The Polymerase Chain Reaction of Genetic Markers for the Detection of Combined Immune Deficiency in Arabian Horses

Scott Seidel

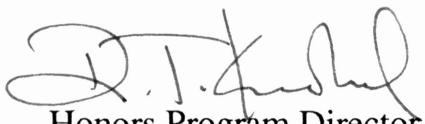
University Undergraduate Fellows Program, 1989 - 1990

Texas A&M University

Department of Biology

APPROVED:

Fellows Advisor



Honors Program Director

INTRODUCTION

Combined immune deficiency, or CID, is a lethal genetic disease that affects humans and a variety of animals. In horses, it is inherited as an autosomal recessive gene, and occurs mainly among members of the Arabian breed.[1,2] However, even though the inheritance pattern has been documented, the etiology of CID is still largely unknown. Several accompanying conditions have been described, including deficiencies in thymus and marrow-derived lymphocytes, and abnormalities in nucleic acid metabolism.[3,4] In addition, studies in other species indicate that several different genetic defects on different chromosomes may cause heritable CID.[5,6]

CID accounts for the death of approximately 2.3% of the North American Arabian foals, and it is estimated that 30% of the population of North American Arabian horses are carriers of this disease and transmit it as asymptomatic carriers.[7] The high frequency of this disease suggests that the elimination of the defective gene from this horse breed would be highly advantageous to Arabian horse breeders. The elimination of this disease requires a reliable, accurate means of detecting asymptomatic CID carriers. This would allow elimination of the disease through selection of non-carriers as breeding stock.

A genetic marker for the CID gene would fulfill the need for a diagnostic test. However, the CID gene itself has not been isolated, and a large-scale search for the CID gene is not feasible at this time. One alternative is to examine genetic sequences on the same chromosome closely linked to the CID gene. These sequences, due to their proximity to the CID gene, would tend to be inherited along with it. Therefore, sequence markers (polymorphisms) linked to the recessive CID gene would be homozygous in affected CID foals, and heterozygous in the carrier parents. In this way, by following the transmission of

the known genetic markers, the transmission of the CID gene may be followed as well. This technique, called "homozygosity mapping," has been proposed for use in detecting recessive genetic disorders in humans.[8]

Identifying closely linked markers to a gene requires information of the genetic loci in the region surrounding the gene. Unfortunately, the genomic map of the horse is one of the least developed of any laboratory or domestic animal. However, the genomic maps of other domestic animals are not as ill-defined as the horse. Therefore, it is possible to use information from these well-mapped genomes and genetically extrapolate the data to other species based on the principle of evolutionally conserved gene linkages. In this way, closely linked markers in one species can be used as homologous markers in another.

One such species of animal is the mouse. Its genetic loci map is extensive compared to that of the horse, with over 1500 loci mapped versus about 23 for the horse. Fortunately, a clinically similar combined immunodeficiency occurs in the mouse. This is the scid recessive mutation, and in the mouse it is located on the centromeric end of chromosome 16.[9] Based on the assumption that the homologous gene is involved in both mouse and horse immunodeficiency diseases (due to evolutionally conserved gene linkages), a marker for the scid gene in the mouse was used as a marker for the detection of the CID gene in horses.
The defective scid gene has not been located in the mouse, but the scid locus is about 1 cM down from the protamine-1 (Prm-1) locus. The frequency with which two linked genes on the same chromosome are inherited together is directly related to the physical distance between them. This frequency is measured in percent recombination or centiMorgans (cM). One centimorgan of DNA represents about $2x10^6$ nucleotide base pairs. Since the scid locus is about 1 cM from the Prm-1 locus, if we assume that the distance between the CID gene

and the protamine gene in the horse is roughly the same, the recombination frequency of these two genes is:

$$\frac{\text{cM} \times \text{number of bases / cM}}{\text{number of bases / genome}} = \frac{(1) \times (2 \times 10^6)}{3 \times 10^9} = .0007 \text{ or } .07\ \%$$

This indicates, if the assumption is correct, that the horse protamine and CID genes will be recombined during meiosis at a frequency of .07%, or in other words, any marker in the horse protamine gene will be associated with the CID gene over 99% of the time.

In addition to its close proximity to the scid locus, Prm-1 has been cloned from the human, mouse, and bovine, and comparison has revealed extensive conservation in many regions.[10] Based on these factors, Prm-1 represents the best available marker for the possible identification of sequence polymorphisms linked to the horse CID gene.

Normally, sequence polymorphisms can be determined by the detection of restriction fragment length poymorphisms, or RFLPs. However, an extensive search for RFLPs in the horse Prm-1 failed to detect polymorphisms necessary for its use as a gene marker.[11]

An alternative method is to sequence the protamine gene from the horse, and determine base by base if any polymorphisms exist in the gene of CID affected animals and heterozygous carriers. Normally this represents a tremendous cost in time and money, due to the methods necessary to clone and isolate the gene. This is a major barrier to the development of a CID diagnostic test.

