3468–3472 *Nucleic Acids Research, 1998, Vol. 26, No. 15*

Highly accurate analysis of heterozygous loci by single cell PCR

Alex M. Garvin, Wolfgang Holzgreve and Sinuhe Hahn*

Department of Obstetrics and Gynaecology and Laboratory for Prenatal Medicine, University of Basel, 46 Schanzenstrasse, Basel CH-4031, Switzerland

Received May 8, 1998; Revised and Accepted June 19, 1998

ABSTRACT

Single cell PCR is a powerful method for determining the genetic properties of individual cells. In the instance of heterozygous loci, however, preferential amplification of one allele can lead to allele drop out (ADO) of the other allele. Fortunately ADO does not occur in all amplifications, and is usually random when it does occur, with both alleles being equally susceptible to drop out. Therefore pooling of results from multiple independently amplified cells should greatly improve the analysis of diallelic loci, and the misdiagnosis rate of diallelic loci should decrease exponentially with the number of cells analysed. We have shown that this is true and that multiplex PCR allows for the simultaneous identification of a cell in a mixture of cells using microsatellite loci known to be informative, and accurate genotyping at other loci. This approach has practical applications to non-invasive prenatal diagnosis where small numbers of fetal cells in the presence of maternal cells must be both identified and genotyped at loci involved in genetic disease.

INTRODUCTION

Single cell multiplex PCR was first used to genotype either haploid cells or diploid cells homozygous at the loci under interrogation (1). Later studies involving diallelic loci have shown that the low template copy number inherent to single cell PCR can lead to preferential amplification of one allele (2–4). Preferential amplification can be explained in terms of the cycling efficiency of PCR. For example, single cell amplification of the β-globin locus was calculated to occur with an average efficiency per cycle of 65% (1). If by chance only one allele in a heterozygote amplifies during the first few cycles, the other allele will be under-represented in the final amplification products. Allele drop out (ADO) is an extreme form of preferential amplification in which the under-represented allele is undetectable. ADO is thus a potential problem when attempting to genotype heterozygous loci in single cells, for example in clinical applications such as forensic medicine, preimplantation genetics, or genetic diagnosis of fetal cells isolated from the peripheral blood of pregnant women (5).

The first step when analyzing single cells taken from a mixture of cells is to unambiguously determine their identity. Previous PCR-based studies on fetal nucleated erythrocytes obtained from maternal blood have relied on the expression of fetal proteins, such as fetal and embryonic hemoglobin, for their identification (6,7). Unfortunately, such proteins may be expressed in maternal cells under some circumstances, therefore their presence is not definitive proof of fetal origin. An alternative approach is to look directly at the DNA of candidate fetal cells and to detect sequences that differ from the maternal sequence at a given locus. The most reliable identification is made by detecting a paternal sequence not found in the mother, since the absence of a maternal sequence could be accounted for by other means such as ADO. Y chromosome sequences can be used to identify fetal cells in the case of a male fetus (8), but a general strategy must use autosomal sequences. One approach is to first genotype the mother and father at a number of autosomal loci in order to find a locus where the parents have no alleles in common. Highly heterogeneous microsatellites, or short tandem repeats (STRs), are convenient for this purpose (9,10). Fetal cells must have one paternal allele at this locus which is not present in maternal cells, thus the presence of a paternal allele in a candidate fetal cell proves fetal origin. The hypothesis is that by performing multiplex PCR on a candidate fetal cell, the cell can be proven fetal and analysed at additional loci.

Some of the most common genetic diseases, such as cystic fibrosis or the β thalassemias, are autosomal recessive and prenatal diagnosis is most useful when both parents are heterozygous carriers for loss of function mutations in the same gene. In these circumstances, the fetus may be heterozygous at the locus of interest, implying that any reliable PCR-based test using single fetal cells must address the problem of ADO. Fortunately, ADO is random for most loci, therefore analysis of multiple cells at a diallelic locus, followed by signal averaging, should allow for accurate diagnosis even if ADO occurs during each amplification. We wanted to investigate whether a correct diagnosis of four out

*To whom correspondence should be addressed at: Laboratory for Prenatal Medicine, University Women's Hospital, 46 Schanzenstrasse, Basel CH-4031, Switzerland. Tel: +41 61 325 9224; Fax: +41 61 325 9399; Email: shahn@uhbs.ch

