Nucleic Acids Research, 1995, Vol. 23, No. 8 1411–1418

PCR bias in amplification of androgen receptor alleles, a trinucleotide repeat marker used in clonality studies

George L. Mutter[1,2,*] and Kevin A. Boynton[1]

[1]Department of Pathology, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115, USA and
[2]Harvard Medical School, Boston, MA 02115, USA

Received November 28, 1994; Revised and Accepted March 8, 1995

ABSTRACT

Trinucleotide CAG repeats in the X-linked human androgen receptor gene (HUMARA) have proved a useful means of determining X chromosome haplotypes, and when combined with methylation analysis of nearby cytosine residues permits identification of non-random X inactivation in tumors of women. Co-amplification of two alleles in a heterozygote generates PCR products which differ in the number of CAG units, and thus their melting and secondary structure characteristics. We have shown that under optimal conditions amplification efficiency of two HUMARA alleles is near-equivalent, generating PCR products in a ratio proportional to that of the genomic template. In contrast, reduction of template quantity, damage of template by ultraviolet irradiation or addition of monovalent salts (sodium chloride, sodium acetate or ammonium acetate) produces highly variable imbalances of allelic PCR products, with a strong tendency to preferentially amplify lower molecular weight alleles. Variability and biasing was diminished by substitution of 7-deaza-2′-dGTP for dGTP during amplification, an intervention which reduces stability of intramolecular and intermolecular GC base pairing. We conclude that DNA which is scanty, damaged or salt contaminated may display amplification bias of GC-rich PCR targets, potentially confounding accurate interpretation or reproducibility of assays which require co-amplification of alleles.

INTRODUCTION

Non-random X chromosome inactivation patterns may be used to infer a monoclonal growth pattern in tumors, and provide an index of the neoplastic phenotype in pathologic states (1,2). Recently several PCR-based assays have been developed to obtain such information from non-viable frozen and archival paraffin-embedded tissues (3–5). Amplification of the polymorphic X-linked androgen receptor gene (HUMARA) with and without preceding digestion by methylation sensitive restriction endonucleases allows quantitation of the distribution of alleles between the active and inactive X chromosomes (6). Following digestion with the endonuclease *Hha*I, only the methylated inactive X chromosome alleles persist as intact PCR template and are visualized as PCR product. Polyclonal tissues that have an admixture of alleles on the inactive X chromosome maintain a balance (non-skewed, or random) of both alleles following *Hha*I pre-digestion, whereas monoclonal proliferations characteristically contain the same inactive X copy in all cells, leading to disappearance (skewed or non-random) of the active allele with *Hha*I pre-digestion (example in Fig. 1).

Accurate interpretation of clonality status from X inactivation patterns assessed through PCR requires control of both biologic and technical parameters. Normal polyclonal tissues may have skewed X inactivation patterns caused by chance inactivation of the same X chromosome in the few component clones (2,7). The extent and frequency of skewing within a population of women has been used as a means of estimating the number of clones present in normal polyclonal tissues (8). If skewing is present in the polyclonal source tissues from which a tumor has arisen, then skewing is not a specific feature of the monoclonal state, and X inactivation is non-informative in ascertainment of clonality.

The hypervariable androgen receptor gene target has a high (65%) GC content, and highly stable secondary structure (Fig. 5) that potentially could alter amplification efficiency in an allele-specific manner, producing 'skewing' of PCR products in the absence of a comparable imbalance of starting template. Precedents for these possibilities include the PCR inhibitory effect of a highly stable secondary structure in the Fragile X gene (9), and allelic divergence of regional melting characteristics of homoduplex template DNA leading to preferential amplification of HLA alleles (10). The androgen receptor gene may be difficult to amplify, with more reliable PCR results obtained by chemical [addition of DMSO (11)], or thermal [combination of 'hot start' PCR (3,12) followed by prolonged initial denaturation intervals (3,13)] modification of cycling conditions. The efficacy of these

* To whom correspondence should be addressed at Department of Pathology, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115, USA

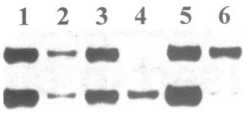

Figure 1. Monoclonal uterine leiomyomata demonstrated by PCR. DNA isolated from paraffin embedded archival uterine myometrium (lanes 1 and 2) and two different smooth muscle tumors (lanes 3–4 and 5–6) in the same patient (#93-99) was used for AR-a/b PCR either without (odd lanes) or with (even lanes) preceding *Hha*I digestion (3). Note that *Hha*I pre-digestion spares HUMARA alleles present on the inactive X chromosome, which in the polyclonal myometrium are approximately evenly distributed between the HMW and LMW alleles (lane 2, balanced or random pattern of X inactivation). The two leiomyomata are independent lesions as evidenced by derivation from different clones which have non-randomly inactivated different X chromosomes. Skewing of PCR signal towards one of the two alleles provides a rationale for inferring the monoclonal growth patterns of these benign tumors. PCR product molecular weights: 233 and 218 bp. Autoradiographic exposure 14 h.

measures in improving PCR product yield and/or specificity is a warning that the efficiency of this amplification is unusually sensitive to the PCR conditions utilized.

Expansion of the androgen receptor CAG repeat has also been associated with a disease phenotype, spinal and bulbar muscular atrophy (Kennedy's disease) (14). In this and other disorders characterized by trinucleotide repeat expansion such as Fragile X syndrome (15) and Huntington's disease (16,17), structural analysis of the repetitive sequence may play a role in patient diagnosis. The fidelity and reliability of PCR amplification of such targets is therefore also of immediate relevance in defining the sensitivity and specificity of PCR based molecular diagnosis of these genotypes.

