Human Reproduction vol.6 no.6 pp.836–849, 1991

Using PCR in preimplantation genetic disease diagnosis

W.Navidi[1] and N.Arnheim[2]

Departments of Mathematics and [2]Molecular Biology, University of Southern California, Los Angeles, CA 90089, USA

[1]To whom correspondence should be addressed

Preimplantation diagnosis of genetic disease can be accomplished by embryo biopsy or polar body analysis using in-vitro gene amplification (PCR). PCR analysis of single cells is subject to a number of errors which decrease the reliability of the diagnosis. Using realistic assumptions about error rates based on experimenal data, we analyse some of the practical consequences to be faced by whose wishing to use this diagnostic procedure. We considered both autosomal dominant and recessive diseases. We calculate the probability of making mistakes in the diagnosis, assuming a realistic range in the magnitude of PCR efficiency, cell transfer, and contamination errors. We conclude that, in general, analysing blastomeres is subject to less mis-diagnosis than polar body analysis, except in the case of dominant diseases which are caused by genes which lie extremely close to the centromere. We also show that typing multiple blastomeres from a single embryo or combining polar body typing with blastomere analysis results in significantly lower levels of mis-diagnosis with unacceptable consequences. The preimplantation diagnosis of X-linked diseases based upon Y chromosome sequence analysis is also discussed.

Key words: polymerase chain reaction/gene amplification/blastomere typing/polar body typing/preimplantation diagnosis

Introduction

The polymerase chain reaction (PCR; Saiki *et al.*, 1985, 1988; Mullis and Faloona, 1987) is a simple method capable of rapidly amplifying DNA sequences *in vitro*. A single DNA segment composed of a few hundred base pairs present in a human genomic background having a complexity of three billion base pairs can be selectively amplified hundreds of millions to billions of times. In this way, the proverbial needle in a haystack is converted to a stack of needles. The fundamental principle of PCR and its applications to biological and medical science have been reviewed elsewhere (White *et al.*, 1989; Erlich, 1989; Arnheim, 1990; Arnheim *et al.*, 1990; Innis *et al.*, 1990; Erlich *et al.*, 1991). The first application of PCR was to the prenatal diagnosis of sickle cell anaemia (Saiki *et al.*, 1985). Since then, PCR has been applied to the prenatal diagnosis of many other genetic diseases using materials obtained from amniocentesis or chorionic villus sampling (CVS) (Kazazian, 1989).

The ability of PCR to be so selective in its amplification is accompanied by an exquisite sensitivity. Thus, a single molecule of DNA present in a single sperm cell can be amplified and analysed (Li *et al.*, 1988, 1990; Cui *et al.*, 1989). The ability to study DNA sequences in a single haploid or diploid cell (Li *et al.*, 1988; Jeffreys *et al.*, 1988) led geneticists and reproductive scientists to propose and experiment with the idea that the diagnosis of genetic disease could be established in human embryo produced by in-vitro fertilization prior to implantation (Li *et al.*, 1988; Handyside *et al.*, 1989; Coutelle *et al.*, 1989). Mouse embryos obtained from mated females have also been biopsied and the DNA analysed (Holding and Monk, 1989; Bradbury *et al.*, 1990; Gomez *et al.*, 1990). Normal pregnancies have been demonstrated after human (Handyside *et al.*, 1990) and mouse (Gomez *et al.*, 1990; Bradbury *et al.*, 1990) embryo analysis. The possibility of diagnosing genetic diseases even before fertilization using eggs and analysing the first polar body has also been examined (Monk and Holding, 1990; Strom *et al.*, 1990).

Successful preimplantation diagnosis of genetic disease depends upon being able accurately to determine the genotype of one or perhaps a few cells using PCR. In addition, the manipulations of the embryo, or unfertilized egg in the case of polar body analysis, must not affect its normal development. In this paper, we consider only the accuracy of DNA analysis by PCR.

For most routine applications of PCR, a sample consisting of the amount of DNA purified from 150 000 diploid cells (1 μg) is typical. In this case, not all of the original 300 000 copies of the target are required to be amplified during every PCR cycle in order to determine the genotype of the DNA accurately. The analysis of DNA in a single cell is an entirely different matter. A single diploid cell or polar body contains only two DNA molecules representing each single copy gene and therefore the accuracy of genotype determination is much more sensitive to random fluctuations in the efficiency with which each individual molecule is amplified during each PCR cycle. Thus, each of the two DNA molecules of the target gene present in a diploid cell must be efficiently amplified to a detectable level, and if a cell is heterozygous for a gene, both alleles need to be capable of being identified.

In the case of X-linked genetic diseases, however, a different approach may reduce the difficulty of single cell analysis. If the mother is a carrier of an X-linked recessive gene, female embryos would not suffer from the disease while male embryos would have a 50% chance of being affected. Handyside *et al.* (1989, 1990) have used human embryo sex-typing to allow implantation of only female embryos. In this procedudre, a DNA sequence repeated many times and specific to the Y chromosome was

analysed by PCR. Highly repeated gene targets increase the probability of detecting PCR product from a single cell.

The analysis of DNA sequences in single diploid cells using PCR has other associated problems. The single cell needs to be reliably transferred to the PCR reaction tube and contamination of the PCR reaction tube with exogenous DNA must be kept to a minimum since the target itself consists of only two molecules. We have analysed the accuracy of DNA diagnosis on one or a few cells taking into consideration the effects of amplification efficiency, the reliability of cell transfer, and contamination by either maternal cell contributions or by PCR products present in the laboratory from previous experiments. We consider the effect of these types of errors on blastomere and polar body analysis of single copy genes in the case of both dominant and recessive autosomal diseases. We also consider errors associated with the diagnosis of X-linked disease.

Assumptions

Denote the two alleles at the locus of interest by A and a, where A is dominant and a is recessive. We consider two cases. In the first case, the disease is recessive and both parents are heterozygous. In this case, an embryo has 25% probability of being of genotype AA, 50% of being of genotype Aa and 25% of being of genotype aa. Since the disease is recessive, only embryos of genotypes AA and Aa are suitable for implantation. Thus, an untyped embryo has 75% probability of being suitable. In the second case, the disease is dominant and one parent is heterozygous while the other is homozygous aa. In this case, an embryo has 50% probability of being of genotype Aa and 50% probability of being of genotype aa. Since the disease is dominant, only embryos of genotype aa are suitable for implantation, so an untyped embryo has 50% probability of being suitable. In either case, using blastomere typing to select an embryo, or using polar body typing to select an oocyte for fertilization, considerably increases the probability of selecting a suitable embryo. The probability does not rise to 100%, however, because of the possibility of typing error. It is of interest to compute the probability that an embryo has a genotype which makes it suitable for use, given that a blastomere cell has been typed as having such a genotype, and the probability that a fertilized oocyte has a genotype making it suitable for implantation given that it has been deduced to be suitable through typing of the corresponding polar body.

We present some calculations of such probabilities, based on some assumptions about the typing procedure. We consider three sources of error. First, the chance that the blastomere or polar body may fail to be placed in the reaction tube, so that the tube contains no relevant DNA. Second, the chance that an allele present in the tube may fail to be amplified to a detectable level. Third, the chance that the reaction may be contaminated, resulting in the false detection of an allele. We make the following assumptions:

(i) Each allele from the polar body or blastomere which is present in the tube has probability r of being amplified to a detectable level, independently of each other allele.

(ii) Each cell has probability d of being placed in the reaction tube, and probability $1-d$ of failing to be placed in the tube.

(iii) There is probability c that the reaction tube is contaminated and that the contaminant will be detected. At most one contaminating allele will be present. Each of the two alleles is equally likely to be a contaminant. When contamination is present, the number of contaminating molecules is approximately the same as the number of target molecules, so that the presence of contamination will not affect the detection of the target molecules. Other models of contamination are also plausible. We might assume, for example, that simultaneous contamination of two distinct alleles is possible, or that the amount of contamination can be so great as to prevent detection of the target molecules. The effect of such alternative contamination assumptions will be discussed below.

In a PCR analysis of >700 single spermatozoa, Cui *et al.* (1989) estimated that the value of c was $<5\%$, the value of d, using micromanipulation, was $\sim 13\%$, and the value of r was $\sim 95\%$. In other unpublished experiments, the value of r was usually found to range from 80 to 95%, c from 0 to 7%, and d from 10 to 20%. Our calculations take these ranges into consideration.

The parameters r, d and c used for calculation are unlikely to be constant from laboratory to laboratory, or even from time to time within the same laboratory. For this reason, we include calculations across a range of values. In any particular laboratory where this work is to be carried out, laboratory-specific parameter values should be estimated empirically. The results given here can then provide a guide to estimating the accuracy which may be expected.

Errors in typing can be divided into three categories, reflecting varying degrees of severity. The least serious error is one which results from using an embryo or oocyte which should have been used anyway, or in not using one which should not have been used. An example of such an error is typing an Aa cell as AA, in a situation where only aa cells should be used. Such errors will be classified as 'acceptable'. Given the limited number of embryos or oocytes available, a somewhat more undesirable sort of error is one which results in not using an embryo or oocyte which could have been used. Since in general no great harm is done by such errors, they are classified as 'tolerable'. The most serious error is one which results in implanting an embryo which should not be implanted. Such errors are classified as 'unacceptable'.

Using assumptions (i–iii), we calculated the probabilities that a cell which has been typed as being of a given genotype will in fact be of some other given genotype, and from these we computed the probabilities of error for each of the three categories mentioned above. The results are given below. Details of the calculations are given in the Appendix.

