80%. Although the exact duration of time that the tissue had been kept before embedding in paraffin wax was not available, we estimated it from the time between surgery and the date of the pathology report. The date of the pathology report was used to calculate how long the samples were stored in paraffin wax before DNA extraction. All 20 breast tissue samples were embedded in paraffin wax for approximately six years. All samples, except those from patient 1, were processed on the same day as the date of surgery. The samples for patient 1 were processed two days after the date of surgery.

### Genetic testing

The samples were tested for a germline C1183T polymorphism in the MnSOD gene (EMBL, accession number S77127)<sup>12</sup> using three methods.

### PCR-RFLP

Initially, all samples were genotyped using a PCR-RFLP method similar to that published previously.<sup>13</sup> The DNA samples were amplified as follows: 2 µl of dNTP (10mM) (GeneAmp; AME Bioscience Ltd, Havnaasveien, Norway), 2 µl of 25 mmol MgCl<sub>2</sub> (Promega, Madison, Wisconsin, USA), 2.5 µl of 10× buffer, 2 µl of each primer (20 pg/µl) (GeneSys, South San Francisco, USA), 0.25 µl of Taq polymerase (5 U/µl) (Promega), and 5–40 ng of sample DNA were added to a 25 µl reaction. The following primers were used in the reaction: MNSOD (forward), 5'-AGCCCAGCCTGCGTAGAC-3' and MNSOD (reverse), 3'-TACTTCTCCTCGGTGACG-5'. The reaction conditions were as follows: denaturation at 94°C for six minutes, then 35 cycles of 94°C for one minute, 60°C for one minute, and 72°C for two minutes, followed by a final extension at 72°C for 10 minutes. The resulting PCR product was a 246 bp fragment. The C1183T polymorphism created a BsaWI restriction site that was detected by digestion with BsaWI at 60°C for two hours. The presence of the SNP yielded two bands of 82 bp and 164 bp. A 3% agarose gel was prepared and stained with ethidium bromide. Agarose gel electrophoresis was then performed and the digested products were read using ultraviolet light.

**Table 2** PCR primer sequences and conditions for 5 DNA repair gene SNPs

Primers	Sequences	PCR conditions
XPA (new) Forward	5'-AGG CGC TCT CAC TCA GAA AGG	94°C/30 sec, 61°C/1.5 min, 72°C/1 min for 35 cycles
Reverse	5'-TCC GCG GGT TGC TCT AAA G	
XPD-23 (new) Forward	5' CCC CCT CTC CCT TTC CTC TG 3'	94°C/30 sec, 60°C/1 min, 72°C/ 1 min for 35 cycles
Reverse	5' AAC CAG GGC CAG GCA AGA C	
XPG-15 Forward	5'-GAA AGA ATA CAT GCG GTG	94°C/30 sec, 59°C/3 min for 30 cycles
Reverse	5'-GGG AGC TTC CTT CAC TGA	
XRCC1-6 Forward	5'-TGA AGG AGG AGG ATG AGA GC	94°C/1 min, 63°C/ 30 sec, 72°C/30 sec for 35 cycles
Reverse	5'-CCC TAC TCA CTC AGG ACC CA	
XRCC3-7 Forward	5'-TGT GAA TTT GAC AGC CAG G	94°C/1 min, 60°C/ 1 min, 72°C/30 sec for 35 cycles
Reverse	5'-AAA ATA CGA GCT CAG GGG TG	

PCR, polymerase chain reaction; SNP, single nucleotide polymorphism.