A technique has been described that allows the amplification of specific regions of DNA to many thousand times their original concentration.[12] This

technique, called the Polymerase Chain Reaction (PCR), involves two oligonucleotide primers that flank the sequence to be amplified in opposite orientation. Multiple cycles of heat denaturing of the DNA, annealing and reannealing of the primers to their complementary sequences, and extension of the primers with *Taq* DNA polymerase, result in the exponential accumulation of the specific target sequence, approximately 2^n where n is the number of cycles.[13] By using the thermostable *Taq* DNA polymerase, the addition of more polymerase enzyme following the heat denaturaton step is avoided.[14]

The target amplified by PCR depends on the selection of gene specific primers. Since the protamine gene sequence in the horse is not known, primers were chosen based on the known sequence of the mouse Prm-1 gene. The purpose of this experiment is to develop PCR for the amplification of the horse protamine gene. Once amplified, the genes from many animals can then be sequenced to determine the presence of genetic polymorphisms.

MATERIALS

Enzyme. *Taq* DNA polymerase, a thermostable, 94kDa polymerase (GeneAmp) purchased from Perkin-Elmer Cetus Instruments, as part of the GeneAmp DNA Amplification Reagent Kit. The enzyme, 250 units, was stored at -20 °C in 50% glycerol (v/v)/100 mMolar KCl/ 20 mMolar Tris-HCl, pH8.0/0.1 mMolar EDTA/1mMolar dithiothreitol/200 µg/ml gelatin.

Nucleotides. 2'-Deoxynucleotides (dNTPs) were purchased from Perkin-Elmer Cetus Instruments as part of the GeneAmp kit described above.

Oligonucleotide primers. The two primers, PRM-1 (5'- CCAC-CTGCTCACAGGTTGGCTGGC) , and PRM-2 (5'-GGCATTGTTCGT-

TAGCAGGCTCCTG), were purchased from R. Van derHarr (Texas A&M University, Department of Biology).

Horse DNA Samples. The horse DNA samples were obtained from L. Perryman, Department of Veterinary Microbiology and Pathology, Washington State University as fibroblast cultures. The cultures were unfrozen and grown according to Fig. 1, and then lysed, incubated with proteinase K, and the DNA collected with organic solvents. The DNA samples include two stallions, Sonny and Bo, 8 mares known to be heterozygous for CID, several other mares, and various foals with and without CID (Table 1). The pedigrees of the two families are shown in Fig. 2.

METHODS

Amplification of Samples. Amplification of the horse samples with *Taq* DNA polymerase took place in 100 µl reaction mixtures containing 1X modified PCR buffer (10 mM Tris-HCl, pH 8.3 at 25 °C / 50 mM KCl / 3.0 mM $MgCl_2$), 200 µM each dNTP, 1.0 µM of each primer, 15-200 ng of whole genomic horse sample DNA, and 2.5 units of *Taq* DNA polymerase. Additionally, a 100 µl overlay of mineral oil was added to prevent water loss during the amplifications. PCR was carried out in a programmable thermal block (TwinBlock, Ericomp). The program used for amplifying the horse samples consisted of an initial denaturation step of 3 minutes at 95°C, and then 35 cycles of: 1) denaturation of the DNA for 5 seconds at 95°C; 2) annealing of the primers for 1 minute 30 seconds at 55°C; and 3) extension of the target region by the polymerase for 3 minutes at 72° (Fig. 3). Each amplification took approximately 6 hours.

Selection of Primers. Since the mouse protamine gene has been sequenced, primers were chosen that flank the majority of the gene. These two primers, PRM-1 and PRM-2, flank a 454 base region in the protamine gene (Fig. 4).[15] The PRM -1 oligonucleotide begins at base 575, and has the following sequence: 5'-CCACCTGCT-CACAGGTTGGCTGGC. PRM-2 begins at base 1029, and has the following sequence: 5'-GGCATTGTTCGTTAGCAGGC-TCCTG.

Agarose Gel Electrophoresis. After amplification by PCR, all of the samples were analyzed by agarose gel electrophoresis. Samples were prepared for electrophoresis by adding 2 µl of stop-buffer(0.5% bromophenol blue/0.5% xylene cyanol/25% Ficoll 400/1% SDS/10mM EDTA) to 20 µl of each sample taken from the reaction mixtures. The samples were then loaded onto .75% agarose gels, prepared from .75 g electrophoresis grade agarose, 100 ml 1X TAE (.04 M Tris-acetate/0.002M EDTA), and 10 µl of ethidium bromide for visualization. Lambda DNA previously cut with HindIII was added as a size standard. The gels were run at 100 volts for 1 hour, the samples were visualized with UV light, and photographed.

Purification of the Samples. After electrophoresis, samples positive for the 454 base band were purified on a DNA exclusion column (NACS PREPAC, Bethesda Research Laboratories). Two buffers were prepared, a loading buffer, buffer A, 0.2 M NaCl in TE (10mM Tris-HCl, pH 7.2, 1 mM EDTA), and an eluting buffer, buffer B (1.0 M NaCl in TE). The samples were purified as follows. The column was first hydrated with 3 mls of buffer B, then equilibrated with 5 mls of buffer A. The sample was then loaded onto the top of the column with 3 mls of buffer A, and expelled by means of a syringe inserted into the top of the column. Elution of the DNA was carried out three times with 3 mls of buffer B, and 1 ml fractions were collected.

