Biochemical Marker Genes for Molecular Genetics
and Plant Breeding in Pisum sativum L.

The copyright of this thesis rests with the author.
No quotation from it should be published without
his prior written consent and information derived
from it should be acknowledged.

A thesis submitted in accordance with
the requirements of the University
of Durham for the degree of
Doctor of Philosophy

By

Sayed Hassan Mahmoud
B.Sc. (Assuit, Egypt)
M.Sc. (El-Minia, Egypt)

Department of Botany

October, 1985
To my Mother
# CONTENTS

<table>
<thead>
<tr>
<th>FIGURES</th>
<th>(vii)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLES</td>
<td>(ix)</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>(x)</td>
</tr>
<tr>
<td>MEMORANDUM</td>
<td>(xi)</td>
</tr>
<tr>
<td>ABBREVIATIONS</td>
<td>(xii)</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>(xiii)</td>
</tr>
</tbody>
</table>

1. **INTRODUCTION**  
   1.1 General introduction  
   1.2 Isoenzymes  
      1.2.1 Amylase  
      1.2.2 Esterase  
      1.2.3 Glutamate oxaloacetate transaminase (GOT)  
   1.3 Storage proteins  
      1.3.1 Vicilin  
      1.3.2 Legumin  
      1.3.3 Convicilin  
      1.3.4 Storage protein synthesis  
      1.3.5 Genetic variation  
   1.4 Plant gene structure  
   1.5 Legumin gene family of pea  
   1.6 Ribosomal RNA  
   1.7 Objectives, rationale and content of the present research  

2. **MATERIALS AND METHODS**  
   2.1 Biological materials  
   2.2 Chemicals
<table>
<thead>
<tr>
<th>Contents (continued)</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3 Recombinant clones</td>
<td>33</td>
</tr>
<tr>
<td>2.4 M13 vectors and bacterial strains</td>
<td>33</td>
</tr>
<tr>
<td>2.5 Genotypes and nomenclatures</td>
<td>33</td>
</tr>
<tr>
<td>2.6 Growth of biological materials</td>
<td>34</td>
</tr>
<tr>
<td>2.7 Preparation of single seed extracts</td>
<td>34</td>
</tr>
<tr>
<td>2.8 Extraction buffers</td>
<td>35</td>
</tr>
<tr>
<td>2.9 Polyacrylamide gel electrophoresis (PAGE)</td>
<td>35</td>
</tr>
<tr>
<td>2.9.1 Isoenzymes</td>
<td>35</td>
</tr>
<tr>
<td>2.9.2 Total protein</td>
<td>35</td>
</tr>
<tr>
<td>2.10 General gel assay methods</td>
<td>36</td>
</tr>
<tr>
<td>2.10.1 Amylase</td>
<td>36</td>
</tr>
<tr>
<td>2.10.2 Esterase</td>
<td>36</td>
</tr>
<tr>
<td>2.10.3 Glutamate oxaloacetate transaminase (GOT)</td>
<td>37</td>
</tr>
<tr>
<td>2.10.4 Total protein</td>
<td>37</td>
</tr>
<tr>
<td>2.10.5 Immunological detection of vicilin polypeptides after &quot;Western&quot; blotting to nitrocellulose filters</td>
<td>37</td>
</tr>
<tr>
<td>2.11 Preparation of pea DNA</td>
<td>38</td>
</tr>
<tr>
<td>2.12 Restriction endonuclease digestion</td>
<td>40</td>
</tr>
<tr>
<td>2.13 Agarose gel electrophoresis</td>
<td>40</td>
</tr>
<tr>
<td>2.14 Recovery of DNA fragments from agarose gels</td>
<td>41</td>
</tr>
<tr>
<td>2.15 Isolation of DNA from low melting point agarose gels</td>
<td>42</td>
</tr>
<tr>
<td>2.16 Transfer of DNA from agarose gels to nitrocellulose filters by &quot;Southern&quot; blotting</td>
<td>43</td>
</tr>
<tr>
<td>2.17 $^{32}$P-labelling of double-stranded DNA (ds-DNA) by nick translation</td>
<td>44</td>
</tr>
<tr>
<td>2.18 $^{3}$H-labelling of single-stranded DNA (ss-DNA) by primer extension on M13 templates</td>
<td>45</td>
</tr>
<tr>
<td>2.19 $^{3}$H-labelling of rDNA by nick translation</td>
<td>45</td>
</tr>
</tbody>
</table>
Contents (continued)

