Journal of Molecular Diagnostics, Vol. 4, No. 3, August 2002

Loss of Heterozygosity Studies Revisited

Prior Quantification of the Amplifiable DNA Content of Archival Samples Improves Efficiency and Reliability

Kathryn Farrand,* Lydija Jovanovic,* Brett Delahunt,* Bryan McIver,† Ian D. Hay,† Norman L. Eberhardt,†‡ and Stefan K. G. Grebe*

From the Department of Pathology and Molecular Medicine, Wellington School of Medicine and Health Sciences, Wellington, New Zealand; and the Departments of Medicine† and Molecular Biology and Biochemistry,‡ Mayo Clinic and Foundation, Rochester, Minnesota*

Polymerase chain reaction (PCR)-based loss of heterozygosity (LOH) studies of archival formalin-fixed, paraffin-embedded (FFPE) tumor tissues have become an important tool in the search for tumor suppressor genes and oncogenes and are also used increasingly in clinical practice. However, FFPE tissue samples may contain little amplifiable DNA, resulting in frequent reaction failures and unreliable LOH data. Using pairs of serial dilutions of reference DNA, we determined the minimum amplifiable DNA concentration necessary for reliable microsatellite-PCR LOH analysis. We then measured the amplifiable DNA content of a selection of frozen and FFPE-derived tumor specimens by real-time quantitative PCR. A minimum input of 600 pg of 100% amplifiable DNA per PCR was required for reliable LOH analysis. While the total DNA concentrations of all samples exceeded this figure, most FFPE-sample-derived DNA was non-amplifiable, with ratios of actually amplifiable DNA to total DNA as low as 1 to 3625. Many FFPE samples therefore contained substantially less than 600 pg/μl of actually amplifiable DNA, making them potentially unsuitable for LOH studies. Real-time quantitative PCR before LOH studies of FFPE tissues allows: identification of samples, which will fail microsatellite-PCR; exclusion of samples, which will yield unreliable results; and optimal adjustment of template input for the remainder. Amplification reactions undertaken without this precaution can result in unreliable LOH data. *(J Mol Diagn 2002, 4:150–158)*

Loss of heterozygosity (LOH) studies are used to identify genomic imbalance in tumors, indicating possible sites of tumor suppressor gene (TSG) deletion or oncogene amplification. In particular, LOH analysis is frequently used to search for novel TSGs. Also, LOH studies, and the related technique of microsatellite instability analysis, are increasingly used in clinical molecular diagnosis, including tumor diagnosis, classification/staging, and detection of residual or recurrent disease.[1–7]

LOH studies were originally based on restriction fragment polymorphism analysis (RFLP), but this approach has been superseded by microsatellite polymerase chain reaction (PCR)-based methods, requiring much less DNA. This has greatly increased the utility and popularity of LOH studies, and has permitted analysis of archival tissues. During the last decade about 100 medline-listed papers have been published per year that have used microsatellite PCR to identify genomic imbalance in formalin-fixed and paraffin-embedded (FFPE) archival tumor samples. However, DNA derived from FFPE samples may be difficult or impossible to amplify, particularly with long microsatellite markers, rendering 10 to 50% of FFPE samples unsuitable for LOH analysis.[8–10] Even samples that amplify and yield copious amounts of PCR product, may exhibit poor reproducibility of observed allele ratios or even random allele drop-outs. Consequently, researchers must use conservative LOH detection thresholds, which will fail to detect a significant proportion of allelic imbalances, or risk detecting spurious LOH. The latter produces confusing and unreliable LOH mapping data, making the detection of accurate minimal common regions of allelic imbalance, that are critical for detecting TSGs or amplified oncogenes, impossible. Similar problems can arise in clinical applications. It is essential to have well established, functioning protocols for obtaining both patients' germline DNA and surgical specimens. Also, portions of the surgical specimens should be snap-frozen and excluded from FFPE treatment. Without such precautions, microsatellite PCR on clinical samples may

Supported by a New Zealand Lotteries Health Research grant (to Stefan K. G. Grebe), National Institutes of Health grant CA80117 (to Norman Eberhardt), and funds from the Wellington School of Medicine and Health Sciences and the Mayo Foundation/Clinic.

Accepted for publication April 28, 2002.

Stefan K. G. Grebe's current address is Department of Pathology and Laboratory Medicine, Mayo Clinic, Rochester, Minnesota, USA.

Address reprint requests to Stefan K. G. Grebe, Endocrine Laboratory, Hilton Building 730, Mayo Clinic and Foundation, 200 1st Street S.W., Rochester, MN 55905. E-mail: grebs@mayo.edu.

result in significant diagnostic errors. For example, if random allele drop-out has occurred in the DNA derived from non-tumorous portions of specimens, serving as "germline-controls," this may lead to the erroneous diagnosis of microsatellite instability and assignment of patients to low tumor recurrence and tumor death risk groups.[7]

The factors leading to poor PCR amplification with FFPE specimen-derived DNA include: sequestration of DNA in protein-DNA complexes, DNA-strand breakage and degradation, DNA-cross-linking or other chemical modification, and co-purification of PCR inhibitors contained in the fixatives.[8–14] As a result, DNA extracted from FFPE archival samples contains low numbers of amplifiable DNA molecules, especially for fragments >300 base pairs (bp). Depending on the number of amplifiable microsatellite targets remaining in a given FFPE tissue-derived DNA sample, results can vary from complete PCR failure to variable allele ratios and random allele drop-outs at slightly higher copy number inputs.

Several approaches have been used to address these problems, including limiting PCR cycle numbers, requiring relatively high minimal PCR product yields and multiple replicate reactions, with or without statistical evaluation of obtained allele ratios.[15] In addition, the DNA concentration of FFPE tissue-derived samples can be measured before PCR. While this is modestly helpful, it fails to identify samples containing degraded DNA or PCR inhibitors. Template degradation can be crudely assessed by visualizing DNA damage using agarose gel electrophoresis or the number of DNA strand breaks can be quantified more precisely by terminal transferase-mediated end-labeling of free DNA ends.[12] However, only the direct determination of amplifiable DNA concentrations addresses all variables related to FFPE-derived PCR template quality and quantity. This has been achieved using PCR amplification of a 110-bp fragment of the *N-ras* single-copy gene on serial dilutions of FFPE tissue-derived DNA samples.[16] Unfortunately, this technique is too labor intensive and time consuming for routine allelotyping and the results are only applicable to PCR targets of similar size to the studied *N-ras* target.

