STRUCTURE AND EVOLUTION OF *b* ALLELES IN MAIZE

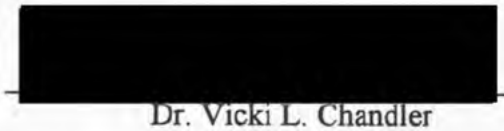
by

THERESA MICHELLE MERRILL

A THESIS

Presented to the Department of Biology
and the Honors College of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Bachelor of Arts

June 1996

APPROVED: 

Dr. Vicki L. Chandler

An Abstract of the Thesis of

Theresa Michelle Merrill for the degree of Bachelor of Arts

in the Department of Biology to be taken June 1996

Title: STRUCTURE AND EVOLUTION OF *b* ALLELES IN MAIZE

Approved: [redacted]

Dr. Vicki L. Chandler

The *b* gene in maize encodes a regulatory protein that activates the anthocyanin biosynthetic pathway. Numerous *b* alleles have been discovered, and induce the production of anthocyanins with varying tissue specificity and pigment levels. To date, three *b* alleles, *B-I, B-Peru,* and *B-Bolivia* have been characterized at the molecular level, and the upstream regions of each allele has been shown to produce the tissue-specific pattern of expression characteristic of each allele. Uncharacterized portions of the *B-Peru* and *B-Bolivia* alleles may be responsible for part of the tissue-specific expression, and could be essential to understanding how the upstream regions regulate the expression of anthocyanins. I have attempted to characterize these sequences. A yeast artificial chromosome (YAC) containing a *b* gene has also been characterized, and compared to the *b* gene found in the strain from which it was taken, LH82. Upstream probes will be developed from the YAC, and used to further characterize the *b* alleles.

ACKNOWLEDGEMENTS

The author would like to express sincere appreciation to Dr. Vicki Chandler for providing the opportunity to work and learn in her lab, as well as her guidance and advising of my research and this thesis. In addition, special thanks are due to Dr. David Selinger and Dr. Kenneth Kubo for their patience, advice, and many hours spent teaching me how to perform the laboratory techniques necessary to complete this thesis. I also thank the other members of the Chandler Lab for their interest, good humor, and helpfulness during my time spent in their laboratory.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

In their book Fundamentals of Molecular Evolution, Li and Graur (1991) propose that DNA, deoxyribonucleic acid, is not only the key to heredity, but an accumulation of historical records, as well. The challenge for geneticists lies in deciphering the many clues evolution has left behind to answer the primary question: How do genes evolve? In many cases, evolution has left a large trail of evidence in the form of recognizable patterns in the DNA sequence. This evidence is collected using laboratory techniques to discover how related alleles, or alternate forms of a gene, in the genome of an organism have undergone changes that result in dramatically different phenotypes. Phenotypes are physical manifestations of the characteristic specified by a gene, and common examples include blue eyes or brown hair. Eventually, this information will lead to an understanding of how evolution occurs at a molecular level.

DNA is an amazing accomplishment of the force of life. It is often referred to as "the master molecule of life", because it is found in every cell and serves as a genetic blueprint for almost everything an organism needs to develop and function. This genetic information is presented in the form of genes. DNA is a long, double-stranded molecule that is wound in a double-helical shape. It is composed of four basic molecules that each

include a phosphate group, nitrogen base, and a sugar. These are adenine, guanine, cytosine, and thymine, which are often depicted by the letters A, G, C, and T. These molecules are called nucleotides and form the building blocks of DNA. One of the central properties of DNA is its observation of complementary base-pairing; adenine and thymine are complementary and base-pair with each other, as do guanine and cytosine. This complementarity allows the two strands to come together in much the same way as one would zip up a zipper. DNA sequence information in one strand is sufficient to dictate the nucleotide sequence in the other strand, as well. The nucleotide sequence of a DNA molecule encodes chains of amino acids which form proteins. Proteins carry out most of the vital functions of the organism. DNA sequences not only encode proteins, however, but contain information specifying when and where that protein should be made.

In order to examine the differences that result from gene evolution, I have characterized portions of three maize alleles that encode the same protein, but express it in different places in the plant. Maize is a fairly large plant, and therefore phenotypes are easy to detect; it has been intensively studied at the molecular level, and many techniques have been developed in order to examine its genome structure and evolutionary patterns. Maize is an excellent organism to study because of the powerful array of genetic tools available and the ease of genetic manipulation. It is a recently evolved organism, as well; it was domesticated within the past 10,000 years (Doebley, 1992) and various traits have been selected for by man. The maize genome is highly complex and is about twice the size of mammalian genomes (Sprague, et al., 1988).

Therefore, the study of maize is definitely worthwhile in that it will further our understanding of how genes evolve in other complex genomes.

The focus of this thesis is the maize *b* gene, which encodes a regulatory protein that activates the anthocyanin biosynthetic pathway. Anthocyanins are purple pigments that form in seed and plant tissues, resulting in easily visible phenotypes (Dooner, et al., 1991). Over thirty *b* alleles have been discovered that induce the anthocyanin pathway in different tissues and stages of development (Coe, 1979). Three of these alleles, *B-I, B-Peru,* and *B-Bolivia*, all encode a functional B protein, which plays an essential regulatory role in the synthesis of anthocyanins. However each allele displays distinct differences in tissue specificity (see Figure 1). *B-Bolivia* is darkly pigmented in the culm and sheath tissues of the plant, and the aleurone of the seed is lightly and variably pigmented. In contrast, *B-Peru* is weakly pigmented in the plant, and strongly expressed in the aleurone (Patterson, et al., 1991). The *B-I* allele, however, shows dramatically enhanced plant pigmentation when compared to *B-Bolivia* and *B-Peru*, but there is no pigmentation in the aleurone of the seed.

What causes the differences in pigmentation conferred by the *B-I, B-Peru*, and *B-Bolivia* alleles? Comparison of the sequences that encode the protein shows that they are essentially identical, with 98% identity (Radicella, et al., 1992, D. Selinger, unpublished data). Initial examination of the nearby upstream regions of these alleles, however, demonstrates that these sequences are quite distinct from each other (see Figure 2). Upstream regions usually contain regulatory sequences that can positively or negatively affect whether the gene is expressed. In fact, a *B-I* upstream sequence has been joined to

a *B-Peru* coding sequence by genetic manipulation, and the pattern of pigment synthesis observed. This hybrid gene produces a *B-I* phenotype. Conversely, attaching an upstream sequence from *B-Peru* to a coding sequence from *B-I* results in expression of the *B-Peru* phenotype. The upstream regions of *B-I* and *B-Peru* are therefore important for determining where pigment will be expressed in the plant (Patterson, et al., 1995). While the upstream region of *B-I* has been well characterized, sequences in *B-Peru* and *B-Bolivia* upstream regions remain unknown (see Figure 2). It is portions of these upstream regions of the *B-Peru* and *B-Bolivia* alleles that I have attempted to compare and characterize through restriction mapping.

A restriction map details where a particular enzyme can cut DNA into smaller fragments. Restriction enzymes are isolated from various bacteria and cut DNA at specific sites, like precise molecular scissors. For example, the enzyme *Eco*RI recognizes the nucleotide sequence G^AATTC and cleaves the DNA after the guanine on CTTAA^G both strands. Depending on the number of sites that the enzyme can recognize, many fragments of DNA will be produced. The fragments are loaded into an agarose gel, where they are subjected to voltage and allowed to differentially migrate by size in a technique called gel electrophoresis. Because DNA carries a negative charge, the fragments will move toward the positive electrode from the negative end of the gel. The smaller fragments will be able to move through the gel faster than the larger fragments, and over the course of a few hours, the fragments will be spread throughout the gel by size. Different *b* alleles will produce a distinct pattern of fragments, according to their

individual sequences. This conventional method of gel electrophoresis can resolve fragments up to 25 kb in length (Ausubel, et al., 1994).

An extension of this technique is called pulsed-field gel electrophoresis, which is used when the DNA fragments are too large to be separated by conventional gel electrophoresis and migrate at the same rate through the gel. A computer program alternates the direction of the voltage so that larger fragments, up to 2000 kb in size, can be resolved (Ausubel, et al., 1994). When DNA fragments subjected to an electric field move through an agarose gel, they assume extended conformations that allow them to squeeze through the pores in the agarose. The variation of the orientation of the electric field used in pulsed-field gel electrophoresis forces larger molecules to assume a different conformation in order to move in a different direction. The larger the molecule, the longer this change of conformation will take, thus allowing pulsed-field techniques to utilize this time difference in order to allow separation of large DNA fragments (Ausubel, et al., 1994). The alternating field also prevents the large fragments from bypassing the pores in the agarose, and snaking through the gel. Instead, the fragments are forced to squeeze through the pores, which also helps to separate them.

Once the fragments have been separated by size through gel electrophoresis, a Southern blot can be done to identify the DNA fragments with which a known sequence of DNA, called a probe, is able to base-pair with. Because maize DNA contains ~3 x 10^9 base pairs, any restriction digest will produce millions of fragments. Southern blotting allows the identification of a specific fragment in the midst of these millions (see Figure 3). The gel containing the DNA fragments is washed in buffers that denature the double-

stranded fragments into single strands of DNA. The fragments are then transferred to a nylon membrane in positions that directly correspond to their placement in the gel. Radioactively labeled DNA probes that allow identification of specific *b* gene sequences are made and allowed to hybridize with the fragments on the membrane. Due to the complementarity of DNA, the probes will seek out sequences that complement their own and base-pair to them. The membrane is then exposed to film, and when the radioactivity in the probe decays, it activates the silver grains in the film and indicates where the probe hybridized. This information allows identification of large fragments of unknown DNA that contain sequences the probes can hybridize to. A restriction map can then be created, which details where each enzyme cuts the DNA of the *b* gene. Differences in DNA structure between *b* alleles are reflected by differences in where the restriction enzymes cut the DNA, and result in different sized bands.

Another allele of the *b* gene that I have studied is found in a stock named LH82. The plant and seed from this stock do not exhibit anthocyanin pigmentation, (Kenneth Kubo, personal communication) suggesting that the B protein encoded in LH82 is non-functional. However, LH82 is not only interesting because it may not have an intact coding region for the B protein, but it is also the strain whose *b* gene was inserted into a YAC library. A YAC is a yeast artificial chromosome into which large portions of DNA (100,000-300,000 bp) containing multiple genes from another organism can be cloned. These YACs replicate faithfully in a living yeast cell, and can serve as vehicles for obtaining many copies of the gene of interest, which is in this case, the *b* gene in maize.

Characterization of the far upstream regions of the LH82 *b* allele found in the YAC may yield probes useful in further characterization of *B-I, B-Peru,* and *B-Bolivia.*

In addition to the ease of experimental manipulation when working with YACs, the yeast strains containing the YAC can be easily grown and maintained, and do not suffer from the large amounts of repetitive DNA that is commonly found in the maize genome. A YAC library consists of the entire maize genome inserted into many YACs. Each yeast cell in a library contains a different artificial chromosome, which accomodates a small portion of the entire maize genome. A YAC from this library was isolated and sent to us by Zeneca seeds, producers of the initial YAC library who kindly shared this clone with us, by probing with sequences from the *b* gene

Before the gene located in the YAC can be regarded as another allele of the *b* gene, however, it must be determined that the *b* gene contained within the YAC has not suffered from rearrangements or mishaps of cloning. In addition to studying the upstream regions of the three *b* alleles, *B-I, B-Peru,* and *B-Bolivia,* I have also created a map detailing both the *b* gene in the YAC, as well as the *b* gene in the stock from which the gene in the YAC was cloned, LH82. Comparison of the Southern blots has assisted in determining whether the gene located in the YAC is truly representative of the *b* gene from the LH82 stock, and has provided the opportunity to study another *b* allele, in addition to *B-I, B-Peru,* and *B-Bolivia.*

CHAPTER 2

RESULTS

B-Peru and *B-Bolivia* Mapping

In order to characterize the unknown regions of *B-Peru* and *B-Bolivia,* restriction analysis that involved the use of previously determined sites, as well as new enzymes, was performed. Construction of a restriction map often requires double-digests to place restriction sites. The use of two enzymes in a double restriction digest can be very helpful in determining whether a site is on the "inside" or "outside" of another site. If the DNA were digested with enzyme A alone, enzyme B alone, and then enzymes A and B together, comparison of the resulting fragments could show that the DNA fragments digested with enzyme A alone and A and B together would be the same size, while fragment B would be larger. Therefore, the second restriction site for enzyme A would be on the *inside* of restriction site B. Conversely, if site A were on the *outside* of site B, the fragments digested with enzyme B alone, as well as A and B together would be the same size, while fragment A would be larger (see Figure 4).

