Highly Accurate Analysis of Heterozygous Loci by Single Cell PCR

Alex M. Garvin, Ph.D., Wolfgang Holzgreve M.D., and Sinuhe Hahn, Ph.D.
Laboratory for Prenatal Medicine, Department of Obstetrics and Gynecology, University of Basel, Switzerland, 46 Schanzenstrasse, Basel, CH 4031 Switzerland

ABSTRACT

Since the cell is the basic unit of life, analysis of a single cell, even if taken from a mixture of cells from more than one person, ensures that the information obtained is from a single individual. Although single cell PCR is a powerful method for determining the genetic properties of individual cells, many useful loci such as STRs, are often heterozygous and preferential amplification of one allele can lead to allele drop out (ADO) of the other allele. Fortunately ADO does not occur in all single cell PCR amplifications, and is usually random when it does occur, with both alleles being 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 analyzed. We have shown that this is true and that multiplex PCR analysis of STRs allows for the accurate identification of a cell in a mixture of cells.

INTRODUCTION

PCR amplification of STRs often allows for the unambiguous identification of a sample of human cells. In most cases large numbers of cells, all from the same individual, are available for analysis. For example, 1 microliter of blood has 10,000-20,000 nucleated cells, while 1 µl of saliva often has more than 100 nucleated cells. Therefore, even when using seemingly small amounts of bodily fluids as starting material for PCR, the number of cells used as template is invariably much greater than unity.

Under certain circumstances, a sample of human cells will consist of a mixture of cells from 2 or more individuals. Analysis of such a mixture presents some difficulties, especially when identification of a minor component of the mixture is required. Fortunately, cells can be readily separated from other cells. Therefore a sample containing DNA from multiple individuals can be broken down into many samples, each consisting of a single cell, that contain the DNA from a single individual. Subsequent single cell STR analysis should allow the cells to be placed into non-overlapping groups corresponding to the individuals that contributed to the parent sample. Furthermore, the collective genotype of the cells in a given group should correspond to the genotype of the individual that contributed the cells.

ISOLATION OF SINGLE CELLS AND MULTIPLEX PCR

A mixture of cells can easily be resuspended in buffer and cytospun onto a glass slide. Using a light microscope and a micromanipulator, single cells can be picked with a drawn glass capillary and placed in PCR microfuge tubes for single cell multiplex PCR amplification. The cellular proteins are first digested in 5 µl SDS/proteinase K at 50 degrees C for 1 hour, followed by heating to 99 degrees C for 30 minutes to denature the Proteinase K and to further denature and expose the chromosomal DNA. Multiplex PCR of STRs can be carried out either by doing 35 cycles of standard PCR, or by doing nested PCR. Standard PCR is faster and easier than nested PCR, but it requires extensive optimization in order to find conditions where all amplicons have similar amplification efficiencies (1). This is done by varying primer concentrations and in some cases by redesigning primers. Nested PCR requires less optimization, because the second stage of amplification can be split up with aliquots from the first stage placed in separate tubes (2). Since only one primer pair is present at the final amplification cycles under these conditions, competition for substrate between the amplicons in not a problem.

ALLELE DROP OUT

Single cell PCR at heterozygous loci suffers from an amplification artifact, termed allele dropout (ADO), not experienced when multiple copies of template are available as starting material (3, 4). As shown in Figure 1, locus D18S535 is heterozygous for the test individual, and both alleles are well represented when 100 ng of genomic DNA are used as template for 35 cycles of standard PCR. On the other hand, when single cells from the same individual are used as template for nested PCR at the same locus, a variety of patterns are observed. Cell 1 has the expected profile for a heterozygous locus, namely 2 peaks of the correct size and of equal intensity. Cell 2 also has 2 peaks of the appropriate size, but the larger allele is under-represented, having only 14% the intensity of the smaller peak. Certainly the chromosome containing the larger allele was present at the beginning

of the amplification process, otherwise no amplification would have occurred. During the first few cycles of PCR amplification, very few copies of template are present and since the cycling efficiency of PCR is less than 100% (not every template is replicated during every cycle) it is likely that the larger allele amplified less well simply by chance. Indeed, preferential amplification of this type is also seen where the smaller allele is under-represented. Cell 3 demonstrates an extreme example of such preferential amplification of the larger allele in which the smaller allele is under-represented to the point where it is undetectable. Cell 4 gives a similar pattern to cell 3, except in this case the larger allele is undetectable. In extreme cases of preferential amplification of this type, allele drop out (ADO) is said to have occurred. Two important points must be made from Figure 1 concerning ADO: firstly ADO does not occur during all single cell amplifications, and secondly it appears to be random when it does occur, with both alleles being susceptible to drop out.

The distinction between preferential amplification and ADO is somewhat arbitrary, since the undetected allele is most likely present at some level, and is only undetectable under a given set of experimental conditions. In order to minimize the risk of scoring background signals as alleles, it is prudent to establish a minimum threshold below which a signal is considered background and is not scored. We have arbitrarily decided that the under-represented allele must have at least 10% the intensity of the well-amplified allele to be scored. Using this criterion, the frequency of ADO will be overestimated but the risk of scoring artefacts as alleles is diminished. Table 1 presents the results of 92 single cell nested PCR amplifications at 8 different heterozygous loci. In 34% of the amplifications the larger allele failed to amplify, and in 26% the smaller allele failed to amplify, giving an allele drop out rate of 62%.