2.20 Hybridization of $^{32}$P-labelled probes to filter-bound DNA 46

2.21 Autoradiography 47

2.22 DNA cloning in M13 vectors 47

2.22.1 Initial restriction of the insert and vector DNAs 48

2.22.3 Preparation of competent cells 49

2.22.4 Transformation and plating 49

2.22.5 Preparation of M13 subclones (single-stranded DNA templates) 49

2.22.6 Media and solutions used in DNA cloning 50

2.23 Screening M13 subclones 51

2.24 The orientation test of M13 subclones 51

2.25 DNA sequencing 52

2.25.1 Annealing primer to template 52

2.25.2 Preparation of deoxy and dideoxy NTP mixes 53

2.25.3 Sequencing reaction 53

2.25.4 Preparation, electrophoresis and autoradiography of sequencing gels 54

2.26 Cytology and in situ hybridization 55

2.27 Physical mapping of ribosomal RNA gene sites 56

2.28 Spectrophotometric determination of DNA 57

2.29 Statistical methods 57

2.30 Stains for chromosomes 58

3. RESULTS 62

3.1 Genetic variation at isoenzyme loci 63

3.1.1 Variation at the amylase (Amy) locus 63

3.1.2 Variation at the esterase (Est) locus 64

3.1.3 Variation at the glutamate oxaloacetate transaminase (Got) locus 64
## Contents (continued)

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2 Inheritance and mapping of isoenzyme genes</td>
<td>65</td>
</tr>
<tr>
<td>3.2.1 Amylase</td>
<td>65</td>
</tr>
<tr>
<td>3.2.2 Esterase</td>
<td>66</td>
</tr>
<tr>
<td>3.2.3 Glutamate oxaloacetate transaminase (GOT)</td>
<td>68</td>
</tr>
<tr>
<td>3.3 Genetic variation at storage protein loci</td>
<td>79</td>
</tr>
<tr>
<td>3.3.1 Variation at the vicilin (Vc) locus</td>
<td>79</td>
</tr>
<tr>
<td>3.3.2 Variation at the legumin (Lg) locus</td>
<td>79</td>
</tr>
<tr>
<td>3.3.2.1 Restriction fragment length polymorphism</td>
<td>79</td>
</tr>
<tr>
<td>3.3.2.2 The nucleotide sequence of legumin gene &quot;legB&quot;</td>
<td>80</td>
</tr>
<tr>
<td>3.4 Inheritance and mapping of storage protein genes</td>
<td>88</td>
</tr>
<tr>
<td>3.4.1 Inheritance of vicilin (50,000 Mr) and legumin (α-subunit) genes</td>
<td>88</td>
</tr>
<tr>
<td>3.4.2 Chromosomal location of vicilin (50,000 Mr) gene locus</td>
<td>89</td>
</tr>
<tr>
<td>3.5 Genetic variation at the ribosomal RNA gene sites</td>
<td>96</td>
</tr>
<tr>
<td>3.6 Inheritance and mapping of ribosomal RNA genes</td>
<td>97</td>
</tr>
<tr>
<td>3.6.1 Inheritance of ribosomal RNA genes</td>
<td>97</td>
</tr>
<tr>
<td>3.6.2 Molecular mapping of ribosomal RNA genes</td>
<td>98</td>
</tr>
<tr>
<td>3.6.3 Cytolocalization of ribosomal RNA genes</td>
<td>99</td>
</tr>
<tr>
<td>3.6.4 Physical mapping of ribosomal RNA genes</td>
<td>99</td>
</tr>
<tr>
<td>4. DISCUSSION</td>
<td>107</td>
</tr>
<tr>
<td>4.1 Isoenzymes</td>
<td>108</td>
</tr>
<tr>
<td>4.2 Storage proteins</td>
<td>114</td>
</tr>
<tr>
<td>4.3 Ribosomal RNA</td>
<td>125</td>
</tr>
<tr>
<td>4.4 Analysis of the genetic variation and mapping of biochemical marker genes</td>
<td>128</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>131</td>
</tr>
</tbody>
</table>
FIGURES

1. PAGE (7.5% acrylamide gel) of amylase isoenzyme banding patterns of pea seeds. 70

2. PAGE (7.5% acrylamide gel) of esterase isoenzyme banding patterns of pea seeds. 71

3. PAGE (7.5% acrylamide gel) of GOT isoenzyme banding patterns of pea seeds. 72

4. SDS-PAGE (12% acrylamide gel) of total protein extracts of seeds of different pea lines (numbered), and of vicilin from var. Feltham First (VcFF) and from lines 360 and 611 (Vc360, Vc611). 83

5. Eco RI digests of total DNA from leaves of different pea lines (numbered) after fractionation on 0.7% agarose gels, autoradiographs of hybridization of $^{32}$P-labelled probe (legumin coding sequence, labelled to a high specific activity of $3 \times 10^8$ cpm/µg) to total genomic DNA after "Southern" blotting to nitrocellulose filter. 84

6. a) Restriction map of $\lambda$Leg2. b) Sequencing map of legB gene. 85

7. Autoradiographs of the nucleotide sequence of DNA using the "dideoxy chain termination" method with M13 sequencing system. 86

8. Nucleotide sequence of the legumin gene legB of pea (Pisum sativum L.). 87

9. SDS-PAGE (12% acrylamide gel) of total protein extracts of seeds of pea lines 1263 and 3080, and F1 and F2 seeds from their cross. 91

10. SDS-PAGE (12% acrylamide gel) of total protein extracts of pea seeds. a) The parental lines 1263 and 3080, and F1 and F2 seeds from their cross and of vicilin from var. Feltham First (VcFF) after "Western" blotting to nitrocellulose filter. b) Isogenic round/wrinkled lines (T), compared with corresponding "Western" blot of vicilin subunit (Vc). 92

11. Eco RI digests of total DNA from pea leaves after fractionation on 0.7% agarose gels, autoradiographs of hybridization of $^{32}$P-rDNA to total genomic DNA after "Southern" blotting to nitrocellulose filter. 101

Figures (continued)

13. The physical map of chromosomes 4 and 7 from root-tip metaphase preparations of pea line 110 (Pisum sativum L.), 2n=14.

14. Relative map positions on the seven chromosomes of Pisum sativum of the loci examined in this study which are shown to the right of the chromosome and the marker loci which are shown to the left of the chromosome.