B-Peru DNA was digested with ~22 different enzyme combinations, and ~32 different enzyme combinations were used on *B-Bolivia* DNA. The resulting fragments

were analyzed using pulsed-field gel electrophoresis, and current working maps of these two alleles are shown in Figure 5 and Figure 6.

A double-digest using *Sna*BI and *Eco*RI revealed the presence of an *Eco*RI site downstream from the coding region in *B-Peru*. *B-Bolivia* was also found to contain an *Eco*RI site beyond the downstream *Xba*I site using *Xba*I and *Eco*RI in double-digests. A newly discovered downstream *Eco*RI site in *B-Bolivia* was used as a point of reference, and double-digests were performed with rare-cutting enzymes and *Eco*RI in order to generate large fragments of *B-Bolivia* DNA. These large fragments were separated by pulsed-field gel electrophoresis and probed with the 550b probe. Southern blotting revealed that the ends of the doubly-digested fragments extended into the first unknown region of *B-Bolivia*, and indicated the existence of a ~35 kb EcoRI fragment.

Subsequent re-probing of the blot with the BIu4 probe, however, did not reveal a smaller *Eco*RI fragment, as one would expect from the *B-Bolivia* map (see Figure 6), but instead hybridized to the same ~35 kb fragment. The double-digests involving the rare-cutters also did not reveal new fragments to correlate with *Eco*RI sites flanking the BIu4 probe, but instead hybridized to the same fragments as the 550b probe. The exception consisted of a digest involving *Apa*I, which did behave as predicted from the earlier experiment and demonstrated hybridization to a fragment of different length when probed with BIu4. These observations could have resulted from a mapping error that placed the *Eco*RI site downstream of the BIu4 probe. Old blots were examined, however, and it appears that the placement of this *Eco*RI site was valid. Another explanation could be that the 550b probe was not thoroughly stripped before reprobing

with BIu4. The intensity of the bands and evidence of hybridization to different *Apa*I fragments after reprobing suggests that incomplete stripping of the 550b probe did not cause the observed results.

The experiment was repeated, however this time both *B-Peru* and *B-Bolivia* DNA were used. The digests involving *Apa*I and *Eco*RI were repeated, and double-digests with rare-cutters and *Xba*I were performed, as well. This second pulsed-field Southern blot was probed with 550b, and the existence of downstream rare-cutter sites, including the *Apa*I, *Nhe*I, and *Sac*II sites in *B-Peru* was revealed (see Figure 5). A subsequent blot was made with *B-Peru* and *B-Bolivia* DNA that had been digested with the rare-cutting enzymes *Apa*I, *Nhe*I, and *Sac*II, and double digests with the rare-cutters and an enzyme called *Bcl*I were also performed. *Bcl*I was chosen because several *Bcl*I sites have been discovered in both *B-Peru* and *B-Bolivia* (see Figures 5 and 6) and they appear to be close to the unknown region I have been attempting to characterize. The blot was probed with 550b with the expectation that the upstream end of fragments produced by rare-cutting enzymes should be found in or near the uncharacterized sequence in *B-Peru,* as well as in the downstream uncharacterized sequence in *B-Bolivia.*

Analysis of this blot allowed the placement of a downstream *Apa*I site in *B-Bolivia*, another *Apa*I site in the downstream uncharacterized region of *B-Bolivia*, and the existence of the downstream *Eco*RI site mentioned previously was also verified (see Figure 6). Unfortunately, some of the rare-cutting enzymes failed to cut the DNA in several of the digests located on this blot, and so another blot was made to determine the upstream ends of the *Apa*I, *Nhe*I, and *Sac*II rare-cutting fragments in *B-Peru.* Additional

digests of *B-Bolivia* DNA involved the rare-cutting enzymes *Pac*I, *Nar*I, *Sma*I, and *Not*I, and were included on this blot in the hopes of gaining additional information. The blot has been probed with 550b and results are pending. Future reprobings of this blot, however, will include the BIu4 probe. This will allow determination of the upstream end of the *Apa*I fragment that has a downstream site in the first unknown region of *B-Bolivia* (see Figure 6). The E/G 700 probe will also be used to reprobe this blot, and determine rare-cutting sites even further upstream. Double-digests can then be performed to determine other sites within the uncharacterized sequences of *B-Peru* and *B-Bolivia*, using the rare-cutting sites as anchors.

b YAC and LH82 Comparison

The LH82 *b* allele in the YAC had not been fully characterized in the past, and so a restriction analysis was undertaken to create maps useful in comparing this allele with the other characterized alleles, *B-I, B-Peru,* and *B-Bolivia.* A Southern blot containing digested YAC DNA was created and probed with the 550b probe. A separate blot containg LH82 maize DNA digested with the same enzymes was created and probed with 550b as well. After exposure to film, the previous probe was melted off and the blots were reprobed with the BIu4, E/G 700, and 1.6 BglII probes (see Figure 2). Comparison of the films from both blots probed with 550b indicates that as expected, the *b* sequences in the YAC and LH82 maize DNA share very similar sequences. For example, in both DNA samples, the *Xba*I, *Hind*III, *Xba*I/*Hind*III, and *Xba*I/*Bam*HI digests demonstrate very similar fragment lengths when probed with 550b (see Figures 7 and 8). This suggests that no rearrangements occurred in the process of cloning the maize DNA into the YAC.

Some differences were noted, however, including a ~14 kb EcoRI band found in the YAC *b* gene, while the LH82 blot shows a weak *Eco*RI band of only ~5 kb, as well as a ~10 kb *Pst*I band in the YAC that is only ~5 kb and weak on the LH82 blot (see Figures 7 and 8). These discrepancies, however, could be attributed to the difficulties of transferring large DNA fragments onto nylon membrane, especially from large genomes such as maize. This would result in the apparent lack of a fragment in maize due to its

inability to transfer, and the weaker fragments observed in maize genomic DNA could represent the *r* gene in the maize genome. The *r* gene has been found to share some homology with the *b* gene in its coding region, and the 550b probe weakly hybridizes to regions of the *r* gene, resulting in the appearance of weaker bands in LH82. This hypothesis could be tested by probing LH82 blots with *r* gene probes. Unfortunately, the LH82 blots were not tolerant of subsequent re-probings. Future experiments, therefore, will include the creation of additional LH82 blots in order to fully compare the two *b* genes, to confirm which fragments can be attributed to the *r* gene, and to verify the placement of the sites on the newly created maps.

It was observed through restriction analysis that a *Bam*HI site appeared to be lacking in the coding sequence of the LH82 allele. This observation coupled with the subsequent analysis of the sizes of other fragments led to the proposal that the LH82 *b* allele had suffered from a deletion in a portion of its coding sequence. For example, in *B-Peru* and *B-Bolivia* the length of the *Pst*I/*Bgl*II fragment that includes exons 2 and 3 is ~900bp, while in the LH82 allele it is only ~200bp (see Figures 5, 6, and 9). To determine the validity of this proposal, and how much of the sequence might be missing, I cloned the interesting portion of the gene using PCR, and tried to determine the actual sequence of the base-pairs through dideoxy DNA sequencing.

PCR, polymerase chain reaction, is a useful procedure that allows the researcher to create multiple copies of a certain gene fragment. Oligonucleotide primers bind to the regions flanking the fragment of interest and serve as the starting point for the polymerase. The triphosphate nucleotides deoxyadenine, deoxyguanine, deoxycytosine,

and deoxythymine serve as building blocks with which the polymerase can construct many copies of the particular fragment bounded by the primers, and template YAC DNA is used as the blueprint. After many rounds of denaturation of the fragment templates, annealing of the primers to the fragments, and synthesis of new fragments, a high concentration of DNA has been obtained. These newly created copies of the YAC DNA fragment of interest were then ligated into plasmids. Plasmids are extrachromosomal pieces of circular DNA which are capable of receiving DNA fragments amplified by PCR. The plasmids are then transformed into *E. coli* bacteria, where one can take advantage of the bacterial DNA replicating machinery to produce large amounts of the plasmid DNA containing the fragment of interest.

After the DNA has been removed from the bacteria and purified in order to remove contaminants such as the bacterial DNA, the fragment can now be sequenced. Sequencing determines the actual "genetic code", comprised of the individual nucleotides, adenine, guanine, cytosine, and thymine (see Figure 10). Primers are annealed to the purified plasmid DNA containing the fragment of interest, and allowed to extend their length with the aid of an enzyme called Sequenase. The enzyme adds nucleotides to the primer, causing it to grow in length, and uses the DNA it is annealed to as a template. One of the nucleotides, however, has been radioactively labeled, and so the fragments that incorporate the labeled nucleotide will later be detected when exposed to film. The enzyme is also provided with dideoxynucleoside triphosphates (ddNTPs) to use as substrates in the growing primer chain, though at a lower concentration than the normal deoxynucleoside triphosphates (dNTPs).

Elongation is terminated when ddNTPs are incorporated into a lengthening chain because the 3' hydroxyl group necessary to form the bond connecting the next deoxynucleoside triphosphate molecule is missing. The addition of the ddNTPs is considered the termination reaction. A termination reaction is set up for each of the four nucleotides; for example, in one reaction, the elongation will stop after a ddNTP specific for guanine has been incorporated, while in the second reaction, it will stop after the addition of ddATP, which contains adenine. However, both reactions are using DNA from the same source as a template. Because the ddNTPs are at a much lower concentration than normal dNTPs, chain termination will occur at different positions on different molecules, rather than solely at the first possible position. Therefore, a series of fragments that differ only by a few base pairs can be generated. Reactions containing termination steps for each of the four ddNTPs are loaded into a polyacrylamide gel, and the resulting fragments are then separated by size using electrophoresis. The gel is dried and exposed to film. The sequence can be "read" directly from the film, starting at the bottom with the smallest fragments, and continuing upwards, base by base. If, for example, the very smallest fragment ended with the addition of ddATP, the next smallest fragment with ddTTP, and the next with ddCTP, the first three nucleotides in the sequence are ATC.

Sequencing reactions were attempted on a portion of the YAC *b* gene that extends from the downstream *Bgl*II site, to the upstream end of the BIu4 probe, however, only portions of the actual sequence have been able to be determined due to overheating of the gel during electrophoresis. This caused the fragments to migrate in a curved manner, and

many are smeared and unable to be distinguished from the other fragment. Approximately 300 nucleotides are distinguishable from the films, and are being analyzed. I will continue to persist in sequencing attempts of this allele in the future, because determining the actual sequence of base-pairs should prove helpful in placing the sites determined by restriction analysis more accurately. In addition, knowing this sequence will be critical for verifying that this allele has indeed lost a portion of its coding region, which is believed to extend from the first *PstI* site in the second exon of other *b* alleles, to the *BglII* site in the third exon. This loss of part of the coding region would explain its inability to produce anthocyanins. In the future, sequencing of the same region in LH82 will be performed as well. This experiment will allow direct comparison of the base-pair sequences between the *b* gene located in the YAC and that of LH82, as well as aid in verifying that the sequence obtained for the YAC *b* gene was correct.