To address these shortcomings, we developed a rapid real-time PCR method for the quantification of amplifiable copy numbers in FFPE tissue-derived DNA samples. We also determined the minimal amplifiable copy number input required for reproducible allelic imbalance ratios following microsatellite PCR. These two measures have allowed us to predict which FFPE tissue-derived DNA samples will be unsuitable for LOH analysis, as well as to determine the minimal sample input required per PCR to assure reliable LOH analysis for those samples with sufficient amplifiable copy numbers.

Materials and Methods

Samples and DNA Extraction

We extracted DNA from the following groups of samples: freshly collected blood samples from a healthy volunteer (screened for being informative at the markers used in this study), five snap-frozen thyroid tumor and matching blood samples, and 10 FFPE archival thyroid tumor samples (manually microdissected by an anatomical pathologist into tumor and non-tumor portions). We assumed that the high quality reference DNA extracted from the healthy volunteer's blood samples was 100% amplifiable and used this as the standard in all experiments. Frozen tumor samples and FFPE archival samples had been examined previously for LOH at several 17p loci. In these studies, we had found that all frozen tumor DNA samples and five of the 10 FFPE samples amplified without problems. For the remaining five FFPE archival samples amplification of the tumor-derived and/or non-tumor-derived DNA had required several PCR attempts or had failed completely.

We extracted DNA from whole blood samples using proteinase K (Roche Diagnostics NZ, Auckland, New Zealand) digestion and ionic detergent lysis,[17] followed by two phenol-chloroform and two chloroform extractions, ethanol precipitation, and re-suspension in 1X TE buffer (10 mmol/L Tris-HCl, 1 mmol/L EDTA, pH 7.6). For frozen tumor DNA extraction, we wrapped snap-frozen thyroid tumor tissue fragments in sterile aluminum foil and pulverized them with a hammer. DNA was extracted and purified from the powdered tissues as described above. We extracted matching control DNA from buffy coat preparations from the same patients, again using the same purification method. The DNA samples were aliquoted and stored at −80°C.

We extracted DNA from archival FFPE thyroid tumor samples following manual microdissection (by an anatomical pathologist) of 10-μm thick unstained sections into neoplastic compartments, containing at least 80% tumor cells, and control compartments, containing no microscopically identifiable tumor tissue. The slides we chose for microdissection all contained approximately equal portions of tumor and control tissue and slides from different specimens were selected so that they also contained similar overall cell numbers. The microdissected tissue fragments were de-waxed through successive xylene and ethanol washes, dried, weighed, and then re-suspended in DNA-extraction buffer (20 μl per 10 mg tissue of 100 mmol/L Tris-HCl, 2 mmol/L EDTA, 2.7 μg/μl proteinase K, pH 8; average total extraction buffer volume used 150 μl) and incubated at 55°C for 24 hours. We then added an additional 2.7 μg/μl proteinase K, and incubated for a further 24 hours at 55°C. The proteinase K was heat-inactivated through 10-minute boiling, and the samples were aliquoted and stored at −80°C.

DNA Quantification following Extraction

We determined the DNA concentrations of all samples on a fluorometer (TD-700, Turner Design, Sunnyvale, CA) using the PicoGreen nucleic acid quantification kit (PicoGreen, Molecular Probes, Eugene, OR), following the manufacturer's instructions (http://www.probes.com/handbook/

Table 1. Marker Characteristics and Primer Sequences

Locus	Marker type	Forward primer sequence (5′-3′)	Reverse primer sequence (5′-3′)	PCR-product size range (bp)
ABR	dinucleotide repeat	FAM*-cga aaa caa aac gca tac acc	tgt gag agg ggt cag agt agg	80–120
D17s695	tetranucleotide repeat	FAM*-ctg ggc aac aag agc aaa att c	ttt gtt gtt gtt cat tga ctt cag tct	175–215
D17s678	dinucleotide repeat	TET*-cag ctt ggc aac aca gcg aaa	tat tct gct cgg cac ata gtg caa	280–322

*, Only forward primers used for LOH experiments were labeled. Unlabeled versions of the same primers were used for the real-time quantitative PCR experiments.

sections/0803.html and http://www.probes.com/media/pis/mp07581.pdf). This assay allows accurate and reproducible quantification of dsDNA at very low concentrations, including DNA derived from archival FFPE tissue samples.[14]

Markers

We selected three microsatellite markers mapping to the short arm of chromosome 17, covering an approximate size range of 100 to 300 bp. Table 1 lists the makers, their characteristics and the primers used for PCR amplification. Both 5′-flurophore labeled (with FAM or TET for GeneScan LOH analysis) and unlabeled versions (for LightCycler quantitative PCR) of each forward primer were used.

All labeled forward primers and all reverse primers were synthesized at the Mayo Clinic oligonucleotide core facility. The unlabelled forward primers were obtained from Invitrogen (Invitrogen NZ, Auckland, New Zealand).

PCR

We performed all PCR using the LightCycler real-time PCR system (Roche Diagnostics NZ, Auckland, New Zealand). We used the labeled primer sets for reactions destined for GeneScan analysis and performed quantitative PCR using the unlabeled primers. Quantitative PCR contained PicoGreen or SybrGreen (Molecular Probes) as reporter dyes (details, see below).

Reactions contained 1X PCR-buffer, 0.2 mmol/L dNTPs, 0.5 μmol/L each primer, 1 mmol/L $MgCl_2$, 0.05 units/μl *Taq*DNA polymerase (all Roche Diagnostics NZ), 0.05 units/μl Platinum *Taq* antibody (Invitrogen NZ), 0.025 μg/μl bovine serum albumin (BSA) (New England Biolabs, Beverly, MA) and 1 μl of template DNA (at various concentrations, see below) in a total reaction volume of 10 μl.