Table 1. Genotyping of mother, fetus and cell picks

Loci to prove fetal identity							Analysis of mono and diallelic loci					
	D18S535			D21S1270			D13S767	D21S1440		D21S1432		
alleles (bp)	129	141	150	182	186	194	162	160	163	139	143	147
Mother (100ng gDNA)	54	46	0	60	40	0	100	52	48	0	54	46
Fetus (100ng gDNA)	0	57	43	0	53	47	100	53	47	53	47	0
Neutrophil	58	42	0	18	81	0	100	NA	NA	0	72	28
NRBC 1	0	0	100	0	0	100	100	50	50	0	100	0
NRBC 2		NA			NA		ND	ND	ND		ND	
NRBC 3	0	86	14	0	15	85	NA	100	0	0	100	0
NRBC 4	0	42	58	0	77	23	100	100	0	0	100	0
NRBC 5	0	100	0	0	100	0	ND	ND	ND		ND	
NRBC 6	0	100	0	0	58	42	100	58	42	100	0	0
NRBC 7	0	100	0	0	51	49	100	51	49	18	82	0
NRBC 8	0	59	41	0	100	0	NA	18	82	58	42	0
NRBC (averages)	0	70	30	0	57	43	100	62	38	29	71	0

Four heterozygous and one homozygous loci were analyzed. The numbers for each allele refer to the % of total signal at a given locus. NA, no amplification; ND, not done.

of four diallelic loci in cells proven to be fetal is possible using only six fetal cells.

MATERIALS AND METHODS

DNA preparation from whole blood

Permission to use blood taken from the patients was obtained from the Ethics Committee at our institution. One ml of maternal blood and 2 ml of fetal cord blood were subjected to EC lysis and DNA from the nucleated cells was isolated using DNAzol (Life Technologies, Basel, Switzerland).

Isolation and preparation of nucleated red blood cells (NRBCs)

A mixture of 2.5 ml maternal blood and 0.5 ml cord blood in 15 ml phosphate buffered saline (PBS) was centrifuged at 1600 *g* for 10 min over ficoll hypac (Pharmacia, Dubendorf, Switzerland) at a density of 1.110 g/ml. The nucleated cell fraction was washed in PBS, after which cells were transferred onto glass slides by cytospin centrifugation (Shandon, Frankfurt, Germany) and stained with hemotoxylin/eosin. NRBCs were identified by their unique morphology, namely a compact darkly staining nucleus smaller in diameter than an erythrocyte. Single cells were picked with glass needles using a Zeiss micromanipulator and inverted microscope at 400× amplification. Cells were individually placed in 0.5 ml microfuge tubes containing 5 µl of 400 ng/µl proteinase K and 17 µM SDS. The solution was overlaid with mineral oil and incubated at 50°C for 1 h, then at 99°C for 30 min.

PCR

Primer pairs for each of the loci listed in Table 1 were purchased from Research Genetics (Huntsville, AL) and contained the Fam fluorescent dye attached to the 5′ primer of each pair. The initial genotyping of fetus and mother was done using 100 ng of DNA in a 20 µl reaction volume cycling 94°C for 30 s, 55°C for 30 s, 30× using standard PCR conditions (2 mM $MgCl_2$).

Semi-nested multiplex PCR was performed on each NRBC using five primer pairs for the external reactions simultaneously. The additional primers which we used for nested PCR were 5′-GCCTGGCAGGCTAAGAGTTA-3′ for D18S535, 5′-AACTCACTCTCTCTCTCTCTT-3′ for D21S1270, 5′-AATATCTCTCCAGACACATGA-3′ for D13S767, 5′-TTCTAAAAGAAATCAAAATGATGC-3′ for D21S1432 and 5′-ATGTGTGTATGCCAGCCTCTG-3′ for D21S1440. PCR mix (50 µl) containing each of the 10 primers at 400 nM, 2 mM $MgCl_2$ and standard PCR buffer (Life Technologies, Basel, Switzerland) were added to each tube containing a cell. The tubes were heated to 94°C before addition of 2 U of *Taq* polymerase. After 30 cycles of amplification as described above, 1 µl aliquots were removed and used as template for the internal reactions using conditions