CHAPTER 3

DISCUSSION

I have characterized portions of the upstream regions of *B-Peru*, and *B-Bolivia*, as well as created maps of the YAC/LH82 *b* allele. These alleles are distinct from each other both in the intensity and tissue specificity of anthocyanin production, as well as in the genomic sequences that regulate this pigmentation.

Close comparison of the *B-I* and *B-Peru* upstream sequences demonstrates that they share similar sequences that hybridize to the 1.6 Bgl, E/G 700, and BIu4 probes. However, in *B-Peru*, it appears that an insertion may have occurred between the E/G 700 and BIu4 hybridizing regions in *B-I* (see Figures 2 and 5). It is probable that another insertion added the three repetitive sequences near the coding region, as well.

Although the evolutionary relationship between *B-Peru* and *B-I* remains unclear, it is interesting to note their phenotypes with reference to their genotypic characteristics. The plant tissues in the *B-I* allele are darkly pigmented, although the aleurone of the seed is not. In contrast, *B-Peru* expresses anthocyanins with high intensity in the seed, as well as in some plant tissues. Pigmentation is manifested as small spots of purple in tissues such as the sheath and culm tissues of the *B-Peru* plant, while total pigmentation is observed in plant tissues homozygous for the *B-I* allele (see Figure 1). Recent

experiments have indicated that the repetitive sequence closest to the coding region in *B-Peru*, as well as a portion of the second repeat sequence, can regulate pigmentation of the seed (D. Selinger, unpublished data). Therefore, the lack of pigmentation of the seed in *B-I* could be due to the absence of such sequences in its upstream region (see Figure 2). Accordingly, the distinctive and intense dark purple pigmentation of *B-I* may be regulated or enhanced by sequences that are not found in the regulatory regions of *B-Peru*, or are displaced thousands of base-pairs away due to the insertions.

Comparison of the upstream sequences of *B-Peru* and *B-Bolivia* indicates that additional insertions may have occurred in *B-Bolivia*, between the first and second repetitive sequences that are found in *B-Peru* (see Figures 2, 4, and 5). The seed color in *B-Bolivia* is dramatically decreased compared to that of *B-Peru*, and the pigmentation in the plant tissues is markedly different as well. While *B-Peru* produces spots of pigmentation in the culm and sheath, *B-Bolivia* demonstrates solid pigmentation except for the margins of the sheath and glumes, and the nodes are not pigmented. Neither allele produces anthocyanins in the leaves (see Figure 1). The decrease in seed pigmentation could be due to the shift of the second (and possibly the third) repetitive sequence upstream when an insertion occurred, resulting in necessary portions of the second repeat sequence to be located too far from the start of transcription to influence seed pigmentation (see Figure 6). This is not likely, however, because transgenic analysis has demonstrated that only one repeat sequence is necessary to confer normal pigmentation (D. Selinger, unpublished data). The differences in plant pigmentation most likely result from either the inserted sequence in *B-Bolivia*, which could contain

negative regulatory sequences, or possibly from consequences of the insertion, such as changing the distance that spans from the coding region to pre-existing upstream sequences that may play a role in regulating pigmentation.

The LH82 allele appears to have suffered from genomic rearrangements, as well. It is missing the *Bam*HI site usually found in the second intron of other *b* alleles (see Figures 5,6, and 9), while the *Bgl*II site usually located in the third exon has shifted upstream. This implies that a portion of the coding region has suffered a deletion, and would produce a non-functional protein if transcribed and translated. These observations will soon be verified through sequencing, and are most likely the reason why the LH82 allele does not produce anthocyanins in either the plant tissue or the seed.

Such rearrangements in the genome of an organism are not unusual, and can occur through several mechanisms. Chromosomal rearrangements can result from recombination, when chromosomes exchange information during the formation of gametes at meiosis (Griffiths, et al., 1993). Spontaneous mutations can result from errors in DNA replication that can cause base-pair substitutions, deletions, and duplications, naturally occurring damage to the DNA can induce mutation, and mutations can result from the actions of transposable elements (Griffiths, et al., 1993). Transposable elements were first discovered in maize, and have been found to be active in most organisms. They are sequences that have the ability to move from one site to another in the genome, and often cause large-scale rearrangements through incomplete excision. Usually, incomplete excision is observed when the element leaves portions of its own sequence while removing portions of the host genome as it leaves. Therefore, it is not difficult to

imagine how the addition of a new sequence in the middle of a gene could affect not only the product of that gene, but the organism itself. A large number of transposable element systems have been characterized in maize, and its evolution could be attributed in part to the mechanisms of transposable elements.

Although the insertions observed in the *B-Peru* and *B-Bolivia* alleles may or may not have been facilitated through transposable elements, the sequences are candidates for future research to address their possible roles in tissue specificity of anthocyanin pigmentation. As stated previously, the first repeat sequence and part of the second in *B-Peru* have been demonstrated to direct pigmentation in the seed when attached to the coding region of the gene (see Figure 5), and the phenotype observed in the manipulated plant (D. Selinger, unpublished data). *B-Peru* also demonstrates unusual pigmentation in the plant, which is exhibited by small purple spots (see Figure 1). The linear mechanisms of cell division that occur in plant tissue would predict stripes or streak-like expression of anthocyanins, instead of the circular spots that are observed in *B-Peru.* It has therefore been suggested that the sequence that will be characterized may play an essential role in conferring the spot-like phenotype upon the *B-Peru* allele, and should be studied further.

Because yeast DNA contains a much lower percentage of repetitive sequences, cloning will be more easily facilitated through use of the YAC, and regions of DNA hundreds of thousands of kilobases in size can be manipulated. Therefore, after the gene in the YAC has been confirmed to be representative of the LH82 allele, it can be used to make probes useful in future research involving the structure of the three *b* alleles, as

well as study another interesting phenomenon, paramutation. Paramutation is an allele-specific interaction that heritably alters transcription, and occurs at the *b* gene when a *B'* allele and a *B-I* allele are heterozygous (Patterson, et al., 1995). *B'* appears to share the same upstream and coding regions that have been characterized in *B-I,* however, pigmentation in *B'* is dramatically decreased. The interaction between the two alleles results in the alteration of *B-I* to *B'* every time, and the different phenotypes are caused by different rates of transcription of the *b* gene (Patterson, et al ., 1993). Extensive mapping of both *B-I* and *B'* was performed to determine what sequences could be responsible for paramutation, and although studies involving recombination determined that these sequences are located upstream of the coding region, no differences have yet been detected between the *B'* and *B-I* alleles (Patterson, et al., 1995). Therefore, extended upstream probes isolated from the YAC could prove useful in discovering differences between the two alleles that might be responsible for the observation of paramutation in *B-I* as a result of *B',* as well as aid in the continuing characterization of the upstream regions of *B-Peru* and *B-Bolivia.*

Characterization of the sequences described in this thesis are a small contribution to the many puzzles that still remain. These include the mechanisms that confer specificity and intensity of anthocyanin expression as a result of sequences present in the upstream regions of the *b* gene, as well as the interactions between *b* alleles that result in paramutation.

CHAPTER 4

MATERIALS AND METHODS

Obtaining the DNA

Maize DNA was extracted from seedlings homozygous for either the *B-I, B-Peru, B-Bolivia,* or LH82 alleles that were grown in the greenhouse for 7-10 days. Leaf tissue was frozen with liquid nitrogen and ground into a fine powder using a mortar and pestle. The powdered leaf tissue was incubated with high-salt extraction buffer at 65°C for ten minutes. 5M potassium acetate was added and the tissue was incubated on ice for twenty minutes in order to allow the proteins and polysaccharides to precipitate. After centrifugation, the DNA was recovered from the supernatant by ethanol precipitation and resuspended in Tris-Cl, EDTA (TE).

Yeast cells containing the LH82 YAC, which was obtained from the Zeneca Seeds library, were grown overnight in media that does not contain uracil. This selects for the cells that contain the YAC, because the YAC encodes an enzyme that allows growth in the absence of uracil. After centrifugation, the cells were washed with water, centrifuged again, and then incubated for an hour at 37°C with DNA prep buffer, zymolyase solution, and β-Mercaptoethanol. Cell lysis buffer and 5M sodium chloride

were added and the lysed cells were incubated on ice for an hour. After centrifugation, the DNA was resuspended in TE and extracted with phenol:chloroform. 3M NaOAc and ethanol were added to precipitate the DNA, and after washing with 70% ethanol, the DNA was resuspended in TE. The presence of the YAC was verified through Southern blotting and the use of specific *b* probes. YACs containing an *Arabidopsis thaliana* gene that does not hybridize to *b* probes were used as controls.

Restriction Digests

The enzymes one uses in a restriction digest will dictate the size of fragments obtained; some enzymes cut DNA more frequently than others, and possess varying sensitivities and strengths. About 3-4 μg of DNA were digested with 1 μL (10-20 units) of the desired restriction enzyme in a reaction containing the appropriate 10X NEBuffer at 1X concentration. The rest of the volume consisted of sterile water. The reaction was incubated for three hours at 37°C.

Gel Electrophoresis

Standard Electrophoresis

In order to separate DNA fragments smaller than 15 kb, standard electrophoresis was employed. Loading dye at 1X concentration was added to the samples, which

contained digested DNA. The samples were loaded into a 0.6% or 0.7% agarose gel containing ethidium bromide, which interacts with the DNA and allows the fragments to be visualized in the gel using UV light. 1X TBE was used as the buffer. Molecular markers made of lambda phage DNA digested to known sizes were also loaded into the gel to serve as a "molecular ruler". The fragments in the gel were then subjected to voltage for several hours (usually 32 or 48 volts overnight), until good separation of the molecular markers had been achieved.

Pulsed-Field Electrophoresis

Pulsed-field electrophoresis is a variation of the standard electrophoresis that allows much larger fragments to be separated by size. A switchback controller that was programmed to separate fragments from 8.3 to 48.5 kb was connected to the power supply and to the gel electrophoresis box. The pulsed-field electrophoresis was conducted in a cold room maintained at 4°C to optimize the sharpness of bands and 0.5X TBE was used as the buffer. The samples containing the DNA fragments and 1X loading dye were loaded into a 1% agarose gel with ethidium bromide and subjected to 150 volts for 20-22 hours with a ramped pulse time of 0.7 to 2.3 seconds.

Southern Blotting

Denaturation and Transfer of DNA

After the DNA fragments have been separated through gel electrophoresis, a picture was taken of the gel and a ruler using a transilluminator that uses UV light to visualize DNA fragments stained with ethidum bromide. This picture was later used to determine how far fragments of a certain size had travelled through the agarose gel. The gel was left exposed to the UV light for two minutes to create nicks in the DNA, facilitating an easier transfer to the nylon membrane. The agarose gel containing the DNA fragments, now separated by size, was washed twice for twenty minutes in denaturing buffer, rinsed with distilled water, and then washed twice for thirty minutes in neutralizing buffer. The DNA fragments were transferred to a nylon membrane the same size as the gel, and 10X SSC was used as the wicking solution. The membrane, which now contained the DNA fragments, was then baked in the vacuum oven for two hours at 80°C and then washed in 0.1X SSC, 0.1% SDS solution for an hour at 65°C.