We have observed dramatic variation in the relative intensity of allelic HUMARA PCR products on repeat analysis of a single DNA sample. While this is infrequent in clean isolates of high molecular weight DNA from fresh or frozen tissues, it is often seen in suboptimal DNAs. In particular, partly degraded DNA isolated in minute quantities from archival paraffin embedded tissue contaminated by blood and metallic impurities frequently shows biased amplification of one of two HUMARA alleles present. Even with optimization of the thermal conditions of PCR through empirical testing of conditions, individual samples may have uncontrollable variation in DNA quantity, quality and contaminants. Titration of DNA amounts or addition of contaminants to a PCR reaction permits delineation of the effects of these variables on PCR results. Short wave ultraviolet light exposure was chosen as a model for DNA damage because it represents a common environmental factor capable of randomly introducing, in a dose dependent fashion, DNA lesions which terminate DNA polymerase action. We here show increased random and systematic amplification bias caused by these factors, and demonstrate that destabilization of GC base pairing through incorporation of 7-deaza-2′-dGTP (18) minimizes their effect.

MATERIALS AND METHODS

Clonality assay using PCR(3)

Wax 7–10 μm thick sections of normal uterine myometrium and uterine leiomyomata were digested at 60°C with intermittent agitation for 48 h in 300 μl proteinase K lysis buffer (200 μg/ml proteinase K, 50 mM Tris pH 8.3, 0.5% Tween 20, 100 μg/ml mussel glycogen) with replenishment of (200 μg/ml) proteinase K after the first 24 h. The lysate was extracted twice with an equal volume of phenol, and 36 μl of 5 mg/ml ethidium bromide was added to the aqueous phase and allowed to incubate 1–2 min at room temperature. One half volume of 7.5 M ammonium acetate was added before extraction with phenol–chloroform (×1) and chloroform (×3). The aqueous phase was treated for 5 min with 1 drop of Chelex-100 (BioRad, Bethesda, MD) bead slurry, and ethanol precipitated. The precipitate was washed with 70% ethanol, and resuspended in 40 μl *Hha*I buffer.

An undigested aliquot of DNA was retained as a control. The remaining DNA was digested in *Hha*I buffer with 20 U *Hha*I for 1–3 h, quenched by phenol–chloroform extraction, and ethanol precipitated prior to PCR with AR-a/b primers.

DNA isolation from frozen human uterine myometrium

Frozen human myometrium (samples 93–99 and 94–157) was pulverized in liquid nitrogen prior to lysis by proteinase K and purification by organic extraction and sequential ethanol precipitation by ammonium acetate and sodium acetate, respectively. The final cold precipitation was followed by careful removal of residual salts through vigorous washing of the pellet in cold 70% ethanol prior to resuspension at a final concentration of 50 μg/ml in 10 mM Tris (pH 8.0), 1 mM EDTA buffer (TE).

Allele titrations

Genomic DNA from frozen human sperm was isolated as previously described (19), and the molecular weight of the HUMARA allele determined by amplification with primers AR-a/b and polyacrylamide gel electrophoresis. Genomic DNA from individuals having alleles of different molecular weight was adjusted to a concentration of 200 μg/ml and volumetrically titrated to generate admixtures of the two genomic DNAs as seen in Figure 2.

Ultraviolet irradiation of DNA

50 ng of genomic DNA in 1 μl TE buffer was placed in the bottom of a 0.5 ml polypropylene microfuge tube mounted in a vertical position directly underneath a short wave (254 nm) ultraviolet bulb and exposed for the time necessary to reach the desired dose as measured by a built-in sensor (Stratalinker apparatus, Stratagene, La Jolla, CA). Following irradiation PCR master mix was added to the tubes.

PCR amplification

Amplification of solubilized DNA with primers AR-a and AR-b (3) (5′-CCGAGGAGCTTTCCAGAATC-3′ and 5′-TACGATG GGCTTGGGGAGAA-3′, respectively), AR-c and AR-d (20mer modifications of primers described by Allen *et al.* (6), 5′-TCCA-GAATCTGTTCCAGAGC-3′ and 5′-GCTGTGAAGGTTG CT-GTTCC-3′, respectively) or AR-e and AR-g (5′-TACCCT-CGGCCGCCGTCCAA-3′ and 5′-CCTCTCTCGGGGTGGC-ACTC-3′, respectively) was performed in a 50 μl PCR reaction mix. Trace amounts (final concentration 50–100 nM) of [^{32}P]TTP (800 Ci/mmol) were added to the reaction mixture [10 mM Tris pH 8.4, 50 mM KCl, 20 μg/ml gelatin, 1.5 mM $MgCl_2$, 0.2 mM dATP, 0.2 mM dGTP (or 0.2 mM 7-deaza-2′-dGTP from Boehringer Mannheim, Indianapolis, IN), 0.2 mM dCTP, 0.05 mM TTP] to label amplified products. Oligonucleotide primers