The amplification protocol consisted of an initial denaturation step at 94°C for 1 minute, followed by 45 cycles at 95°C (ramp rate 20°C/second) for 0 seconds, 61°C (ramp rate 20°C/second) for 0 seconds (66°C for 0 seconds for *D17S678*) and 74°C (ramp rate 20°C/second) for 6 seconds, measuring fluorescence at each extension stage. At completion of PCR amplification, we performed a melting curve analysis; 94°C (ramp rate 20°C/second) for 0 seconds, 40°C (ramp rate 20°C/second) for 20 seconds, 94°C (ramp rate 0.1°C/second) for 0 seconds, collecting fluorescence continuously at each step. Since we used intercalating dyes for PCR product detection, rather than sequence specific probes, the final melting curve step was crucial to allow us to confirm that the PCR had yielded genuine amplification products rather than primer-multimers.

LOH Reliability Analysis

To determine the minimal amplifiable DNA input necessary for reproducible allelic imbalance/LOH analysis we assumed that our high quality reference DNA samples were 100% amplifiable (see above). We then set up several duplicate PCR reactions of reference DNA at different DNA-template inputs and calculated allelic imbalance ratios for each matched template input pair. As the same DNA at the same input concentration was used in each matched template input pair, one would expect that the allele ratios of the two reactions would be very similar, resulting in an allelic imbalance ratio of approximately 1 for each matched sample pair.

We conducted all these experiments in quadruplicate. For each marker, we PCR-amplified matched pairs of samples of 100% amplifiable reference DNA at the following template inputs per reaction: 6 ng, 600 pg, 420 pg, 240 pg, 60 pg, 42 pg, 24 pg, and 6 pg.

Following PCR, we mixed 1.5 μl of each PCR product with 2.5 μl formamide, 0.3 μl loading dye, and 0.3 μl TAMRA-500 size standard (Applied Biosystems, Foster City, CA). The mixture was denatured at 95°C for 3 minutes, quenched on ice, loaded onto a 4% polyacrylamide/6 mol/L urea gel, and electrophoresed on a 377 Automated Sequencer (Applied Biosystems).

We analyzed the gel data using the GeneScan 3.1 software package (Applied Biosystems). For each pair of matched duplicate samples we calculated an allelic imbalance ratio from the maximum allele peak heights:

$$\frac{\text{duplicate\_allele1/duplicate1\_allele2}}{\text{duplicate2\_allele1/duplicate2\_allele2}}.$$

We defined an arbitrary normal range of 0.6 to 1.67 for the allelic imbalance ratios. In a tumor sample, allelic imbalance ratios outside this range would correspond to greater than 50% of tumor cells in an 80% "pure" tumor sample displaying allelic imbalance. For each combination of markers and DNA concentrations we recorded the number of replicate experiments yielding ratios outside this normal range as well as all complete allelic drop-outs and complete reaction failures. The average allelic imbalance ratios and the coefficients of variation (CV) of the allelic imbalance ratios in each group were calculated from the four replicate sets of experiments.

Table 2. Sample Raw DNA Concentrations and Proportion of Amplifiable DNA

DNA samples		Raw DNA concentration (ng/μl)	Amplifiable fraction of raw DNA input		
			ABR	*D17S695*	*D17S678*
DNA derived from frozen tumors and matching buffy coat samples	Frozen-1-normal	170.4	1:1	1:3	1:2
	Frozen-1-tumor	327.8	1:1	1:5	1:2
	Frozen-2-normal	21	1:1	1:3	1:2
	Frozen-2-tumor	41.2	1:1	1:3	1:1
	Frozen-3-normal	1.8	1:1	1:1	1:2
	Frozen-3-tumor	5.4	1:1	1:3	1:2
	Frozen-4-normal	54.8	1:1	1:1	1:2
	Frozen-4-tumor	305.8	1:1	1:3	1:2
	Frozen-5-normal	79	1:2	1:1	1:2
	Frozen-5-tumor	115.6	1:1	1:3	1:1
FFPE-derived samples without PCR-amplification problems in previous LOH studies	FFPE-1-normal	29	1:2,417	1:3,625	FTA*
	FFPE-1-tumor	12.6	1:20	1:325	FTA*
	FFPE-2-normal	2.6	1:6	1:25	1:78
	FFPE-2-tumor	28.2	1:46	1:243	FTA*
	FFPE-3-normal	5.6	1:40	1:58	1:199
	FFPE-3-tumor	13.4	1:16	1:77	1:353
	FFPE-4-normal	7.2	1:1,800	1:153	FTA*
	FFPE-4-tumor	9.4	1:58	1:41	1:855
	FFPE-5-normal	7.4	1:12	1:31	1:58
	FFPE-5-tumor	10.8	1:14	1:46	1:77
FFPE-derived samples with PCR-amplification problems in previous LOH studies	FFPE-6-normal	0.8	FTA*	FTA*	FTA*
	FFPE-6-tumor	2.6	FTA*	FTA*	FTA*
	FFPE-7-normal	2	FTA*	FTA*	FTA*
	FFPE-7-tumor	2.8	1:7	1:110	FTA*
	FFPE-8-normal	2.4	1:27	FTA*	FTA*
	FFPE-8-tumor	7.6	FTA*	FTA*	FTA*
	FFPE-9-normal	5.2	1:28	1:166	FTA*
	FFPE-9-tumor	1	1:5	1:56	FTA*
	FFPE-10-normal	4.2	1:6	1:500	FTA*
	FFPE-10-tumor	6.6	1:1,031	FTA*	FTA*

*, FTA = failed to amplify.

Quantification of Amplifiable Copy Numbers

We used PicoGreen (final concentration 1:5000) or SybrGreen (Molecular Probes; final concentration 1:200,000) as reporter dyes in all LightCycler real-time quantitative PCR. Assay reproducibility was assessed through quantitation runs of 10 replicates of the same reference DNA at three different concentrations for each primer pair.