15. Comparison of the legumin genes; legA, legB and legC of pea (Pisum sativum L.).
# TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Numbers, genotypes and karyotypes of pea lines.</td>
<td>59</td>
</tr>
<tr>
<td>2.</td>
<td>Key to gene symbols and their locations on pea chromosomes.</td>
<td>60</td>
</tr>
<tr>
<td>3.</td>
<td>Analysis of the F2 segregation of amylase band patterns on PAGE (Amy) and the selected marker genes studied in a cross between two pea lines (110 x 1238).</td>
<td>73</td>
</tr>
<tr>
<td>4.</td>
<td>Segregation of amylase band patterns on PAGE (Amy) in the F2 generation of a cross between two pea lines (110 x 1238) with respect to the selected marker genes.</td>
<td>74</td>
</tr>
<tr>
<td>5.</td>
<td>Analysis of the F2 segregation of esterase band patterns on PAGE (Est) and the selected marker genes studied in a cross between two pea lines (102 x 110).</td>
<td>75</td>
</tr>
<tr>
<td>6.</td>
<td>Segregation of esterase band patterns on PAGE (Est) in the F2 generation of a cross between two pea lines (102 x 110) with respect to the selected marker genes.</td>
<td>76</td>
</tr>
<tr>
<td>7.</td>
<td>Analysis of the F2 segregation of GOT banding patterns on PAGE (Got) and the selected marker genes studied in a cross between two pea lines (1238 x 1263).</td>
<td>77</td>
</tr>
<tr>
<td>8.</td>
<td>Segregation of GOT band patterns on PAGE (Got) in the F2 generation of a cross between two pea lines (1238 x 1263) with respect to the selected marker genes.</td>
<td>78</td>
</tr>
<tr>
<td>9.</td>
<td>Comparison between the predicted amino acid composition of legA and legB gene products.</td>
<td>93</td>
</tr>
<tr>
<td>10.</td>
<td>Analysis of the F2 segregation of vicilin 50,000 Mr subunit (Ve-1) and legumin α-subunit (Lg-1) band patterns on SDS-PAGE and round/wrinkled seed surface (R/r) in a cross between two pea lines (1263 x 3080).</td>
<td>94</td>
</tr>
<tr>
<td>11.</td>
<td>Segregation of vicilin 50,000 Mr subunit (Ve-1) and legumin α-subunit (Lg-1) band patterns on SDS-PAGE in the F2 generation of a cross between two pea lines (1263 x 3080) with respect to the selected marker gene (R/r) for round/wrinkled seed surface.</td>
<td>95</td>
</tr>
<tr>
<td>12.</td>
<td>Analysis of the F2 segregation of Eco RI restriction fragments of rDNA (rRNA) and the selected marker genes studied in a cross between two pea lines (110 x 1238).</td>
<td>104</td>
</tr>
<tr>
<td>13.</td>
<td>Segregation of Eco RI restriction fragments of rDNA (rRNA) in the F2 generation of a cross between two pea lines (110 x 1238) with respect to the selected marker genes.</td>
<td>105</td>
</tr>
<tr>
<td>14.</td>
<td>Average length (in microns) of the different parts of chromosomes 4 and 7 from root-tip metaphase preparations of pea (Pisum sativum L.) line 110.</td>
<td>106</td>
</tr>
</tbody>
</table>
Three isoenzyme systems (amylase, esterase and glutamate oxaloacetate transaminase) were examined in seeds of pea (Pisum sativum L.) and showed clear variations in their band patterns on gel electrophoresis between different lines. The inheritance of these isoenzyme systems, and the location of their structural genes on the pea genome were investigated. Reciprocal crosses were made between lines, F2 seeds were analysed for segregation in the band patterns of the isoenzymes, and F2 plants were investigated to find linkage between the genes for these isoenzymes and genes for selected morphological markers. The results obtained showed that each of the investigated isoenzyme systems is genetically controlled by co-dominant alleles at a single locus. The gene for amylase (Amy) was found to be on chromosome 2, linked to the loci k and wb (wb...9...k...25...Amy). The gene for esterase (Est) was found to be linked to the gene Br (chromosome 4) but the exact location is uncertain because of a lack of morphological markers. The gene for glutamate oxaloacetate transaminase (Got) was found to be on chromosome 1 linked to the loci a and d (a...24...Got...41...d).

Gel electrophoresis techniques have also been used to investigate genetically controlled variation in the major subunits (50,000 Mr) of vicilin, a storage protein of Pisum sativum L. The F1 protein band patterns were shown to be additive with respect to those of the parental lines and to be identical in reciprocal crosses. Genetic analysis of the F2 plants indicated that the 50,000 Mr vicilin subunits band pattern is controlled by a pair of co-dominant genes at a single locus. The F2 data were used to locate this major vicilin gene locus (Ve-1) to chromosome 7, closely linked to the r locus (for round and wrinkled seed surface).

A third member of pea legumin gene family, denoted legB, has been sequenced using the "dideoxy chain termination" method with the M13 sequencing system. The complete nucleotide sequence showed that this gene has a general form typical of an eukaryotic gene. The homologies between this gene and the previously published gene "Leg A" were estimated and showed strong homology between the two genes with eight amino acid substitutions and deletion of 14 bp in the third intron (IVS-3).

The inheritance of ribosomal RNA (rRNA) genes in (Pisum sativum L.) was investigated in a cross between two different lines, where length variation in rDNA fragments of Eco RI digests was observed. The results obtained showed that the rRNA genes are controlled by simple Mendelian system with "co-dominance" between alleles. In order to locate the rRNA gene sites to positions on the chromosomes, the segregation of Eco RI restriction fragments of rDNA from F2 plants with respect to genes for selected morphological markers on chromosomes 4 and 7 (the chromosomes known to have nucleolus organizer regions) were tested. The F2 data showed no linkages between the selected markers and rRNA genes, therefore, in situ hybridization using rDNA radioactive probe (1H-labelled rDNA clone, pHAI) and physical mapping procedures were used. The results obtained have located the rRNA gene sites to nucleolus organizer regions (satellite constrictions) at 138 and 60 map units from the centromeres of chromosomes 4 and 7, respectively.
MEMORANDUM

Parts of the work described in this thesis have previously been presented in the following publications:


ABBREVIATIONS

Abbreviations are used as recommended in the "Biochemical Journal Instructions to Authors" (Biochemical Society, 1975) with the additions listed below.

bp: Base pairs
kbp: Kilobase pairs
cDNA: Complementary DNA
ds-DNA: Double-stranded DNA
ss-DNA: Single-stranded DNA
rDNA: Ribosomal DNA
rRNA: Ribosomal RNA
d-NTP: Deoxynucleoside triphosphate
dd-NTP: Dideoxynucleoside triphosphate
cpm: Counts per minute
EtBr: Ethidium bromide
poly (A): Polyadenylic acid
SDS: Sodium dodecyl sulphate
PAGE: Polyacrylamide gel electrophoresis
BSA: Bovine serum albumin
SSC: Saline sodium citrate buffer
SSCP: Saline sodium citrate phosphate buffer
IPTG: Isopropyl-β-D-thio-galactopyranoside
X-Gal: 5-bromo-4-chloro-3-indolyl-β-galactoside
5': 5' terminal phosphate in a DNA molecule
3': 3' terminal hydroxyl in a DNA molecule
IVS: Intervening sequence
ITS: Internal transcribed spacer
NOR: Nucleolus organizer
ACKNOWLEDGEMENTS

I wish to thank my supervisor, Professor Donald Boulter, for his guidance, encouragement and support throughout this study.

I am extremely grateful to Dr. John Gatehouse who has been an invaluable source of help, inspiration and stimulating discussions during the past four years. The present work owes much to his careful and constant guidance.