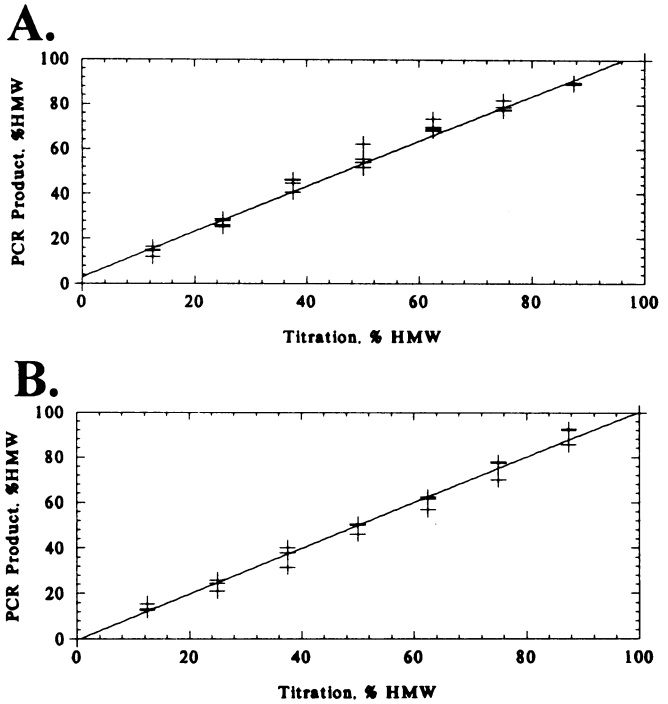

Figure 2. Titrations of HUMARA alleles demonstrates a linear PCR signal response. Genomic DNA from two males with different androgen receptor alleles was mixed in varying amounts and 200 ng used as template for 24 cycles of AR-a/b PCR, the products resolved by PAGE and visualized by autoradiography. Densitometry of the resultant signal, expressed as percentage of all signal present as HMW allele. Replicate determinations are shown as crosses, and the regression line drawn for results using either dGTP (**A**) or 7-deaza-2′-dGTP (**B**) during PCR. Both conditions produce a strong linear relationship between genomic template allele ratios (horizontal axis) and PCR product ratios (vertical axis).

were added to the reaction mix to a concentration of 0.3 μM each of AR-a/b or AR-e/g, or 0.9 μM each for AR-c/d. Annealing temperature was empirically optimized at 55°C by comparison of results obtained over a range of temperatures. Thermal cycling in an MJ PTC-100 (MJ Research, Inc., Watertown, MA) thermal cycler followed the sequence: (i) 'Hot start' pre-heating of reaction mixture without *Taq* polymerase at 97°C for 5 min; (ii) 10 min dwell at 85°C, during which 1.25 U *Taq* polymerase (Amplitaq, Perkin Elmer Cetus, Norwalk, CT) was added to each reaction tube; (iii) three cycles with prolonged denaturation consisting of 95°C × 4 min, 55°C × 45 s and 72°C × 90 s; (iv) 21–31 cycles at 95°C × 30 s, 55°C × 45 s and 72°C × 90 s, and (v) final extension at 72°C for 7 min. PCR products were electrophoresed at 200–500 V in an 8% non-denaturing 0.5 mm thick polyacrylamide gel made in 1 × TBE (45 mM Tris-borate, 1 mM EDTA). Gels were dried and autoradiography performed.

Modifications of amplification reagent conditions were made by supplementing the PCR master mix with additional materials prior to addition of the DNA template. The number of PCR cycles needed for template amounts smaller than 50 ng (0.5 ng) were estimated by adding one additional cycle for each approximate halving of input DNA quantity.

Autoradiography and densitometry

Preflashed Kodak XAR film with intensifying screens at –70°C was used for autoradiographic detection of radiolabelled PCR products in a dried polyacrylamide gel. Autoradiogram optical density was measured with an EC model 910 optical densitometer (EC Apparatus Corporation, St Petersburg, FL) and the resultant plot integrated using the GS365W Electrophoresis Data System, v.2.0 (Hoeffer Scientific Instruments, San Francisco, CA). The signal intensity of both (high molecular weight, HMW and low molecular weight, LMW) androgen receptor alleles was measured in each lane, and the percentage of all specific signal contributed by the HMW allele calculated as follows: percentage HMW signal = (100 × HMW signal)/(HMW signal + LMW signal). Numerical data were imported into the statistical graphics program Systat (Systat, Inc., Evanston, IL), v.5.0W for analysis and graphing.

Densitometry was not performed in cases judged to be PCR 'failures'. These are of two types: (i) absence of any discernible autoradiographic signal in the position expected; or (ii) non-specific bands overlapping with, or in close proximity to, the expected PCR products.

PCR product regional melting and secondary structure prediction

The molecular weight of AR-a/b PCR products generated from DNA 93-99 using dGTP (not 7-deaza-2′-dGTP) was carefully calculated on several gels by comparison with a molecular weight ladder. The HMW allele measured 233 bp, a size expected for an allele containing 17 trinucleotide repeat units. The LMW allele was 15 bases shorter, thereby containing 12 repeat units. DNA sequences for the HUMARA locus were obtained from GenBank (compiled by NIH, Bethesda, MD), and the number of trinucleotide repeat units adjusted to correspond to the measured molecular weights of the test DNA (93-99). Regional melting profiles for double stranded DNA PCR products were calculated using the program MELT 87 (20). Secondary structure predictions of single stranded DNA were performed with DNAsis v.7.0 (Hitachi Software Engineering; Brisbane, CA).

RESULTS

Allele titrations

There was high concordance between titrations of allelic template and relative abundance of AR-a/b PCR products, confirming that under optimal conditions the PCR product abundance does reflect that of the starting template (Fig. 2). Use of 7-deaza-2′-dGTP yielded slightly reduced incorporation into the HMW allele than dGTP, seen as a marginally lower percentage of measured HMW PCR products. We have used titrated sperm DNAs to confirm linearity of allelic amplification under a range of conditions beyond that of Figure 2, including varying DNA input from 20 to 500 ng, and changing cycle number from 22–26 cycles (data not shown).