For the quantification assays of frozen and FFPE tissue-derived tumor DNA, serial dilutions (10 ng, 4 ng, 400 pg, and 40 pg per reaction) of PicoGreen-quantitated high quality (100% amplifiable) reference DNA served as external standards. For each experiment, we diluted frozen tumor-derived DNA and matching white blood cell DNA into the concentration range covered by the standard curve, and added 1 μl of each to the PCR reagent mix. For DNA extracted from FFPE samples, we added 1 μl undiluted to each reaction, as the raw DNA content of these samples was significantly lower compared with that of the frozen samples.

We performed all data analysis using the LightCycler data analysis software. To minimize experimental variability, we used the automated second derivative maximum estimation method to determine amplification crossing-points, both to generate the standard curve and to determine the amplifiable copy numbers of the samples. For each sample, melting curve analysis was also performed to confirm that the amplification curves represented specific PCR products rather than primer-multimers.

Statistical Analysis

We compared continuous data between groups by analysis of variance, using the Bonferroni/Dunn post-hoc test to determine significant differences between groups, corrected for multiple comparisons.

Results

DNA Quantification following DNA Extraction

For DNA extracted from frozen tumors and matched buffy coat samples, concentrations averaged 112.28 ng/μl, which is significantly higher than the averages for both the easily amplifiable FFPE DNA samples (12.62 ng/μl; $P = 0.0033$), and the difficult to amplify FFPE DNA samples (3.52 ng/μl, $P = 0.0016$). The difference between the means of the two groups of FFPE tissue-derived samples was not significant (Table 2).

LOH Reliability Analysis

Figure 1 shows a typical example of allele patterns observed in duplicate reactions at different DNA-input con-

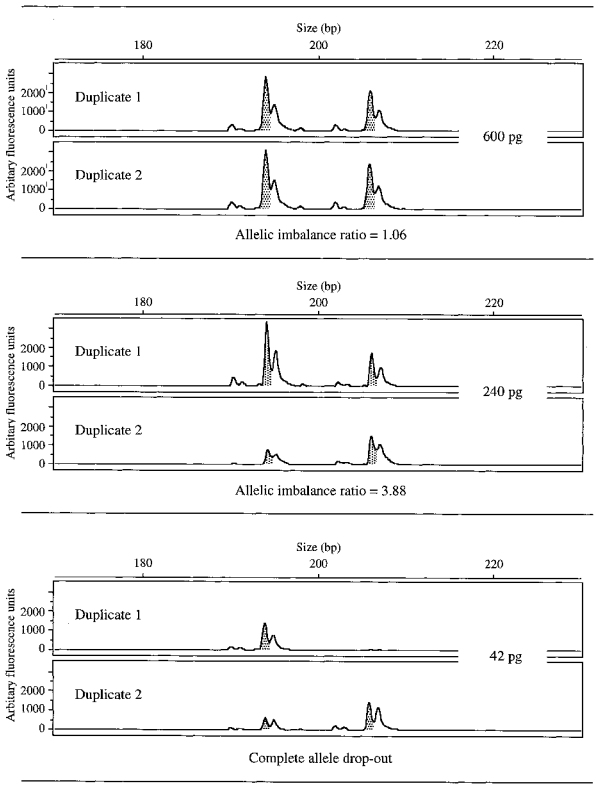

Figure 1. GeneScan electropherograms of three duplicate microsatellite PCR (D17S695) of the same, 100% amplifiable reference DNA sample at template inputs of 600 pg (**top**), 240 pg (**middle**) and 42 pg (**bottom**). Alleles are indicated by **shading**, other peaks represent stutter and a-tails. The abscissas of each **panel** show the allele sizes in bp, the ordinates show the allele peak heights in arbitrary fluorescence units. Allele peak heights fall slightly with decreasing template input, but remain satisfactory in all three experiments. However, only the 600 pg template input samples display no differences in allele height between the two duplicate samples. At 240 pg DNA input, significant allelic imbalance is evident with the peak height of the shorter allele in duplicate 2 being decreased. At 42 pg template input, a complete allelic drop-out of the longer allele of duplicate 1 is observed.

centrations. With decreasing DNA inputs, an increasing number of duplicate reactions displayed abnormal allelic imbalance ratios and complete allelic drop-outs and PCR reaction failures occurred. If we had used control-tumor pairs of reactions instead of identical template-input control-control pairs of reactions, these reactions might have been interpreted as LOH or even homozygous allele loss.

At DNA-template inputs of at least 600 pg of 100% amplifiable, high quality reference DNA per reaction, no allelic imbalance ratios outside the defined normal range were observed. At 420 pg per reaction DNA input, *D17S695* displayed a ratio outside this range for one of the four replicate reaction pairs. At 240 pg DNA input, one of the *ABR* duplicate pairs yielded an abnormal ratio, and at 60 pg DNA input or less, ratios outside the normal range, allele drop-outs, and complete reaction failures were observed for all three markers. The replicate averages and CVs mirrored these results. Table 3 summarizes the results of the LOH reliability analysis. Interestingly, as shown in Figure 1, for those samples which did not suffer complete PCR failure, fluorescent signal strength remained high (greater than 1000 arbitrary fluorescent units), even when complete allele drop-outs occurred.