Equal volumes of 100% ethanol at -20 °C were added to the first fraction of each sample, mixed well, and frozen overnight at -80°C. Afterwards, the samples were centrifuged in a microfuge for 10 minutes, and the pellets were collected and resuspended in 50 µl of autoclaved, distilled H_2O. Concentrations were checked on a DNA fluorometer [2 µl sample diluted into 2 ml of dye from a stock solution of 10 µl Hoechst 3358 dye stock in 100 µl TNE (TE, pH 8.0, containing 0.1 M NaCl)].

RESULTS

Initial Amplification. The initial amplification of the all of the horse samples by the method described above produced several different bands that varied between the samples, instead of only one band as might be expected. While some samples had only one band, others had four (Fig. 5) Most had a band roughly corresponding to 454 bases, but no correlation between the band pattern and the presence of CID was found.

Low Melting Point Amplification. To isolate the multiple bands, selected family samples (Sonny, mares 101 and 105, and foals 1962, 1974, 1776, and 1885) were analyzed on a low melting point (LMP) agarose gel (1% electrophoresis grade low melting agarose in TAE) at 50 volts for approximately 14 hours, with HindIII cut Lambda DNA as a size standard. The low melting point agarose allowed recovery of the specific fragment of DNA in the band cut out from the gel. The gel was washed for 30 minutes with a solution of 500 ml H_2O and 50 µl ethidium bromide with shaking at room temperature to stain the bands. Each lane was then labelled with a letter designation, and the bands in each lane were numbered consecutively from the top (origin) of the gel.(Fig. 6)

Following the band labelling, the band or bands in each of the sample lanes were cut out with a razor blade under UV light, and placed into capped tubes. The number of bands cut out varied from 1 band from the lane of sample 1885, to 4 from sample 1974.(Table 2) At this point, due to the great number of new samples, five cutouts from three different horses were selected for further analysis. These included the band from 1885, labelled Q1, two bands from the stallion, Sonny, labelled K3 and K4, and two bands from 1962, N3 and N4. These samples were heated at 60°C for 15 minutes, diluted by a volume, and then run through the NACS DNA purifying column described above.

Unbalanced PCR. The fragments which were closest to 454 bases were then reamplified using a modification of PCR in which one primer is made limiting, while the other primer is in excess. This results in the production and accumulation of single-stranded product. This was done for the purpose of future sequencing. The same amplification program was run, and after the samples had amplified, they were loaded onto a gel, as described above. The fragments produced in each case corresponded to their size in the low melting point gel (Fig. 7) These included a band a little larger than 400 bases. This fragment is presumed to be the amplified horse protamine gene.

DISCUSSION

The presence of extra bands may have been the result of several factors, such as multiple copies of the protamine gene in the horse or non-specific annealing of the primers during PCR. However, the specific band pattern of each sample was consistent over several different amplifications, so experimental error in the PCR procedure was probably not the cause. In the case of the sigle-stranded DNA amplification products, when the samples were heated before

loading onto the gel, the amount of multiple bands decreased. In this case, the multiple bands were probably due to the annealing of the single-standed DNA products, forming long strands of varying length.

The amplificaton of gene sequences in species with primers prepared for homologous genes from other species should prove extremely useful in the amplification of targets where the type of gene is known, but not the sequence.

Specifically, in the case of CID, I have shown, at least tentatively, that the protamine gene in the horse may be amplified by using primers developed from the mouse. Since some degree of variation is expected due to the slight differences in DNA surrounding conserved gene sequences, the amplification product may not be exactly the same length, or occur in the same number of copies. This fact alone may account for the presence of multiple bands in the initial amplified sequences.

Until the fragments I have amplified can be sequenced, their positive identity will not be known. Sequencing has been tried repeatedly, but has not succeeded. The final amplification using unbalanced PCR should facilitate sequencing using the Sanger Dideoxy method, which requires single-stranded template.

I have demonstrated the amplification of what I assume to be copies of the equine protamine gene. This gene, once the sequence has been determined, may be used as a marker for the detection of combined immune deficiency in the Arabian horse, if polymorphisms associated with CID gene are found. The technique of using PCR primers developed from the analysis of homologous genes in other species to detect markers for disease may also prove useful in the sequencing of homologous genes.

Further study on this topic will include the sequencing of the present amplified fragments and the optimization of a method for sequencing the remainder of the amplified fragments from the horse fibroblast samples.

In conclusion, the technique of PCR sequencing with primers specific for a closely linked marker to the CID gene in Arabian horses may provide a diagnostic test for Combined Immune Deficiency. The elimination of this disease from the breed would be extremely beneficial to the horses, to the breeders, and obviously to the breed as a whole.

I wish to acknowlege the kind and helpful involvement of my advisor, Dr. Loren Skow in the inception, planning, and execution of this project. Without his help and encouragement, this project could not have taken place.