Other members of Botany Department, University of Durham, especially Dr. Ronald Croy, Dr Angharad Gatehouse and Dr Phillip Gates are to be thanked for their interest, useful discussions and practical help throughout this work. I am very grateful to Mr David Bown for his considerable help and introducing me to a wide range of techniques. I thank Mr. Paul Preston for his valuable help during DNA sequencing.

My thanks are also due to Dr R. Flavell of the Plant Breeding Institute, Cambridge, for his useful discussion and suggestions on physical mapping of ribosomal RNA genes, Dr. John Gatehouse for critically reading the manuscript and providing Fig.15 and Dr. A. Shirsat for providing Fig.6a. The assistance of the staff of the greenhouse in the Botanic Garden, Durham University, especially Mr. Stewart Grey is gratefully acknowledged.

This work was supported by a grant from the Egyptian Government and an Overseas Research Students Award from the Committee of Vice-Chancellors and Principals, which are gratefully acknowledged.

Finally, I am grateful to my wife, Salwa and my son Sherief for their support, encouragement and patience during my study.
1. INTRODUCTION
1.1 General Introduction

Due to the growing necessity for improving the quantity and quality of food production, research is needed in fields related to plant breeding with the purpose of improving the precision of the methods by which new genotypes can be constructed to fit identified objectives. In most plant breeding programmes in the past, the synthesis of new and improved genotypes relied upon the processes of recombination and segregation that occur in the progenies of heterozygous individuals within a crop. Alternatively, where attempts are made to modify a simple genetic locus, or a single plant character, the breeder may resort to induced mutation. In both of these situations the entities being manipulated by the breeder are single genes that are potentially, within the limitations of linkage, capable of independent segregation from other genes. Occasionally there is advantage in manipulating not single genes but large arrays of associated genes. Such arrays may consist of all of the genes in a set of chromosomes, or on a single chromosome, or on a substantial segment of chromosome. Where these procedures are followed, plant breeders are applying different cytological techniques i.e. the reduction from the normal chromosome number "2n" to the haploid number "n" (Kasha and Reinbergs, 1980; Hermsen, 1984a); the increase, within a species, from the normal diploid number "2n" to a higher multiple, such as triploid "3n" or tetraploid "4n" (Hermsen, 1984b); the combination in a synthetic allopolyploid or amphiploid, of the full complement of chromosomes of two distinct species.

At the cell level, plant tissue culture technology has been available to the plant breeder for more than two decades (Vasil and
Hildebrandt, 1965a, 1965b; Vasil et al., 1979; Chaleff, 1981; Krikorian, 1982; James, 1984; Maliga, 1984), yet its impact on plant breeding has been minimal. Recent methods of plant tissue culture involving the isolation and genetic manipulation of plant protoplasts, embryo culture, cell and callus culture and clonal propagation have greatly stimulated interest in this area. It is however, clear that such tissue-culture methods, although well established as methods for the plant breeder, are in fact only rarely used in plant breeding. As a result there is little accumulation of expertise of current results from these methods and this in itself tends to detract from their usefulness in plant breeding.

One way of obtaining a plant species with a new character would be through the direct incorporation of the appropriate gene or genes in the plant genomes. Using genetic engineering techniques, it would be possible to transfer to crop plants the desired genes, responsible for desirable phenotypes. The transferred genes might originate from plants or even animals and micro-organisms. From a technical point of view it may be concluded that the first steps on the long road leading to successful genetic engineering in higher plants through the integration of new genes have been taken (Horsch et al., 1984; DeBlock et al., 1984). It can be expected that in the near future "artificial" plants will be obtained with selected new genes. What types of genes should be incorporated and how can these genes be isolated? Since many plant characters are polygenic (quantitative) with poorly characterized genes, the simple Mendelian (qualitative) characters are the first target for applying the new techniques e.g. (i) insect, pathogen and herbicide resistance; (ii) nitrogen fixation;
(iii) heat or cold resistance; (iv) stress tolerance; and (v) improving the nutritional quality of plant proteins. Item (v) is the most promising prospect for many reasons; (i) most of these proteins are under the control of small multigene families e.g. 7 and <10 genes for pea legumin and vicilin storage proteins, respectively (Shirsat, 1984; Gatehouse et al., 1983; Lycett et al., 1983b), (ii) the availability of cloned storage protein genes and (iii) the accumulation of a considerable amount of information of storage protein biochemistry as reviewed by Larkins, 1983; Boulter, 1984; Casey and Domoney, 1984; and Sorenson, 1984.

In bacteria, DNA transformation is the classical way of incorporating new genes into the genome. In higher plants, studies on the plant crown gall tumour caused by Agrobacterium tumefaciens have demonstrated that in this material there is a natural case of "genetic engineering" of plant cells by a bacterium. By transferring a particular type of DNA (T-DNA), the bacterium induces tumour cells, and these then produce compounds that can be used efficiently by the bacterium as a sole source of nitrogen. In the last few years much knowledge has been gained regarding the molecular basis of crown gall formation. This knowledge is now finding applications which can be of benefit to plant breeding. The recombinant DNA technology e.g. cloning and sequencing of genes, provides the simplest and quickest means of determining the primary structures of large numbers of proteins, and the employment of these techniques therefore constitutes the first step in any project aimed at protein engineering.

Different strategies were given by Delauney (1984) for protein engineering to improve the nutritional value of seed proteins i.e. (i)
substitute codones for existing amino acids in the protein genes with codones for the deficient amino acids e.g. methionine and cysteine; (ii) enhancing the transcription of the nutritionally well balanced protein which occur in small amounts in the seed; and (iii) amplification of genes coding for high-sulphur subunits and silencing the low-sulphur protein genes.

Once the amino acid substitutions have been successfully engineered in a chosen gene, the predicted effects of these alternations on the structural and functional properties of the encoded polypeptide can be directly investigated on samples of the modified product synthesized in bacteria or yeasts before the modified gene is introduced into a plant. Delauney (1984) found that vicilin signal peptide from pea can be correctly processed in the bacteria. The modified gene can be integrated first in Ti-plasmid derived vehicle or any other convenient vehicle (Bolivar and Backman, 1979; Howell, 1982; Barton and Chilton, 1983; Shaw et al., 1983) for transfer to the plant genome. The modified genes can be introduced into the new cell by the co-cultivation of the gene carrying vehicle e.g. Agrobacterium with a single cell protoplast on suitable media, followed by selection of the calli which carry the new phenotype and regeneration of the transformed plants (Marton et al., 1979). However, there have been several attempts recently to introduce the gene directly into the plant cell through "microinjection" (Russul and Tilton, 1983; Flavell and Mathias, 1984; and ongoing research in Dept. of Botany, Univ. of Durham).