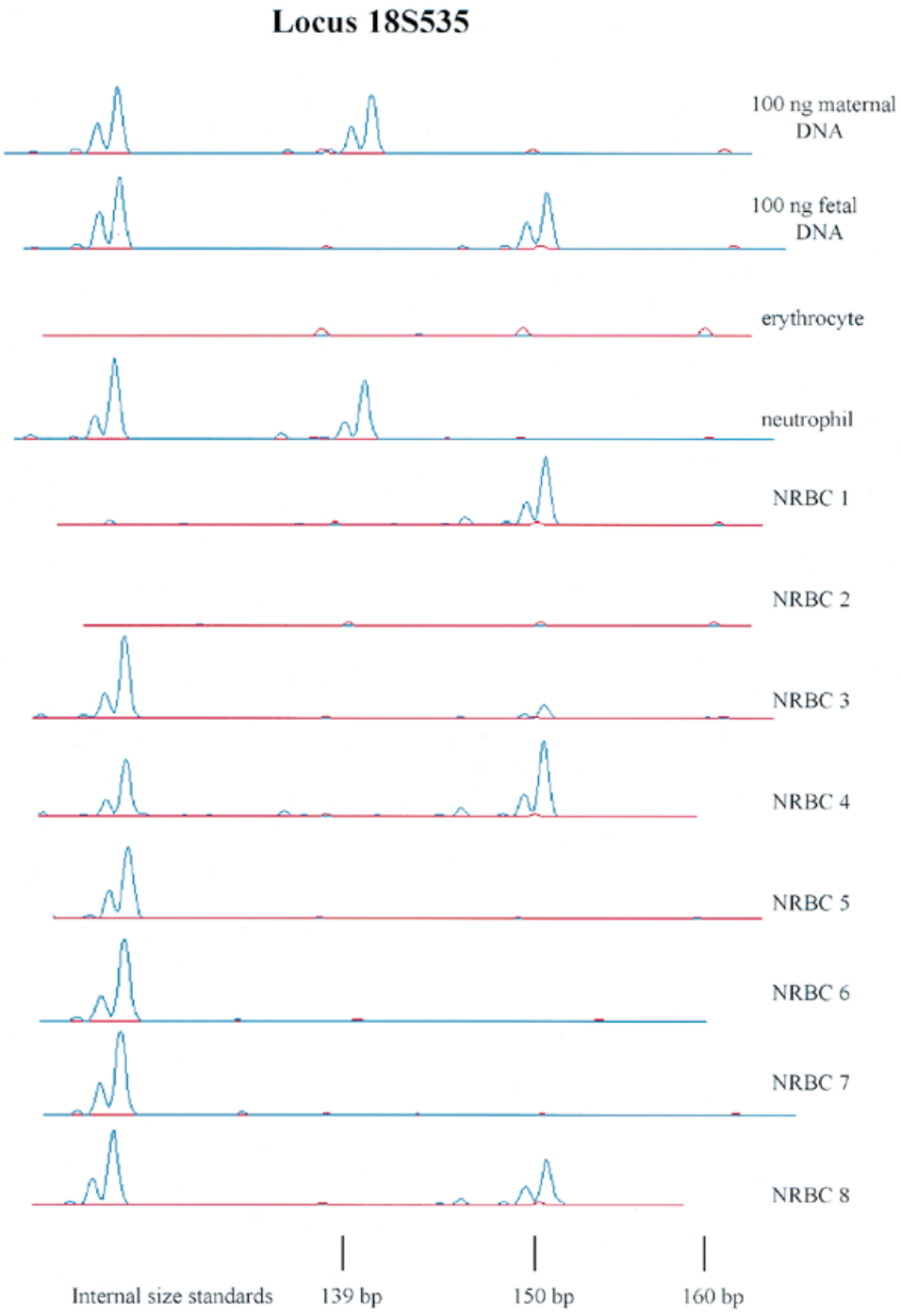

Figure 1. Electropherograms of locus 18S535 for maternal DNA, fetal DNA and 10 cell picks.

identical to those used with 100 ng of DNA except that hot start was performed.

PCR products were first checked on a 2% agarose gel and then analysed on an ABI310 prism automated sequencer using the Genescan software program (Perkin Elmer, Foster City, CA). Tamra labeled GeneScan 500 molecular weight standards were included as internal size standards for each run.

RESULTS

Identification of informative loci

Fetal DNA isolated from cord blood and maternal DNA from maternal blood were used to genotype the two individuals at five microsatellite loci. The allele sizes are shown in Table 1. Two loci

(D18S535 and D21S1270) were used to prove fetal origin in candidate fetal cells by virtue of the unique paternal allele present in the fetus. The other three loci were used to test the ability of single cell PCR to genotype monoallelic (D13S767) and diallelic (D21S1440 and D21S1432) loci.