Table 1. Fibroblasts from Arabian Foals at Washington State University:

CID foals	Dam	Sire
A-1817	137†	Bo†
A-1824	128	Bo†
A-1825	101	Bo†
A-1863*	136†	Bo†
A-1871	112	Bo†
A-1889	8	Sonny
A-1885	105	Bo†
A-1974	101	Sonny
A-1977	121	Sonny

Non-CID foals	Dam	Sire
A-1766	135†	Sonny
A-1776	132†	Sonny
A-1826*	136†	Bo†
A-1887	123	Sonny
A-1898	144	Bo†
A-1971	146	Sonny

† = deceased
* Histocompatible full siblings

Known Heterozygous Mares:

H-101
H-105
H-107
H-108
H-121
H-128
H-146
H-155

Table 2. Low Melting Point Gel Key. The low melting point (LMP) agarose gel (1% electrophoresis grade low melting agarose) was run at 50 volts for approximately 14 hours, with HindIII cut Lambda DNA as a size standard. Then the gel was washed with a solution of 500 ml H_2O and 50 µl ethidium bromide with shaking for 30 minutes at room temperature to stain the bands. Each lane was labelled with a letter designation, and the bands in each lane were numbered consecutively from the top (origin) of the gel.

LMP Horse Samples:	CID	Sonny	101	105	1962	1974	1776	1885
Fragments:	J1	K1	L1	M1	N1	O1	P1	Q1
	J2	K2	L2	M2	N2	O2		
	J3	K3	L3	M3	N3	O3		
	J4	K4	L4	M4	N4	O4		

Fig. 1. Fibroblast protocol for thawing and expanding fibroblast cultures provided by Lance E. Perryman, DVM, PhD, Washington State University Deptartment of Veterinary Microbiology and Pathology:

Procedure for thawing fibroblasts:

1. Thaw vials quickly in 37° water bath
2. Transfer cells to 25cm^2 flask and add approximately 5 ml DMEM (always containing 10mM hepes, 100 U/ml pennicillin, 100 µg streptomycin) with fetal calf serim
3. Incubate at 37° C in 5% CO_2 for 18 hours
4. Pour off supernatant anf replace with DMEM containing 15% fetal calf serum. (If cells recover very quickly and become confluent overnight, serum concentration should be reduced to 10%).

To expand fibroblast cultures:

1. Pour off supernatant and rinse flask with a small volume of HBSS without Ca and Mg.
2. Add 1 ml trypsin solution /25cm^2 flask or 2 ml/75cm^2 flask and let sit at room temperature for 2-5 minutes.
3. Knock or flush cells loose.
4. Transfer cells to larger or multiple flasks and add DMEM with 10 % fetal calf serum, 5 ml/25 cm^2 flask.