### INVADER reaction

The 20 samples from five patients with breast cancer were also genotyped, without amplification, by the INVADER assay (fig 1) (Third Wave Technologies Inc, Madison, Wisconsin, USA).<sup>14</sup> This assay is based on the activity of a group of enzymes known as Cleavase enzymes.<sup>15</sup> These enzymes cleave DNA at particular sites by recognising substrates formed by the addition of two oligonucleotide probes to DNA. All INVADER reactions require a reporter oligonucleotide probe, the FRET probe (with internal quencher), MgCl<sub>2</sub>, and the Cleavase enzyme. In addition, specific wild-type and mutant probes are used to detect the SNPs. The INVADER assay comprises two reactions—a primary reaction, which depends on the amount of DNA present in the reaction, and a secondary reaction, which depends on the amount of “flaps” that are released from the primary reaction. Both reactions were carried out near the melting temperature of the wild-type/mutant probes (65°C).

The INVADER reactions were performed using the monoplex format in two wells (fig 1). The first well contained an oligo probe and the wild-type probe and the second well contained the oligo probe and the mutant probe. In the presence of the appropriate bases in the DNA sequence, these probes combine with DNA to form a particular tertiary

structure, which the Cleavase enzyme recognises and cleaves. The flap that is released by the cleavage then enters a secondary reaction. In the secondary reaction, it interacts with the FRET probe to form a structure recognised by the Cleavase enzyme. The Cleavase enzyme then cleaves this structure and the released FRET probe emits fluorescence, which is measured by a fluorescence detector. The secondary reaction is self-amplificatory because many FRET probes are released from a single flap produced by the primary reaction. Thus, instead of amplifying the target DNA, and increasing the chances of carryover contamination, this procedure accumulates a signal molecule only when the specific target DNA sequence is present.

The reaction was performed in 96 well plates, which contained 100 ng of Cleavase enzyme and 0.33 $\mu$ M of FRET probe in a dried down format. For each reaction, 5  $\mu$ l of 22.5 mmol MgCl<sub>2</sub>, 5  $\mu$ l of appropriate probe, and 5  $\mu$ l of DNA (20  $\mu$ g) were added. Each sample was added to two wells, one containing the wild-type probe and the other containing the mutant probe. This reaction mixture was initially denatured at 94°C for five minutes and then incubated at 65°C for four hours. A Cytoflour 4400 fluorescence microtitre plate reader was used to measure fluorescence after the four hour incubation period (excitation 485/20, emission 530/25, and gain 36). MS Excel was used to capture and analyse the data and genotypes were calculated based on the ratio of the strength of fluorescence of alleles 1 and 2. Ratios  $\geq$  5 were considered wild-type homozygous, ratios less than 0.2 were read as mutant homozygotes, and ratios between 0.3 and 3.5 were considered heterozygotes. If the ratios were between 0.2 and 0.3 or between 3.5 and 5.0 then genotypes could not be determined for the particular sample.

#### PCR-INVADER reaction

All the samples were genotyped using the PCR-INVADER reaction. The PCR-INVADER reaction consisted of a short cycle PCR, followed by dilution of the PCR product, and then use of the diluted PCR product as a template for the INVADER reaction. Initially, a limited PCR reaction for MnSOD (20 cycles) was performed in a 12.5  $\mu$ l reaction volume. All the PCR conditions and reagents were identical to the those described earlier. A smaller reaction volume was

used in this case, so that both the PCR reaction and the subsequent dilution could be carried out in a single well of a 96 well plate. The PCR product obtained was diluted 10 fold and this served as the sample for the INVADER reaction. The INVADER reaction was performed in 96 well plates, as described above, with 5  $\mu$ l of the diluted PCR product as the DNA template. This reaction mixture was initially denatured at 94°C for five minutes and then incubated at 65°C for one hour. The Cytoflour 4400 fluorescence microtitre plate reader was used to measure fluorescence after the one hour incubation period (excitation 485/20, emission 530/25, and gain 36). MS Excel was used to analyse the data and genotypes were calculated based on the ratio of the strength of fluorescence of alleles 1 and 2.