Quantification of Amplifiable Copy Numbers

In initial experiments we used both SybrGreen and PicoGreen as reporter dyes in the quantitative real-time PCR experiments. At the commonly used LightCycler SybrGreen concentrations between 1:10,000 and 1:80,000 a significant number of FFPE tissue-derived DNA samples displayed high baseline fluorescence, making accurate quantification of amplifiable targets impossible. Using lower SybrGreen concentrations of 1:200,000 overcame these problems, but resulted in lesser sensitivity when compared with PicoGreen at a 1:5000 final concentration (data not shown). Therefore, PicoGreen was used for all further experiments. Real-time PCR data obtained with PicoGreen as reporter dye were in agreement with PCR data from previous LOH experiments (data not shown), and allowed reliable detection of specific targets in less than 40 pg total of 100% amplifiable genomic DNA input. Over the assayed genomic DNA input range of 40 pg to 10 ng, the CVs for real-time

Table 3. Reproducibility* of Microsatellite PCRs of Identical Duplicate High Quality DNA Samples at different Input Concentrations: Average Allelic Imbalance Ratio (AIR), Coefficient of Variation (CV), N of AIRs Outside the Normal Range (NR)‡ and N of Allelic Drop-Outs or PCR Failures

	Markers											
	ABR				*D17S695*				*D17S678*			
DNA input	Av. AIR	CV (%)	AIR outside NR† (N)	Drop-outs or PCR-fail. (N)	Av.AIR	CV (%)	AIR outside NR† (N)	Drop-outs or PCR-fail. (N)	Av. AIR	CV (%)	AIR outside NR† (N)	Drop-outs or PCR-fail. (N)
6 ng	0.99	7.37	0	0	1.02	10.04	0	0	0.94	14.18	0	0
600 pg	1.34	10.4	0	0	0.98	7.64	0	0	1.13	15.77	0	0
420 pg	1.07	22.9	0	0	1.57	69.52	1	0	1.13	19.69	0	0
240 pg	1.3	45.83	1	0	1.99	68.61	2	0	1.02	5.57	0	0
60 pg	1.64	43.96	2	0	0.2‡	–‡	1‡	3‡	1.78	68.97	2	0
42 pg	2.18	115.33	3	0	42.86‡	–‡	1‡	3‡	1.68	150.44	4	0
24 pg	0.95‡	–‡	0‡	3‡	1.45‡	–‡	0‡	3‡	3.14	130.49	3	0
6 pg	–§	–§	–§	4§	–§	–§	–§	4§	–§	–§	–§	4§

*, Based on four independent replicate experiments.
†, NR = normal range for allelic imbalance ratios: 0.6–1.67 (see main text for details)
‡, All but one of the four duplicate reactions suffered at least one allelic drop-out or PCR failure in at least one of the reactions of a duplicate pair.
§, All four duplicate reactions suffered at least one allelic drop-out or PCR failure in at least one of the reactions of a duplicate pair.

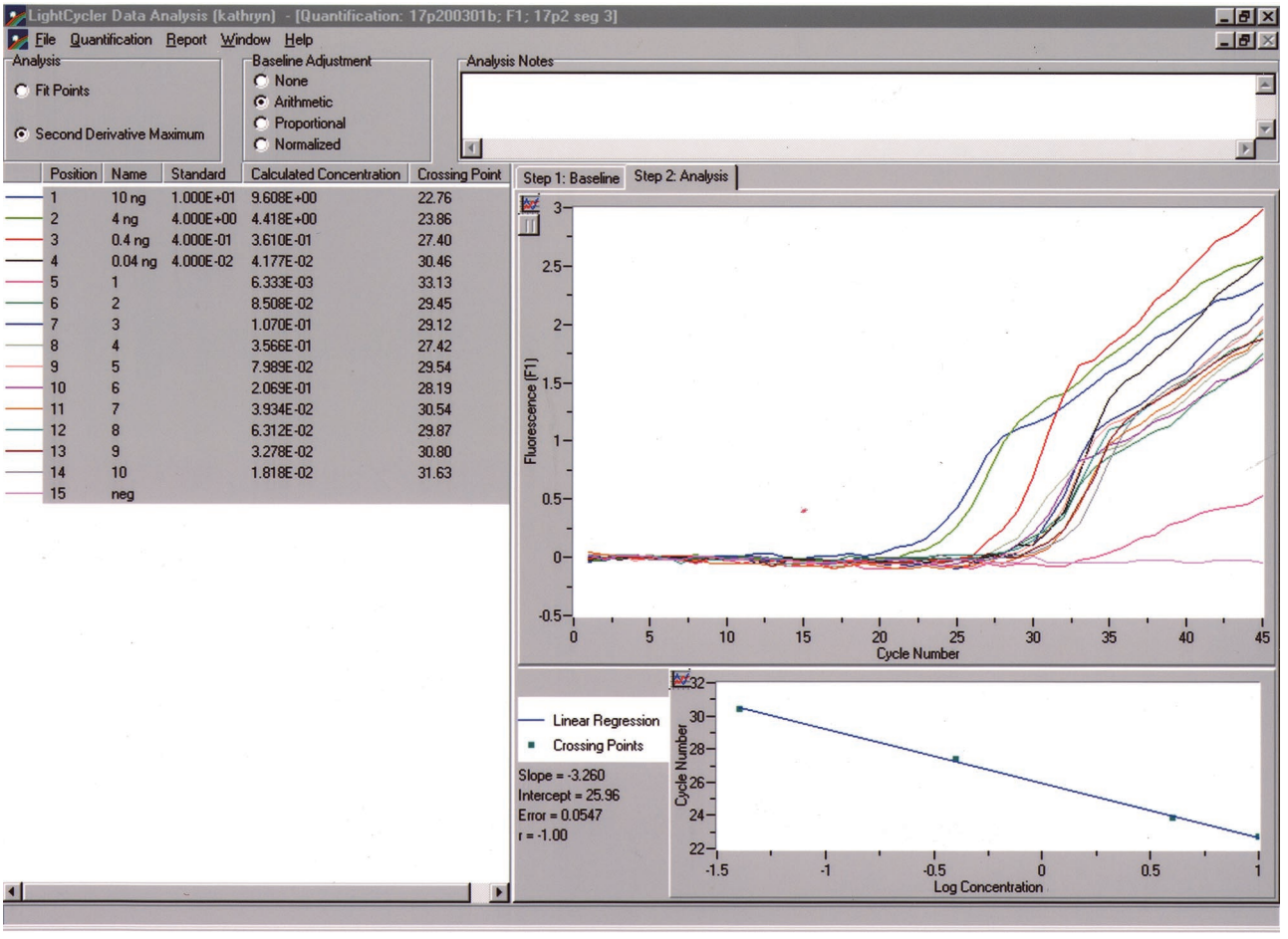

Figure 2. Figure 2 shows the data analysis screen of a typical LightCycler experiment to determine the amplifiable DNA content of thyroid tumor samples. The **uppermost portion** shows the analysis settings. The **table** on the **left**, below the settings, contains the input DNA standards (10 ng, 4 ng, 400 pg, and 40 pg of 100% amplifiable reference DNA), the calculated amplifiable DNA content of the standards and of 10 unknown samples (based on the generated standard curve) and the crossing points (PCR cycle number when the exponential phase was reached) for all standards and samples. The **last row** contains a negative control sample. The actual PCR fluorescence profiles for the standards and samples are depicted in the **panel** to the **right** of the **table**. Beneath the fluorescence profiles, the standard curve and its associated parameters are shown.