Results and discussion

Blastomeres, autosomal recessive disease

When the disease is recessive, embryos with genotypes AA and Aa are usable, while those with genotype aa are not. We assume both parents are heterozygous, so any embryo has probability 75% of being usable. Table I categorizes the possible typing errors in this case. The column labelled 'primary cause(s) of

Table I. Errors on blastomere typing for autosomal recessive diseases

True genotype	Observed genotype	Use?	Error category	Primary cause(s) of error	Conditional probability of error							
					(1) $r = 0.8$ $d = 0.8$ $c = 0.05$	(2) $r = 0.8$ $d = 1$ $c = 0$	(3) $r = 0.9$ $d = 0.8$ $c = 0.5$	(4) $r = 0.9$ $d = 0.9$ $c = 0.5$	(5) $r = 0.9$ $d = 1$ $c = 0.5$	(6) $r = 0.9$ $d = 1$ $c = 0$	(7) $r = 1$ $d = 0.8$ $c = 0.5$	(8) $r = 1$ $d = 1$ $c = 0$
AA	Aa	Yes	Acceptable	c	0.018	0	0.015	0.015	0	0	0.012	0
AA	aa	No	Tolerated	d,c	0.0057	0	0.0056	0.0026	0.00022	0	0.0063	0
Aa	AA	Yes	Acceptable	r	0.26	0.25	0.16	0.16	0.15	0.15	0.013	0
Aa	aa	No	Tolerated	r	0.26	0.25	0.16	0.16	0.15	0.15	0.013	0
aa	AA	Yes	Unacceptable	d,c	0.0057	0	0.0056	0.0026	0.00022	0	0.0063	0
aa	Aa	Yes	Unacceptable	c	0.018	0	0.015	0.015	0.015	0	0.012	0

error' tells which of the parameters r (amplification efficiency), d (cell placement) and c (contamination) plays the greatest role in determining the frequency at which the given error occurs. For example, the last row of the table refers to the error of typing a cell Aa when its true genotype is aa. For this error to occur, the reaction must be contaminated with the A allele. Thus the value of c has more effect on the frequency of this error than the values of r and d. The second to last row refers to the error of typing a cell AA when its true genotype is aa. For this to occur, not only must the reaction be contaminated with the A allele, but both a alleles must escape detection. The likelihood of the former event depends on the parameter c. The latter event may be due either to failure to place the cell in the tube, or to failure of PCR to amplify either of the two copies of the allele. For reasonable values of d and r, failure to place the cell in the tube is more likely to happen than failure of PCR to amplify either allele. Thus in general, the likelihood of both alleles escaping detection depends primarily on the parameter d, with r exerting a mild influence. Thus the parameters d and c are most important in determining the frequency of this error. As a final example, the third row of the table refers to the error of typing a cell AA when the true genotype is Aa. In general, this happens when the A allele is amplified to a detectable level but the a allele is not. Thus the value of r is most important in determining the frequency of this error. It is also possible for this error to be caused by a failure to place the cell in the tube, combined with contamination by a fragment containing the A allele. This combined error is probably much less likely than a lack of amplification, so the effect of the parameters c and d is relatively minor.

The right-most columns of Table I (labelled Conditional probability of error') give probabilities for each possible error in the case of blastomeres where the disease is recessive, for several values of r, d and c. The probabilities are conditional on the predicted genotype. Thus, for values of $r = 0.9$, $d = 0.8$ and $c = 0.05$, then a blastomere which has been typed Aa has probability of $\sim 1.5\%$ of in fact being of genotype aa, thus producing an unacceptable result. A blastomere which has been typed AA, however, has probability of only $\sim 0.56\%$ of having true genotype aa. Thus under these conditions we can expect that $\sim 1.5\%$ of implantations of blastomeres typed Aa and $\sim 0.56\%$ of implantations of blastomeres typed AA will yield unacceptable results. Examination of Table I shows that in each row the parameters primarily responsible for fluctuations in the probabilities are indeed the ones listed as 'primary causes of error'. For example, in the last row, comparing columns 3, 4 and 5 shows that increasing the value of d from 0.8 to 1 causes little or no improvement in the frequency with which cells typed Aa will turn out actually to be aa. Comparing columns 1, 3 and 7 shows that increasing the efficiency of r from 0.8 to 0.9 to 1 results in only a small improvement. Comparing the columns where the contamination rate c is 0.05 with those where $c = 0$ shows that reducing the contamination rate dramatically reduces the error rate. A similar analysis can be made in each row of the table.

Careful perusal of the tables shows what appears to be an anomalous result. In the second line of Table I, comparing columns 3 and 7 shows that if $d = 0.8$, $c = 0.05$, the probability that a cell typed aa will actually be of type AA increases slightly as the efficiency r increases from 0.9 to 1. This is discussed further in the Appendix. This anomaly appears again in line 5 of Table I and also in Tables III and VII.

Polar bodies, autosomal recessive disease

When the disease is recessive, only oocytes with genotype AA are suitable for fertilization. The polar bodies corresponding to these oocytes have genotype aa. The proportion of oocytes which are suitable depends on the recombination fraction θ between the centromere and the locus of interest. The proportion of oocytes of genotypes AA, Aa and aa is $\frac{1}{2}-\theta$, 2θ and $\frac{1}{2}-\theta$ respectively. We assume both parents are heterozygous, so any oocyte has probability 25% of being usable. Table II categorizes the possible typing errors in this case. The errors in the second and fourth rows are classified as 'potentially unacceptable'. In these cases, an oocyte containing the disease allele a is used, and whether the resulting embryo is of the unacceptable genotype aa is a matter of chance. If the oocyte is of genotype aa, the probability that the embryo will have genotype aa is 50%, and if the oocyte is of genotype Aa, the corresponding probability is 25%, which is identical to the probability if no preimplantation diagnosis is attempted.

Table III gives probabilities of unacceptable results in polar body typing for recessive diseases for several values of the parameters, and lists the parameters whose values must be improved in order to improve noticeably the error rate. In polar body typing, the probability of error depends on the recombination fraction θ as well as on r, d and c. The probabilities for the polar body errors are found by calculating the probabilities of each of the two errors which are potentially

Table II. Errors in polar body typing for autosomal recessive diseases

Polar body		Oocyte		Use?	Error category	Primary cause(s) of error
True genotype	Observed genotype	True genotype	Deduced genotype			
AA	Aa	aa	Aa	No	Acceptable	c
AA	aa	aa	AA	Yes	Potentially unacceptable	d,c
Aa	AA	Aa	aa	No	Acceptable	r
Aa	aa	Aa	AA	Yes	Potentially unacceptable	r
aa	AA	AA	aa	No	Tolerated	d,c
aa	Aa	AA	Aa	No	Tolerated	c

Table III. Conditional probabilities of unacceptable errors in polar body typing for autosomal recessive diseases

	Primary cause(s) of error	(1) $r = 0.8$ $d = 0.8$ $c = 0.5$	(2) $r = 0.8$ $d = 1$ $c = 0$	(3) $r = 0.9$ $d = 0.8$ $c = 0.5$	(4) $r = 0.9$ $d = 0.9$ $c = 0.5$	(5) $r = 0.9$ $d = 1$ $c = 0.5$	(6) $r = 0.9$ $d = 1$ $c = 0$	(7) $r = 1$ $d = 0.8$ $c = 0.5$	(8) $r = 1$ $d = 1$ $c = 0$
Polar body, aa observed, $\theta = 0$	d,c	0.0038	0	0.0033	0.0016	0.00013	0	0.0032	0
Polar body, aa observed, $\theta = 0.10$	r	0.023	0.019	0.015	0.013	0.011	0.011	0.0039	0
Polar body, aa observed, $\theta = 0.20$	r	0.050	0.045	0.031	0.029	0.027	0.027	0.0052	0
Polar body, aa observed, $\theta = 0.25$	r	0.067	0.063	0.043	0.041	0.039	0.038	0.0063	0
Polar body, aa observed, $\theta = 0.30$	r	0.088	0.083	0.059	0.056	0.054	0.054	0.0078	0
Polar body, aa observed, $\theta = 0.40$	r	0.15	0.14	0.11	0.11	0.11	0.11	0.015	0

Table IV. Number of oocytes or embryos needed to have 95% probability of finding required number both usable and typed usable: autosomal recessive diseases

	$r = 0.8$, $d = 0.8$, $c = 0.05$ Number required			$r = 0.9$, $d = 0.9$, $c = 0.05$ Number required			$r = 1$, $d = 1$, $c = 0$ Number required		
	1	2	3	1	2	3	1	2	3
Blastomere	5	8	10	4	6	8	3	5	6
Polar body, $\theta = 0$	7	11	15	6	9	13	5	8	11
Polar body, $\theta = 0.10$	9	14	19	7	12	16	6	10	14
Polar body, $\theta = 0.20$	12	19	26	10	17	22	9	14	19
Polar body, $\theta = 0.25$	15	24	32	13	20	27	11	18	23
Polar body, $\theta = 0.30$	19	30	40	16	26	34	14	22	30
Polar body, $\theta = 0.40$	39	61	82	33	53	71	29	46	61

unacceptable (Table II), multiplying each by its probability of leading to the implantation of an unacceptable embryo (50 and 25%), and summing the results. Table III shows that the greater the value of θ, the greater the frequency of unacceptable error. This is due to the fact that when the recombination fraction is large, a large proportion of oocytes do not have the usable genotype. Thus it will more often happen that a non-usable oocyte will mistakenly be typed as being usable, and less often happen that a usable oocyte will be correctly typed. It follows that when an oocyte is typed as being usable, it is more likely to be the result of a typing error.

Table III shows that the primary cause of error in polar body typing is PCR inefficiency, except when the recombination fraction is quite small, in which case contamination and failure to place a cell in the reaction tube are the primary causes. The reason for this is as follows. As shown in Table II, potentially unacceptable results in polar body typing for recessive diseases come about when an oocyte of genotype aa or Aa is mistakenly typed as having genotype AA. In general, given reasonable values of r, d and c, it is much more common for an oocyte of genotype Aa to be mistyped as AA than for an oocyte of genotype aa to be so mistyped. Thus most unacceptable errors result from the use of an oocyte of genotype Aa. When the recombination fraction is quite small, however, very few oocytes of genotype Aa exist, so most unacceptable errors result from use of oocytes with genotype aa. As shown in the second and fourth rows of Table II, the frequency of potentially unacceptable errors involving mistyping oocytes of genotype aa is determined primarily by the values of d and c, while the frequency of unacceptable errors involving mistyping oocytes of genotype Aa is determined primarily by the value of r. Thus the values of d and c exert primary influence over the frequency of unacceptable errors when the recombination fraction is very small, while the value of r is most important otherwise.