Identification of fetal cells

Cord blood (0.5 ml), containing NRBCs at a frequency of 4% of total nucleated cells, was mixed with 2.5 ml maternal blood that had no NRBCs detectable by hemotoxylin/eosin staining of whole blood. Nucleated cells were purified by density centrifugation, placed on slides and stained. Since adult blood has a higher leukocyte count than cord blood, NRBCs represented ~0.3% of nucleated cells on these slides, which is similar to the frequency of fetal NRBCs in the blood of pregnant women after enrichment with anti-CD71 coated magnetic beads (11,12). Eight NRBCs, one neutrophil and two erythrocytes were picked and individually subjected to multiplex PCR at the five loci listed in Table 1. The ten primer pairs required for the external amplification step of semi-nested PCR were added to each cell and 30 cycles of PCR were performed. Two 1 µl aliquots of the external PCR products were removed and used as template for two separate amplifications of the two loci used for identification. Those cells proven to be fetal were then analysed at the other three loci.

Electropherograms of locus D18S535 for all eight NRBCs, the neutrophil and one erythrocyte are presented in Figure 1. The neutrophil had the expected maternal genotype, while the erythrocytes were negative for all five loci (Fig. 1 and data not shown). One of the eight NRBCs (NRBC 2) did not amplify at either of the two loci used for identification and was not analysed further. The other seven NRBCs amplified at both loci used for identification, and as can be seen in Figure 1 for locus D18S535, either one or both fetal alleles were amplified from each of these cells.

By examining only one locus (18S535), three NRBCs could be proven to be fetal either by virtue of the amplification of both alleles (NRBCs 4 and 8) or only the paternally inherited fetal allele (NRBC 1). On the other hand, by examining both loci, six of the NRBCs could be shown to be fetal (Table 1). No conclusive identification could be made for NRBC 5, since the paternal allele failed to amplify for both loci. This cell was not analysed further. Cells 1, 3 and 4 had the definitive paternal allele at both loci, while cells 6, 7 and 8 possessed the paternal allele at only one locus, and had ADO of the paternal allele at the other locus. ADO occurred in five out of the 12 amplifications (42%) used to prove fetal origin in these six cells.

Genotyping monoallelic and diallelic loci

We next examined three other loci in these six NRBCs, which we had shown to be fetal based on the presence of a unique paternal allele. Of these three loci, locus D13S767 is homozygous in the fetus, while loci D21S1440 and D21S1432 are heterozygous. All three loci were correctly genotyped using the combined data from the six cells proven to be fetal (Table 1). Furthermore, the two loci used to prove fetal origin can also be considered as test loci if cells are used which are proven fetal at other loci. For example, NRBCs 1, 3, 4, 6, 7 and 8 can be proven fetal using loci D21S1270 and D21S1432. Analysis of D18S535 in these cells clearly demonstrates that the locus is heterozygous. Similarly, locus D21S1270 can be proven heterozygous using cells 1, 3, 5, 6, 7 and 8 proven to be fetal due to paternal alleles at loci D18S535 and D21S1432. Thus all four heterozygous loci tested were correctly shown to be heterozygous by averaging the signals from the six cells analyzed.

Frequency of ADO

Although these results showed that a correct diagnosis can be made despite the problem of ADO, provided that multiple cells are examined by single cell PCR, we wished to quantitate the frequency of ADO by analyzing a sufficiently large number of heterozygous loci amplified from single cells. In Table 2 the results are presented for 92 amplifications done at eight different loci in a total of 50 cells. ADO was considered to have occurred when one allele had <10% the intensity of the other allele, and hence was too close to background to score reliably. This arbitrary 10% minimum may result in an overestimation of the ADO frequency, but it should also reduce the likelihood of scoring background noise as an allele. These data show that the frequency of ADO is 62%, and that ADO is indeed random, with the smaller alleles dropping out with roughly the same frequency as the larger alleles at a given locus.

Table 2. ADO in single cell PCR

Locus	Only smaller allele amplifies	Only larger allele amplifies	Both alleles amplify	Number of cells examined
D18S535	3	1	4	8
D21S1432	6	11	7	24
D21S1437	5	3	4	12
D21S1270	2	1	5	8
D21S1440	3	2	5	10
GATA129D11	6	2	2	10
GATA71H10	3	1	3	7
DXS6785	3	5	5	13
Totals	31 (34%)	26 (28%)	35 (38%)	50[a]

[a]Multiplex PCR was done on some of the cells, thus the total number of cells examined is less than the total number of loci examined (92).