Types of PCR 'failures' under non-standard PCR conditions

The desired endpoint for all experiments was quantitation of relative abundance of the two androgen receptor allele PCR products. We were unable to do this under several circumstances (Figs 3 and 4), including absence of PCR products (a blank lane), or obscuring of specific products by a ladder of non-specific bands. For example, we were unable to quantitate results under conditions of increasing Mg^{2+} concentration (data not shown) due to

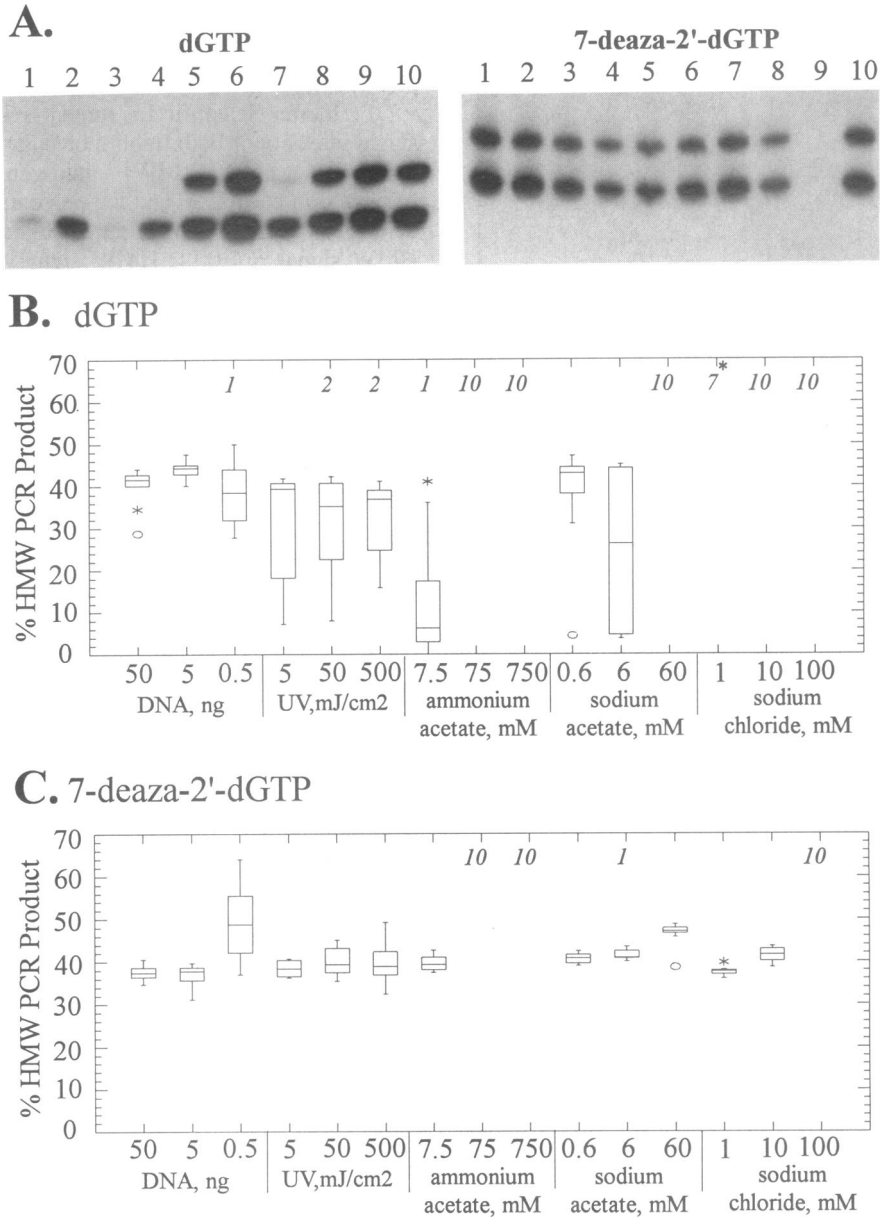

Figure 3. 7-deaza-2′-dGTP reduces allelic biasing during PCR. Co-amplification of 17 and 12 CAG repeat unit alleles in uterine myometrium DNA (93-99) remains stable under a wide variety of conditions when using 7-deaza-2′-dGTP, whereas amplification with dGTP produces variable results, under-representing the allele with most CAG repeats (HMW allele). **(A)** Shows an autoradiogram of 233 and 218 bp AR-a/b PCR products generated in presence of 6 mM sodium acetate, representative of results quantitated for inclusion in (B) and (C). The dGTP amplified DNAs variably demonstrate biasing towards the LMW allele whereas 7-deaza-2′-dGTP amplified samples have equivalent signal intensity of each allele except for lane 9, a 'failure' with no signal. Eighteen hour exposure. **(B)** and **(C)** The PCR products of AR-a/b amplified DNA (93-99) were quantitated, and results expressed as percentage of specific signal (HMW + LMW) present in the HMW allele. PCR was in the presence of dGTP (B) or 7-deaza-2′-dGTP (C). 50 ng template DNA and 26 PCR cycles used except for DNA titrations where 5 ng and 0.5 ng DNA were amplified for 30 and 34 amplification cycles, respectively. Amounts of listed electrolytes are supplemental additions to 1 × PCR buffer. Each bar contains data from all 10 reactions performed for each group, except in cases where the number of failed reactions is recorded in italics above the bar. The median is indicated by a horizontal line within the box which delineates mid-range data points, and whiskers extend an additional 1.5 times the interquartile range. Outliers are shown as circles or asterisks. The 1 mM NaCl dGTP group had seven (*) tubes with no signal, and three with all signal in the LMW allele: the three cases with 0% HMW do not appear as a visible box in this representation.

overwhelming loss of PCR specificity. Blank lanes were the usual pattern of failures noted in Figures 3 and 4, although at very low quantities of input DNA non-specific bands appeared.