Comparing blastomere typing with polar body typing, autosomal recessive disease

Comparing the probabilities in the last two rows of Table I with the probabilities in Table III shows that for recessive diseases,

Table V. Errors in blastomere typing for autosomal dominant diseases

True genotype	Observed genotype	Use?	Error category	Primary cause(s) of error	Conditional probability of error							
					(1) $r = 0.8$ $d = 0.8$ $c = 0.05$	(2) $r = 0.8$ $d = 1$ $c = 0$	(3) $r = 0.9$ $d = 0.8$ $c = 0.05$	(4) $r = 0.9$ $d = 0.9$ $c = 0.05$	(5) $r = 0.9$ $d = 1$ $c = 0.05$	(6) $r = 0.9$ $d = 1$ $c = 0$	(7) $r = 1$ $d = 0.8$ $c = 0.05$	(8) $r = 1$ $d = 1$ $c = 0$
Aa	aa	Yes	Unacceptable	r	0.15	0.14	0.088	0.086	0.084	0.083	0.0063	0
aa	Aa	No	Tolerated	c	0.037	0	0.033	0.029	0.027	0	0.030	0

Table VI. Errors in polar body typing for autosomal dominant diseases

Polar body		Oocyte		Use?	Error category	Primary cause(s) of error
True genotype	Observed genotype	True genotype	Deduced genotype			
AA	Aa	aa	Aa	No	Tolerated	c
AA	aa	aa	AA	No	Tolerated	d,c
Aa	AA	Aa	aa	Yes	Potentially unacceptable	r
Aa	aa	Aa	AA	No	Acceptable	r
aa	AA	AA	aa	Yes	Unacceptable	d,c
aa	Aa	AA	Aa	No	Acceptable	c

and the same values for r, d and c, unacceptable results occur less frequently with blastomere typing than with polar body typing unless the recombination fraction is quite small, or unless both PCR efficiency and the contamination rate are rather high.

Two factors are at work here, one favouring blastomere typing and one favouring polar body typing. Favouring blastomere typing is the fact that 75% of all embryos are usable, while the percentage of oocytes which are usable is 50% or less, depending on the recombination fraction between the centromere and the locus of interest. The factor favouring polar body typing is that selection of an oocyte which should not be used does not always result in an embryo of an unacceptable genotype. Fertilizing an oocyte of genotype Aa will yield an unacceptable embryo only 25% of the time, and selecting an oocyte of genotype aa will yield an unacceptable embryo 50% of the time.

When the recombination fraction is small, the percentage of usable oocytes is nearly 50%, and the two factors combine to provide a lower frequency of error for polar body typing. When the recombination fraction is larger than ~ 0.1, and both the PCR efficiency r and the contamination rate c are both not high, blastomere typing has a lower error rate. If r and c are both large, the error rate for polar body typing is lower. This is because large values of r tend to decrease the frequency of polar body typing errors, and large values of c tend to increase the frequency of blastomere typing errors. Thus typing procedures with high levels of PCR efficiency and high contamination rates may be more accurate with polar body typing, while under other conditions blastomere typing might be more accurate.

Number of oocytes or embryos needed, autosomal recessive disease

Table IV gives the number of oocytes or embryos which must be typed in order to have at least 95% probability of finding 1, 2 or 3 which are usable and are typed as being usable. For example, when $r = 0.9$, $d = 0.9$ and $c = 0.05$, if six blastomeres are typed, the probability is at least 95% that at least two of the corresponding embryos will in fact be usable (i.e. of genotype AA or Aa) and will be typed as such. On the other hand, when the recombination fraction is 0.1, 12 oocytes must be typed to be equally confident of finding two usable ones typed as such. The reason that fewer blastomeres need to be typed is that 75% of oocytes are usable, while only 50% or fewer of oocytes are. When θ is large, the proportion of oocytes which are usable is quite small. As shown in Table IV, one cannot be confident of finding a usable oocyte unless a very large number are available. This is a clear advantage for blastomere typing.

Blastomere typing, autosomal dominant disease

In the case of dominant disease, only embryos of genotype aa are usable. We assume that one parent is of genotype Aa, and the other is of genotype aa. An embryo has 50% probability of being of genotype aa and 50% probability of being of genotype Aa. Table V categorizes the two typing errors that are possible in this situation. The probabilities of error are noticeably greater than in the case of recessive disease, unless the PCR efficiency r is very high. This is partly because only 50% of embryos are usable, compared with 75% when the disease is recessive. Also, unacceptable errors when the disease is dominant are usually due to lack of PCR efficiency, while unacceptable errors when the disease is recessive usually involve contamination (cf. Tables I and V). Unless the PCR efficiency is very high, efficiency errors are more common than contamination errors.

Polar body typing, autosomal dominant disease

In the case of dominant disease, only oocytes of genotype aa are usable. The proportion of oocytes which are usable is $\frac{1}{2} - \theta$. We assume that the female is heterozygous and that the male is homozygous aa. Table VI categorizes the various typing errors in this case. Line 3 of Table VI indicates that using an oocyte of genotype Aa is 'potentially unacceptable'. This is because the

Table VII. Conditional probabilities of unacceptable errors in polar body typing for autosomal dominant diseases

	Primary cause(s) of error	(1) $r = 0.8$ $d = 1$ $c = 0.5$	(2) $r = 0.8$ $d = 1$ $c = 0$	(3) $r = 0.9$ $d = 0.8$ $c = 0.5$	(4) $r = 0.9$ $d = 0.9$ $c = 0.5$	(5) $r = 0.9$ $d = 1$ $c = 0.5$	(6) $r = 0.9$ $d = 1$ $c = 0$	(7) $r = 1$ $d = 0.8$ $c = 0.5$	(8) $r = 1$ $d = 1$ $c = 0$
Polar body, AA observed, $\theta = 0$	d,c	0.0076	0	0.0066	0.031	0.00025	0	0.0063	0
Polar body, AA observed, $\theta = 0.10$	r	0.047	0.038	0.029	0.025	0.022	0.022	0.0079	0
Polar body, AA observed, $\theta = 0.20$	r	0.099	0.091	0.063	0.058	0.054	0.054	0.010	0
Polar body, AA observed, $\theta = 0.25$	r	0.013	0.13	0.086	0.081	0.077	0.77	0.013	0
Polar body, AA observed, $\theta = 0.30$	r	0.18	0.17	0.12	0.11	0.11	0.11	0.016	0
Polar body, AA observed, $\theta = 0.40$	r	0.29	0.29	0.22	0.22	0.21	0.21	0.030	0

Table VIII. Number of oocytes or embryos needed to have 95% probability of finding required number both usable and typed usable: autosomal dominant diseases

	$r = 0.8$, $d = 0.8$, $c = 0.05$ Number required			$r = 0.9$, $d = 0.9$, $d = 0.05$ Number required			$r = 1$, $d = 1$, $c = 0$ Number required		
	1	2	3	1	2	3	1	2	3
Blastomere	7	11	15	6	9	13	5	8	11
Polar body, $\theta = 0$	7	11	15	6	9	13	5	8	11
Polar body, $\theta = 0.10$	9	14	19	7	12	16	6	10	14
Polar body, $\theta = 0.20$	12	19	26	10	17	22	9	14	19
Polar body, $\theta = 0.25$	15	24	32	13	20	27	11	18	23
Polar body, $\theta = 0.30$	19	30	40	16	26	34	14	22	30
Polar body, $\theta = 0.40$	39	61	82	33	53	71	29	46	61

Table IX. Conditional probabilities of unacceptable errors when two or three blastomeres are typed in one tube: autosomal diseases

	(1) $r = 0.8$ $d = 0.8$ $c = 0.05$	(2) $r = 0.8$ $d = 1$ $c = 0$	(3) $r = 0.9$ $d = 0.8$ $c = 0.05$	(4) $r = 0.9$ $d = 0.9$ $c = 0.05$	(5) $r = 0.9$ $d = 1$ $c = 0.05$	(6) $r = 0.9$ $d = 1$ $c = 0$	(7) $r = 1$ $d = 0.8$ $c = 0.05$	(8) $r = 1$ $d = 1$ $c = 0$
Recessive, AA observed, 2 cells	0.0013	0	0.0011	0.00029	0.0000025	0	0.0011	0
Recessive, Aa observed, 2 cells	0.014	0	0.013	0.013	0.012	0	0.012	0
Dominant, aa observed, 2 cells	0.075	0.037	0.036	0.024	0.0098	0.0098	0.0011	0
Recessive, AA observed, 3 cells	0.00030	0	0.00023	0.00033	0.00000003	0	0.00021	0
Recessive, Aa observed, 3 cells	0.013	0	0.013	0.012	0.012	0	0.012	0
Dominant, aa observed, 3 cells	0.034	0.0079	0.013	0.0055	0.0010	0.0010	0.00021	0

use of such an oocyte yields the unacceptable result of an embryo carrying the A allele with 50% probability. Line 5 of the table indicates that using an oocyte of genotype AA is 'unacceptable'. Use of such an oocyte always results in an embryo of unacceptable genotype.

Table VII gives conditional probabilities of unacceptable errors. The errors are exactly twice as large as in the recessive case. This is because while the proportion of oocytes which are usable, $\frac{1}{2}-\theta$, is the same for both dominant and recessive diseases, the frequencies with which oocytes with non-usable genotypes result in embryos with unacceptable genotypes are twice as great when the disease is dominant, i.e. 50 and 100% versus 25 and 50%.

Comparing blastomere typing with polar body typing, autosomal dominant disease

The two factors affecting the relationship between blastomere typing errors and polar body typing errors when the disease is recessive are relevant when the disease is dominant, but they are less pronounced. The proportion of embryos which are usable is 50%, while the proportion of oocytes which are usable is always $<50\%$. As mentioned above, selection of a non-usable oocyte results in an embryo of unacceptable genotype twice as

Table X. Number of embryos needed to have 95% probability of finding required number both usable and typed usable when two or three blastomeres are typed in one tube: autosomal diseases

	$r = 0.8$ $d = 0.8$ $c = 0.05$ No. required			$r = 0.9$ $d = 0.9$ $c = 0.05$ No. required			$r = 1$ $d = 1$ $c = 0$ No. required		
	1	2	3	1	2	3	1	2	3
Recessive, 2 cells	3	5	7	3	5	6	3	5	6
Dominant, 2 cells	5	9	12	5	8	11	5	8	11
Recessive, 3 cells	3	5	7	3	5	6	3	5	5
Dominant, 3 cells	5	8	11	5	8	11	5	8	11

Table XI. Conditional probabilities of unacceptable errors when two or three blastomeres are typed in separate tubes: autosomal diseases

	(1) $r = 0.8$ $d = 0.8$ $c = 0.05$	(2) $r = 0.8$ $d = 1$ $c = 0$	(3) $r = 0.9$ $d = 0.8$ $c = 0.05$	(4) $r = 0.9$ $d = 0.9$ $c = 0.05$	(5) $r = 0.9$ $d = 1$ $c = 0.05$	(6) $r = 0.9$ $d = 1$ $c = 0$	(7) $r = 1$ $d = 0.8$ $c = 0.05$	(8) $r = 1$ $d = 1$ $c = 0$
Recessive, A observed, 2 cells	0.0049	0	0.0041	0.0022	0.00041	0	0.0035	0
Dominant, A not observed, 2 cells	0.076	0.037	0.037	0.024	0.0098	0.0098	0.0021	0
Recessive, A observed, 3 cells	0.0018	0	0.0013	0.00040	0.000016	0	0.0010	0
Dominant, A not observed, 3 cells	0.034	0.0079	0.014	0.0056	0.0010	0.0010	0.00060	0

often as it does when the disease is recessive, but still with $< 100\%$ frequency. Comparing the probabilities in the first row of Table V with those in Table VII shows that blastomere typing errors are less frequent than polar body typing errors if the PCR efficiency is very high or if the recombination fraction θ is greater than ~ 0.25. Otherwise, polar body typing errors are less frequent.