At this point several questions will arise e.g. to which part of the genome, the new gene is going to integrate, what will be the
effect of the new gene on the physiological balance of the cell, and is the new gene going to undergo normal transcription and translation processes. To answer the first question, much more work is needed regarding the in situ hybridization techniques in order to produce methods accurate enough to detect single copy genes and localization at present must rely on mapping many biochemical marker genes onto chromosomal locations either by in situ hybridization techniques or by the classical linkage methods. Recently much work has been undertaken in different places on mapping different biochemical genes onto the genomes of different plant species (Blixt et al., 1980; Burger and Knäumann, 1980, 1983; Goodman et al., 1980; Shewry et al., 1980, 1983; Matta and Gatehouse, 1982; Brown 1983; Hart, 1983; Rick, 1983; Pollak et al., 1984; Weeden and Marx, 1984).

1.2 Isoenzymes

Electrophoretically variant enzymes have provided suitable material for studying genetic variation in higher plants. These variations have been detected with several enzyme systems in different plant species. When isoenzymes are separated on starch or acrylamide gels, differential migration of the protein depends on net electrophoretic charge and to some extent the size or configuration of the molecules involved. Although changes in electrophoretic mobility can occur in a number of ways, perhaps the most important is through amino acid substitutions to produce changes in net charge (Ingram, 1957; Hunt and Ingram, 1958). Another possible cause of electrophoretic variation affecting net charge of protein is non-homologous crossing over, resulting in deletions and duplications. Isoenzyme variants have
been found to be under the control of major genes with simple Mendelian segregation in nearly all instances and are often controlled by co-dominant alleles, so that the phenotypes of the heterozygotes are readily distinguishable from those of homozygotes.

Electrophoretic techniques of isoenzymes have been widely used in plant breeding. One of the most extensive application of the technique has been towards the cultivar identification in different plant species. Gates (1978) lists a wide range of plant species whose cultivars have been identified in this way. Isoenzymes have also been used as biochemical markers in different studies e.g. Saghai-Maroof et al. (1984) used esterase isoenzyme variants in investigating the chromosomal location of spacer-length variation in barley rDNA; and Al-Helal (1985) used GOT banding pattern of pea seeds in testing the competition between pollen grains of different pea inbred lines. In the field of cytogenetics, isoenzymes have proved to be useful chromosome markers, when their genes can be mapped on specific chromosomes. In this context, Upadhya (1968) demonstrated the peroxidase isoenzyme associated with the Aegilops umbellulata chromosome segment transferred to Chinese Spring wheat; Tang and Hart (1975) were able to use isoenzymes as chromosome markers in wheat-rye addition lines; Kobrehel (1978) identified the chromosome segments of wheat controlling peroxidase synthesis and used this to identify transferred lines with Agropyron elongatum; and Morikawa (1978) identified monosomic and nullisomic oats using leaf peroxidase isoenzymes.

Isoenzymes can be used as biochemical markers in plant molecular biology and plant breeding schemes because the alleles at most isoenzyme loci are co-dominant in nature and cause no deleterious
changes in plant phenotype through recessiveness or pleiotropy. This co-dominance, as mentioned earlier, allows heterozygotes to be distinguished from homozygotes. In contrast, morphological markers have several disadvantages when used as markers e.g. the recessive alleles of genes for morphological characters may be deleterious when homozygous and epistatic and/or pleiotrophic effects of such genes may limit the number of markers which can be assembled in a single stock. The equipment and materials needed for screening the isoenzyme banding patterns (zymograms) of plants is relatively inexpensive and it is possible to screen large numbers of plants rapidly. The process is nondestructive since only small amounts of plant tissue are needed, moreover, different plant tissue can be analyzed. Finally, it is often possible to screen plants at the seedling stage and retain only desirable genotypes which may lead to large savings in time.

1.2.1 Amylase

Amylase (E.C.3.2.1.1.2) of higher plants exists as two distinctive forms α- and β- amylase (Frydenberg and Nielsen, 1965; Yamamoto, 1975). The existence of genetic variation in amylase patterns in plants has been established in barley (Frydenberg and Nielsen, 1965); maize (Scandalios, 1966); wheat (Nishikawa and Nobuhara, 1971); Vicia (Yamamoto, 1975); soybean (Gorman and Kiang, 1977); and pea (Przybylska et al., 1982). Both the study of Frydenberg and Nielsen (1965) in barley and of Scandalios (1966) in maize strongly suggest simple Mendelian control of the amylases in higher plants.

The differentiation between α- and β-amylose and their role in
starch degradation have been reported by Swain and Dekker (1966), Juliano and Varner (1969), Bilderback (1973), and Al-Helal (1985). It was found that $\alpha$-amylase is the main starch-degrading enzyme in the cotyledons, involved in the initial degradation of starch into more soluble forms while $\beta$-amylase assists in the further conversion to free sugars. The pathway for the degradation of starch reserves to glucose involving the sequential action of amylases and a glucosidase is called "hydrolytic pathway" which is as follows:

\[
\text{starch} \xrightarrow{\alpha\text{-amylase}} \text{soluble oligosaccharides} \xrightarrow{\beta\text{-amylase}} \text{maltose} \xrightarrow{\alpha\text{-glucosidase}} \text{glucose}
\]

1.2.2 Esterase

Esterase (E.C.3.1.1.2) of higher plants has been used in genetic studies (Kahler and Allard, 1970; Rudin and Rasmuson, 1973; Bassiri and Rouhani, 1977; Gates, 1978; Hart et al., 1980). West and Garber (1967b) determined the inheritance of two sites of esterase in the genus *Phaseolus* as a Mendelian monogenic inheritance. In *Pisum* Frankel and Garber (1965) extracted six detectable esterases from the germinating seeds of twelve varieties and found that the presence or absence of two of these esterases is determined by monogenic differences. Weeden and Marx (1984) found that pea leaf esterase-4 was controlled by co-dominant alleles and lies half way between the genes $tl$ and $Bt$ on chromosome 7. Brown and Allard (1969) studied the inheritance of five enzyme systems including esterase in maize and found that each of these systems was governed by a single locus. In *Cucurbita*, Wall and Whitaker (1971) stated that there are two
forms of esterase controlled by co-dominant alleles.