ACCURATELY GENOTYPING CELLS FROM A MIXTURE

After performing multiplex PCR on a number of single cells taken from a mixture of cells, the STR profiles should allow the cells to be placed into non-overlapping groups of cells with each group being derived from a single individual. Once this has been accomplished, the genotype of the individuals who contributed to the parent sample can be determined from the genotype of the single cells. The problem of ADO can be overcome by summing the signals from all the cells in a particular group. In order to misdiagnose a heterozygous locus as homozygous, not only must ADO occur for each cell, but it must also occur for the same allele each time as well. The frequency of misdiagnosing a heterozygous locus as homozygous will equal $2(ADOx.5)^n$ where ADO is the frequency of allele drop out and "n" is the number of cells analyzed. When 6 cells are used and the ADO frequency is 62%, a heterozygous locus will be correctly genotyped 99.8% of the time. Accuracy at this level is sufficient for many clinical applications.

SINGLE CELL FORENSIC ANALYSIS OF SPERM

Sperm DNA taken from a victim of multiple rape is difficult to analyse because it consists of a mixture from 2 or more individuals. However, single spermatocytes contain DNA from only a single individual, and can be subjected to multiplex single cell PCR. Since males only have 1 X chromosome, all X chromosome-containing sperm from an individual must have the same X chromosome, and hence the same STR profile for X chromosome loci. Furthermore, X chromosomes from different men have different STR patterns. Thus the genotype of loci on the X chromosome can be used to place single sperm from a mixture into discrete non-overlapping sets corresponding to the individuals who contributed the cells to the parent sample. X chromosome containing sperm can be readily isolated from the victim's diploid epithelial cells and Y chromosome containing sperm by fluorescent cell sorting based on DNA content. Summation of the signals obtained at autosomal loci in all of the sperm from a given set should provide an accurate genotype of the individual who is the source of the cells. The frequency of misdiagnosing a heterozygous locus as homozygous will be $2(.5)^n$ where "n" equals the number of sperm analyzed from a given individual.

REFERENCES

1. Findlay I., Taylor A., Quirke P., Frazier R., and Urquhart A. (1997) DNA Fingerprinting from Single Cells. *Nature* **389**: 555-556.

2. Li H. H., Gyllensten U. B., Cui X. F., Saiki R. K., Erlich H. A., and Arnheim N. (1988) Amplification and Analysis of DNA Sequences in Single Human Sperm and Diploid Cells. *Nature*, **335**:414-417.

3. Ray P. F., Winston R. M., and Handyside A. H. (1996) Reduced Allele Dropout in Single-Cell Analysis for Preimplantation Genetic Diagnosis of Cystic Fibrosis. *J. Assist. Reprod. Genet.*, **13**: 104-106.

4. Ray P. F. and Handyside A. H. (1996) Increasing the Denaturation Temperature During the First Cycles of Amplification Reduces Allele Dropout from Single Cells for Preimplantation Genetic Diagnosis. *Mol. Hum. Reprod*, **2**: 213-218.

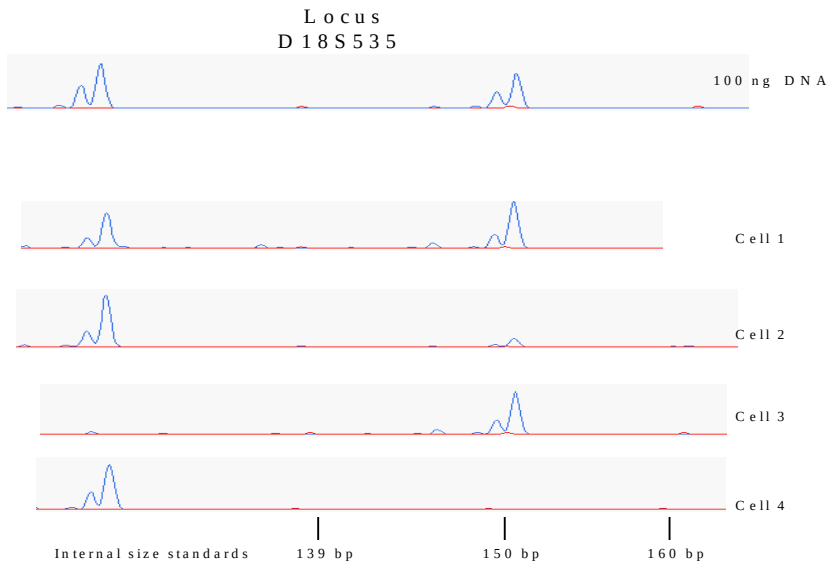

Figure 1. 100 ng of genomic DNA and 4 cells were used as template for PCR amplification of the D18S535 locus.

Table 1 Allele Drop Out 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*

*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).