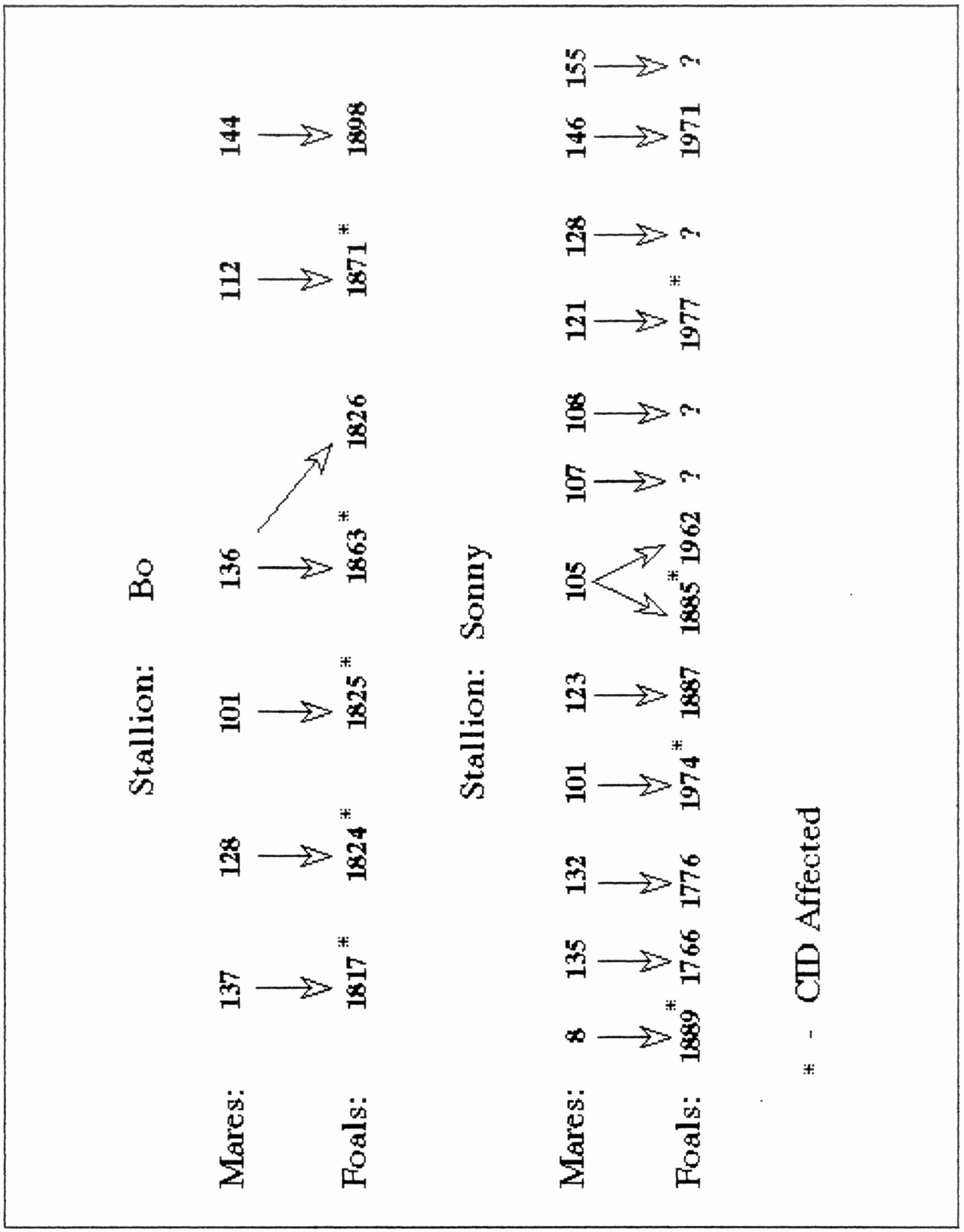


Fig. 2. Pedigrees of CID samples. The horse DNA samples were obtained from L. Perryman, Department of Veterinary Microbiology and Pathology, Washington State University as fibroblast cultures.

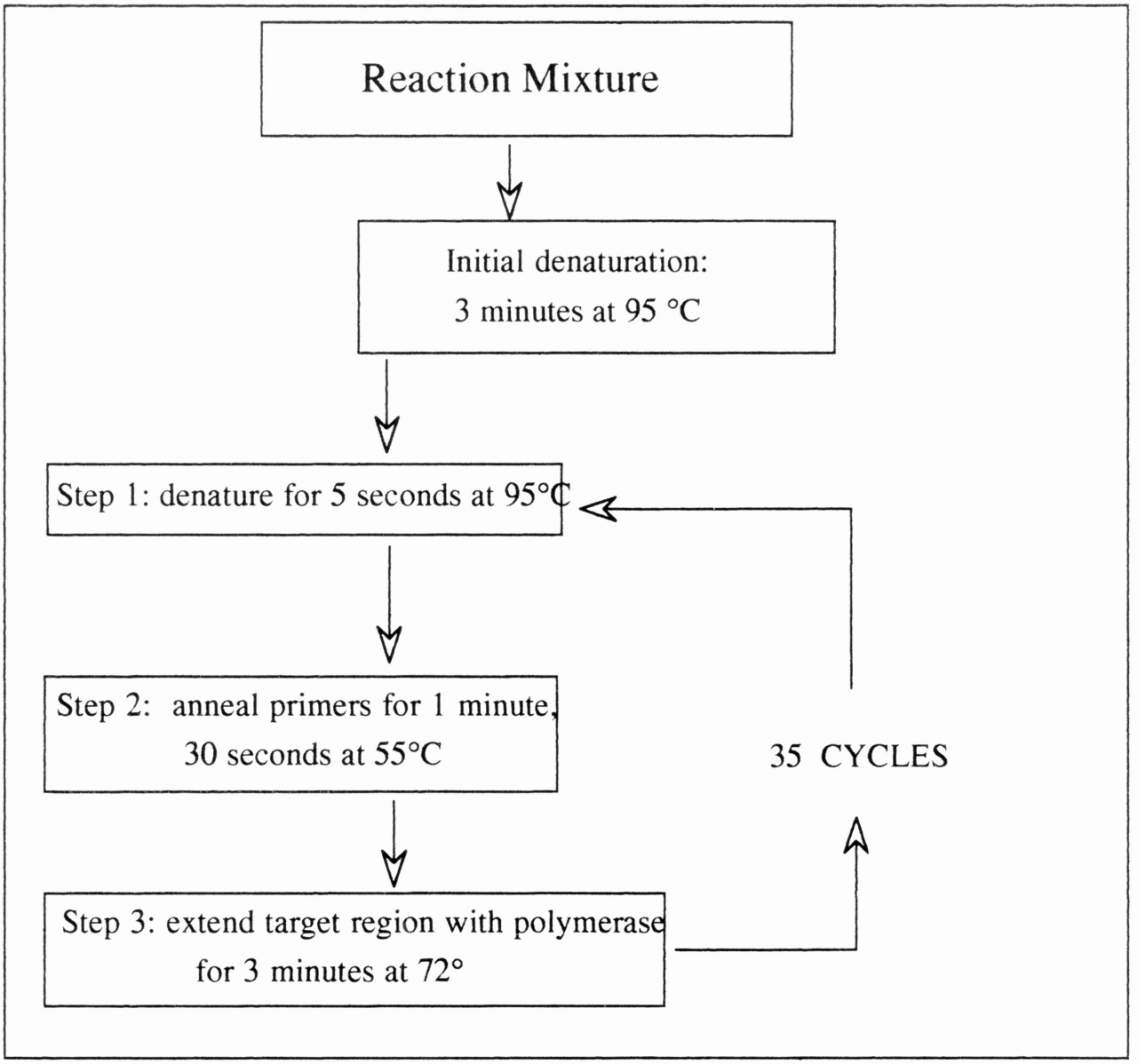


Fig. 3. Polymerase chain reaction flowchart of program used to amplify horse samples. PCR was carried out in a programmable thermal block (TwinBlock, Ericomp). The program used for amplifying the horse samples consisted of an initial denaturation step of 3 minutes at 95°C, and then 35 cycles of: 1) denaturation of the DNA for 5 seconds at 95°C; 2) annealing of the primers for 1 minute 30 seconds at 55°C; and 3) extension of the target region by the polymerase for 3 minutes at 72°. Each amplification took approximately 6 hours.