1.2.3 Glutamate oxaloacetate transaminase (GOT)

Glutamate oxaloacetate transaminase or asparatate aminotransferase (GOT) (E.C.2.6.1.1) in higher plants has been widely investigated e.g. MacDonald and Brewbaker (1972) studied the genetic control and dimeric nature of transaminase hybrid maize isoenzymes. In Pinus sylvestris, Rudin (1975) identified two clearly separated regions in gels after staining for GOT and an additional hybrid band in a heterozygote appeared which indicate that the isoenzyme bands were a dimeric product. He also reported the occurrence of monohybrid Mendelian inheritance of the GOT structural genes. Guries and Ledig (1978) observed three zones of GOT activity in pitch pine. In Vicia faba, Suso and Moreno (1982) found that there were two zones of GOT banding pattern, zone "A" with two bands and zone "B" with different single bands. The F1 hybrid between two different patterns of zone "B" had three bands indicating that the enzyme is at least a dimer and the segregation ratios in F2 showed that zone "B" variants are controlled by a locus with two alleles. Przybylska et al., (1982) studied the electrophoretic patterns of several enzyme systems in Pisum and identified two independently varying zones of GOT activity, "GOT-2" and "GOT-1" which correspond to "GOT-A" and "GOT-B" of Suso and Moreno (1982).
1.3 Storage proteins

During the development of legume seeds, protein is synthesized at a variable rate through the developmental period and subsequently stored in membrane-bound protein bodies, which also appear during seed development. It has been shown with several legumes that this protein is broken down in germination and is used as a source of nitrogen for the new nitrogen compounds synthesized in developing seedlings. Protein fractions isolated from legumes on the basis of solubility were termed legumin and vicilin by Osborne and Campbell (1898). The term legumin is now used for principal protein of the legumin fraction (11S) and the term vicilin is retained for the solubility fraction, whose major proteins are 7S proteins.

1.3.1 Vicilin

The vicilin fraction of the seeds of pea (\textit{Pisum sativum} L.) contains major polypeptides of Mr 71,000, 50,000 and 33,000 with minor components of lower Mr (Croy et al., 1980b). The 71,000 Mr polypeptides were found to be distinct from the rest of the vicilin fraction (Croy et al., 1980c). This protein named "convicilin" has an Mr approx. 280,000 (i.e. four 71,000 Mr subunits). Only polypeptides of Mr about 50,000 and 71,000 were immunoprecipitated from translation products by anti-vicilin antibodies (Evans et al., 1979); no smaller vicilin polypeptides were observed in immunoprecipitates, or were apparent in the total translation products. Spencer et al. (1980) found that no radioactivity was incorporated into vicilin polypeptides of Mr less than 50,000 after pulse labelling in vivo of developing pea cotyledons with radioactive amino acids. Therefore, Gatehouse et
al. (1981) investigated the synthesis of vicilin polypeptides in vitro and post-translational modification of vicilin, in comparison with the structure of accumulated vicilin. They found that vicilin subunits were synthesized as polypeptides of Mr ~50,000, then undergo two distinct types of proteolytic modification (i) co-translational removal of a small polypeptide (Mr less than 1000), (ii) "nicking" of polypeptide chains in assembled vicilin molecule. Finally, they concluded that vicilin molecule is a multimer, probably a trimer of 50,000 Mr subunits but that these subunits may be "nicked" by post-translational proteolysis without significantly altering the structure of the molecule as a whole.

The following structural model for vicilin, showing the molecule as a trimer of Mr 50,000 subunits, has been put forward by Gatehouse et al. (1984). The generalized vicilin precursor polypeptide has two sites at which post-translation proteolysis may occur. Cleavage at both these sites gives three subunits as products, polypeptides of 19,000 Mr \( (a) \), 13,500 \( (\beta) \) and 12,500 Mr or 16,000 Mr, if glycosylated \( (\gamma) \). If neither of the potential cleavage sites undergoes proteolysis the resulting intact 50,000 Mr polypeptide is denoted \( (a+\beta+\gamma) \). Finally, cleavage at one potential processing site but not the other leads to two possibilities; cleavage at \( \beta:\gamma \) site gives \( (a+\beta) \), which is the 33,000 Mr polypeptide, and \( \gamma \), while cleavage at the \( a:\beta \) site gives \( a \) and \( (\beta+\gamma) \) which has been tentatively identified as a minor vicilin polypeptide of 31,000 Mr.
Model of derivation of vicilin polypeptide from the 50,000 Mr precursors. The top line represents the vicilin messenger RNA.

Unlike legumin, vicilin contains no measurable amounts of sulphur amino acids and thus can contain no disulphide-linked polypeptides, however, it contains small amounts of covalently bound carbohydrate (Basha and Beevers, 1976; Gatehouse et al., 1980b, 1984). The comparison between a predicted complete amino acid sequence for vicilin and the partial or complete amino acid sequences of the different vicilin subunits, shows that the vicilin subunits of Mr less than approx. 50,000 are very similar in sequence to parts of the 50,000 Mr precursor (Boulter, 1984).
1.3.2 Legumin

Casey (1979a) and Croy et al., (1979) determined the molecular weight of pea legumin to be 395,000 and 390,000 respectively. Polyacrylamide gel electrophoresis using non-reducing denaturants like SDS, resulted in the dissociation of native legumin into subunits of approx. Mr 60,000. On addition of reducing agents such as 2-mercaptoethanol, the Mr 60,000 subunits dissociated into smaller polypeptides of Mr 40,000 and 20,000 (Croy et al., 1979). The study of Gatehouse et al., (1980a) has shown that pea legumins have molecular weight between 380,000 and 410,000 and consist of six subunit pairs, each of which is made up of a larger, relatively acidic subunit (α) and a smaller, relatively basic subunit (β) joined by a disulphide bond. Cell-free translation of poly (A)-mRNA preparations from developing pea seeds has demonstrated that the subunit pairs are synthesized covalently joined together and that they are subsequently cleaved to yield the two subunits (Croy et al., 1980a). This was later confirmed by cDNA sequence information which showed the absence of any initiation or termination codones in the region immediately upstream of the β-subunit coding sequence (Croy et al., 1982; Lycett et al., 1984b). The acidic subunits were richer in aspartic, glutamine and arginine, while the basic subunits were richer in alanine, valine and leucine. Gatehouse et al. (1984) presented the following model which proposed that legumin consists of a hexamer of disulphide-bonded subunit pairs i.e. six pairs of polypeptides, each pair containing one large and one small subunit held together by one or more disulphide bonds.
Model of derivation of legumin α- and β-polypeptides from the 60,000 Mr precursors. The top line represents the legumin messenger RNA.