Failed PCR reactions became commonplace under extreme ionic conditions. Although changes in target DNA melting and secondary structure were probably significant, a direct effect on the activity of *Taq* polymerase, or primer–template interactions is also possible.

Isolated 'blank lane' failures were occasionally seen under conditions where the remaining samples within the group not only had strong signal, but tightly reproducible allelic ratios (for

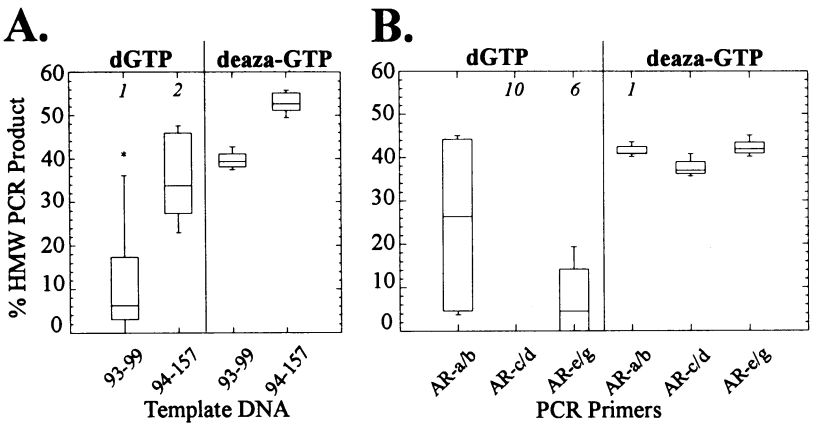

Figure 4. Biased co-amplification of androgen receptor alleles is seen with a variety of template DNAs and PCR primers. Substitution of 7-deaza-2′-GTP ('deaza-GTP') for dGTP ('dGTP') during PCR consistently reduces biased amplification of HUMARA alleles, independent of the template genotype or PCR primers used. (**A**) When using primers AR-a/b under 7.5 mM ammonium acetate PCR conditions, generation of HMW PCR products is variably diminished in DNAs in which the two HUMARA alleles differ by either five (93-99) or two (94-157) CAG repeat units, of which the HMW allele PCR product measures 233 bp and 236 bp, respectively. (**B**) Results of amplification of a single DNA (93-99) using three different PCR primer pairs, each of which flank the CAG repeat region of the HUMARA gene. The HMW PCR product generated from DNA 93-99 using primers AR-a/b, AR-c/d and AR-e/g was 233, 273 and 344 bp, respectively. Under the 6.0 mM sodium acetate PCR conditions used, there is strong bias against amplification of the HMW allele when using dGTP. Both panels used 50 ng template DNA, 26 PCR cycles and autoradiogram quantitation as described in the text. Figure format is as described in the Figure 3 legend.

example, Fig. 3A, 7-deaza-2′-dGTP lane 9). This suggests a low frequency of random failure such as might be caused by faulty pipetting, or thermal transfer (loose tubes) during PCR.

Effect of DNA quantity

Genomic DNA template quantities of 5 ng or more generate consistent PCR product allelic signal ratios, whereas lesser amounts (0.5 ng) are associated with deterioration of reproducibility manifest as large coefficients of variation (Figs 3B and 4C). The variation seen with small template amounts (0.5 ng) is further increased by non-specific products that obscure precise densitometry.

Effect of DNA integrity

DNA damaged by low to moderate levels of ultraviolet exposure was much more consistently amplified by 7-deaza-2′-dGTP (Fig. 3C) than dGTP (Fig. 3B). When using 7-deaza-2′-dGTP for PCR, non-irradiated, and 5 or 50 mJ/cm^2 UV irradiated DNA have similar means (37.6, 38.4 and 40%, respectively) and only slightly increased coefficients of variation (4.6, 4.5 and 8.0%, respectively) of measured HMW allele signal. In contrast, when using dGTP a sporadic biasing towards the LMW allele was notable along with a 4-fold increase in the coefficient of variation at the lowest exposure level.

Effect of added salts

Irrespective of the type of GTP utilized for PCR, increasing concentrations of ammonium acetate, sodium acetate and sodium chloride eventually diminish overall signal intensity to undetectable levels creating PCR 'failures' unsuitable for densitometry (Fig. 3). At low salt concentrations 7-deaza-2′-dGTP consistently maintained a reproducible allele signal ratio (Fig. 3C) similar to that of standard PCR buffer. dGTP amplified DNAs all showed systematic but highly variable biasing towards the LMW allele in the presence of low concentrations of the salts evaluated (Fig. 3B).

Effect of number of trinucleotide repeat (CAG) units

Salt-induced biasing of PCR amplification, reversed upon use of 7-deaza-2′-dGTP, was seen in DNAs with alleles differing by either five (patient 93-99, Fig. 4) or two CAG repeat units (patient 94-157, Fig. 4).