Number of oocytes and embryos needed, autosomal dominant disease

Table VIII gives the number of oocytes or embryos which must be typed in order to have at least 95% probability of finding 1, 2 or 3 which are usable and are typed as being usable. The number of oocytes needed is the same whether the disease is dominant or recessive (cf. Table IV), because the proportion of usable oocytes is the same, $\frac{1}{2} - \theta$ in each case. When $\theta = 0$, the number of oocytes needed is the same as the number of embryos, because the proportion of oocytes which are usable, 50%, is equal to the proportion of embryos which are usable. When $\theta > 0$, the proportion of oocytes which are usable is $< 50\%$, so on average, a greater number of oocytes than embryos will need to be typed to find a given number which are usable and typed as such.

Typing two or three blastomere cells in one reaction tube, autosomal disease

In polar body analysis, only one diploid cell can be studied. Blastomere analysis is potentially capable of using more than one cell. If two or more blastomeres are available for typing, greater accuracy can be achieved. We make assumptions (i–iii) as before, and assume in addition that the cells are independent of each other with regard to being placed in the tube.

Conditional probabilities of error when two or three cells are typed in one reaction tube are given in Table IX. The frequency of error is greatly reduced when more than one cell is typed. Comparing Table IX with Tables III and VII shows that for either autosomal dominant or autosomal recessive disease, typing two blastomere cells yields fewer unacceptable errors than polar body typing when $\theta > 0.1$, except in some cases where the value of d is quite low. When the disease is autosomal recessive, and the cells are typed AA, the probability of unacceptable error is always less than for polar body typing regardless of the recombination fraction. Typing three cells reduces the error rate still further.

Comparing Table X with Tables IV and VIII shows that typing two or three cells tends to reduce the number of embryos which will need to be typed before finding a given number which are usable and typed as such.

Table XII. Number of embryos needed to have 95% probability of finding required number both usable and typed usable when two or three blastomeres are typed in separate tubes: autosomal diseases

	$r = 0.8$ $d = 0.8$ $c = 0.05$ No. required			$r = 0.9$ $d = 0.9$ $c = 0.05$ No. required			$r = 1$ $d = 1$ $c = 0$ No. required		
	1	2	3	1	2	3	1	2	3
Recessive, 2 cells	4	6	9	3	5	7	3	5	6
Dominant, 2 cells	5	9	12	5	9	12	5	8	11
Recessive, 3 cells	4	7	9	3	6	8	3	5	6
Dominant, 3 cells	5	9	12	5	9	12	5	8	11

Typing two or three blastomere cells in separate reaction tubes, autosomal disease

An alternative to the procedure discussed above is to type the individual cells from the same embryo in separate tubes. In this case, we assume that the reactions are independent of each other, and that assumptions (i–iii) hold for each reaction. In particular, this implies that the contamination of any one reaction neither increases nor decreases the likelihood that another reaction is contaminated. When two or three cells are typed in separate tubes, it will often be the case that the reactions give conflicting evidence about the genotype of the embryo. This difficulty can be dealt with by noticing that the decision whether to implant an embryo does not require complete knowledge of its genotype, but only whether it carries the dominant A allele. When the disease is autosomal recessive, embryos carrying the A allele are the usable ones. Thus an embryo may be typed as usable if each tube gives a signal for the A allele. When the disease is autosomal dominant, embryos not carrying the A allele are the usable ones, so an embryo may be typed if no tube gives a signal for the A allele, and at least one tube gives a signal for the a allele.

Table XI gives conditional probabilities of unacceptable errors when two or three blastomere cells are typed in separate tubes. When the disease is autosomal dominant, the frequency of error is about the same as for the one tube procedure. When the disease is autosomal recessive, the probability of error is less with the separate tubes procedure than with the one tube procedure when the one tube procedure types a blastomere as Aa, but greater than when the one tube procedure types a blastomere as AA. Since Aa cells occur twice as often as AA cells, on balance the separate tubes procedure will result in unacceptable results less often than the one tube procedure.

Table XII gives the number of embryos needed in order to have at least 95% probability of finding 1, 2 or 3 which are usable

Table XIII. Conditional probabilities of unacceptable errors in blastomere typing for X-linked diseases

	Primary cause(s) of error	(1) $r = 0.8$ $d = 0.8$ $c = 0.5$	(2) $r = 0.8$ $d = 1$ $c = 0$	(3) $r = 0.9$ $d = 0.8$ $c = 0.5$	(4) $r = 0.9$ $d = 0.9$ $c = 0.5$	(5) $r = 0.9$ $d = 1$ $c = 0.5$	(6) $r = 0.9$ $d = 1$ $c = 0$	(7) $r = 1$ $d = 0.8$ $c = 0.5$	(8) $r = 1$ $d = 1$ $c = 0$
Recessive	r,d	0.13	0.083	0.11	0.080	0.045	0.045	0.083	0
Recessive, X amplified	r	0.074	0.071	0.044	0.043	0.042	0.042	0.002	0
Dominant	c	0.037	0	0.033	0.030	0.027	0	0.030	0
Dominant, X amplified	c	0.036	0	0.029	0.029	0.029	0	0.024	0

and are typed as such when the separate tube procedure is used. In some cases the values are larger by one or two than with the one tube procedure.

In summary, when more than one blastomere cell is available for typing, the frequency of unacceptable error is reduced. Typing more than one blastomere cell results in a lower frequency of error than polar body typing except in some cases in which the recombination fraction is quite small. When the disease is autosomal dominant, the one tube procedure and the separate tubes procedure are about equally effective. When the disease is autosomal recessive, the separate tubes procedure results in fewer errors at the cost of perhaps having to type one or two more embryos to find a given number which are usable and typed as such.

Combining polar body and blastomere analysis

By combining these two typing methods, the frequency of unacceptable errors can be reduced dramatically. The procedure for combining them is as follows. Each available oocyte is typed by polar body analysis, then those typed as usable are fertilized. The resulting embryos are then typed using the blastomere method, and only those typed as being usable are considered for implantation. The effect of this combined procedure is that unacceptable errors can occur only when both typing methods fail. To illustrate the accuracy of this combined procedure, we have calculated that when the combined procedure is used with $r = 0.8$, $d = 0.8$, $c = 0.05$ and $\theta = 0.2$, the probability of an unacceptable error when the disease is recessive is 0.071% when the blastomere is typed AA, and 0.35% when the blastomere is typed Aa. In contrast, Table I shows that with blastomere typing alone, the corresponding probabilities are 0.57% and 1.8%. Table III shows that with polar body typing alone, the probability is 5.0%. In the case of dominant disease, the improvement is even greater. Using the parameter values above, we have calculated that the probability of unacceptable error is 1.9% using the combined procedure. Table V shows that the probability when blastomere typing alone is used is 15%, and Table VII shows that when polar body typing alone is used, the probability is 9.9%.

Recessive X-linked disease

We consider the case when the disease gene is recessive and located on the X chromosome, the mother carries the disease gene on exactly one of her X chromosomes, and the father does not have the disease. In this case, blastomere typing can be used to determine the sex of the embryo and PCR of the disease-causing gene itself is not required. None of the female embryos will have the disease, but 50% of the males will. Thus only female embryos are considered for implantation. In this typing procedure, a DNA sequence on the Y chromosome is amplified. For this purpose, Y-specific repeated genes have been used (Handyside, 1989, 1990; Bradbury *et al.*, 1990) to enhance the efficiency, but single copy genes could be used as well. If no PCR product is detected, the embryo is typed XX and considered suitable for implantation. An unacceptable error occurs if an embryo typed XX is in fact XY, with the X chromosome carrying the disease. It turns out that if an embryo typed XX is implanted, the conditional probability of unacceptable error is:

$$P(\text{unacceptable error}) = \frac{1-rd}{4-2rd}$$

(See the Appendix for details.)

Table XIV. Number of embryos needed to have 95% probability of finding required number both usable and typed usable: X-linked diseases

	$r = 0.8$ $d = 0.8$ $c = 0.05$ No. required			$r = 0.9$ $d = 0.9$ $c = 0.05$ No. required			$r = 1$ $d = 1$ $c = 0$ No. required		
	1	2	3	1	2	3	1	2	3
Recessive	4	7	9	4	7	10	5	8	11
Dominant	8	13	18	6	10	14	5	8	11

The first line of Table XIII gives the conditional probability of unacceptable error for selected values of r, d and c. The values depend only on r and d, and not on c.