Table 3. The frequency of misdiagnosing a heterozygote as homozygous as a function of cells analyzed with an ADO rate of 62%

Cells analysed	% misdiagnosed
1	62
2	19.2
3	6.0
4	1.8
5	0.57
6	0.18
7	0.055
8	0.017

DISCUSSION

The accurate diagnosis of heterozygous loci that we describe requires the averaging of a small number of single cell PCR, hence it is technically not a single cell procedure. However, for most forensic and prenatal diagnosis applications a small number of cells are usually available. Genetic diagnosis of the preimplanted embryo is exceptional in this respect in that only a single cell is available. Primer extension preamplification (PEP) as method for whole genome amplification should allow multiple amplifications of the same locus from a single cell. Unfortunately, we have found that ADO and spurious products of the incorrect size are even more of a problem when preamplified template from a single cell is used as template for PCR.

Our four out of four success rate in correctly identifying a locus as heterozygous in Table 1 may seem high, especially when considering the problem of ADO. However, it is to be expected given the number of cells analysed and the fact that the frequency of ADO is <100%. The likelihood that all cells suffer ADO at a given locus decreases exponentially as the number of cells tested increases. Furthermore, ADO is random for most loci, hence the chance of the same allele dropping out in all amplifications suffering ADO also decreases exponentially with the number of cells tested. The rate of misdiagnosis equals 2× (frequency of ADO × 5)^n , where n equals cell number. Table 3 shows the calculated rate of misdiagnosing a heterozygous locus as homozygous when the ADO rate is 62%. When six cells are analysed, the diagnosis will be correct at 99.82% of all loci tested. Therefore, when using six cells, we can expect to correctly genotype four out of four loci 99.28% of the time. Accuracy at this level should be sufficient for many clinical applications.

ACKNOWLEDGEMENTS

This work was supported in part by Swiss National Science Foundation Grant Number: 3200-047112.96 and NIH (USA) Contract Number: N01-HD-4-3202.

REFERENCES

1 Li,H.H., Gyllensten,U.B., Cui,X.F., Saiki,R.K., Erlich,H.A. and Arnheim,N. (1988) *Nature*, **335**, 414–417.
2 Findlay,I., Ray,P., Quirke,P., Rutherford,A. and Lilford,R. (1995) *Hum. Reprod.*, **10**, 1609–1618.
3 Ray,P.F., Winston,R.M. and Handyside,A.H. (1996) *J. Assist. Reprod. Genet.*, **13**, 104–106.
4 Ray,P.F. and Handyside,A.H. (1996) *Mol. Hum. Reprod.*, **2**, 213–218.
5 Holzgreve,W., Garritsen,H.S. and Ganshirt-Ahlert,D. (1992) *J. Reprod. Med.*, **37**, 410–418.
6 Cheung,M.C., Goldberg,J.D. and Kan,Y.W. (1996) *Nature Genet.*, **14**, 264–268.
7 Zheng,Y.L., Carter,N.P., Price,C.M., Colman,S.M., Milton,P.J., Hackett,G.A., Greaves,M.F. and Ferguson-Smith,M.A. (1993) *J. Med. Genet.*, **30**, 1051–1056.
8 Bianchi,D.W., Flint,A.F., Pizzimenti,M.F., Knoll,J.H. and Latt,S.A. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 3279–3283.
9 Weber,J.L. and May,P.E. (1989) *Am. J. Hum. Genet.*, **44**, 388–396.
10 Risch,N., de Leon,D., Ozelius,L., Kramer,P., Almasy,L., Singer,B., Fahn,S., Breakefield,X. and Bressman,S. (1995) *Nature Genet.*, **9**, 152–159.
11 Ganshirt-Ahlert,D., Borjesson-Stoll,R., Burschyk,M., Dohr,A., Garritsen,H.S., Helmer,E., Miny,P., Velasco,M., Walde,C., Patterson,D. *et al.* (1993) *Am. J. Reprod. Immunol.*, **30**, 194–201.
12 Ganshirt,D., Garritsen,H.S. and Holzgreve,W. (1995) *Curr. Opin. Obstet. Gynecol.*, **7**, 103–108.