Effect of changing PCR primers flanking the repetitive HUMARA target

In addition to the extensively tested PCR primer pair AR-a/b (Fig. 3), two different external flanking PCR primer sets (AR-c/d, AR-e/g, Fig. 4) show salt-induced biasing or failure of PCR amplification, reversed upon use of 7-deaza-2′-dGTP. The ability to generate specific PCR products under salt-spiked conditions varies with the primers used, with a higher frequency of 'failures' observed when using primers (AR-c/d or AR-e/g) which generate amplification products longer than those of primers AR-a/b. We initially attempted to use 7.5 mM ammonium acetate PCR conditions to generate Fig. 4B, but in that case only primers AR-a/b generated any specific products capable of analysis (data not shown).

PCR product melting and secondary structure predictions

Thermodynamic properties of HUMARA alleles encountered in the primary test DNA (patient 93-99, used for Figs 1 and 3) are shown in Figure 5. Homoduplex DNA melting profiles were essentially identical throughout conserved portions of the sequence (panel A). The expanded repeat region in the HMW allele resulted in an elongated CAG melting domain without an apparent increase in the melting temperature itself. Regional melting temperature predictions do not, therefore, explain a different amplification efficiency between the two alleles.

Secondary structure calculations (panel B) of single stranded DNA PCR products show a high degree of self-complementarity, with thermal stability of the LMW and HMW alleles of −127.6 and −138.8 kcal/mol, respectively. Intramolecular base pairing overwhelmingly involved GC pairs, which comprised 79.7% (59

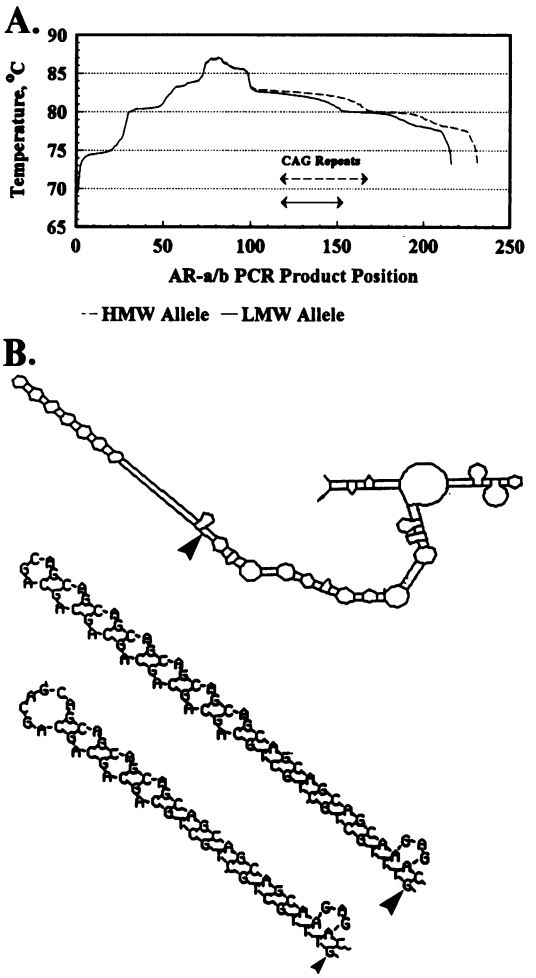

Figure 5. Regional melting profile and secondary structure of AR-a/b PCR products. (**A**) Regional melting temperatures of double stranded PCR products generated with primers AR-a/b from alleles with 12 (LMW allele, 218 bp) or 17 (HMW allele, 233 bp) CAG trinucleotide repeats, as present in DNA from patient 93-99, used to generate Figures 1 and 3. The melting profiles have comparable melting temperatures in trinucleotide repeat areas (indicated by arrows) and similar flanking regions of maximal stability (highest melting temperature). (**B**) Secondary structure predictions of single stranded PCR products. The upper line drawing is of the complete sequence, with details of areas of divergence between the HMW (large arrowhead) and LMW (small arrowhead) allele PCR products shown below.

of 74) of all bases paired in the folded single stranded HMW allele. The LMW allele differed only in containing six fewer paired bases, all of which were G-C pairs (Fig. 5B).

DISCUSSION

We have shown a tendency for allelic biasing in PCR amplification of the GC-rich androgen receptor gene under conditions of scanty, damaged or salt contaminated DNA. These effects, which are independent of the primer design or template genotype, might be attributed to the inherent structure of the HUMARA gene target, the most characteristic feature of which is a hypervariable CAG trinucleotide repeat region. Much of the biasing is resolved by substitution of 7-deaza-2′-dGTP during PCR.

The basis of allelic biasing

As the template DNA quantity drops, more PCR cycles are required to detect signal, and the resultant allelic ratios are less consistent. Some bias may be due to molecular sampling error caused by a small cohort of potential templates. The smallest quantity of genomic DNA tested, 0.5 ng, contains only ~78 diploid cell equivalents. Additionally, at these low template amounts there was some loss of specificity of PCR, with numerous unexpected bands appearing. Some of these non-specific products migrated close to, or with, the expected alleles, complicating accurate densitometry. For example, an apparent increase in HMW allele PCR products measured using 0.5 ng of genomic template and 7-deaza-2′-dGTP reagents may be erroneous for this reason (Fig. 3C). In any case, the variation in results using dGTP or 7-deaza-2′-dGTP dropped dramatically (to 5.1 nd 7.5% coefficient of variation, respectively) at a level of 5 ng input DNA. We chose to use 50 ng input DNA for additional experiments, due to the high stability of these baseline results.