Making a Probe

The sequences used as probes had been previously inserted into plasmids, and so the probe DNA was isolated from the plasmid by digesting a large quantity of the

plasmid with the appropriate restriction enzymes. The chosen enzymes should cut on both sides of the inserted DNA fragment. The digested plasmid was loaded into a low melting point agarose gel, and run at 100 volts for several hours, until the small fragment of probe DNA had clearly separated from the larger plasmid fragments. The band containing the probe DNA was cut out of the gel, and the gel was melted until it liquified. The probe DNA was removed from the melted agarose gel and purified through a series of phenol, phenol:chloroform, and chloroform extractions. Ethanol and 3M NaOAc were added to precipitate the probe. After centifugation, the supernatant was removed and the pellet containing the probe DNA was resuspended in TE.

Probe Labeling and Hybridization

The nylon membrane, which contains the DNA fragments in the same positions that electrophoresis placed them in the agarose gel, was incubated in pre-hybridization buffer at 42°C for two to three hours. The buffer contains salmon sperm DNA that blocks sites on the nylon membrane that the probe could bind to non-specifically. 0.1 to 0.2 μg of probe DNA was denatured at 90-95°C and then cooled immediately on ice. The probe DNA was then added to the labeling reaction, which included a labeling solution composed of DTM-C, 1M Hepes 6.6, and Hexamer O.D. 90, BSA, Klenow Polymerase, and 50 μCi of ^{32}P d-CTP. In this reaction, the polymerase incorporates a radioactively labeled cytosine into copies of the probe DNA, which is used as a template. After three to five hours, the radioactively labeled probe DNA was precipitated with 5M

NH_4OAc and ethanol and placed on dry ice for several minutes. After centrifugation at 4°C for ten minutes, the radioactive supernatant was removed, and the probe DNA was resuspended in TE. After denaturing the probe DNA at 90-95°C a second time, it was added to the nylon membrane in the pre-hybridization buffer. The newly-labeled probe binds with much greater specificity than the salmon sperm DNA to fragments in the membrane containing the complementary sequence. After the probe has been allowed to hybridize for 24-36 hours, the membrane was washed in 0.2X SSC, 0.1% SDS at 58°C, wrapped in plastic wrap, and exposed to film at -80°C for three to five days. After the film was developed, dark bands indicated what fragments the probe hybridized to, and the size of these fragments was determined from their position on the membrane.

Cloning

Polymerase Chain Reaction

The DNA sequence in the YAC strain extending from the downstream *Bgl*II site to the 5' end of the BIu4 probe was amplified using two different PCR reactions. One reaction used the *Dde*I and MA1 primers to amplify the upstream third of the fragment of interest, while the GIP1 and *Bgl*II primers were used to create numerous copies of the remaining two-thirds of the fragment. The reactions consisted of TAQ Polymerase, buffer, $MgCl_2$, water, and deoxynucleotide triphosphates, as well as the appropriate primers and denatured YAC DNA, which was used as a template. A programmable

thermal controller was used to facilitate the denaturation, annealing, and synthesis of the fragments of interest.

Ligation

Plasmids pTZ18u and pTZ19u were cut with enzymes that created either staggered or shear cuts in the DNA, depending on the fragment to be inserted. This allows the DNA fragments obtained through PCR to fit into the open plasmid like a puzzle piece. The fragments were ligated into the plasmid in a reaction that involved T4 DNA ligase, buffer, the PCR fragment, and the open plasmid. The ligation reactions were incubated for several hours at 16°C.

Transformation

The plasmids containing the DNA fragment of interest were introduced into *E. coli* bacterial cells through electroporation. The cells were then incubated in SOC media and then plated with X-Gal and IPTG. The plasmids contain genes that allow resistance to certain antibiotics, therefore bacteria that had the plasmid successfully introduced into the cell will be able to grow on plates or in media containing the antibiotic, which was in this case ampicillin. The success of ligation was measured by testing whether the *lac Z* gene in the plasmid remained intact. If the PCR fragment was successfully incorporated into the plasmid, the *lac Z* gene would be disrupted and the bacteria would be unable to

produce β-galactosidase, resulting in white colonies. Cells that contain a plasmid that failed to ligate the PCR fragment into it are identified by the blue colonies that form as a result of the cleavage of X-Gal by β-galactosidase.

DNA Sequencing

DNA Preparation

The *E. coli* strains that contained a plasmid with the PCR fragment successfully ligated into it were grown overnight and then lysed with salt solutions to obtain the DNA. Centrifugation removed large impurities including cell walls or cells that did not lyse, and the supernatant was treated with RNAse A and incubated at 37°C for 30-60 minutes. The DNA was precipitated with ethanol, and centrifuged after extracting with equal volumes of phenol, phenol:chloroform, and chloroform. The supernatant was removed, and the pellet was resuspended in TE. A solution of 1.6M NaCl, 13% PEG 8,000 was added to the DNA and incubated on ice for one hour. The DNA was isolated through centrifugation, washed with 70% ethanol, and resuspended in TE.

Sequencing Reaction

Purified DNA was denatured with 1M NaOH, and incubated at room temperature for five minutes. Both universal and reverse primers were added, in separate reactions,

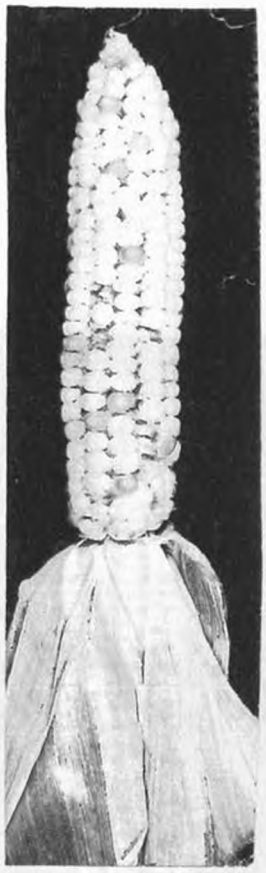
and annealed to the now single-stranded DNA. The DNA was precipitated with ethanol and 3M NaOAc, and pelleted in a microcentrifuge at 4°C. After the removal of the supernatant, the DNA was washed with 70% ethanol, centrifuged again, and the supernatant removed. A mixture containing 0.1M DTT, dGTP lab mix, water, and ^{35}S-dATP was added to the DNA, as well as the Sequenase enzyme. The DNA mixture was added to each of the termination mixes, ddGTP, ddATP, ddTTP, and ddCTP, after incubating at room temperature for five minutes. The termination reaction was incubated at 37°C for five minutes, and then stop mix was added. The DNA was denatured at 100°C for 3-4 minutes and then placed on ice.

Gel Analysis of the Sequencing Reaction

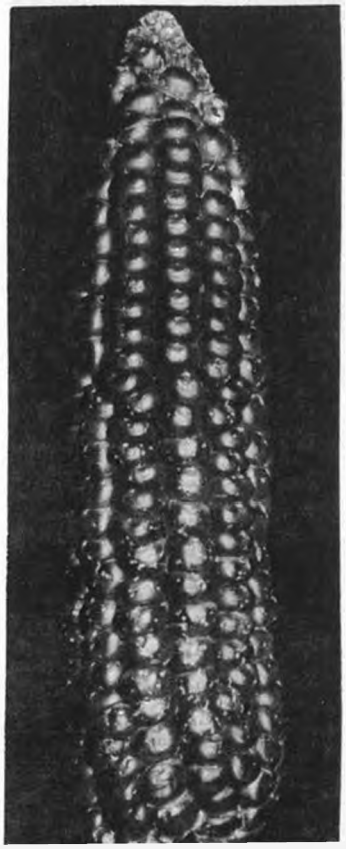
A 6% polyacrylamide/urea gel was poured between glass plates with a thin separation between them and allowed to polymerize. The denatured DNA samples were loaded into the gel, and subjected to 1750 volts for three hours. Buffers were utilized to create a salt gradient, with 0.5X TBE in the top well, and a 1X TBE, 3M NaOAc solution in the bottom well. The gel was then fixed in a 10% acetic acid, 12% methanol solution to remove the urea from the gel, and then dried under vacuum onto blotting paper. The paper with the gel dried onto it was exposed to film for two days, and the sequence was read from the film.

Figure 1: *Tissue and Seed Phenotypes of B-Bolivia, B-Peru, and B-I.* These pictures demonstrate the differences in pigmentation between the three alleles. The culm and sheath tissues of *B-Bolivia* are darkly pigmented, while the leaves do not produce anthocyanins. The aleurone of the seed is lightly and variably pigmented. *B-Peru* produces spot-like pigmentation in the culm, sheath, and husk tissues, however, the leaves are not pigmented. *B-Peru* colors the aleurone layer of the seed dark purple. *B-I* displays dark purple pigmentation in all vegetative plant tissues, however in the aleurone, it does not induce anthocyanin synthesis.

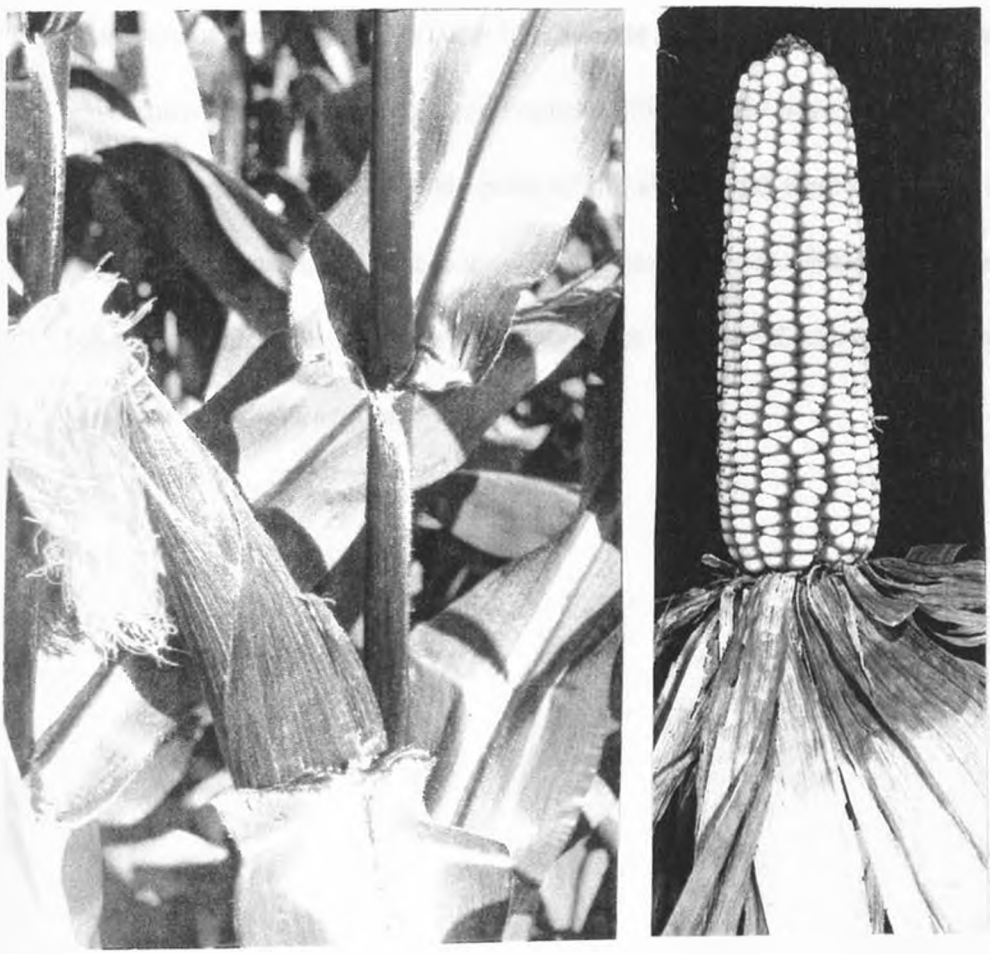
Tissue and Seed Phenotypes of B-Bolivia, B-Peru, and B-I



B-Bolivia



B-Peru



B-I

Figure 2: *Divergence of the Upstream Region of Three b Alleles.* In this diagram, the coding sequence is depicted on the right of the vertical line. The open boxes in this region correspond to known exons, which are the parts of the DNA sequence that end up in the mRNA, most of which is translated into protein. The upstream region is on the left of the vertical line. The boxes indicate known sequences, and the lines beneath the boxes correspond to probes that are known to hybridize to these regions. The jagged lines in between the known sequences indicate the parts of the alleles that were unknown, and repetitive sequences are indicated by the larger boxes with arrows. Notice that the alleles share similar sequences, but contain additional DNA insertions in their upstream sequences.