5' -AGGGTGCTGGCTCCCAGGCCACAGCCCACAAAATTCCACCTGCTCA - 587

CAGGTTGGCTGGCTCGACCCAGGTGGTGTCCCCTGCTCTGAGCCAGCTC - 636

CCGGCCAAGCCAGCACCATGGCCAGATACCGATGCTGCCGCAGCAAAAG - 685

CAGGAGCAGATGCCGCCGTCGCAGGCGAAGATGTCGCAGACGGAGGAG - 733

GCGATGCTGCCGGCGGAGGAGGCGAAGTAAGTAGAGGGCTGGGCTGGGCT -782

GTGGGGGGTGTGGCCTGCGGGACTTGGGCATGTCTGGGAGTCCCTCTCAC -832

CACTTTTCTTACCTTTCTAGGATGCTGCCGTCGCCGCCGCTCATACACCAT -883

AAGGTGTAAAAAATACTAGATGCACAGAATAGCAAGTCCATCAAAACTC -933

CTGCGTGAGAATTTTACCAGACTTCAAGAGCATCTCGCCACATCTTGAAA -981

AATGCCACCGTCCGATGAAAAACAGGAGCCTGCTAAGGAACAAT -1026

GCCCCTGTCAATAAATGTTGAAA- 1050

Fig. 4. Primer placement in the mouse protamine gene. Since the mouse protamine gene has been sequenced, primers were chosen that flank the majority of the gene. These two primers, PRM-1 and PRM-2, flank a 454 base region in the protamine gene. The PRM -1 oligonucleotide begins at base 575, and has the following sequence: 5'-CCACCTGCTCACAGGTTGGCTGGC. PRM-2 begins at base 1029, and has the following sequence: 5'-GGCATTGTTCG-TTAGCAGGCTCCTG. The primers are shown outlined, and the 454 target sequence is underlined.

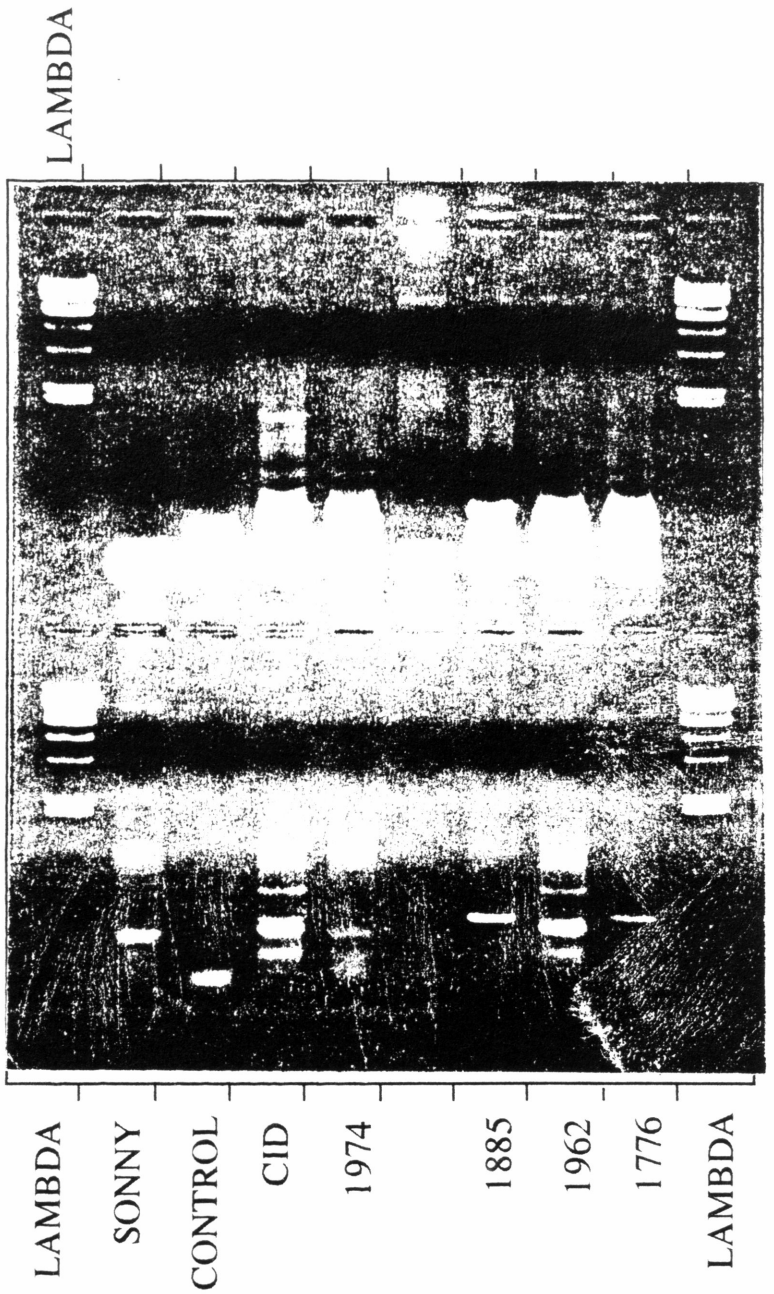


Fig. 5. Initial amplification of the horse DNA samples. The variety of bands may be seen in the bottom lanes. Sonny, 1885, and 1962 were later taken for further analysis.