PCR determination of amplifiable copy numbers of the three examined markers were between 1.3% and 2.9%.

Figure 2 shows the results of a typical PicoGreen real-time PCR reaction. Melting curves confirmed that all amplification curves represented specific PCR products (data not shown). There was wide variability between samples in the proportion of input DNA that was amplifiable. In the DNA samples derived from frozen tumors and matched buffy coat DNA, most input DNA was amplifiable with the ratios ranging from 1:1 to 1:5 (Table 2). By contrast, in the two groups of FFPE tissue-derived DNA samples only a much smaller proportion of input DNA molecules could be amplified with the ratios ranging from 1:5 to 1:3625. For the longer markers, many samples failed to amplify, particularly in the group of FFPE tissue-derived DNA samples that had previously been problematic during a LOH project (Table 2). As demonstrated in Figure 2, most samples which could be amplified completed the exponential PCR amplification phase within 30 to 35 cycles, by which time they had generally achieved more than 50% of their final total product yield.

Discussion

We determined that the DNA template input of 100% amplifiable DNA for LOH experiments has to be greater than 600 pg per reaction for reproducible results. We then developed a rapid, highly reproducible assay to determine the amplifiable DNA fraction of FFPE tissue-derived DNA samples. Using this method before microsatellite PCR analysis identifies samples that are likely to fail PCR amplification or yield unreliable results and allows optimization of input DNA concentrations for the remaining samples. For example, if the raw DNA concentration of a FFPE tissue sample is determined at 10 ng/μl but only 1 in 100 DNA molecules is amplifiable, then at least 6 μl of template input per reaction are required for reliable LOH results

LOH analysis of tumors was introduced as a technique to identify genomic regions of allelic imbalance using comparative RFLP-Southern blots of tumor and corresponding normal tissues.[18] Allelic imbalance scoring in these assays relies on the relative hybridization intensity

of the individual alleles. The main drawbacks of the method are that it requires large quantities of high quality DNA, excludes the use of archival paraffin-embedded specimens, and limits the number of markers that can be tested simultaneously. While PCR-based microsatellite LOH analysis has overcome these limitations ,[19] this approach also has significant problems. Because PCR reactions are easily saturable, accurate LOH scoring cannot be performed by comparison of band intensities, but relies on comparison of allele ratios of tumor and corresponding normal control samples.[20] With sufficient high quality DNA inputs, these allelic imbalance ratios are usually highly reproducible.[21,22] However, even relatively large quantities of high quality DNA samples may sometimes display significant variation in allele ratios.[15] This problem is magnified in FFPE tissue samples, where DNA recoveries are often low and DNA quality is poor, leading to low amplifiable copy numbers. Based on our dilution studies of duplicate pairs of identical reference DNA samples, allelic imbalance ratios become increasingly irreproducible with falling DNA-template inputs. A minimum template input of 600 pg of 100% amplifiable DNA is necessary to ensure reliable allelic imbalance ratios for all markers. Lesser inputs may result in artifactually abnormal allelic imbalance ratios, mimicking LOH.

There are several reasons why minimum template inputs are necessary. At very low concentrations of DNA-template, there is an increased likelihood that a PCR reaction will lack sufficient starting material for amplification.[16,23] In the case of a semi-quantitative method, such as microsatellite PCR-based LOH analysis, the problem is exacerbated because the template input has to accurately represent the allele ratios in the starting material. At high template dilutions, random deviations from this ideal will be magnified by the exponential PCR amplification process, resulting in artifactual allelic imbalances. For most informative microsatellite markers amplification of the longer allele is less efficient than amplification of the shorter allele, a fact reflected in our data, which show largely allelic imbalance ratios greater than one. The observed degree of this bias exceeds the theoretical predictions for two nearly identical DNA sequences, which differ by only a few bp.[24] As a result, microsatellite PCR-based LOH analysis requires comparatively large DNA inputs for accurate and reproducible results, even under optimal PCR conditions. Our data show that some markers are more tolerant of low template inputs than others, but unless all markers are tested in advance, it seems prudent to maintain minimal PCR-template inputs of 600 pg of 100% amplifiable DNA per marker. The less efficient multiplex reactions, often used to allow detection of homozygous deletions, may require even higher inputs.

LOH studies using fresh or frozen tumor tissues should not have problems meeting the minimal template input requirements, as essentially all their DNA is amplifiable. By contrast, our data indicate that FFPE tissue-derived samples typically contain lower concentrations of DNA of which only a fraction is amplifiable, resulting in partial or complete PCR failure. Moreover, even for many samples without obvious amplification problems, allelic imbalance ratios may be unreliable, due to low amplifiable DNA input. Similar difficulties may also be encountered in applications that analyze very small numbers of fresh or frozen cells, including clinical diagnosis of malignancy based on small cytological samples, minimal residual disease screening in solid malignancies using microsatellite LOH analysis, and prenatal diagnosis of numerical chromosome aberrations using microsatellite PCR.[4,5,25]

While many investigators are aware of these problems, review of published LOH studies reveals that these problems are frequently ignored. Often investigators trust their data as long as relatively unbiased amplification of two alleles is achieved in FFPE-derived non-tumorous DNA and amplification of FFPE-derived tumor DNA does not fail completely. Based on our dilution studies this approach is false and will likely result in poor quality, unreliable, or confusing LOH data. Even using increased numbers of markers from the same chromosomal region does not overcome this problem. If the starting amplifiable DNA input is too low the result of such an approach will simply be the creation of poorly reproducible, patchy interstitial LOH patterns. A lack of appreciation of the problem of inadequate amplifiable DNA input into microsatellite PCR reactions may therefore well be one of the factors leading to the commonly observed phenomenon of contradictory LOH results in different papers analyzing the same tumor types.