Divergence of the Upstream Region of Three b Alleles

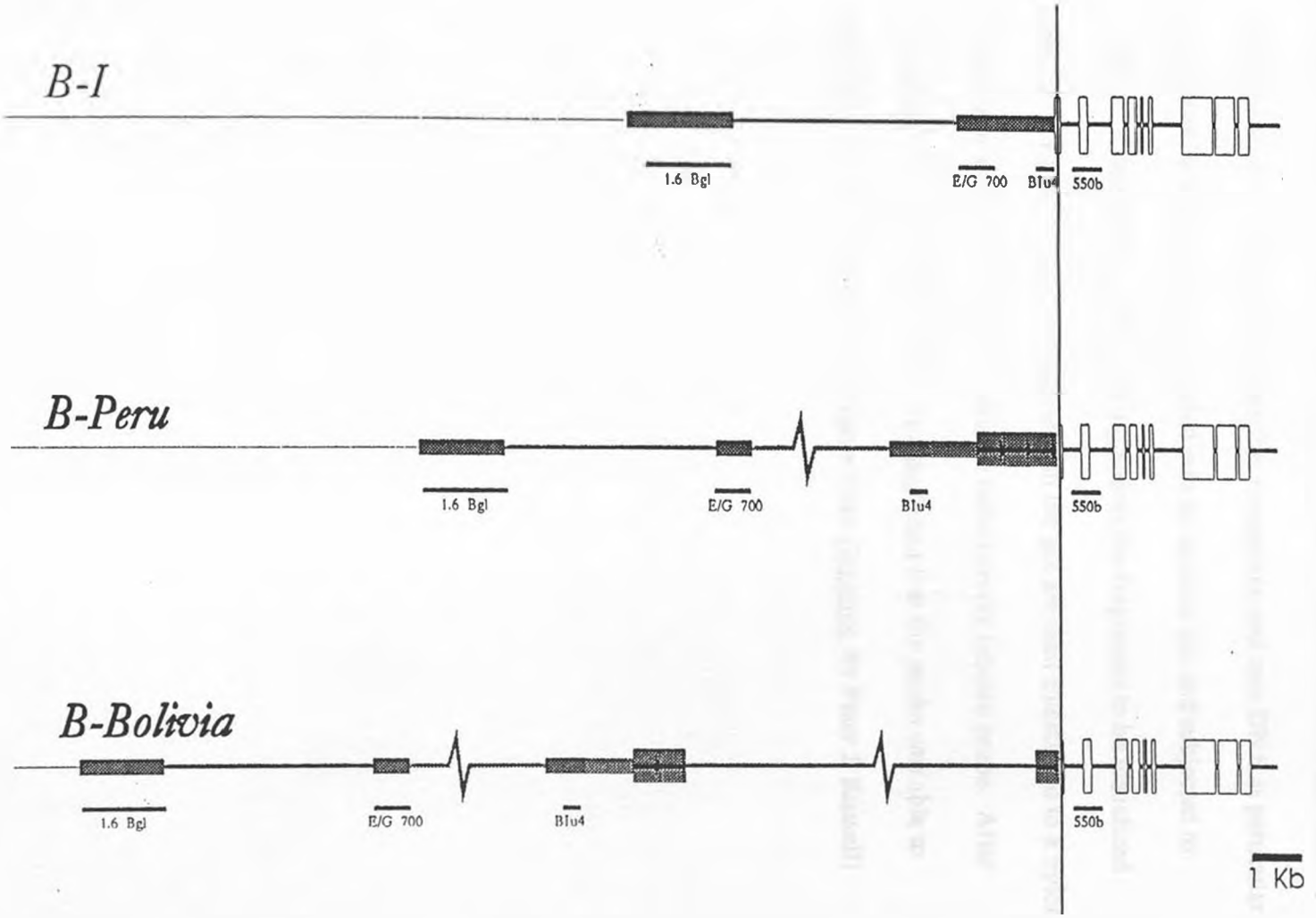


Figure 3: *Diagram Detailing Southern Blotting Procedure.* After the DNA has been digested with a restriction enzyme that recognizes and cuts DNA at particular sequences, the fragments are loaded into an agarose gel and subjected to voltage. Ethidium bromide staining allows the fragments to be visualized using UV light. The DNA fragments in the gel are then transferred to a nylon membrane and allowed to hybridize to radioactively labeled probe. After exposing the membrane to film, the fragments that the probe was able to hybridize to can be observed. (Figure from Genetics, by Peter J. Russell)

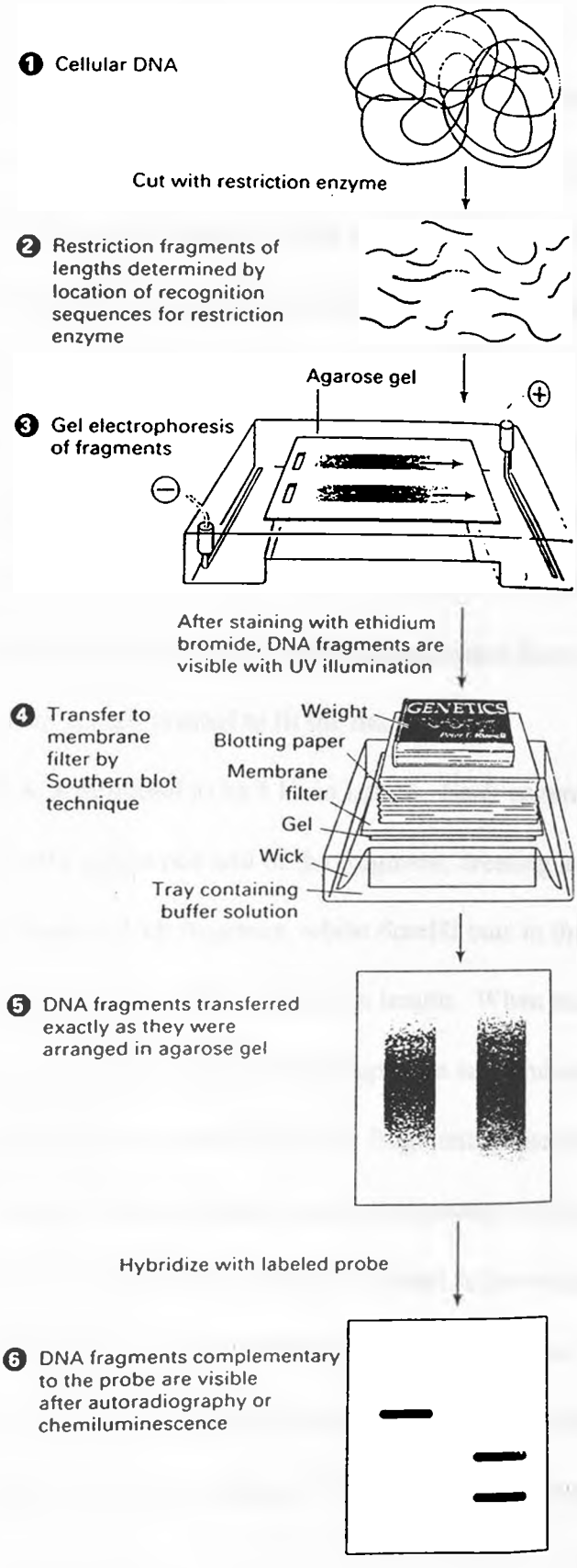

1 Cellular DNA
Cut with restriction enzyme
2 Restriction fragments of lengths determined by location of recognition sequences for restriction enzyme
Agarose gel
3 Gel electrophoresis of fragments
After staining with ethidium bromide, DNA fragments are visible with UV illumination
4 Transfer to membrane filter by Southern blot technique
Weight
Blotting paper
Membrane filter
Gel
Wick
Tray containing buffer solution
5 DNA fragments transferred exactly as they were arranged in agarose gel
Hybridize with labeled probe
6 DNA fragments complementary to the probe are visible after autoradiography or chemiluminescence

Figure 4: *Construction of a Restriction Map.* An agarose gel has been loaded with a marker that consists of particular DNA fragments cut to known sizes, serving as a "molecular ruler". An uncut fragment of the DNA in question is loaded into the gel, as well, in order to determine its total length. The other samples include the same DNA digested with *Eco*RI alone, *Bam*HI alone, and double-digested with *Eco*RI and *Bam*HI. Electrophoresis separates the fragments and a calibration curve is created using the distances that the known fragments in the marker were able to migrate. The size of the uncut DNA, as well as the DNA digested with *Eco*RI and *Bam*HI is determined from the calibration curve. A map is then created to fit the data.

The uncut DNA is determined to be 5 kb in length. Each enzyme cuts the DNA one time. *Eco*RI cuts at one end of the fragment, creating a small 0.5 kb fragment and a larger 4.5 kb fragment, while *Bam*HI cuts in the middle of the fragment, creating fragments 2 kb and 3 kb in length. When the two enzymes digest the DNA at the same time, three fragments are produced, such that it appears the *Bam*HI site lies within the 4.5 kb fragment created by *Eco*RI. Model B is chosen to demonstrate the correct placement of the restriction sites, because although the *Bam*HI site in model A lies within the 4.5 kb *Eco*RI fragment, it does not result in the correct fragment sizes.

Imagine that a radioactively labeled probe were allowed to hybridize to a Southern blot made from the gel in this figure. The 4.5 kb *Eco*RI fragment,

the 3 kb *Bam*HI fragment, and the 2.5 kb *Eco*RI/*Bam*HI fragment were visualized after exposing the blot to film. One can now determine that the region that the probe is able to hybridize to must be contained within the 2.5 *Eco*RI/*Bam*HI fragment, because this fragment is the smallest region that is common to all three fragments visualized on the film. (Figure from Genetics, by Peter J. Russell)

Construction of a restriction map for *Eco*RI and *Bam*HI in a DNA fragment.