#### Controls

The blinded duplicate samples and blank samples were genotyped using only the PCR-INVADER methodology. Bplex PCR-INVADER was used to genotype the blinded duplicate samples. All reaction conditions for the bplex invader were identical to the monoplex PCR-INVADER assay except that the bplex PCR-INVADER reactions were performed in a single well whereas the monoplex PCR-INVADER assay required two separate wells. Individual PCR primers were designed for each of the five SNPs in the DNA repair genes. The individual PCR primers and the PCR conditions are listed in tables 1 and 2. For the blinded duplicate samples a 30–35 cycle PCR was performed because sufficient amplification could not be obtained with a 20 cycle PCR. Individual INVADER probes were designed for each of the seven different polymorphisms and a PCR-INVADER assay was performed for each genotype. Duplicate pairs were unblinded only after the genotypes were obtained for all samples.

#### RESULTS

The amounts of DNA isolated from 10  $\times$  10  $\mu$ m sections ranged from 1.3 to 215  $\mu$ g (mean, 57.4; SD, 58). The highest to lowest mean yield of DNA to tissue was obtained from normal lymph node (97.7  $\mu$ g), breast cancer (84.3  $\mu$ g), nipple/areola (34.3  $\mu$ g), diseased but non-cancerous (inflammatory/fibrotic) breast (31.2  $\mu$ g), and normal breast

**Table 3** Comparison of genotypes for a germline C1183T polymorphism in the MnSOD gene obtained from paraffin wax embedded tissue samples by three genotyping methods

Sample	Patient	Tissue type	Total DNA ( $\mu$ g)	260/280 ratio	Cellularity*	% Tissue in block	Genotype by analysis method		
							PCR-RFLP	INVADER assay	PCR-INVADER assay
1	1	Lymph node	36.3	1.95	5	<20	N/A	N/A	N/A
2	1	Nipple	28.7	1.82	2	<20	N/A	N/A	C
3	1	Normal breast	11.6	1.7	1	<20	N/A	N/A	C
4	1	Breast disease	62.8	1.8	3	60–79	N/A	N/A	N/A
5	2	Lymph node	84.1	1.77	5	20–39	HET	N/A	HET
6	2	Nipple	41.9	1.58	1	20–39	HET	N/A	HET
7	2	Cancer	38.2	1.53	3	40–59	HET	N/A	HET
8	2	Breast disease	45.5	1.35	2	60–79	HET	N/A	HET
9	2	Breast disease	9.4	1.41	1	60–79	HET	N/A	HET
10	3	Lymph node	55.3	1.93	4	<20	C	N/A	C
11	3	Nipple	62.9	1.63	2	40–59	C	N/A	C
12	3	Breast disease	45.2	1.91	2	60–79	C	C	C
13	3	Breast disease	22.9	1.9	2	<20	C	N/A	C
14	4	Cancer	194.8	1.68	4	20–39	HET	N/A	HET
15	4	Cancer	100.7	1.41	2	40–59	HET	N/A	HET
16	4	Cancer	2.65	1.78	1	60–79	HET	N/A	HET
17	5	Lymph node	215	1.69	5	20–39	T	N/A	T
18	5	Nipple	3.8	1.44	2	40–59	T	N/A	T
19	5	Breast disease	1.3	1.39	2	20–39	T	N/A	T
20	5	Cancer	85	1.69	4	20–39	T	C	T

C, homozygous for the C allele; HET, heterozygous; N/A, not available; T, homozygous for the T allele.

\*1, low; 2, low to medium; 3, medium; 4, medium to high; 5, high.