In those cases where reliability problems of allelic imbalance analysis are acknowledged, the main problem-solving approaches have been the use of larger numbers of markers from the same chromosomal region, limitation of PCR cycle numbers coupled with the requirement for certain minimal PCR yields/signal strengths, experimental replication, and utilization of alternative techniques to confirm LOH results. The first and second approach both seem simple and uncomplicated, but our experiments suggest that they are flawed. As indicated above, increased marker density simply increases the complexity of the data and is only helpful if all that is needed is exclusion or confirmation of any LOH; in that case it represents a form of experimental replication. As to limitation of cycle number and requirements for minimal PCR yield, we observed high PCR product yields even at very low template input concentrations, when artifactual allelic imbalances occurred. At the same time, our real-time PCR data suggest that 50% or more of this yield was achieved within 30 to 35 cycles. Therefore, limiting cycle numbers will still result in strong GeneScan signals. The signal strength threshold for accepting results as valid would therefore have to be set very high with attendant rejection of many legitimate results. While experimental replication is not fundamentally flawed, it may result in exclusion of data from the analysis and increases workload and costs. Unless every experiment is repeated several times, apparently reproducible allelic imbalance ratios may still be observed by chance. Using very conservative allelic imbalance criteria for LOH diagnosis solves this problem only partly, since allelic drop-outs may still be misclassified as LOH, and cases of legitimate LOH may be excluded if they fall outside the defined LOH range. Finally, while using alternative techniques such as fluorescence *in situ* hybridization (FISH) will help in confirming LOH data, this again results in additional work.

Accurate quantification of amplifiable copy numbers is the most reliable way to address the problem of sufficient template input. As outlined in the introduction, all other methods have serious shortcomings. In particular, our data show that although simple measurement of DNA concentrations is somewhat helpful, the wide variation in amplifiable DNA fraction makes this a very unreliable means of ensuring adequate amplifiable template input. Provided an accurately quantified high quality DNA standard is available, quantitative PCR is the most direct way to determine the concentration of amplifiable copies in DNA samples. However, traditional quantitative PCR methods suffer from several draw-backs. Because of the sensitivity of PCR to minor variations in starting conditions, quantitative PCR requires multiple replicates per sample, each containing an internal standard at different concentrations. To cover the wide range of amplifiable DNA-input concentrations expected with FFPE tissue-derived DNA samples, a large number of combinations of different input and internal standard concentrations would be required. In this context, the design of suitable internal standards can be challenging, as even minor variations in target length or sequence may result in differences in amplification efficiency.[26] Finally, PCR reflects accurately DNA-target inputs only during the exponential phase of amplification.[27] This increases the error of endpoint-based quantification, which is used in most conventional quantitative PCR assays. Alternative solutions, using serial dilutions of samples and Poisson statistics of failed *versus* successful PCR to estimate copy numbers, overcome some of these problems, but require very large replicate numbers per sample to generate accurate statistical estimates.[23] Consequently, quantitative PCR methods have not been used widely to assess amplifiability of FFPE-derived DNA samples.

Real-time PCR overcomes these problems, as it monitors PCR-product amplification continuously during the reaction, either through hydrolysis of fluorophore-quencher probes, fluorescence resonance energy transfer between neighboring reporter probes, or through fluorescence increase of dsDNA binding dyes. Quantification is achieved by comparing PCR cycle numbers at which unknown samples enter the exponential PCR phase with corresponding PCR cycle numbers of external standards of a range of DNA-input concentrations, run in separate tubes in the same PCR. On a capillary thermal cycler, accurate and reproducible quantification can be achieved within 30 to 40 minutes. Using dsDNA binding dyes rather than probes makes the assay versatile, cheap, and applicable to all markers. By using PicoGreen even relatively crude, non-purified, FFPE tissue-derived DNA samples can be analyzed. By contrast, the usually used dye, SybrGreen, displays too much background fluorescence with such crude samples when used at concentrations which allow good sensitivity. As an alternative to PicoGreen, consensus probes for di-, tri- and tetranucleotide repeats can be used,[28] but this may add experimental complexity and increase cost. For either approach, it is necessary to examine amplifiability across the entire size range of markers, which are to be used in the LOH studies. The use of microsatellite markers, which are actually going to be used in the LOH experiments, rather than other genomic targets, may give the best assurance that the results will be valid for LOH experiments.

In conclusion, quality and productivity of microsatellite PCR-based LOH analysis of FFPE tissue-derived samples can be improved if amplifiable DNA reaction inputs are increased above the threshold for reproducible allele ratios. Using a forced air-driven capillary thermal cycler for amplification and an automated sequencer for fragment detection, this threshold appears to lie at about 600 pg of 100% amplifiable DNA per reaction. We found that real-time quantitative PCR on the LightCycler is a rapid and reliable means of determining, whether the actually amplifiable DNA concentration (*versus* total DNA concentration) of FFPE tissue-derived DNA samples exceeds this threshold. This permits the exclusion of unsuitable samples from subsequent LOH analysis and allows the DNA input of suitable samples to be optimized. While quantitative PCR before LOH-PCR adds some additional work, it is probably sufficient to determine the amplifiable DNA fraction of each sample by using no more than three or four markers covering the size-range of the probes which are going to be used in subsequent LOH experiments. For all but the smallest LOH studies, and certainly most relevant clinical applications where accurate diagnosis is paramount, this seems a minimal effort for the downstream increases in efficiencies and productivity and the improved data quality and reliability.

Acknowledgments

We thank Ann Thornton and Joan Nicoll for technical support.