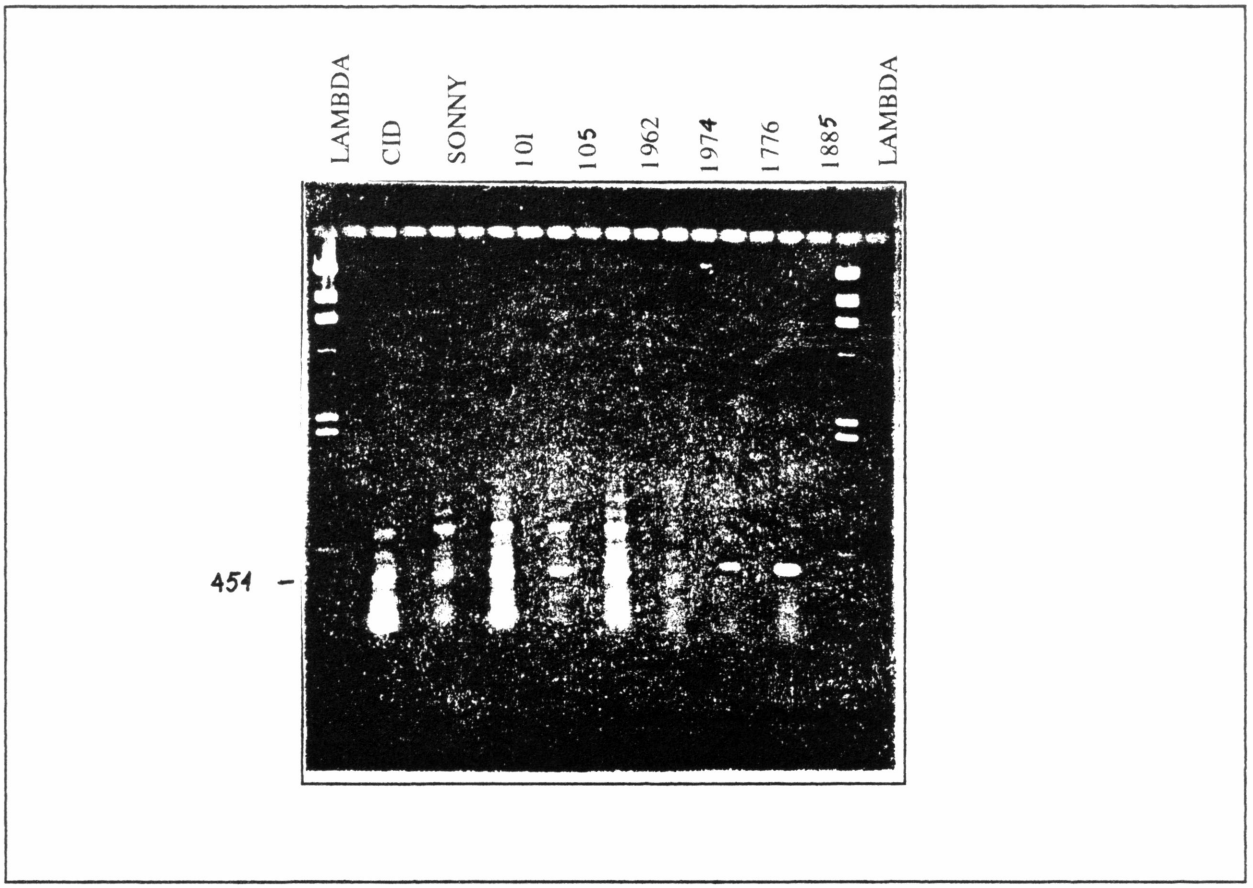


Fig. 6. Low melting point agarose gel of selected horse samples.The gel was washed for 30 minutes with a solution of 500 ml H_2O and 50 µl ethidium bromide with shaking at room temperature to stain the bands. Each lane was then labelled with a letter designation, and the bands in each lane were numbered consecutively from the top (origin) of the gel. Following the band labelling, the band or bands in each of the sample lanes were cut out with a razor blade under UV light, and placed into capped tubes.