**Table 4** Comparison of the genotypes of 5 DNA repair gene SNPs across 8 pairs of blinded duplicate samples

Samples	XPA Ala23Gly	XPD-23 Lys751Gln	XPG-15 Arg1104His	XRCC3-7 Thr241Met	XRCC1-6 Arg194Trp	DNA quantity (µg)	260/280
Pair 1	HET	G	HET	C	C	108.20	1.72
Pair 1	HET	G	HET	C	C	55.40	1.78
Pair 2	HET	G	G	C	C	5.60	1.88
Pair 2	HET	G	G	C	C	101.90	1.81
Pair 3	G	G	G	HET	C	25.73	1.47
Pair 3	G	G	N/A	HET	C	11.13	1.57
Pair 4	G	HET	HET	HET	C	96.06	1.30
Pair 4	G	HET	C	HET	C	134.64	1.71
Pair 5	N/A	G	C	T	C	25.53	1.86
Pair 5	G	G	C	T	C	17.13	1.58
Pair 6	HET	HET	HET	C	C	28.40	1.71
Pair 6	HET	HET	HET	C	C	18.86	1.61
Pair 7	HET	HET	G	C	C	76.38	1.80
Pair 7	G	N/A	N/A	N/A	C	150.81	1.30
Pair 8	HET	G	C	HET	C	82.68	1.85
Pair 8	N/A	N/A	N/A	N/A	N/A	N/A	N/A

C, homozygous for the C allele; G, homozygous for G allele; HET, heterozygous; N/A, not available; SNP, single nucleotide polymorphism; T, homozygous for the T allele.

(11.6 µg). The purity of DNA as determined by the A260/280 ratio was high (1.8, highly pure) for lymph nodes and slightly lower (1.6) for all other tissue types.<sup>16</sup>

Fourteen of the 20 specimens had low to medium cellularity (table 3). The remaining specimens had high cellularity and were primarily lymph node tissue sections. For all tissue types, the amount of DNA increased with increasing cellularity and increasing amount of tissue present in the section. In this limited set of 20 samples, tissue cellularity was a stronger predictor of DNA quantity than the amount of tissue present in the section.

Sixteen of the 20 samples were amplified sufficiently for PCR-RFLP analysis (table 2). The remaining samples did not yield sufficient amounts of DNA (samples 1, 2, 3, and 4); three of these samples had tissue in less than 20% of the section surface (samples 1, 2, and 3). When the INVADER assay was used without prior PCR amplification, genotypes were obtained for only two of the 20 samples (samples 12 and 20). One of the genotypes obtained did not correspond to the genotype obtained by the PCR-RFLP method (sample 20). Eighteen of the 20 samples yielded genotypes when the DNA samples were first amplified using 20 cycles of PCR, the amplified product diluted, and then used as a DNA template for the INVADER reaction. Two of the samples that could not be amplified by PCR-RFLP were genotyped by the PCR-INVADER assay (samples 2 and 3). The two samples that could not be genotyped did not amplify in the PCR-RFLP reaction (samples 1 and 4). One of these samples had tissue in less than 20% of the section surface (sample 1). Results obtained by both the PCR-INVADER assay and the PCR-RFLP method were 100% concordant, indicating that the PCR step was essential and that PCR-INVADER is a reliable method to genotype DNA from tissue samples.

The amount of DNA obtained from the blinded duplicate samples (eight pairs) was similar to that obtained from the 20 samples listed above (table 4). An adequate amount of DNA was obtained from 15 of the 16 samples (seven pairs). DNA could not be obtained from one sample even after multiple extraction attempts. The amounts of DNA obtained from 10 × 10 µm sections ranged from 5.6 to 150.8 µg (mean, 62.6; SD, 47.9). The 260/280 ratio ranged from 1.6 to 1.8, indicating good quality DNA. Comparisons between blinded duplicates were performed on seven pairs of samples. Individual genotypes were not obtained for some samples among the seven pairs, so that the number of pairs compared varied by genotype (tables 4, 5). There was good concordance between the blinded duplicate samples for each of five genotypes. There were no discordant pairs for the XRCC3-7, XPD-23, and XRCC1-6 genotypes (tables 4, 5). There was only one discordant pair for XPA and XPG-15. In both cases where discordance was seen within duplicate pairs (XPA and XPG-15), one sample from a pair was genotyped as a heterozygote whereas the other sample from the pair was genotyped as a homozygous variant (table 4). Overall, 35 blinded sample pairs were available for comparison (five SNPs were analysed in seven blinded sample pairs). Individual genotypes were not obtained for five of these sample pairs. The overall concordance rate for the blinded duplicate samples across all five genotypes was 93.3% (28 of 30 sample pairs).