If the X chromosome can be identified by amplification in the same reaction and with the same efficiency as the Y chromosome, the chance of error can be greatly reduced. Either repeated X and Y sequences or single copy X and Y genes (but not one of each) could be analysed. In this situation, an embryo would only be considered for implantation if the typing procedure yielded a positive signal for the X chromosome and no signal for the Y chromosome. This situation is very much like the case of blastomere typing when the disease is autosomal dominant, with the Y chromosome assuming the role of the dominant (A) allele, and the X chromosome assuming the role of the recessive (a) allele. The only difference between the two situations which needs to be taken into account is that if an XY embryo is mistakenly typed as XX, the probability of an unacceptable error is not 1 but 0.5 since only one half of the males will be affected. Therefore the conditional probabilities of unacceptable error in

Table XV. Conditional probabilities of unacceptable errors when two or three blastomeres are typed in one tube: X-linked diseases

	(1) $r = 0.8$ $d = 0.8$ $c = 0.05$	(2) $r = 0.8$ $d = 1$ $c = 0$	(3) $r = 0.9$ $d = 0.8$ $c = 0.05$	(4) $r = 0.9$ $d = 0.9$ $c = 0.05$	(5) $r = 0.9$ $d = 1$ $c = 0.05$	(6) $r = 0.9$ $d = 1$ $c = 0$	(7) $r = 1$ $d = 0.8$ $c = 0.05$	(8) $r = 1$ $d = 1$ $c = 0$
Recessive, 2 cells	0.057	0.019	0.036	0.017	0.0050	0.0050	0.019	0
Dominant, 2 cells	0.028	0	0.026	0.025	0.025	0	0.025	0
Recessive, 3 cells	0.022	0.0040	0.011	0.0034	0.00050	0.00050	0.0040	0
Dominant, 3 cells	0.026	0	0.025	0.025	0.024	0	0.025	0

this case are exactly one-half as great as the case of blastomere typing when the disease is autosomal dominant. Comparing the first two lines of Table XIII shows that amplifying an X chromosome marker in addition to the Y marker results in a considerable decrease in the frequency of unacceptable error when d is low. When d is high, the decrease is smaller.

Dominant X-linked disease

We consider the case when the disease gene is dominant and located on the X chromosome. If the mother has one disease gene, 50% of the male embryos and 50% of the female embryos will have the disease, so sex typing by itself is of no use. If the male has the disease gene on his X chromosome, then all of the female embryos and none of the male embryos will have the disease. In this case, embryos for which a Y chromosome is detected are considered suitable for implantation. An unacceptable error occurs if a Y chromosome is detected in an embryo which is in fact of genotype XX. It turns out that if an embryo typed XY is implanted, the conditional probability of unacceptable error is:

$$P(\text{unacceptable error}) = \frac{c}{2c + rd(2-c)}$$

(See the Appendix for details.) The third line of Table XIII gives the conditional probability of unacceptable error for selected values of r, d and c. These probabilities depend on c more than on r and d.

Consider an X-linked dominant disease where an X chromosome marker can be amplified in the same reaction and with the same efficiency as the Y chromosome marker, and only embryos where both X and Y chromosomes are detected are implanted. If an XX cell fails to be deposited in the reaction tube, and if there is Y chromosome DNA contamination, the embryo will not be implanted. This is offset by the fact that XY embryos where the X chromosome is not detected because of lack of efficiency will not be implanted, although they could be. Comparing the third and fourth lines of Table XIII shows that in general there does not seem to be much gain by amplifying an X chromosome marker when the disease is dominant. Details of the calculations are given in the Appendix.

Number of embryos needed, X-linked disease

Table XIV gives the number of embryos which must be typed in order to have 95% probability of finding 1, 2 or 3 which are usable and typed as such for X-linked diseases. When the disease is recessive, these numbers depend only on c, and are about the same for all reasonable values of that parameter. When the disease is dominant, the number of embryos which must be typed depends on r, d and c, and is much higher than when the disease is recessive.

Table XVI. Number of embryos needed to have 95% probability of finding required number both usable and typed usable when two or three blastomeres are typed in one tube: X-linked diseases

	$r = 0.8$ $d = 0.8$ $c = 0.05$ No. required			$r = 0.9$ $d = 0.9$ $c = 0.05$ No. required			$r = 1$ $d = 1$ $c = 0$ No. required		
	1	2	3	1	2	3	1	2	3
Recessive, 2 cells	5	8	10	5	8	11	5	8	11
Dominant, 2 cells	6	9	13	5	8	11	5	8	11
Recessive, 3 cells	5	8	11	5	8	11	5	8	11
Dominant, 3 cells	5	8	11	5	8	11	5	8	11

Typing two or three cells in one reaction tube, X-linked disease

Table XV gives conditional probabilities of unacceptable errors for X-linked diseases when two or three cells are typed in the same tube. When the disease is recessive, the error rate decreases significantly as the number of available cells increase. This is because an unacceptable error results when the embryo has a Y chromosome which is not detected. When several cells are typed, the chance that none of the Y chromosomes will be detected is lower. When the disease is dominant, the improvement in the error rate from typing additional cells is negligible. This is because unacceptable errors result when the reaction is contaminated with DNA from a Y chromosome. Typing two or three cells rather than one cell does not, under our assumptions, noticeably alter the probability that the typing results will be affected by contamination.

Table XVI gives the number of embryos needed to have at least 95% probability of finding 1, 2 or 3 which are usable and are typed as such when two or three cells are typed in one tube. Notice that when the disease is recessive, increasing the values of r and d can sometimes result in a slight increase in the number of embryos which need to be typed. This seemingly anomalous result is due to the fact that even though 50% of the male embryos are usable, all are discarded except those for which typing errors are made. When the typing procedure is less accurate, more of these embryos will be typed as being usable, so fewer embryos may need to be typed before finding a usable one.

Typing two or three cells in separate tubes, X-linked disease

We also consider the possibility of typing two or three cells in separate tubes. When the disease is recessive, we type an embryo

Table XVII. Conditional probabilities of unacceptable errors when two or three blastomeres are typed in separate tubes: X-linked diseases

	(1) $r = 0.8$ $d = 0.8$ $c = 0.05$	(2) $r = 0.8$ $d = 1$ $c = 0$	(3) $r = 0.9$ $d = 0.8$ $c = 0.05$	(4) $r = 0.9$ $d = 0.9$ $c = 0.05$	(5) $r = 0.9$ $d = 1$ $c = 0.05$	(6) $r = 0.9$ $d = 1$ $c = 0$	(7) $r = 1$ $d = 0.8$ $c = 0.05$	(8) $r = 1$ $d = 1$ $c = 0$
Recessive, 2 cells	0.057	0.019	0.036	0.017	0.0050	0.0050	0.019	0
Dominant, 2 cells	0.015	0	0.0012	0.00094	0.00077	0	0.00096	0
Recessive, 3 cells	0.022	0.0040	0.011	0.0034	0.00050	0.00050	0.0040	0
Dominant, 3 cells	0.0026	0	0.0023	0.0020	0.0019	0	0.0020	0

as being usable if no Y chromosomes are detected. When the disease is dominant, and two cells are typed, we type an embryo as being usable if Y chromosomes are detected in both tubes. When three cells are typed, we type an embryo as being usable if Y chromosomes are detected in at least two of the three tubes.

Table XVII gives conditional probabilities of unacceptable errors in this case. When the disease is recessive, the error rates are the same as for the one tube procedure, because the probability of failing to detect two or three Y chromosomes is the same under our assumptions whether the chromosomes are in the same tube or in separate tubes. When the disease is dominant, error rates are much lower than in the one tube procedure, because more than one tube must be contaminated for an unacceptable error to occur.

Table XVIII gives the number of embryos needed to have at least 95% probability of finding 1, 2 or 3 which are usable and are typed as such when two or three cells are typed in separate tubes. When the disease is recessive, the numbers are about the same as for the one tube procedure. When the disease is dominant, the numbers are much higher when two cells are typed, or when the values of r and d are low. This is because it will happen fairly often that an embryo will fail to be typed due to conflicting results among the tube.

Alternative contamination assumptions

In the case of autosomal diseases, the effect of contamination on the frequency of unacceptable error is determined almost entirely by the contamination rate of the non-disease allele in the case of blastomere typing, because spurious detection of a disease allele will not lead to a diseased embryo being typed as disease-free. By the same reasoning, the frequency of unacceptable error in polar body typing is determined by the contamination rate of the disease allele. If the predominant cause of contamination is diploid cellular contamination in which two alleles are co-contaminants, a contamination rate of $c/2$ would produce essentially the same result as a contamination rate of c under our assumptions. In the case of X-linked diseases, the same reasoning applies, with the rate of contamination from an X chromosome the important rate in the recessive case, and the rate from a Y chromosome important in the dominant case. The contamination rate does not matter in the X-linked recessive case unless the amount of contamination is great enough to prevent detection of the target, a condition which is discussed below.

If the number of contaminating molecules is much greater than the number of target molecules, preventing detection of the target, then when the disease is autosomal dominant and blastomere typing is used, contamination from the non-disease allele will cause a diseased embryo to be typed as disease-free, leading to an unacceptable error. In the same way, a large amount of contamination from the disease allele can lead to an unacceptable error when polar body typing is used. The frequency of unacceptable error will increase by an amount approximately equal to the frequency with which a large amount of contamination from a given allele occurs. In the X-linked recessive case, a large amount of contamination from the X chromosome will cause an XY embryo to be typed XX. Since 50% of XY embryos are diseased, this leads to a 50% chance of unacceptable error. Large amounts of contamination thus have somewhat less effect in the X-linked recessive case than in the autosomal dominant case. The frequency of unacceptable error will increase, but by an amount much less than the frequency with which a large degree of contamination from the X chromosome occurs. When contamination occurs in the autosomal recessive or X-linked dominant case, the amount of X chromosome or disease allele contaminant has little effect on the frequency of unacceptable error. In these situations, spurious detection of a non-disease allele or Y chromosome will cause an unacceptable error, whether or not the target molecules are also detected.

Table XVIII. Number of embryos needed to have 95% probability of finding required number both usable and typed usable when two or three blastomeres are typed in separate tubes: X-linked diseases

	$r = 0.8$ $d = 0.8$ $c = 0.05$ No. required			$r = 0.9$ $d = 0.9$ $c = 0.05$ No. required			$r = 1$ $d = 1$ $c = 0$ No. required		
	1	2	3	1	2	3	1	2	3
Recessive, 2 cells	5	8	11	5	8	11	5	8	11
Dominant, 2 cells	14	22	29	8	13	17	5	8	11
Recessive, 3 cells	5	8	11	5	9	12	5	8	11
Dominant, 3 cells	7	12	16	5	9	12	5	8	11

Conclusion

Blastomere and polar body DNA typing by PCR for preimplantation diagnosis of genetic disease has been considered for autosomal dominant and recessive genes. We find that blastomere analysis is the most generally applicable method. Under most circumstances, it results in lower levels of unacceptable errors when compared to polar body analysis, given the same frequencies of PCR efficiency, contamination and cell transfer errors. Polar body analysis yields lower rates of mis-diagnosis than blastomere analysis only in special circumstances, for

example when the disease-causing gene is recessive and is located virtually at the centromere or when the disease-causing gene is dominant and lies within 25 cM of the centromere. Since it is unlikely that many disease-causing genes are located so close to the centromere of a chromosome, polar body typing is less generally applicable. We also show that the levels of unacceptable errors can be significantly decreased by typing more than one blastomere from a single embryo or by combining polar body analysis with blastomere typing. If such multiple procedures are consistent with normal implantation and development, they are considered the method of choice.