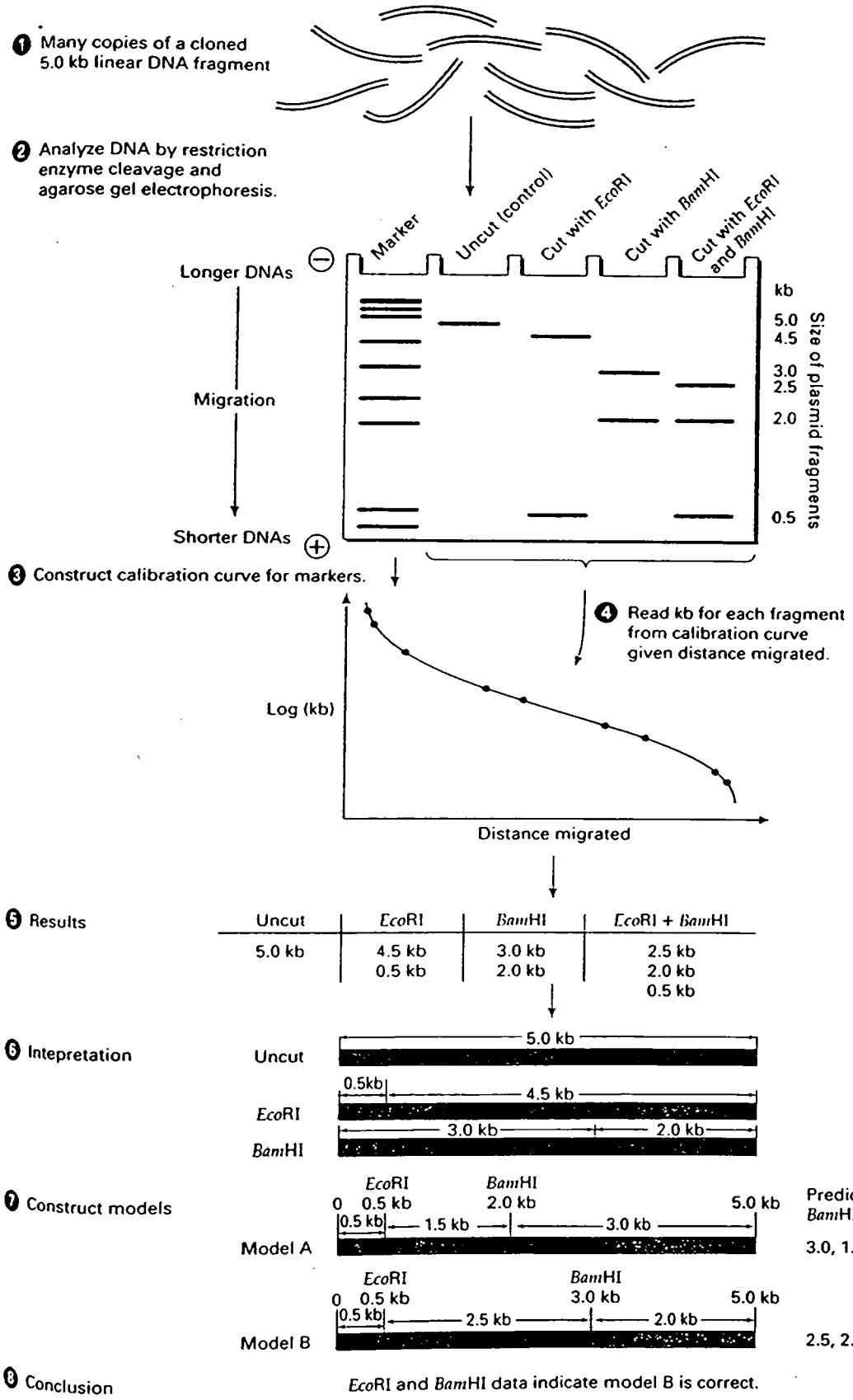


Figure 5: *Restriction Map of B-Peru.* The coding region and nearby DNA sequences are indicated by a rectangular box, with the exons represented by darkly shaded regions within the rectangle. The probes are depicted by the dark lines underneath the sequence. The jagged line indicates the unknown sequence that is being mapped using rare-cutting enzymes. Repetitive sequences within the *B-Peru* promoter are depicted by the rectangles containing arrows. Asterisks indicate restriction sites located as a result of research detailed within this thesis.

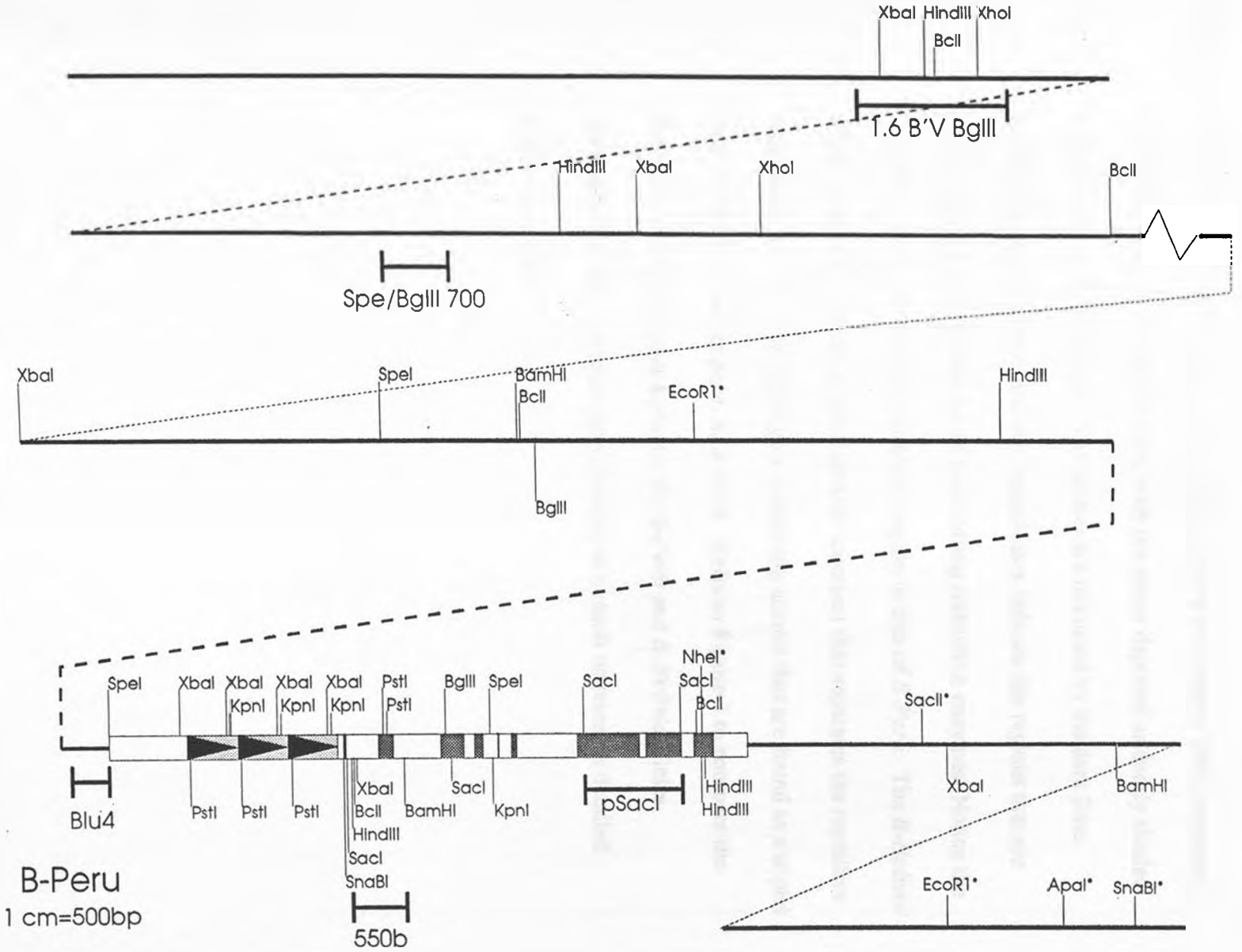

XbaI HindIII XhoI
BclI
1.6 B'V BglII
HindIII XbaI XhoI BclI
Spe/BglII 700
XbaI SpeI BamHI BclI EcoR1* HindIII
BglII
SpeI XbaI XbaI XbaI XbaI PstI BglII SpeI SacI NheI* SacI BclI SacII*
KpnI KpnI KpnI PstI
Blu4
PstI PstI PstI XbaI BclI HindIII SacI SnaBI BamHI SacI KpnI pSacI HindIII HindIII XbaI BamHI
B-Peru
1 cm=500bp
550b
EcoR1* ApaI* SnaBI*

Figure 6: *Restriction Map of B-Bolivia.* The coding region and nearby DNA sequences are indicated by a rectangular box, with the exons depicted as darkly shaded regions within the rectangle. The probes are indicated by the dark lines underneath the sequence, and the jagged lines indicate the regions that are being mapped through the use of rare-cutting restriction enzymes. Notice the similarity in structure of the upstream region to that of *B-Peru.* The *B-Bolivia* allele, however, appears to have another insertion that separates the repetitive sequences, indicated by rectangles containing arrows that are found as a triplet near the start of transcription in *B-Peru.* Refer to Figure 1 to compare the differences in phenotypes between the *B-Peru* and *B-Bolivia* alleles. Asterisks indicate restriction sites located as a result of research detailed within this thesis.

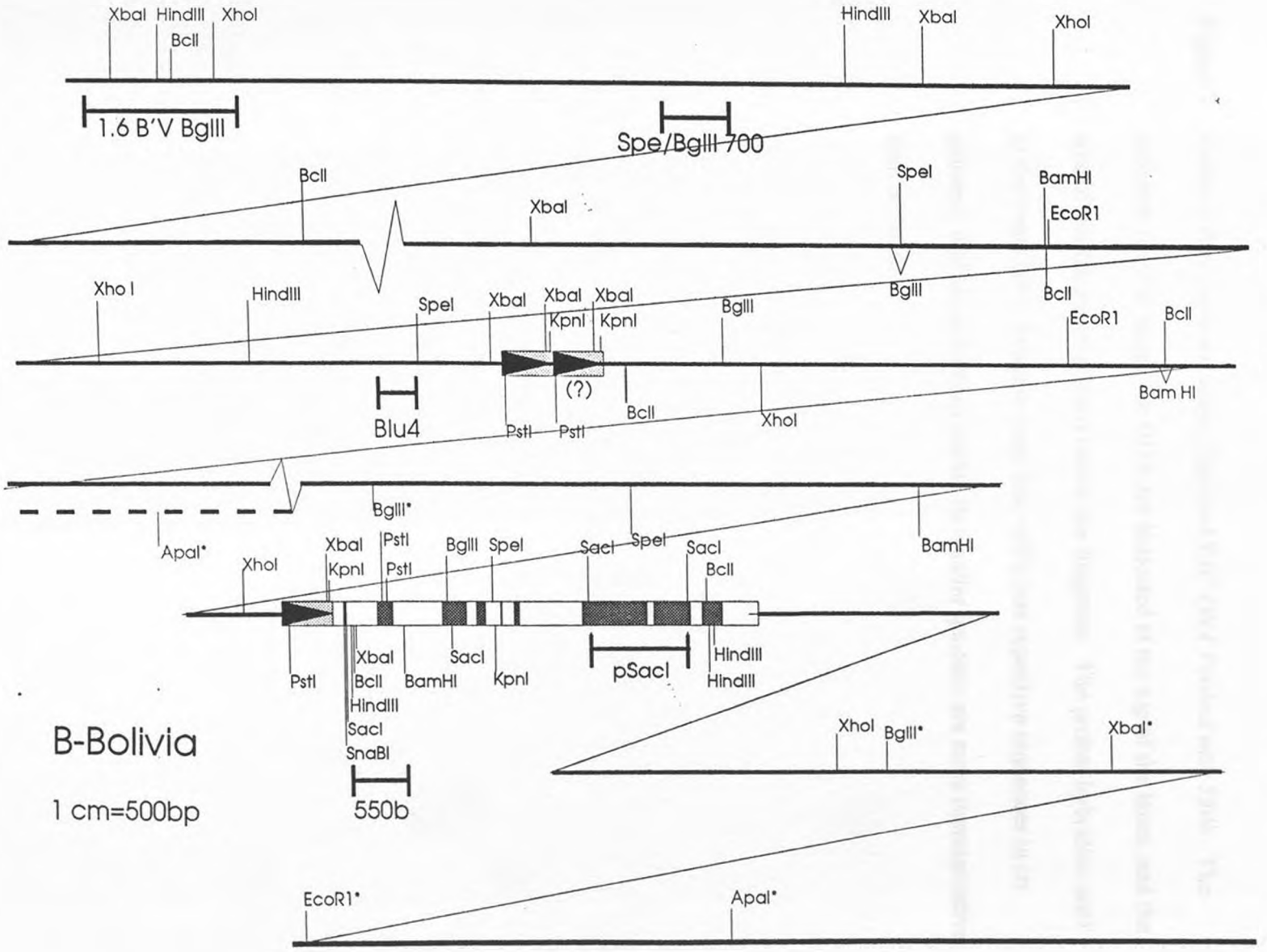

XbaI HindIII XhoI
BclI
HindIII XbaI XhoI
1.6 B'V BglII
Spe/BglII 700
BclI
XbaI
SpeI
BamHI
EcoR1
Xho I
HindIII
SpeI
XbaI
XbaI
XbaI
KpnI
KpnI
BglII
BglII
BclI
EcoR1
BclI
(?)
Blu4
PstI
PstI
BclI
XhoI
Bam HI
BglII*
ApaI*
XhoI
XbaI
KpnI
PstI
PstI
BglII
SpeI
SacI
SpeI
SacI
BclI
BamHI
PstI
XbaI
BclI
BamHI
HindIII
SacI
SnaBI
SacI
KpnI
pSacI
HindIII
HindIII
B-Bolivia
1 cm=500bp
550b
XhoI
BglII*
XbaI*
EcoR1*
ApaI*

Figure 7: *Southern Blot Demonstrating Digested YAC DNA Probed with 550b.* The enzymes used to digest the DNA are indicated at the top of the lanes, and the size of the fragment is listed below the fragment. The probes hybridize well to the yeast DNA because yeast has ~60% less repetitive sequences in its genome, and the sequences within its smaller genome are more representative than in maize.