No DNA and hence no genotypes were obtained from the blank samples.

**DISCUSSION**

High quality DNA was obtained from each of the five breast tissue types. The amount of tissue present in sections was lowest in lymph nodes. However, tissue cellularity was highest in lymph nodes. The highest amounts of DNA were obtained from benign lymph nodes and the lowest amounts of DNA were obtained from normal breast tissue. These results support the feasibility of using different types of breast tissue for genotyping and, when given a choice, lymph node tissue would be the preferred source of germline DNA for genetic epidemiology studies involving tissue removed for breast cancer.

Although both the INVADER assay and the PCR-INVADER assay have been used to genotype DNA from blood samples previously,<sup>17-19</sup> this is the first time that the INVADER assay or the PCR-INVADER assay has been used to genotype tissue samples stored in paraffin wax blocks. The INVADER assay when used alone was able to genotype only one of 20 samples

**Table 5** Concordant and discordant pairs for 5 genotypes\*

SNP	Concordant	Discordant	Missing
XPA	5	1	1
XPD-23	6	0	1
XPG-15	4	1	2
XRCC3-7	6	0	1
XRCC1-6	7	0	0

\*One pair (pair 8) was excluded from these analyses because DNA could not be obtained from one of the samples.  
SNP, single nucleotide polymorphism.

correctly. Most genotypes were obtained by both PCR-RFLP and PCR-INVADER. A possible explanation for the vastly improved detection of genotypes after using PCR-INVADER (compared with the INVADER assay alone) is that compounds interfering with the INVADER assay would be diluted after amplification. The PCR-INVADER failure rate, with this limited number of samples, was 10%. Subsequent studies, performed with thousands of paraffin embedded tissue samples have yielded a failure rate of 5–7% (data not shown). This failure rate is higher than that reported for blood samples by other studies, which have reported a failure rate of 2.3%.<sup>18</sup> Regardless of the technology, DNA from blood generally has a lower failure rate than DNA from paraffin wax blocks. The difference in failure rate may be a result of interfering substances in DNA from blocks not found in DNA from blood samples. The failure rate for the PCR-INVADER assay has also been reported to be higher than that found using Taqman PCR (0%)<sup>9</sup> in a study using paraffin wax embedded tissue. The failure rates for genotyping paraffin wax embedded tissue samples have not been published for other high throughput technologies such as Sequenom (MassARRAY matrix assisted laser desorption time of flight mass spectrometry) and pyrosequencing.

The effectiveness of the extraction procedure was one possible factor in our failure rates (5–7%) with the PCR-INVADER technology and we considered alternative procedures. Our initial objective was to develop a method that could be used readily by many laboratories, that would be easy to implement, and that would have the capacity for high throughput genotyping. Thus, we tested commonly used methods of extraction and found that the PureGene extraction method, with our minor modifications, meets these criteria, although failure rates may be lower with alternative procedures. A primary consideration is the technical complexity of alternative methods and the feasibility of using them in population studies. A method by Inadome and Noguchi involves high performance liquid chromatography purification of high and low molecular weight DNA.<sup>20</sup> This method would probably improve the quality of some DNA samples. However, it is rather laborious for large numbers of samples, still depends on an effective extraction step, is costly, and requires a considerable amount of DNA. For population studies, the cost–benefit ratio may be favourable only in a limited application, wherein samples that require larger amplification products or are difficult to amplify would undergo this procedure. A second procedure by Shi *et al* uses heating and basic pH conditions for DNA extraction.<sup>6</sup> This method may provide larger amounts of better quality DNA, and should be explored further for population studies. Our initial testing of this method found that it was feasible for population studies, but did not show a significant improvement over our current method (data not shown). Only a small difference would be expected because the major effect of this method results from differences in pH from 2 to 9 and the pH of the cell lysis solution used in our method is 7–8, which is only slightly lower than the optimal pH (9) in the Shi *et al* procedure. Thus, we believe that the method reported here is a cost effective method suitable for most tissue culture samples, which possibly could be improved by additional minor modifications.