Diminution of salt-induced allelic biasing during PCR with 7-deaza-2′-dGTP compared to dGTP is strongly suggestive of a reversible inhibitory effect caused by intramolecular or intermolecular GC bonding. Others have reported regional homoduplex melting temperatures (10) and single-stranded template secondary structure (9) as significant determinants of PCR amplification efficiency. In a system where two different templates are competitively co-amplified, small divergences of amplification efficiency may be manifest by dramatically altered PCR product ratios. Biased amplification consistently favors the LMW allele, which contains fewer trinucleotide CAG elements than its HMW counterpart. Homoduplex melting temperatures (Fig. 5) are not significantly different between the two modeled alleles, but secondary structure stability increases rapidly with increasing numbers of repeat elements, with a 9% increase in secondary structure stability conferred by the additional six GC base pairs which distinguish the two alleles in the test DNA (93-99). The extent of biasing diminishes as the number of repeats become more similar between the alleles co-amplified (Fig. 4A), offering additional evidence that it is divergence within the CAG repeat region which is primarily responsible for the divergence of allele amplification efficiency. We conclude that secondary structure of the PCR template is a probable major determinant of the observed effect, although we cannot rule out participation of other factors.

Damage of DNA by UV irradiation leads to more variable results using dGTP than 7-deaza-2′-dGTP, probably for multiple reasons. First, chemical modification or breakage of template reduces the number of intact PCR targets, increasing sampling bias of remaining copies. The minimum UV dose used for these experiments (5 mJ/cm^2), however, is only 0.25% of that used for selective UV destruction (2 J/cm^2) (3) of 90% of the HUMARA (AR-a/b) PCR signal from paraffin sections (data not shown). Secondly, the biasing effect of preferential amplification is additive with molecular sampling error. We propose this preferential amplification is diminished by use of 7-deaza-2′-dGTP, and may explain maintenance of a consistent result in spite of reduced template numbers. When using dGTP for PCR, dilution of non-irradiated DNA produces progressive loss of amplification specificity, whereas irradiation of larger quantities of DNA leads to diminished signal without substantial changes in specificity.

There must be additional factors which reduce PCR specificity in scanty, compared to damaged, DNA template.

Use of 7-deaza-2′-dGTP in PCR

7-deaza-2′-dGTP may be stoichiometrically substituted for its isomer dGTP in PCR reactions using *Taq* polymerase, and has been shown to increase PCR efficiency of GC rich targets (18) with extensive single stranded hairpin loop formation such as the human FMR-1 (9) and murine ODC (21) genes. 7-deaza-2′-dGTP preserves GC base pairing albeit at a 28% [from –50 to –36 kJ/mol of stack (22)] reduced stability level compared to dGTP. This produces measurable reductions in homoduplex melting temperatures (22), and predicts lower stability of single stranded DNA secondary structure. PCR products synthesized with 7-deaza-2′-dGTP tend to have slightly reduced labelling efficiency ([^{32}P]TTP incorporation) and higher molecular weight than those with dGTP.

Baseline relative signal intensities remain slightly in favor of the LMW androgen receptor allele with either dGTP or 7-deaza-2′-dGTP PCR (40.1 or 37.6% HMW allele signal, respectively using 50 ng starting DNA (93-99) and normal PCR conditions). Relative to the LMW allele, the HMW allele either incorporates [^{32}P]TTP less efficiently, or fewer PCR products are made. Although the explanation for this has eluded us, the consistency of 7-deaza-2′-dGTP PCR, and resistance to biasing under the conditions noted, do not press for investigation of this point.

Relevance of the parameters studied

The conditions evaluated approximate several problems inherent in utilization of archival paraffin-embedded tissues, or forensic samples. Archival and forensic templates may be unsuitable for PCR due to DNA quality or quantity, or contaminants which co-isolate with the DNA. In both cases a semi-quantitative assessment of amplified alleles is crucial: for determining the pattern of X inactivation in clonality studies, and for assessment of admixed DNAs in forensic studies.

DNA quantities recovered from optimally fixed and paraffin embedded human tissues may be as much as 2 μg per section (23), but small lesions such as precursor or early stage lesions may occupy only a fraction of a larger tissue block (3). For DNAs isolated by selective UV irradiation (24) or microdissection, technical limitations of DNA quantity and quality may define the lower limits of application.

UV irradiation is a well characterized model for dose-dependant introduction of lesions which prevent polymerase extension through formation of stable thymine dimers (25). This is in many ways functionally equivalent to the strand breakage created during processing of paraffin embedded tissues, a problem further exacerbated by DNA extraction protocols which favor smaller fragments (23). While the wavelength spectrum of solar UV irradiation at sea level is more diverse and of longer wavelength than that generated by the germicidal lamps used in these experiments, substantial UV exposure by ambient natural light is a common event for forensic specimens collected from outdoor sites.

The amounts and types of salts studied here are approximations of those retained following typical pre-PCR precipitation and concentration of DNA. These may persist by several routes following ethanol precipitation; crystals which co-precipitate with the DNA, incomplete decanting of salt-containing supernatants, and solubilized salts trapped in the dead space of the DNA pellet. Washing of the pellet with 70% ethanol helps to remove these retained compounds, but may not always be completely efficient. The lowest salt levels studied for ammonium acetate, sodium acetate and sodium chloride (7.5, 0.6 and 1 mM, respectively), represent only 1/10 000 of those added pre-precipitation and later retained in a 20-fold concentrated resuspension. If only solubilized salts are considered, a 1 ml aqueous solution precipitated under these conditions would contain this amount of salt in 0.3 μl of carried-over supernatant, either within the pellet dead space or tube sidewalls. Even with small DNA quantities, added carrier can contribute substantially to pellet dead space, further compromising washing.