References

1. Naidoo R, Chetty R: The application of microsatellites in molecular pathology. Pathol Oncol Res 1998, 4:310–315
2. Arzimanoglou II, Gilbert F, Barber HRK: Microsatellite instability in human solid tumors. Cancer 1998, 82:1808–1820
3. Bugert P, Kovacs G: Molecular classification of renal cell carcinomas by microsatellite analysis. Am J Pathol 1996, 149:2081–2087
4. Nawroz H, Koch W, Anker P, Stroun M, Sidransky D: Microsatellite alterations in serum DNA of head and neck cancer patients. Nat Med 1996, 2:1035–1037
5. Mao L, Schoenberg MP, Scicchitano M, Erozan YS, Merlo A, Schwab D, Sidransky D: Molecular detection of primary bladder cancer by microsatellite analysis. Science 1996, 271:659–662
6. Störkel S, Eble JN, Adlakha K, Amin M, Blute ML, Bostwick DG, Darson M, Delahunt B, Iczkowski K: Classification of renal cell carcinoma. Cancer 1997, 80:987–989
7. Gryfe R, Kim H, Hsieh ETK, Aronson MD, Holowaty EJ, Bull SB, Redstone M, Gallinger S: Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer. N Engl J Med 2000, 342:69–77
8. Iwamoto KS, Mizuno T, Ito T, Akiyama M, Takeichi N, Mabuchi K, Seyama T: Feasibility of using decades-old archival tissues in molecular oncology/epidemiology. Am J Pathol 1996, 149:399–406
9. Coombs NJ, Gough AC, Primrose JN: Optimization of DNA and RNA extraction from archival formalin-fixed tissue. Nucleic Acids Res 1999, 27:e12i–iii
10. Baisse B, Bian Y-S, Benhattar J: Microdissection by exclusion and DNA extraction for multiple PCR analysis from archival tissue sections. Biotechniques 2000, 28:856–862
11. Greer CE, Peterson SL, Kiviat NB, Manos MM: PCR amplification from

paraffin-embedded tissues: effects of fixative and fixation time. Am J Clin Pathol 1991, 95:117–124
12. Peng L, Brisco MJ, Morley AA: A method for assessing strand breaks in DNA. Anal Biochem 1998, 262:9–16
13. Masuda N, Ohnishi T, Kawamoto S, Monden M, Okubo K: Analysis of chemical modification of RNA from formalin-fixed samples and optimization of molecular biology applications for such samples. Nucleic Acids Res 1999, 27:4436–4443
14. Serth J, Kuczyk MA, Paseslack U, Lichtinghagen R, Jonas U: Quantitation of DNA extracted after micropreparation of cells from frozen and formalin-fixed tissue sections. Am J Pathol 2000, 156:1189–1196
15. Liloglou T, Maloney P, Xinarianos G, Fear G, Field JK: Sensitivity and limitations of high throughput fluorescent microsatellite analysis for the detection of allelic imbalance: application in lung tumors. Int J Oncol 2000, 16:5–14
16. Skyes PJ, Neoh SH, Brisco MJ, Hughes E, Condon J, Morley AA: Quantitation of targets for PCR by use of limiting dilution. Biotechniques 1992, 13:444–449
17. Sambrook J, Fritsch EF, Maniatis T: Molecular Cloning. Edited by J Sambrook, EF Fritsch, T Maniatis. Cold Spring Harbor, NY, Cold Spring Harbor Press, 1989, pp 916–919
18. Vogelstein B, Fearon ER, Kern SE, Hamilton SR, Preisinger AC, Nakamura Y, White R: Allelotype of colorectal carcinomas. Science 1989, 244:207–211
19. Louis DN, von Deimling A, Seizinger BR: A (CA)n dinucleotide repeat assay for evaluating loss of allelic heterozygosity in small and archival human brain tumor specimens. Am J Pathol 1992, 141:777–782
20. Cawkell L, Bell SM, Lewis FA, Dixon MF, Taylor GR, Quirke P: Rapid detection of allele loss in colorectal tumours using microsatellites and fluorescent DNA technology. Br J Cancer 1993, 67:1262–1267
21. Hampton GM, Larson AA, Baergen RN, Sommers RL, Kern S, Cavenee WK: Simultaneous assessment of loss of heterozygosity at multiple microsatellite loci using semi-automated fluorescence-based detection: subregional mapping of chromosome 4 in cervical carcinoma. Proc Natl Acad Sci USA 1996, 93:6704–6709
22. Niederacher D, Picard F, van Roeyen C, An HX, Bender HG, Beckmann MW: Patterns of allelic loss on chromosome 17 in sporadic breast carcinoma detected by fluorescent-labeled microsatellite analysis. Genes Chromosomes Cancer 1997, 18:181–192
23. Wang Z, Spadoro J: Determination of target copy numbers of quantitative standards used in PCR-based diagnostic assays. Gene Quantification (Advanced Biomedical Technologies). Edited by F Ferre. New York, Springer Verlag, 1998, pp 31–44
24. Liu J, Zabarovska VI, Braga E, Alimov A, Klein G, Zabarovsky ER: Loss of heterozygosity in tumor cells requires re-evaluation: the data are biased by the size-dependent differential sensitivity of allele detection. FEBS Lett 1999, 462:121–128
25. Toth T, Findlay I, Papp C, Toth-Pal E, Marton T, Nagy B, Quirke P, Papp Z: Prenatal detection of trisomy 13 from amniotic fluid by quantitative fluorescent polymerase chain reaction. Prenat Diagn 1998, 18:669–674
26. Raeymaekers L: Quantitative PCR: theoretical considerations with practical implications. Anal Biochem 1993, 214:582–585
27. Lubin MB, Elashoff JD, Wang SJ, Rotter JI, Toyoda H: Precise gene dosage determination by polymerase chain reaction: theory, methodology, and statistical approach. Mol Cell Probes 1991, 5:307–317
28. Nigro JM, Takahashi MA, Ginzinger DG, Law M, Passe S, Jenkins RB, Aldape K: Detection of 1p and 19q loss in oligodendroglioma by quantitative microsatellite analysis, a real-time quantitative polymerase chain reaction assay. Am J Pathol 2001, 158:1253–1262