In our study, smaller amounts of DNA could be used to genotype the samples by the PCR-INVADER assay than by the conventional INVADER assay. We were able to genotype tissue samples using 2–40 ng of DNA, whereas 100 ng was needed for the conventional INVADER assay. Because DNA obtained from formalin fixed, paraffin wax embedded tissues is usually degraded, PCR amplification works well only for small size fragments, usually under 200–300 bp.<sup>5</sup> Thus, primer pairs need to be designed to amplify small fragment

lengths. In addition, we found that longer incubation times or, in some cases, an increased number of PCR cycles (as seen in the duplicate samples) may be required for accurate genotyping of tissue samples by the PCR-INVADER assay.

**“The PCR-INVADER method is also highly suited for automation because it is performed entirely in a microtitre format”**

The results from our blinded duplicate samples indicate that the PCR-INVADER assay is a reliable genotyping platform. Overall, the 93.3% concordance rate indicates that genotyping error using DNA obtained from tissue blocks is low (7%). However, the presence of discordant samples (0–20%) indicates that genotypes obtained from paraffin wax embedded samples are not as reliable as those obtained from blood samples. Genotypes obtained using the PCR-INVADER assay on blood samples were 98–100% reproducible (data not shown). The lower reproducibility rate for some genotypes found in our study can explain the discrepancy seen between genotypes obtained from paraffin wax embedded tissue and genotypes obtained from fresh tissue.<sup>8</sup> The two discordant pairs were different sample pairs for XPA and XPG-15, indicating that the genotyping error was random, rather than a result of sample contamination. Genotyping error is a potential source of misclassification error in molecular epidemiological studies. The low rate of discrepancy seen in our pilot study indicates that this genotyping error may not lead to substantial misclassification bias in molecular epidemiological studies, particularly when the allele frequencies are highly prevalent (> 15%),<sup>21</sup> although it could potentially bias the results of those studies in which the allele frequencies are less than 15%.<sup>21</sup> However, the PCR-INVADER assay does not work well for SNPs at all loci. SNPs in XRCC1 Asn399Gln and XPD Asp312Asn could not readily be genotyped using the PCR-INVADER assay even after adequate amplification was obtained by PCR. At present, we are attempting to genotype the paraffin wax embedded tissues for these two SNPs using other genotyping platforms.

The PCR-INVADER assay can be performed relatively quickly compared with PCR-RFLP or the traditional INVADER assay. The PCR-INVADER method is also highly suited for automation because it is performed entirely in a microtitre format by adding the INVADER reagents to the diluted PCR products. Gel electrophoresis and manual data entry are not required, reducing the amount of labour required and the error rate significantly.

The PCR-INVADER assay appears to be suitable for high throughput genotyping of DNA from paraffin wax blocks. However, before using the PCR-INVADER assay in a high throughput genotyping platform, several safeguards need to be established to ensure that genotypes are identified accurately. In addition to running validated controls with each sample run, it is also necessary to check the accuracy of the PCR-INVADER assay at each locus. To address this problem it is necessary to incorporate blinded duplicate samples into the study during study design so that they can be used to evaluate the reproducibility of the PCR-INVADER assay at each locus. In addition, with increases in the number of genotypes being studied, multiple genotyping platforms may be needed to genotype all loci accurately.