We also considered several cases of preimplantation diagnosis for X-linked diseases based on the analysis of Y-specific genes. We examined the consequences of (i) matings between normal males and females heterozygous for a recessive disease gene and (ii) matings between normal females and males with a dominant disease. In the recessive case, we show a significant improvement if the blastomere can be typed for both an X-specific and Y-specific marker rather than the Y-marker alone. We did not analyse any of the possible matings which could be studied by PCR of the X-linked disease-causing locus itself. The errors in such cases would be similar but not identical to those for autosomal dominant and recessive genes, and slightly different calculations would be needed.

Mistakes in typing single cell DNA are a function of errors in PCR efficiency, cell transfer and contamination. For our calculations we used some realistic values which came from experimental data (Cui *et al.*, 1989; H.Li, X.Cui, R.Hubert and N.Arnheim, unpublished data). It seems clear that groups wishing to carry out preimplantation diagnosis should make an attempt to estimate the magnitude of the errors (r, c and d) using a single cell system in their own laboratories. Such a system could involve the use of single sperm, oocytes or tissue culture cells, using appropriate mathematical models for analysing the data. [For single sperm data analysis see Cui *et al.* (1989).] In this way, realistic values can be determined for the risk that a fetus originally diagnosed as a disease-free embryo might be found to have the disease, thereby warranting termination of the pregnancy. It is self-evident that in its current state, preimplantation analysis should be confirmed by CVS or amniocentesis.

Acknowledgements

This research was partially supported by NSF grant DMS90-05833 (W.N.), a grant from the University of Southern California Faculty Research and Innovation Fund (W.N.) and NIH grants GM36745 (N.A.) and GM42942 (N.A.).

References

Arnheim,N. (1990) The polymerase chain reaction. In Setlow,J. (ed.), *Genetic Engineering, Principles and Methods*. Vol. 12. Plenum Press, New York, pp. 115–138.

Arnheim,N., White,T. and Rainey,W. (1990) Application of the polymerase chain reaction to organismal and population biology. *Bioscience*, **40**, 174–182.

Bradbury,M.W., Isola,L.M. and Gordon,J.W. (1990) Enzymatic amplification of a Y chromosome repeat in a single blastomere allows identification of the sex of preimplantation mouse embryos. *Proc. Natl. Acad. Sci. USA*, **87**, 4053–4057.

Coutelle,C., Williams,C., Handyside,A., Hardy,K., Winston,R. and Williamson,R. (1989) Genetic analysis from DNA from single human oocytes: a model for preimplantation diagnosis of cystic fibrosis. *Br. Med. J.*, **299**, 22–24.

Cui,X., Li,H., Goradia,T.M., Lange,K., Kazazian,H.H., Galas,D.J. and Arnheim,N. (1989) Single sperm typing: determination of genetic distance between the G-gamma globin and parathyroid hormone loci. *Proc. Natl. Acad. Sci. USA*, **86**, 9389–9393.

Erlich,H. (1989) (ed.) *PCR Technology: Principles and Applications for DNA Amplification*. Stockton Press, New York.

Erlich,H., Gelfaud,D. and Sminsky,J. (1991) Recent advances in the polymerase chain reaction. *Science*, **252**, 1643–1651.

Feller,W. (ed.) (1970) *An Introduction to Probability Theory and its Applications*. Vol. I, 3rd edn. Wiley, New York.

Gomez,C.M., Muggleton-Harris,A.L., Whittingham,D.G., Hood,L.E. and Readhead,C. (1990) Rapid preimplantation detection of mutant (shiverer) and normal alleles of the mouse myelin basic protein gene allowing selective implantation and birth of live young. *Proc. Natl. Acad. Sci. USA*, **87**, 4481–4484.

Handyside,A.H., Pattinson,J.K., Penketh,R.J.A., Delhanty,J.D.A., Winston,R.M.L. and Tuddenham,E.G.D. (1989) Biopsy of human preimplantation embryos and sexing by DNA amplification. *Lancet*, **i**, 347–349.

Handyside,A.H., Kontogianni,E.H., Hardy,K. and Winston,R.M.L. (1990) Pregnancies from biopsied human preimplantation embryos sexed by Y-specific DNA amplification. *Nature*, **344**, 768–770.

Holding,C. and Monk,M. (1989) Diagnosis of beta-thalassemia by DNA amplification in single blastomeres from mouse preimplantation embryos. *Lancet*, **ii**, 532–535.

Innis,M.A., Gelfand,D.H., Sninsky,J.J. and White,T.J. (eds) (1990) *PCR Procotols. A Guide to Methods and Applications*. Academic Press, New York.

Kazazian,H. (1989) Use of PCR in the diagnosis of monogenic diseases. In Erlich,H. (ed.), *PCR Technology: Principles and Applications for DNA Amplification*. Stockton Press, New York, pp. 153–169.

Jeffreys,A.J., Wilson,V., Neumann,R. and Keyte,J. (1988) Amplification of human minisatellites by the polymerase chain reaction: towards DNA fingerprinting of single cells. *Nucleic Acids Res.*, **16**, 10953–10971.

Li,H., Gyllensten,U., Cui,X., Saiki,R., Erlich,H. and Arnheim,N. (1988) Amplification and analysis of DNA sequences in single human sperm. *Nature*, **335**, 414–417.

Li,H., Cui,X. and Arnheim,N. (1990) Direct electrophoretic detection of the allelic state of single DNA molecules in human sperm using PCR. *Proc. Natl. Acad. Sci. USA*, **87**, 4580–4584.

Monk,M. and Holding,C. (1990) Amplification of a β-haemoglobin sequence in individual human oocytes and polar bodies. *Lancet*, **335**, 985–988.

Mullis,K.B. and Faloona,F.A. (1987) Specific synthesis of DNA in vitro via a polymerase catalyzed chain reaction. *Methods Enzymol.*, **155**, 335–351.

Saiki,R., Scharf,S., Faloona,F., Mullis,K., Horn,G., Erlich,H. and Arnheim,N. (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, **230**, 1350–1354.

Saiki,R.K., Gelfand,D.H., Stoffel,S., Scharf,S.J., Higuchi,R., Horn,G.T., Mullis,K.B. and Erlich,H.A. (1988) Primer directed amplification of DNA with a thermostable DNA polymerase. *Science*, **239**, 487–491.

Received on December 2, 1990; accepted on February 27, 1991

Appendix

If we allow the letters x and y to represent arbitrary genotypes, it is straightforward to calculate the probability under assumptions (i–iii) that the typing procedure applied to a polar body or blastomere cell will result in predicted

genotype x given that the true genotype of the cell is y, in terms of the parameters r, d, c and, in the case of polar body analysis, θ.

Let p_0 be the probability that neither allele in the cell is detected, whether through failure to be deposited in the tube or through failure of PCR to amplify sufficiently. Let p_1 be the probability that exactly one of the two alleles is detected, and let p_2 be the probability that both are detected. Then

$$p_0 = d(1 - r)^2 + 1 - d \tag{1}$$

$$p_1 = 2dr(1 - r) \tag{2}$$

$$p_2 = dr^2 \tag{3}$$

The conditional probabilities that a cell is typed x given that its true genotype is y are as follows:

$$P(\text{cell is typed AA} \mid \text{true genotype is AA}) = (p_1 + p_2)(1 - c/2) + p_0c/2 \tag{4}$$

$$P(\text{cell is typed AA} \mid \text{true genotype is Aa}) = (p_1/2)(1 - c/2) + p_0c/2 \tag{5}$$

$$P(\text{cell is typed AA} \mid \text{true genotype is aa}) = p_0c/2 \tag{6}$$

$$P(\text{cell is typed Aa} \mid \text{true genotype is AA}) = (p_1 + p_2)(c/2) \tag{7}$$

$$P(\text{cell is typed Aa} \mid \text{true genotype is Aa}) = p_2 + p_1(c/2) \tag{8}$$

$$P(\text{cell is typed Aa} \mid \text{true genotype is aa}) = (p_1 + p_2)(c/2) \tag{9}$$

$$P(\text{cell is typed aa} \mid \text{true genotype is AA}) = p_0c/2 \tag{10}$$

$$P(\text{cell is typed aa} \mid \text{true genotype is Aa}) = (p_1/2)(1 - c/2) + p_0c/2 \tag{11}$$

$$P(\text{cell is typed aa} \mid \text{true genotype is aa}) = (p_1 + p_2)(1 - c/2) + p_0c/2 \tag{12}$$

In addition there is a small probability that the cell will remain untyped through a failure to detect any alleles. This quantity is not needed in our calculations.

The probability that a blastomere or polar body has any given genotype is known. Denote by p_{AA} the probability that a blastomere cell or polar body has genotype AA, and define p_{Aa} and p_{aa} similarly. When the female is heterozygous, the probabilities of a polar body having true genotype AA, Aa or aa is $p_{AA} = \frac{1}{2} - \theta$, $p_{Aa} = 2\theta$ and $p_{aa} = \frac{1}{2} - \theta$. The corresponding probabilities for a blastomere are $p_{AA} = 0.25$, $p_{Aa} = 0.5$ and $p_{aa} = 0.25$ when both parents are heterozygous, and $p_{AA} = 0$, $p_{Aa} = 0.5$ and $p_{aa} = 0.5$ when one parent is heterozygous and the other is homozygous. The conditional probability that a polar body or blastomere cell has true genotype y given that it has been typed as x can now be computed using Bayes's rule (see e.g. Feller, 1970, vol. 1, p. 124), as follows:

$$P(\text{cell has genotype } y \mid \text{typed as } x) = \frac{P(\text{typed } x \mid \text{true is } y)P(\text{true is } y)}{P(\text{typed } x \mid \text{true is AA})p_{AA} + P(\text{typed } x \mid \text{true is Aa})p_{Aa} + P(\text{typed } x \mid \text{true is aa})p_{aa}} \tag{13}$$

To compute the number of oocytes or embryos which need to be tested to have 95% probability of finding a given number both usable and typed as such, we first compute the probability that an oocyte or embryo chosen at random will be both usable and typed as such. Let p_u denote this probability. When the disease is recessive,

$$\begin{aligned} p_u = {} & P(\text{cell is typed AA} \mid \text{true genotype is AA})P(\text{true genotype is AA}) \\ & + P(\text{cell is typed Aa} \mid \text{true genotype is AA})P(\text{true genotype is AA}) \\ & + P(\text{cell is typed AA} \mid \text{true genotype is Aa})P(\text{true genotype is Aa}) \\ & + P(\text{cell is typed Aa} \mid \text{true genotype is Aa})P(\text{true genotype is Aa}). \end{aligned}$$

When the disease is dominant,

$$p_u = P(\text{cell is typed aa} \mid \text{true genotype is aa})P(\text{true genotype is aa}).$$

The number of oocytes or embryos needed to be typed in order to have 95% probability of finding at least k which are usable and typed as such is the smallest value of n for which the 5th percentile of the binomial distribution with parameters n and p_u is greater than k. For example, if $k = 1$, we find the smallest value of n satisfying $(1 - p_u)^n \leq 0.05$. If $k = 2$, we find the smallest value of n satisfying $np_u(1 - p_u)^{n-1} + (1 - p_u)^n \leq 0.05$. If $k = 3$, we find the smallest value of n satisfying $[n(n-1)/2]p_u^2(1 - p_u)^{n-2} + np_u(1 - p_u)^{n-1} + (1 - p_u)^n \leq 0.05$.