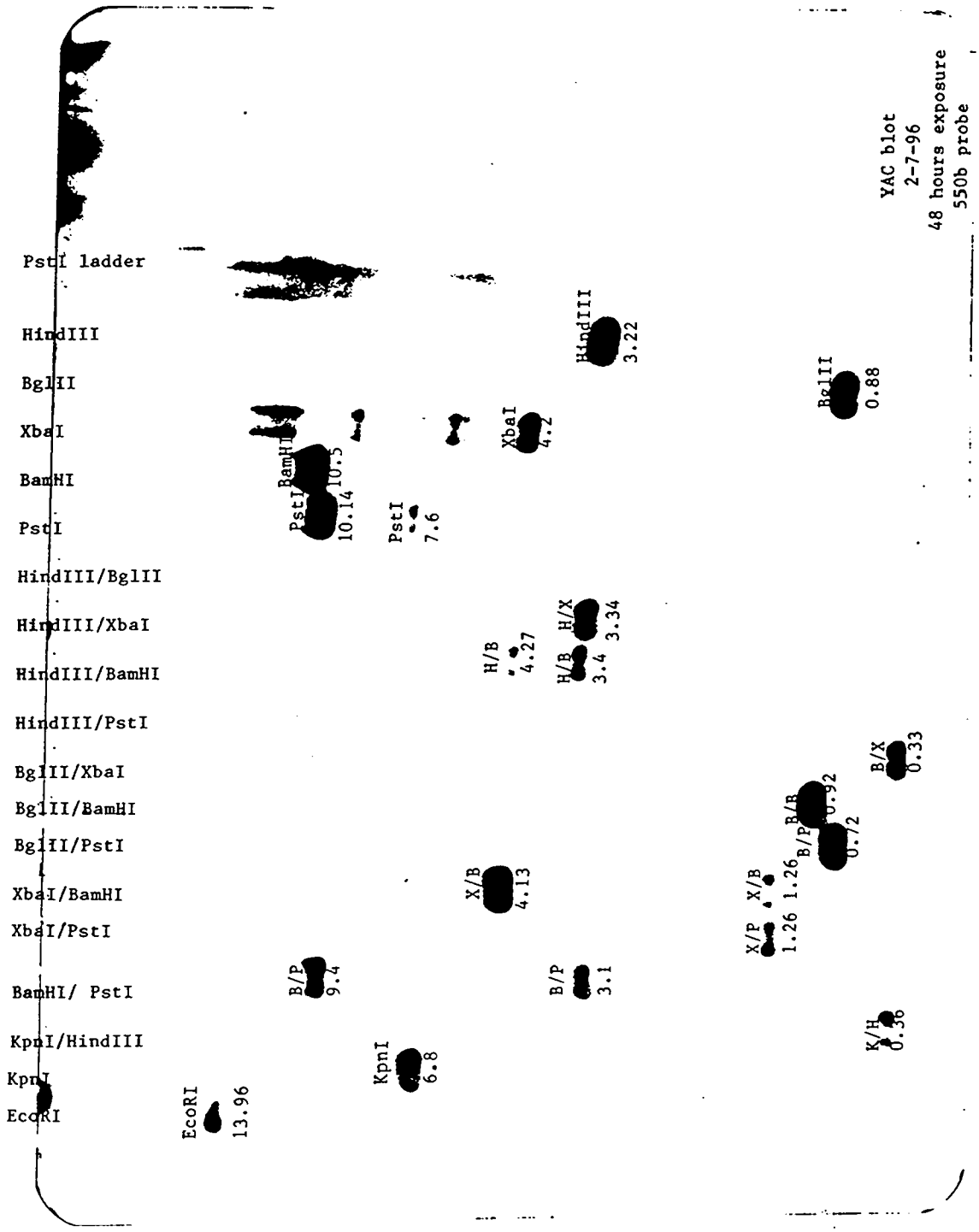

YAC blot
2-7-96
48 hours exposure
550b probe
PstI ladder
HindIII
BglII
XbaI
BamHI
PstI
HindIII/BglII
HindIII/XbaI
HindIII/BamHI
HindIII/PstI
BglII/XbaI
BglII/BamHI
BglII/PstI
XbaI/BamHI
XbaI/PstI
BamHI/ PstI
KpnI/HindIII
KpnI
EcoRI
EcoRI 13.96
KpnI 6.8
B/P 9.4
PstI 10.14
BamHI 10.5
PstI 7.6
X/B 4.13
H/B 4.27
XbaI 4.2
B/P 3.1
H/B 3.4
H/X 3.34
HindIII 3.22
X/P 1.26
X/B 1.26
B/P 0.72
B/B 0.92
BglII 0.88
B/X 0.33
K/H 0.36

Figure 8: *Southern Blot Demonstrating Digested LH82 DNA Probed with 550b.* This blot contains maize DNA, and therefore some double bands can be attributed to DNA encoding the functionally duplicate R protein. The *r* gene shares homology with the *b* gene in certain parts of the coding region, and weakly hybridizes with the 550b probe.

Hind
BglII
BglII/BamHI
BglII/PstI
XbaI/BamHI
XbaI/PstI
BamHI/PstI
KpnI/HindIII
KpnI
EcoRI
KpnI
EcoRI
5.49
5.18
B/P
5.39
K/H
3.9
B/P
3.37
X/B
4.5
B/P
4.79
B/B
4.98
3.91
B/B
B/P
1.14
B/B

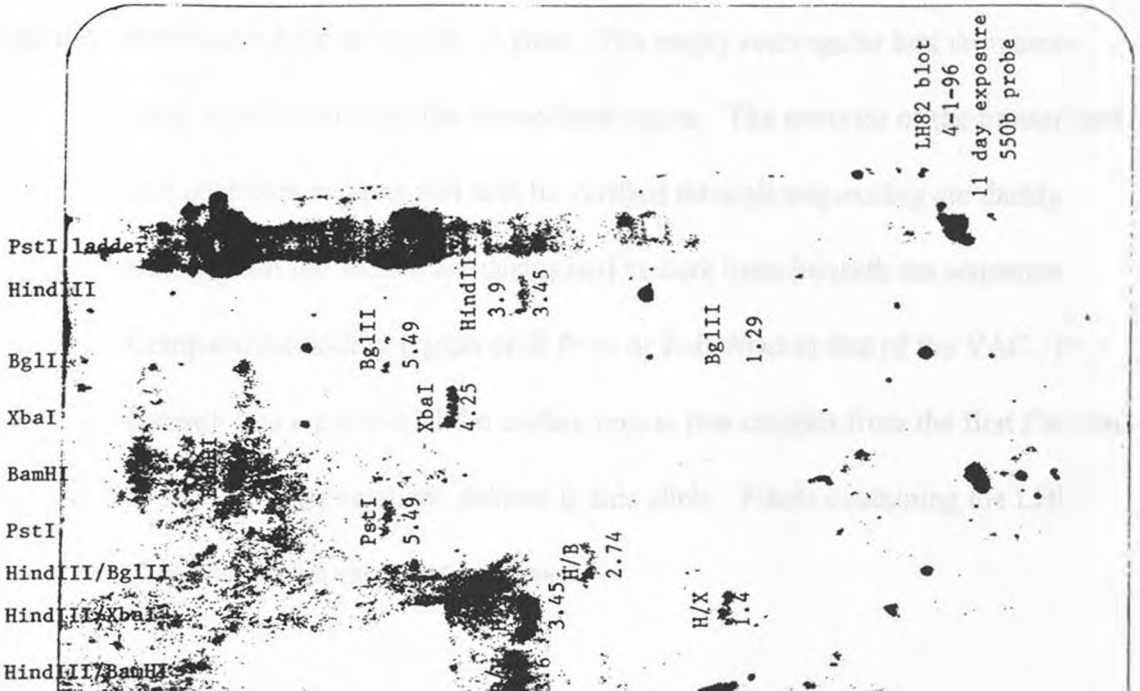
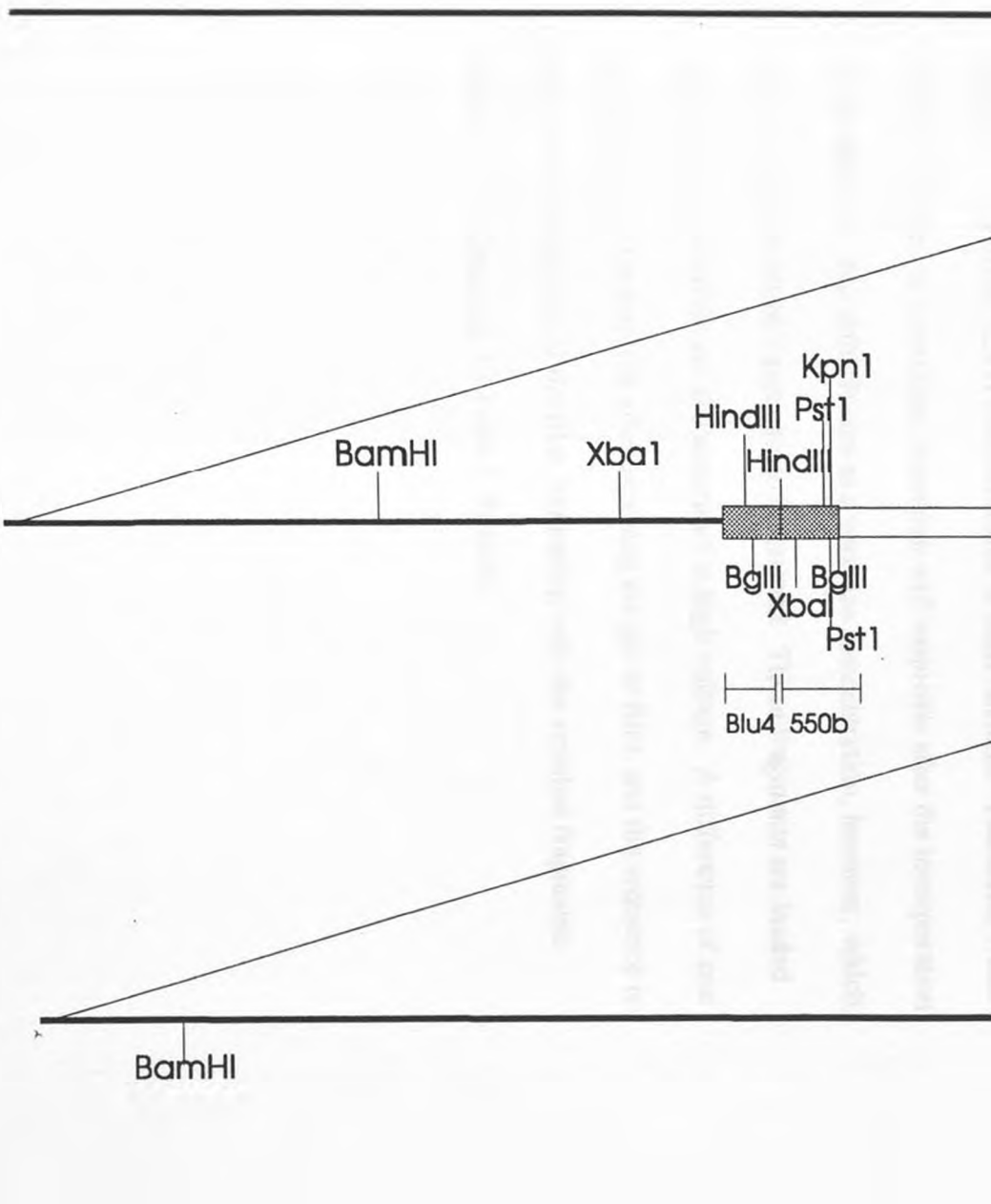