Lastly, it is possible that GC-rich targets similar to, but different from, the HUMARA target might demonstrate similar amplification bias. While we have no data on other genes, our experience suggests that it is the CAG repeat itself which is the predominant determinant of the effect. Since this motif is shared by a number of other cloned genes, such as the Huntington's disease locus (16,17), it will be interesting to see if they demonstrate similar amplification properties.

Conclusions

PCR co-amplification of GC-rich allelic targets may preferentially favor one of the two templates under a variety of suboptimal conditions including small amount of available template, damage to template, and contaminants which are carried over into the PCR reaction itself. This bias is not detectable when using large quantities of pure high molecular weight DNA, a pre-condition which cannot always be met. Implementation of several precautions may improve the quantitative fidelity of co-amplification of these templates. First, use of 7-deaza-2′-dGTP during PCR minimizes many of the confounding effects. Secondly, the effect of contaminating salts can be minimized by judicious selection of salt type used for pre-PCR precipitation (sodium acetate instead of ammonium acetate), and careful removal by washing of the DNA pellet. Alternatively, concentration by a method such as ultrafiltration avoids use of salts entirely. Repeat, or multiple analysis of a single sample will further reduce the likelihood of misinterpretation due to random variation.

Finally, the conditions under which reproducible results are obtained must consider all potential confounding factors, and be validated for each experimental system. For example, reproducibility data garnered on DNA isolated from fresh tissue is difficult to extrapolate to partially degraded archival paraffin isolated DNA. This is of special concern when defining the lower limits of resolution of a complicated clonality PCR assay such as that utilizing the HUMARA marker system. In this case, reproducible results must control not only analytical bias conferred by the kinetics of PCR, but also potential errors in sampling a subset of cells from a complex tissue.

ACKNOWLEDGEMENT

This work was supported by a research grant from the American Cancer Society, # EDT-86 (GLM).

REFERENCES

1 Vogelstein,B., Fearon,E.R., Hamilton,S.R., Preisinger,A.C., Willard,H.F., Michelson,A.M., Riggs,A.D. and Orkin,S.H. (1987) *Cancer Res.*, **47**, 4806–4813.
2 Fialkow,P. (1972) *Adv. Cancer Res.*, **15**, 191–226.

3 Mutter,G.L., Chaponot,M. and Fletcher,J. (1995) *Am. J. Pathol.*, **146**, 501–508.
4 van Kamp,H., Jansen,R., Willemze,R., Fibbe,W.E. and Landegent,J.E. (1991) *Nucleic Acids Res.*, **19**, 2794.
5 Shroyer,K. and Gudlaugsson,E. (1994) *Hum. Pathol.*, **25**, 287–292.
6 Allen,R., Zoghbi,H., Moseley,A., Rosenblatt,H. and Belmont,J. (1992) *Am. J. Hum. Genet.*, **51**, 1229–1239.
7 Gale,R., Wheadon,H. and Linch,D. (1991) *Br. J. Haematol.*, **79**, 193–197.
8 Fialkow,P. (1973) *Ann. Hum. Genet.*, **37**, 39–48.
9 Pergolizzi,R., Erster,S., Goonewardena,P. and Brown,W. (1992) *Lancet*, **339**, 271–272.
10 Walsh,P., Erlich,H. and Higuchi,R. (1992) *PCR Methods Appl.*, **1**, 241–250.
11 Zhang,L., Leeflang,E., Yu,J. and Arnheim,N. (1994) *Nature Genet.*, **7**, 531–535.
12 Chou,Q., Russell,M., Birch,D.E., Raymond,J. and Bloch,W. (1992) *Nucleic Acids Res.*, **20**, 1717–1723.
13 Singer–Sam,J., LeBon,J.M., Tanguay,R.L. and Riggs,A.D. (1990) *Nucleic Acids Res.*, **18**, 687.
14 La Spada,A.R., Wilson,E.M., Lubahn,D.B., Harding,A.E. and Fischbeck,K.H. (1991) *Nature*, **352**, 77–79.
15 Richards,R.I. and Sutherland,G.R. (1992) *Cell*, **70**, 709–712.
16 Warner,J., Barron,L. and Brock,D. (1993) *Mol. Cell. Probes*, **7**, 235–239.
17 Trottier,Y., Biancalana,V. and Mandel,J. (1994) *J. Med. Genet.*, **31**, 377–382.
18 Innis, M. (1990) edited by Innis, M., Gelfand, D., Sninsky, J. and White, T. *PCR Protocols*. Academic Press, Boston, MA, pp.54–59.
19 Lobel,S., Pomponio,R. and Mutter,G.L. (1993) *Fertil. Steril.*, **59**, 387–392.
20 Lerman,L. and Silverstein,K. (1987) *Methods Enzymol.*, **155**, 482–501.
21 McConlogue,L., Brow,M. and Innis,M. (1988) *Nucleic Acids Res.*, **16**, 9869.
22 Seela,F. and Driller,H. (1989) *Nucleic Acids Res.*, **17**, 901–910.
23 Jackson,D.P., Lewis,F.A., Taylor,G.R., Boylston,A.W. and Quirke,P. (1990) *J. Clin. Pathol.*, **43**, 499–504.
24 Shibata,D., Hawes,D., Zhi–Hua,L., Hernandez,A., Spruck,C. and Nichols,P. (1992) *Am. J. Surg. Path.*, **141**, 539–543.
25 Sarkar,G. and Sommer,S.S. (1991) *Biotechniques*, **10**, 590–594.