Example 1

We compute the probability that an embryo typed AA with blastomere typing will be unacceptable for implantation when the disease is recessive, both parents are heterozygous, and $r = 0.9$, $d = 0.8$ and $c = 0.05$. In this case, only embryos typed aa are unacceptable. Therefore we must compute

$$P(\text{unacceptable error}) = P(\text{cell has genotype aa} \mid \text{typed as AA}).$$

Using equations (1), (2) and (3), we compute $p_0 = 0.208$, $p_1 = 0.144$, $p_2 = 0.648$. Using equations (4), (5) and (6) we compute

$$\begin{aligned} P(\text{cell is typed AA} \mid \text{true genotype is AA}) &= 0.7774 \\ P(\text{cell is typed AA} \mid \text{true genotype is Aa}) &= 0.0754 \\ P(\text{cell is typed AA} \mid \text{true genotype is aa}) &= 0.0052. \end{aligned}$$

The values of p_{AA}, p_{Aa} and p_{aa} are 0.25, 0.5 and 0.25 respectively. Therefore equation (13) yields

$$P(\text{cell has genotype aa} \mid \text{typed as AA}) = \frac{(0.0052)(0.25)}{(0.7774)(0.025) + (0.0754)(0.5) + (0.0052)(0.25)} = 0.0056 \tag{14}$$

which is the probability of an unacceptable error. This value appears in line 5, column (3) of Table I.

Example 2

To explain the anomaly on line 5 of Table I, we compute the probability of an unacceptable error under the same conditions as in example 1, except that the value of r is increased from 0.9 to 1. The anomaly is that the probability of unacceptable error increases slightly.

We use equations (1), (2) and (3) to compute $p_0 = 0.2$, $p_1 = 0$, $p_2 = 0.8$, and equations (4), (5) and (6) to compute

$$\begin{aligned} P(\text{cell is typed AA} \mid \text{true genotype is AA}) &= 0.785 \\ P(\text{cell is typed AA} \mid \text{true genotype is Aa}) &= 0.005 \\ P(\text{cell is typed AA} \mid \text{true genotype is aa}) &= 0.005. \end{aligned}$$

Comparing these results to the corresponding ones in example 1 shows that increasing the value of r from 0.9 to 1 results in more accurate typing. The probability of typing a cell AA when the true genotype is AA increases from 0.7774 to 0.785, while the probability of typing a cell AA when the true genotype is Aa or aa decreases. Notice, however, that the probability of typing a cell AA when the true genotype is Aa decreases greatly, from 0.0754 to 0.005, whereas the probability of typing a cell AA when the true genotype is aa decreases only slightly, from 0.0052 to 0.005.

The values of p_{AA}, p_{Aa} and p_{aa} are 0.25, 0.5 and 0.25 respectively. Therefore equation (13) yields

$$P(\text{cell has genotype aa} \mid \text{typed as AA}) = \frac{(0.005)(0.25)}{(0.785)(0.025) + (0.005)(0.5) + (0.005)(0.25)} = 0.0063 \tag{15}$$

which is the probability of an unacceptable error. This value appears in line 5, column (7) of Table I.

Comparing equations (14) and (15) reveals the source of the anomaly. The numerators of the fractions, along with the first and third terms of the denominator, have changed only slightly, but the second term in the denominator is much smaller in equation (15) than in (14). Thus the value of the fraction is greater in equation (15).

Example 3

We compute the probability of implanting an unacceptable embryo using blastomere typing when the disease is dominant, one parent is heterozygous while the other

is homozygous aa, and $r = 0.9$, $d = 0.8$ and $c = 0.05$. In this case, only embryos typed aa will be implanted. Therefore we must compute

$$P(\text{unacceptable error}) = P(\text{cell has genotype AA} \mid \text{typed as aa}) + P(\text{cell has genotype Aa} \mid \text{typed as aa}).$$

We compute $p_0 = 0.208$, $p_1 = 0.144$ and $p_2 = 0.648$, as in example 1. Using equations (10), (11) and (12) we compute

$$\begin{aligned} P(\text{cell is typed aa} \mid \text{true genotype is AA}) &= 0.0052 \\ P(\text{cell is typed aa} \mid \text{true genotype is Aa}) &= 0.0754 \\ P(\text{cell is typed aa} \mid \text{true genotype is aa}) &= 0.7774 \end{aligned}$$

The values of p_{AA}, p_{Aa} and p_{aa} are 0, 0.5 and 0.5 respectively. Therefore two applications of equation (13) yield

$P(\text{cell has genotype AA} \mid \text{typed as aa}) = 0$ and $P(\text{cell has genotype Aa} \mid \text{typed as aa}) = 0.088$, so $P(\text{unacceptable error}) = 0 + 0.088 = 0.088$. This value appears in line 1, column (3) of Table V.

Example 4

We compute the probability of implanting an unacceptable embryo using polar body typing when the disease is recessive, both parents are heterozygous, and $r = 0.9$, $d = 0.8$, $c = 0.05$ and $\theta = 0.2$. In this case, only oocytes typed AA will be fertilized. Therefore we must compute $P(\text{cell has genotype AA} \mid \text{typed as aa})$ and $P(\text{cell has genotype Aa} \mid \text{typed as aa})$.

As in example 1 we use equations (1), (2), (3), (10), (11) and (12) to compute $p_0 = 0.208$, $p_1 = 0.144$ and $p_2 = 0.648$, and

$$\begin{aligned} P(\text{cell is typed AA} \mid \text{true genotype is AA}) &= 0.7774 \\ P(\text{cell is typed AA} \mid \text{true genotype is Aa}) &= 0.0754 \\ P(\text{cell is typed AA} \mid \text{true genotype is aa}) &= 0.0052 \end{aligned}$$

The values of p_{AA}, p_{Aa} and p_{aa} are 0.3, 0.4 and 0.3 respectively. Therefore two applications of equation (13) yield

$$P(\text{cell has genotype aa} \mid \text{typed as AA}) = \frac{(0.0052)(0.3)}{(0.7774)(0.3) + (0.0754)(0.4) + (0.0052)(0.3)} = 0.0059$$

and

$$P(\text{cell has genotype Aa} \mid \text{typed as AA}) = \frac{(0.0754)(0.4)}{(0.7774)(0.3) + (0.0754)(0.4) + (0.0052)(0.3)} = 0.114$$

If an oocyte has genotype AA the probability of an unacceptable error is 0.5, while if an oocyte has genotype Aa the probability of an unacceptable error is 0.25. Therefore the probability of an unacceptable error is $(0.5)(0.0059) + (0.25)(0.114) = 0.0031$. This value appears in line 3, column (3) of Table III.

Example 5

We compute the probability of implanting an unacceptable embryo using polar body typing when the disease is dominant, the female is heterozygous while the male is homozygous aa, and $r = 0.9$, $d = 0.8$, $c = 0.05$ and $\theta = 0.2$. In this case, only oocytes typed aa will be fertilized. Therefore we must compute $P(\text{cell has genotype AA} \mid \text{typed as aa})$ and $P(\text{cell has genotype Aa} \mid \text{typed as aa})$. These two quantities are computed exactly as in example 4. If an oocyte has genotype AA the probability of an unacceptable error is 1, while if an oocyte has genotype Aa the probability of an unacceptable error is 0.5. Therefore $P(\text{unacceptable error}) = (1)(0.0059) + (0.5)(0.114) = 0.063$. This value appears in line 3, column (3) of Table VII.