LH82 blot
4-1-96
11 day exposure
550b probe
PstI ladder
HindIII
BglII
XbaI
BamHI
PstI
HindIII/BglII
HindIII/XbaI
HindIII/BamHI
BglII
5.49
XbaI
4.25
HindIII
3.9
3.45
BglII
1.29
PstI
5.49
H/B
2.74
3.45
H/X
1.4

Figure 9: *Restriction Map of the YAC b gene.* The empty rectangular box represents what is believed to be the transcribed region. The portions of the transcribed and upstream regions that will be verified through sequencing are darkly shaded, and the probes are designated as dark lines beneath the sequence. Compare the coding region of *B-Peru* or *B-Bolivia* to that of the YAC. It appears that a portion of the coding region that extends from the first *Pst*I site to the *Bgl*II site has been deleted in this allele. Plants containing the LH82 allele do not produce anthocyanins.


Kpn1
Pst1
HindIII
HindIII
BamHI
Xba1
BglII
BglII
Xba1
Pst1
Blu4
550b
BamHI

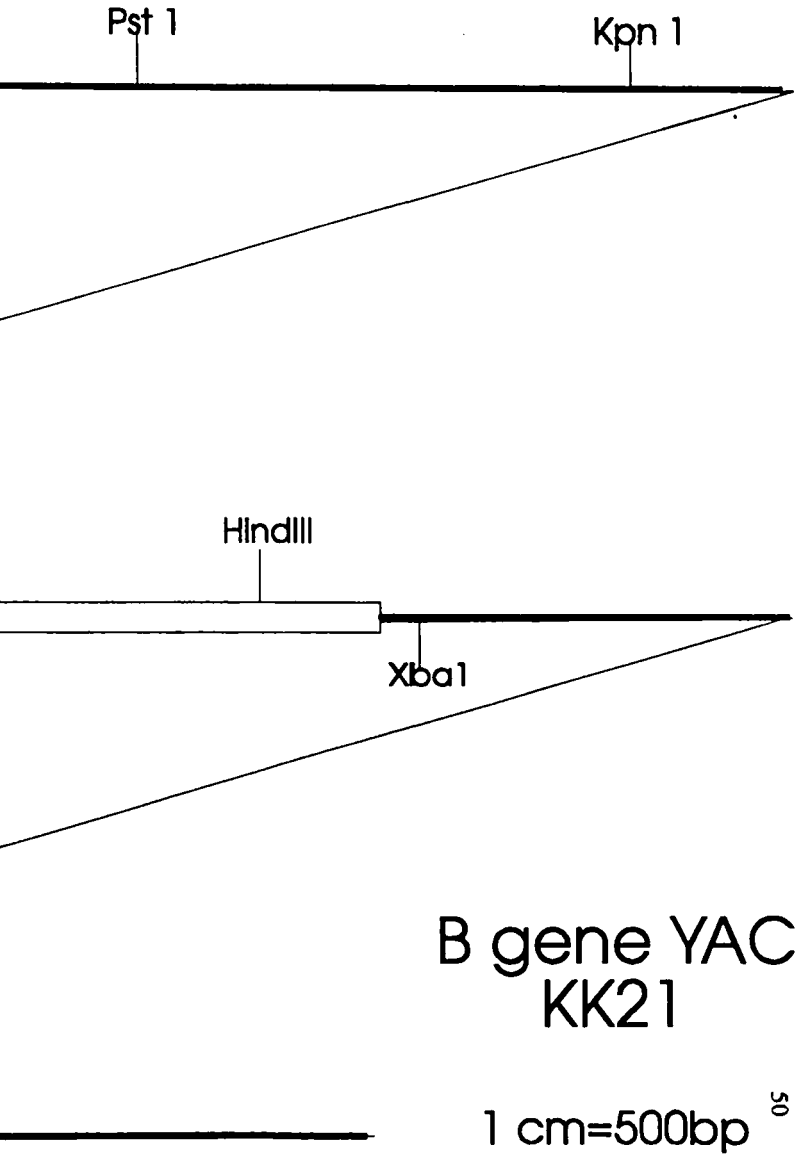
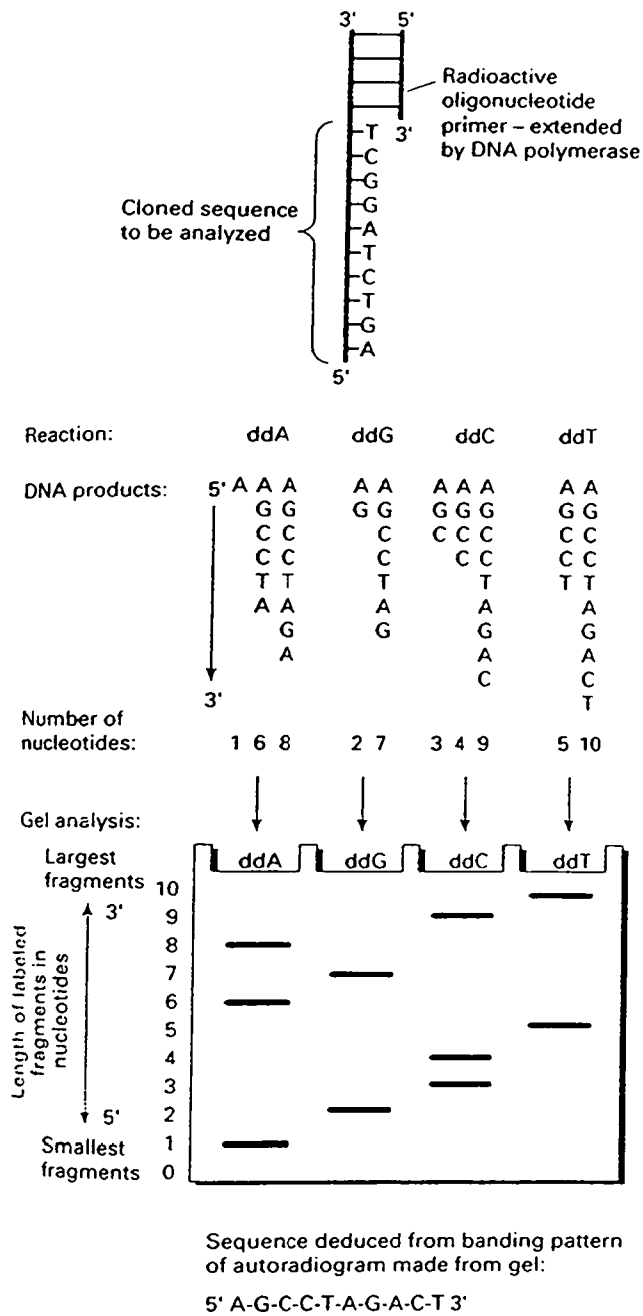

Pst 1
Kpn 1
HindIII
Xba1
B gene YAC
KK21
1 cm=500bp

Figure 10: *Diagram Detailing Dideoxy Sequencing.* A primer is annealed to a cloned DNA fragment to be sequenced, and an enzyme is allowed to extend the sequence using the fragment as a template. A radioactively labelled nucleotide is allowed to become incorporated into the extending chain, as well. After a dideoxynucleoside triphosphate (ddNTP) has been incorporated, however, growth of the fragment is terminated. There are four different nucleotides, adenine, guanine, cytosine and thymine, and so four identical reactions are occurring to extend the sequence of the template DNA from the primer. A different ddNTP will be added to each reaction. Therefore, when ddATP is added to a reaction, extension will terminate after the incorporation of the ddATP. The ddNTPs are at a very low concentration, however, which allows different sized fragments to be created. These fragments are loaded into a polyacrylamide gel and separated at high voltage. A difference of one nucleotide can be detected after exposing the gel to film, and the sequence is read from the bottom of the film, beginning with the smallest fragments. (Figure from Genetics, by Peter J. Russell)

Dideoxy DNA sequencing of a theoretical DNA fragment.



Sequence deduced from banding pattern of autoradiogram made from gel:

5' A-G-C-C-T-A-G-A-C-T 3'

LITERATURE CITED

Ausubel, F.K., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl, eds. Current Protocols in Molecular Biology. New York: Current Protocols, 1994.

Coe, E.H., Jr. "Specification of the anthocyanin biosynthetic function by B and R in maize." Maydica XXIV (1979): 49-58.

Doebley, J. "Mapping the genes that made maize." Trends in Genetics 8 (1992): 302-307.

Dooner, H.K., T.P. Robbins, and R.A. Jorgensen. "Genetic and Developmental Control of Anthocyanin Biosynthesis." Annual Review of Genetics 25 (1991): 173-199.

Griffiths, A.J.F., J.H. Miller, D.T. Suzuki, R.C. Lewontin, and W.M. Gelbart. An Introduction to Genetic Analysis. New York: W.H. Freeman & Co., 1993.

Li, W.H., and D. Graur. Fundamentals of Molecular Evolution. Sunderland: Sinauer Assoc. Inc., 1991.

Patterson, G.I., L.J. Harris, V. Walbot, and V.L. Chandler. "Genetic Analysis of *B-Peru*, a Regulatory Gene in Maize." Genetics 126 (1991): 205-220.

Patterson, G.I., K.M. Kubo, T. Shroyer, and V.L. Chandler. "Sequences Required for Paramutation of the Maize *b* Gene Map to a Region Containing the Promoter and Upstream Sequences." Genetics 140 (1995): 1389-1406.

Patterson, G.I., C.J. Thorpe, and V.L. Chandler. "Paramutation, an Allelic Interaction, is

Associated with a Stable and Heritable Reduction of Transcription of the Maize *b* Regulatory Gene." Genetics 135 (1993): 881-894.

Radicella, J.P., D. Brown, L.A. Tolar, and V.L. Chandler. "Allelic diversity of the maize B regulatory gene: different leader and promoter sequences of two B alleles determine distinct tissue specificities of anthocyanin production." Genes & Development 6 (1992): 2152-2164.

Russell, P.J. Genetics. Fourth edition. New York: Harper Collins, 1996.

Sprague, G.F. and J.W. Dudley, eds. Corn and Corn Improvement. Madison: American Society of Agronomy, Inc., 1988.