In pea legumin subunit pairs the larger or α-subunits have been shown to have pI values in the range of 4.8 to 6.1, whereas the smaller or β-subunits have higher pI values in the range of 6.2 to 8.0 (Krishna et al., 1979; Matta et al., 1981). The amino acid composition of legumin (Gatehouse et al., 1984) shows low levels of methionine and cysteine. Finally, legumin has been reported to contain no covalently attached carbohydrate (Casey, 1979a; Gatehouse et al., 1980a).

1.3.3 Convicilin

Convicilin, the third storage protein of pea has a molecular weight of 210-280,000 and is dissociated by denaturants (e.g. SDS) into subunits of 71,000 Mr (Croy et al., 1980c). On this basis it has been suggested that convicilin is a simple multimer, possibly a tetramer. The close antigenic relationship of this protein to vicilin
has already been noted; however, convicilin contains no significant amounts of carbohydrate, but does contain sulphur amino acids (Gatehouse et al., 1984).

1.3.4 Storage protein synthesis

Gatehouse et al. (1982a) found that storage-protein synthesis commenced 8-9 days after flower opening (d.a.f.) and ceased at 21 d.a.f. Vicilin and legumin mRNAs, which are present in very low concentrations during earlier stages of seed development become major components of total mRNA after approx. two thirds of the interval from flower opening to cessation of synthetic activity in the seed (14 d.a.f.). Legumin mRNA remained at relatively high levels until the end of this period (approx. 21 d.a.f.) whereas the mRNA encoding vicilin 50,000 Mr polypeptides declined to lower levels before storage-protein deposition ceased (i.e. by 19 d.a.f.). A second vicilin mRNA, encoding the 47,000 Mr polypeptide, became a detectable component in total mRNA 2-3 days earlier in seed development than the other mRNAs and also declined earlier (by 16 d.a.f.). These results demonstrate developmental regulation of mRNA concentration and it was suggested that control was likely to occur at the level of transcription. More recently, the regulation of the transcription of storage-protein mRNA in nuclei isolated from developing pea cotyledons was described by Evans et al. (1984). They found that the proportion of vicilin to legumin transcripts changed during development; vicilin transcripts were predominant at 9 and 11 d.a.f. and were similar in amount to legumin at 14 d.a.f., whereas at 18 d.a.f., legumin transcripts predominated and little vicilin transcription was observed.
1.3.5 Genetic variation

Genetic variation in band pattern of the seed proteins of pea (*Pisum sativum* L.) on gel electrophoresis has been observed for both the major storage proteins, legumin (Davies, 1973) and vicilin (Hynes, 1968). In the case of legumin, Thomson and Schroeder (1978) found that acidic subunits were each controlled by a single pair of co-dominant genes, whereas the basic subunits were under multigenic control. Casey (1979b) used two-dimensional electrophoresis analysis to study the inheritance of legumin subunits and showed that the major acidic subunits of legumin followed a 1:2:1 segregation in the F2 generation i.e. were controlled by a single pair of co-dominant genes. Both the results of Thomson and Schroeder (1978) and Casey (1979b) were confirmed by the findings of Matta and Gatehouse (1982) that the major legumin acidic subunits behaved as products of a single Mendelian gene with at least five different possible alleles.

In the case of the other major storage protein "vicilin", it was shown that F1 seed protein band patterns arose from the additive patterns of the two parental types (i.e. that vicilin protein molecules were inherited additively) and that the patterns were controlled by alternative alleles at a single locus. Further vicilin loci were identified by Thomson and Schroeder (1978) by carrying out gel electrophoresis under differing conditions and the use of denaturing conditions allowed some of the vicilin loci to be identified with the "subunits" or component polypeptides of the protein. These results showed that in reciprocal crosses between two lines of peas, band patterns of vicilin from F2 seeds segregated in a 1:2:1 parent A: additive: parent B ratio on gel electrophoresis,
indicating that vicilin band patterns were controlled by pairs of co-dominant genes at single loci. The loci were not identified with the one demonstrated by Hynes (1968), although they were claimed to correlate with vicilin structural genes on the pea genome. Of the three loci identified, one "Ve" was subsequently identified with a related storage protein, convicilin (Matta and Gatehouse, 1982) while the others were identified with component polypeptides of vicilin of Mr < 35,000. None of the loci were mapped.

Elucidation of the synthesis of vicilin has shown that the initial products of the vicilin structural genes are 50,000 Mr precursors, of differing amino acid sequences, some of which are proteolytically cleaved after translation is complete to yield the "subunits" or component polypeptides of Mr < 35,000 observed in the protein as isolated (Gatehouse et al., 1981, 1982b). Variation in the 50,000 Mr subunits of vicilin is thus most likely to reflect variation in the structural genes for the protein rather than to be a result of post-translational modification (which may not be simply related to protein structure). However, variation in vicilin 50,000 Mr subunit band patterns on SDS-PAGE has not been reported. In contrast, the well-established genetic variation in the α-subunits of legumin on SDS-PAGE (Casey, 1979b) has led to the mapping of a legumin locus, Lg-1, to chromosome 7 in the pea genome, near the r-tl where r (sometimes called ra) is the locus determining round (R) or wrinkled (r) seeds, and tl is the locus determining the presence of tendrils (Tl) or leaflets (tl) at the ends of the side-branches (r...5...tl...17...Lg-1) (Davies, 1980; Matta and Gatehouse, 1982). The latter authors also reported the mapping of the genetic locus for
the storage protein related to vicilin, convicilin (\textit{Cvc}) , to a segment of chromosome 2 on the pea genome (\textit{s...Cvc...3...k}) where \textit{s} is the locus determining "sticky" (\textit{s}) or "non-sticky" (\textit{S}) seeds ("sticky" seeds adhere together in the pod), and \textit{k} is the locus determining normal (\textit{k}) or keel-like (\textit{k}) wings on the flowers. The above data shows that the storage protein loci mapped are independent.