When two or more blastomeres are typed in one tube, the calculations are the same as in examples 1–3, except that equations (4)–(12) are replaced as follows: let n be the number of cells removed from the blastomere and let t_k be the probability that exactly k cells are deposited in the tube. If $n = 2$,

$$t_0 = (1-d)^2,\ t_1 = 2d(1-d),\ t_2 = d^2,\ t_3 = 0. \tag{16}$$

If $n = 3$

$$t_0 = (1-d)^3,\ t_1 = 3d(1-d)^2,\ t_2 = 3d^2(1-d),\ t_3 = d^3. \tag{17}$$

Let p denote the probability that no alleles are detected in the tube, whether by failure to deposit, failure to amplify, or a combination of the two. Then $p = t_3(1-r)^6 + t_2(1-r)^4 + t_1(1-r)^2 + t_0$. The equations which replace (4)–(12) are:

$$P(\text{cell is typed AA} \mid \text{true genotype is AA}) = (1-p)(1-c/2) + pc/2$$

$$P(\text{cell is typed AA} \mid \text{true genotype is Aa}) = [t_3(1-r)^3(1-(1-r)^3 + t_2(1-r)^2(1-(1-r)^2) + t_1r(1-r)](1-c/2) + pc/2$$

$$P(\text{cell is typed AA} \mid \text{true genotype is aa}) = pc/2$$

$$P(\text{cell is typed Aa} \mid \text{true genotype is AA}) = (1-p)(c/2)$$

$$P(\text{cell is typed Aa} \mid \text{true genotype is Aa}) = t_3[1-(1-r)^3]^2 + t_2[1-(1-r)^2]^2 + t_1r^2 + c[t_3(1-r)^3[1-(1-r)^3] + t_2(1-r)^2[1-(1-r)^2] + t_1r(1-r)]$$

$$P(\text{cell is typed Aa} \mid \text{true genotype is aa}) = (1-p)(c/2)$$

$$P(\text{cell is typed aa} \mid \text{true genotype is AA}) = pc/2$$

$$P(\text{cell is typed aa} \mid \text{true genotype is Aa}) = [t_3(1-r)^3[1-(1-r)^3] + t_2(1-r)^2[1-(1-r)^2] + t_1r(1-r)](1-c/2) + pc/2$$

$$P(\text{cell is typed aa} \mid \text{true genotype is aa}) = (1-p)(1-c/2) + pc/2$$

When two or three cells are typed in separate tubes and the disease is recessive, an embryo is considered suitable for implantation if every tube which gives a signal gives a signal for the A allele. When the disease is dominant, an embryo is considered suitable for implantation if no tube gives a signal for the A allele, and at least one tube gives a signal for the a allele. In the first case we shall say that the embryo is typed A, and in the second case we shall say that the embryo is typed a. When the disease is dominant, the conditional probability of unacceptable error is the quantity

$$P(\text{true genotype is AA} \mid \text{typed a}) + P(\text{true genotype ia Aa} \mid \text{typed a}),$$

and when the disease is recessive the corresponding value is

$$P(\text{true genotype is aa} \mid \text{typed A}).$$

To compute these quantities, let p_0, p_1 and p_2 be as in equations (1), (2) and (3), and let n represent the number of tubes. We compute

$$P(\text{cell is typed A} \mid \text{true genotype is AA}) = (1-p_0c/2)^n - p_0^n(1-c)^n \tag{18}$$

$$P(\text{cell is typed A} \mid \text{true genotype is Aa}) = [1-(p_1/2)(1-c/2) - p_0c/2]^n - p_0^n(1-c)^n) \tag{19}$$

$$P(\text{cell is typed A} \mid \text{true genotype is aa}) = [p_0(1-c/2) - (1-p_0)c/2]^n - p_0^n(1-c)^n \tag{20}$$

$$P(\text{cell is typed a} \mid \text{true genotype is AA}) = p_0^n(1-c/2)^n - p_0^n(1-c)^n \tag{21}$$

$$P(\text{cell is typed a} \mid \text{true genotype is Aa}) = (p_0 + p_1/2)^n(1-c/2)^n - p_0^n(1-c)^n \tag{22}$$

$$P(\text{cell is typed a} \mid \text{true genotype is aa}) = (1-c/2)^n - p_0^n(1-c)^n \tag{23}$$

In addition there is a small probability that the cell will remain untyped through a failure to detect any alleles, or because of conflicting results in the tubes, e.g. only A is detected in one tube, only a in another. This quantity is not needed in our calculations.

The calculation of the conditional probability of unacceptable error is carried out as before, using Bayes's rule.

Example 6

Let $r = 0.9$, $d = 0.8$, $c = 0.05$, $n = 3$. We compute the probability of an unacceptable error if the disease is recessive and both parents are heterozygous. Using equations (18), (19) and (20) we compute

$$\begin{aligned} P(\text{cell is typed A} \mid \text{true genotype is AA}) &= 0.977 \\ P(\text{cell is typed A} \mid \text{true genotype is Aa}) &= 0.783 \\ P(\text{cell is typed A} \mid \text{true genotype is aa}) &= 0.00331. \end{aligned}$$

The values of p_{AA}, p_{Aa} and p_{aa} are 0.25, 0.5 and 0.25 respectively. Using Bayes's rule, we find that the probability of unacceptable error is

$$P(\text{true genotype is aa} \mid \text{typed A}) = \frac{(0.00331)(0.25)}{(0.977)(0.25) + (0.783)(0.5) + (0.00331)(0.25)} = 0.0013$$

This quantity appears in Table X1, line 3, column (3).

In the X-linked case where only the Y chromosome is amplified, the calculations are slightly different from the ones in the examples above. To compute the conditional probability of such an event, we first compute

$$P(\text{Y chromosome is detected} \mid \text{true genotype is XX}) = c/2 \quad (24)$$

$$P(\text{Y chromosome is detected} \mid \text{true genotype is XY}) = rd(1-c/2) + c/2 \quad (25)$$

$$P(\text{no Y chromosome detected} \mid \text{true genotype is XX}) = 1-c/2 \quad (26)$$

$$P(\text{no Y chromosome detected} \mid \text{true genotype is XY}) = (1-rd)(1-c/2) \quad (27)$$

In the recessive case, an unacceptable error occurs when an embryo typed as having no Y chromosome does in fact have one. Bayes's rule yields

$$P(\text{true genotype is XY} \mid \text{no Y detected}) = \frac{P(\text{no Y detected} \mid \text{true is XY})P(\text{true is XY})}{P(\text{no Y detected} \mid \text{true is XY})P(\text{true is XY}) + P(\text{no Y detected} \mid \text{true is XX})P(\text{true is XX})} \quad (28)$$

If 50% of embryos are XX and 50% are XY, we obtain

$$P(\text{true genotype is XY} \mid \text{no Y chromosome detected}) = \frac{1-rd}{2-rd} \quad (29)$$

Since only 50% of embryos with true genotype XY have the disease, the conditional probability of unacceptable error is one-half the quantity on the right hand side of equation (29). That is,

$$P(\text{unacceptable error}) = \frac{1-rd}{4-2rd}$$

In the dominant case, an unacceptable error occurs when an embryo typed as having a Y chromosome does not in fact have one. Bayes's rule yields

$$P(\text{true genotype is XX} \mid \text{Y detected}) = \frac{P(\text{Y detected} \mid \text{true is XX})P(\text{true is XX})}{P(\text{Y detected} \mid \text{true is XY})P(\text{true is XY}) + P(\text{Y detected} \mid \text{true is XX})P(\text{true is XX})} \quad (30)$$

If 50% of embryos are XX and 50% are XY, we obtain

$$P(\text{true genotype is XX} \mid \text{Y chromosome is detected}) = \frac{c}{2c + rd(c-2)}$$

Since all embryos with true genotype XX have the disease, the conditional probability of unacceptable error is equal to this quantity as well.

In the case of dominant X-linked disease, if an X chromosome marker is amplified and if only embryos for which both X and Y chromosomes are detected are considered for implantation, then the probability of unacceptable error involves the following quantities, where p_0, p_1 and p_2 are as in equations (1), (2) and (3):

$$P(\text{cell is typed XY} \mid \text{true genotype is XX}) = (p_1 + p_2)(c/2)$$

$$P(\text{cell is typed XY} \mid \text{true genotype is XY}) = p_2 + p_1(c/2)$$

The conditional probability of unacceptable error is given by Bayes's rule as:

$$P(\text{true genotype is XX} \mid \text{typed XY}) = \frac{P(\text{typed XY} \mid \text{true is XX})P(\text{true is XX})}{P(\text{typed XY} \mid \text{true is XY})P(\text{true is XY}) + P(\text{typed XY} \mid \text{true is XX})P(\text{true is XX})}$$

If 50% of embryos are XX and 50% are XY, we obtain

$$P(\text{true genotype is XX} \mid \text{typed XY}) = \frac{(p_1 + p_2)c/2}{(p_1 + p_2)c/2 + p_1 c/2 + p_2}$$

Since all embryos with true genotype XX have the disease, the conditional probability of unacceptable error is equal to the quantity as well.

When two or three blastomere cells are typed in a single tube and the disease is X-linked, the conditional probability of unacceptable error is computed as in the one cell case, except that equations (24)−(27) are replaced by the following, where t_0, t_1, t_2, t_3 are as in equations (16) and (17):

$$P(\text{Y chromosome is detected} \mid \text{true genotype is XX}) = c/2$$

$$P(\text{Y chromosome is detected} \mid \text{true genotype is XY}) = 1-t_3(1-r)^3(1-c/2)-t_2(1-r)^2(1-c/2)-t_1(1-r)(1-c/2)-t_0(1-c/2)$$

$$P(\text{no Y chromosome detected} \mid \text{true genotype is XX}) = 1-c/2$$

$$P(\text{no Y chromosome detected} \mid \text{true genotype is XY}) = t_3(1-r)^3(1-c/2)+t_2(1-r)^2(1-c/2)+t_1(1-r)(1-c/2)+t_0(1-c/2)$$

The probability that an embryo will be of genotype XY when no Y chromosome is detected is computed using Bayes's rule, as in equation (28). Since only 50% of embryos of genotype XY are not usable, the conditional probability of unacceptable error is found by multiplying the right hand side of equation (28) by 0.5.

When two blastomere cells are typed in separate tubes and the disease is X-linked, the conditional probability of unacceptable error is computed as in the one cell case, except that equations (24)−(27) are replaced by the following:

$$P(\text{Y chromosome is detected in both tubes} \mid \text{true genotype is XX}) + (c/2)^2$$

$$P(\text{Y chromosome is detected in both tubes} \mid \text{true genotype is XY}) = [c/2 + rd(1 - c/2)]^2$$

$$P(\text{no Y chromosome detected in either tube} \mid \text{true genotype is XX}) = (1 - c/2)^2$$

$$P(\text{no Y chromosome detected in either tube} \mid \text{true genotype is XY}) = (1 - rd)^2(1 - c/2)^2.$$

When three blastomere cells are typed in separate tubes and the disease is X-linked, equations (24)−(27) are replaced by the following:

$$P(\text{Y chromosome is detected in two or more tubes} \mid \text{true genotype is XX}) = 3(c/2)^2(1 - c/2) + (c/2)^3$$

$$P(\text{Y chromosome is detected in two or more tubes} \mid \text{true genotype is XY}) = 3[c/2 + rd(1 - c/2)]^2(1 - rd)(1 - c/2) + [c/2 + rd(1 - c/2)]^3$$

$$P(\text{no Y chromosome detected in any tube} \mid \text{true genotype is XX}) = (1 - c/2)^3$$

$$P(\text{no Y chromosome detected in any tube} \mid \text{true genotype is XY}) = (1 - rd)^2(1 - c/2)^3.$$

When either two or three cells are typed in separate tubes, the conditional probability of unacceptable error is given by equation (28) when the disease is recessive, or by equation (30) when the disease is dominant.