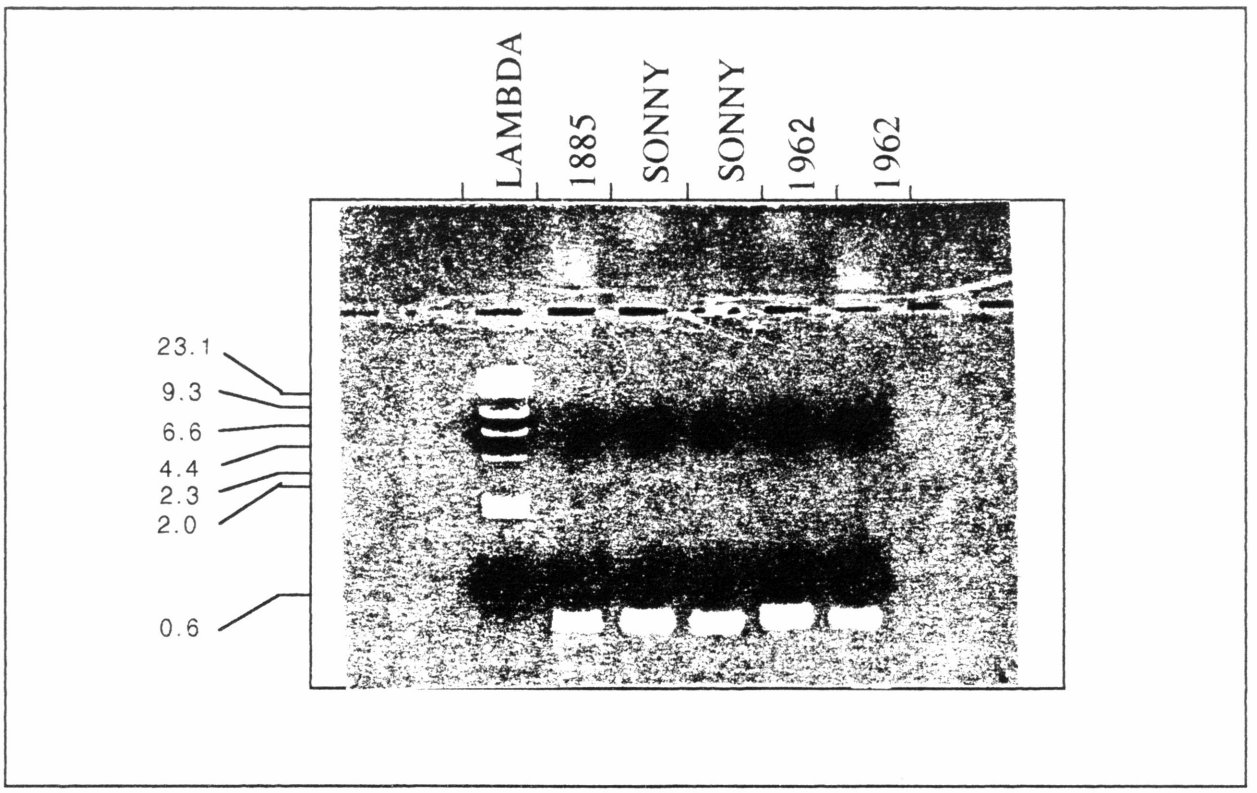


Fig. 7. Unbalanced PCR of selected horse samples. The fragments from the low melting point agarose gel which were closest to 454 bases were reamplified using a modification of PCR in which one primer is made limiting, while the other primer is in excess. This results in the production and accumulation of single-stranded product. This was done for the purpose of future sequencing.

REFERENCES

1 McGuire, T.C. and Poppy, M. G.: Hypoammaglobulinemia and thymic hypoplasia in horses: A primary combined immunodeficiency syndrome. Infec. Immun. 8:272-277, 1973.

2 Perryman L. E. and Torbeck, R. L.: Combined immunodeficiency of Arabian horses. Confirmation of the autosomal recessive mode of inheritance. J. Am. Vet. Med. Assoc. 176(11):1250-1251,1980.

3 McGuire, T. C., Banks, K. L. and Poppie , M. J.: Combined immunodeficiency in horses: Characterization of the lymphocyte defect. Clin. Immuno. Immunopath. 3:555-566, 1975.

4 Magnuson, N. S., Perryman, L. E., Wyatt, C. R., Ishizaka, T., Mason, P. H., Norman, A. E., Banks, K. L. and Magnuson, J. A.: Continuous cultivation of equine lymphocytes: Evidence for occasional T-cell-like maturation events in horses with hereditary severe combine immunodeficiency. J. Immuno. 133:2518-2524, 1984.

5 Chandra, R. F.: WHO Report: Immunodeficiency. Clin. Immuno. Immunopath. 13:296-359, 1979.

6 Sordat, B.: 4th International Workshop on Immuno-deficient Animals in Experimental Research. Expl. Cell. Biol. 50:305-349, 1982.

7 Poppie, M. J., and McGuire, T. C.: Combined immunodeficiency in foals of Arabian breeding: Evaluation of mode of inheritance and estimation of prevalence of affected foals and carrier mares and stallions. J. Am. Vet. Med. Assoc. 170: 31-33, 1977.

8 Lander, E. W. and Botstein, D.: Homozygosity mapping: A way to map human recessive traits with the DNA of inbred children. Science 236:1567-1570, 1987.

9 Skow, L. C.: The mouse gene map: Implications for the genomes of domestic animals. (unpublished) 1990.

10 Krawetz, S. A. and Dixon, G. H.: Sequence similarities of the protamine genes: Implications for regulation and evolution. J. Mol. Evol. 27: 291-297, 1988.

11 Skow, L. C.: personal communication.

12 Faloona, F., Mullis, K. B.: Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. Meth. Enzym. 155:335-350, 1987.

13 Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. and Arnheim, N.: Enzymatic amplification of ß-globin genomic sequences and restriction site analysis for the diagnosis of sickle cell anemia. Science 230:1350-1354, 1985.

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

15 Krawetz, S. A. and Dixon, G. H.: Sequence similarities of the protamine genes: Implications for regulation and evolution. J. Mol. Evol. 27: 291-297, 1988.