Potential limitations of the PCR-INVADER assay include its sensitivity to the amount of DNA used in the INVADER reaction and the accuracy of the PCR-INVADER assay at different loci. The sensitivity of the PCR-INVADER assay to the amount of DNA is demonstrated by varying parameters such as the dilution and length of incubation. At incubation times longer than one hour, some samples identified as

### Take home messages

- We found that several types of paraffin wax embedded breast tissues could be used as a source of DNA for germline genetic analysis, and that lymph nodes provided the highest yield
- Combining a common DNA extraction method with a high throughput single nucleotide polymorphism analysis method (the polymerase chain reaction-INVADER assay) provided a suitable method for high throughput genotyping of DNA from paraffin wax blocks

homozygous for the mutation by PCR-RFLP were read as heterozygotes. This was mainly the result of the increased non-specific reactivity of one of the probes (mutant/wild-type) at longer incubation times. Diluting the samples less (for example, 1/3 or 1/5) resulted in similar problems, whereas diluting them more did not yield genotypes for many samples. The standard parameters used to check the validity of the reaction (controls supplied by Third Wave) were not useful in detecting incorrect genotypes at longer incubation times. To ensure that all the genotypes are interpreted correctly, validated samples need to be assayed along with the tissue samples. The genotypes of the validated samples were determined previously by another genotyping methodology, such as PCR-RFLP. Adjusting various parameters, such as incubation time and dilution, and using validated samples as controls allowed for the accurate determination of genotypes in tissue samples. In spite of these limitations, the six PCR-INVADER reactions analysed using a standard dilution procedure and a one hour incubation time have produced consistent and accurate results, indicating that the PCR-INVADER methodology is robust over a range of DNA concentrations.

In summary, we have demonstrated the feasibility of using paraffin wax embedded tissues as a source of DNA for genetic studies. We have also shown that DNA obtained from paraffin wax blocks can be used to genotype samples using PCR-INVADER, a high throughput assay.

### ACKNOWLEDGEMENTS

Thanks to Dr A Folsom (principal investigator of IWHS) for providing study samples, advice on analysis, and manuscript preparation.

### Authors' affiliations

**B Thyagarajan, Kristin E Anderson, F Kong**, University of Minnesota, Division of Epidemiology, Suite 300, West Bank Office Building, Minneapolis MN-55454, USA  
**F R Selk, C F Lynch**, Department of Epidemiology, University of Iowa, Iowa City, IA 52242, USA

**Myron D Gross**, Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, MN-55455, USA

### REFERENCES

- 1 **Zheng WAC**, Deitz DR, Campbell WQ, *et al*. N-acetyltransferase 1 genetic polymorphism, cigarette smoking, well-done meat intake, and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 1999;**8**:233-9.
- 2 **Deitz ACW**, Zheng MA, Leff M, *et al*. N-acetyltransferase-2 genetic polymorphism, well-done meat intake, and breast cancer risk among postmenopausal women. *Cancer Epidemiol Biomarkers Prev* 2000;**2**:905-10.
- 3 **Sachidanandam RD**, Weissman SC, Schmidt JM, *et al*. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 2001;**409**:928-33.
- 4 **Risch N**, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996;**273**:1516-17.
- 5 **Ohara Y**, Honma M, Iwasaki Y. Sensitivity of the polymerase chain reaction for detecting human T-cell leukemia virus type I sequences in paraffin-embedded tissue. Effect of unbuffered formalin fixation. *J Virol Methods* 1992;**37**:83-8.
- 6 **Shi SR**, Cote RJ, Wu L, *et al*. DNA extraction from archival formalin-fixed, paraffin-embedded tissue sections based on the antigen retrieval principle: heating under the influence of pH. *J Histochem Cytochem* 2002;**50**:1005-11.
- 7 **Sato YR**, Sugie B, Tsuchiya T, *et al*. Comparison of the DNA extraction methods for polymerase chain reaction amplification from formalin-fixed and paraffin-embedded tissues. *Diagn Mol Pathol* 2001;**10**:265-71.
- 8 **Blomeke B**, Bennett WP, Harris CC, *et al*. Serum, plasma and paraffin-embedded tissues as sources of DNA for studying cancer susceptibility genes. *Carcinogenesis* 1997;**18**:1271-5.
- 9 **Paris PL**, Kupelian PA, Hall JM, *et al*. Association between a CYP3A4 genetic variant and clinical presentation in African-American prostate cancer patients. *Cancer Epidemiol Biomarkers Prev* 1999;**8**:901-5.
- 10 **Carlson CS**, Newman TL, Nickerson DA. SNPping in the human genome. *Curr Opin Chem Biol* 2001;**5**:78-85.
- 11 **Miller SA**, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988;**16**:1215.
- 12 **Wan XS**, Devalaraja MN, St Clair DK. Molecular structure and organization of the human manganese superoxide dismutase gene. *DNA Cell Biol* 1994;**13**:1127-36.
- 13 **Grasbon-Frodl EM**, Kosel S, Riess O, *et al*. Analysis of mitochondrial targeting sequence and coding region polymorphisms of the manganese superoxide dismutase gene in German Parkinson disease patients. *Biochem Biophys Res Commun* 1999;**255**:749-52.
- 14 **de Arruda M**, Lyamichev VI, Eis PS, *et al*. Invader technology for DNA and RNA analysis: principles and applications. *Expert Rev Mol Diagn* 2002;**2**:487-96.
- 15 **Lyamichev V**, Mast AL, Hall JG, *et al*. Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes. *Nat Biotechnol* 1999;**17**:292-6.
- 16 **Glaser JA**. Validity of nucleic acid purities monitored by 260nm/280nm absorbance ratios. *Biotechniques* 1995;**18**:62-3.
- 17 **Olivier M**, Chuang LM, Chang MS, *et al*. High-throughput genotyping of single nucleotide polymorphisms using new biplex invader technology. *Nucleic Acids Res* 2002;**30**:e53.
- 18 **Mein CA**, Barratt BJ, Dunn MG, *et al*. Evaluation of single nucleotide polymorphism typing with invader on PCR amplicons and its automation. *Genome Res* 2000;**10**:330-43.
- 19 **Ohnishi Y**, Tanaka T, Ozaki K, *et al*. A high-throughput SNP typing system for genome-wide association studies. *J Hum Genet* 2001;**46**:471-7.
- 20 **Inadome Y**, Naguchi M. Selection of higher molecular weight genomic DNA for molecular diagnosis from formalin-fixed material. *Diagn Mol Pathol* 2003;**12**:231-6.
- 21 **Rothman N**, Stewart WF, Caporaso NE, *et al*. Misclassification of genetic susceptibility biomarkers: implications for case-control studies and cross-population comparisons. *Cancer Epidemiol Biomarkers Prev* 1993;**2**:299-303.



## New approaches for genotyping paraffin wax embedded breast tissue from patients with cancer: the Iowa women's health study

B Thyagarajan, K E Anderson, F Kong, et al.

*J Clin Pathol* 2005 58: 955-961  
doi: 10.1136/jcp.2004.023374

---

Updated information and services can be found at:  
<http://jcp.bmj.com/content/58/9/955.full.html>

---

	<i>These include:</i>
<b>References</b>	This article cites 21 articles, 9 of which can be accessed free at: <a href="http://jcp.bmj.com/content/58/9/955.full.html#ref-list-1">http://jcp.bmj.com/content/58/9/955.full.html#ref-list-1</a>
	Article cited in: <a href="http://jcp.bmj.com/content/58/9/955.full.html#related-urls">http://jcp.bmj.com/content/58/9/955.full.html#related-urls</a>
<b>Email alerting service</b>	Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

---

<b>Topic Collections</b>	Articles on similar topics can be found in the following collections <a href="#">Immunology (including allergy)</a> (1279 articles)
--------------------------	--

---

### Notes

---

To request permissions go to:  
<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:  
<http://journals.bmj.com/cgi/reprintform>

To subscribe to BMJ go to:  
<http://group.bmj.com/subscribe/>