In French bean (\textit{Phaseolus vulgaris}), there are three groups of genes encoding phaseolin storage protein. Each group is inherited as a co-dominant allelic alternative (Brown \textit{et al.} 1981).

1.4 Plant gene structure

Many plant genes have been studied and showed a general form for plant gene structure. Like other eukaryotic genes, plant genes have intervening sequences (IVS or introns). All plant genes so far investigated obey the boundary rules of Breathnach \textit{et al.} (1978) i.e. the intron junction begins with the two bases GT at the 5' end and ends with AG at the 3' end. The number of the introns varies between different plant genes e.g. soybean nodulin-23 gene contains one intron which is 530 bp long, whereas the coding sequence is 669 bp long (Mauro \textit{et al.}, 1985). Pea legumin genes contain three introns (IVS-1, IVS-2 and IVS-3), two (IVS-1 and IVS-2) within the sequence encoding the legumin \textit{a}-subunit and one (IVS-3) within the sequence encoding the \textit{b}-subunit. In legumin gene "\textit{legA}" the three introns are 88, 88 and 99 bp long, respectively, (Lycett \textit{et al.}, 1984b). Soybean genes, actin (Shah \textit{et al.}, 1982), glycinin (Nielsen, 1984) and leghemoglobin (Brisson and Verma, 1982; Hyldig-Nielsen \textit{et al.}, 1982) each contains three introns. French bean phaseolin gene contains five introns (Slightom \textit{et al.}, 1983). In contrast maize zein genes (Messing \textit{et al.}, 1983), wheat gliadin gene (Anderson \textit{et al.}, 1984), pea nuclear gene "\textit{AB 80}" encoding the chlorophyll AB (Cashmore, 1984) and french
bean lectin gene (Hoffman, 1984) were found to contain no introns.

The consensus sequence AATAAA was found in the 3'-end of all plant genes studied except some members of the zein family in maize which have the sequence AATAAT instead of the normal AATAAA (Messing et al., 1983). There are at least two polyadenylation signals in each 3' untranslated region in plant genes. Rather than the simple AATAAA, these often take the form AATAA(A)TAAA, and sometimes a G residue is substituted for one of the nucleotides in the sequence (Lycett et al., 1983a). In pea legumin gene "legA" there are three polyadenylation signals which are AATAAGAAAA, AATAAATAAAA and AATAAATAAA, respectively downstream of the stop codon (Lycett et al., 1984b). Mauro et al. (1985) found that soybean nodulin-23 gene contains three poly A addition signals "ATAAA" located 16,57 and 61 bp from the 3' end of the message. The position of the polyadenylation signals was found to be different in plant genes with respect to the stop codon and the poly A tail. Soybean leghemoglobin and maize alcohol dehydrogenase and zein genes have AATAAA close to the stop codon and another variant of this sequence close to the poly A tail, i.e. leghemoglobin has GATAAA (Brisson and Verma, 1982; Hyldig-Nielsen et al., 1982), alcohol dehydrogenase has AATGAG (Gerlach et al., 1982) and zein genes have (A/G)ATAA1-3 (Messing et al., 1983). In contrast, pea legumin gene "legA" contains AATAAG close to the stop codon and AATAAA near the poly A tail (Lycett et al., 1984b).

The sequencing results of higher plant genes show the presence of the sequence TATA or "TATA box" at the 5'-end. Soybean leghemoglobin and actin genes (Brisson and Verma, 1982; Shah et al., 1982) and pea legumin gene "legA" (Lycett et al., 1984b) have the dinucleotide TC
preceding the "TATA box". The zein gene family in maize have variable pairs of nucleotides preceding the "TATA box" i.e. TC, CC and TG (Pedersen et al., 1982; Messing et al., 1983). The transcription start point (cap site) of several plant mRNAs lie 18-33 bp downstream from the "TATA box" (Messing et al., 1983; Vodkin et al., 1983; Lycett et al., 1984b). In pea legumin and soybean glycinin genes, this distance was found to be 25 bp (Lycett et al., 1984b; Nielsen, 1984). The translation start or the "ATG" initiation codon was found 11-68 and 30 bp downstream from the "cap site" in the zein gene family in maize and pea legumin gene "legA", respectively (Messing et al., 1983; Lycett, et al., 1984b). Cashmore (1984) found that pea nuclear gene (AB 80) which encodes for the chlorophyll AB polypeptides contains two TATA sequences, the first occurs 31 nucleotides 5' from the "cap site" and the second was found 7 nucleotides on the 5' side of the initiation methionine codon and the surrounding sequences of both "TATA boxes" were found to be strikingly similar.

Other sequences may be involved in regulation of transcription of eukaryotic genes; an example is the consensus sequence GG(C/T)CAATCT or the "CAAT box". This sequence appears 80-100 bp upstream of the "cap site" (Messing et al., 1983; Lycett et al., 1984b; Nielsen, 1984). However, this sequence does not appear to be present in some plant genes (e.g. zein) whereas it is in others (e.g. legA). A further sequence was described by Messing et al., (1983) as (C/T)A 2-5 (G/T) NGA 2-4 (C/T)(C/T) or "AGGA box" at 36-59 bp upstream from the "TATA box". This sequence was found adjacent to the "CAAT box" on the 5' side of soybean leghemoglobin and actin genes (Hyldig-Nielsen et al., 1982; Shah et al., 1982), zein gene family in maize (Pedersen et
al., 1982) and pea legumin gene *legA* (Lycett et al., 1984b).

### 1.5 Legumin gene family of pea

Several legumin genes were isolated from pea genomic bank of pea *Pisum sativum L.* (Shirsat, 1984; Croy et al., paper in preparation). One of these genes *legA* has been completely sequenced, revealing typical plant gene features (Lycett et al., 1984b). These features were discussed in detail earlier. This gene encodes a legumin precursor which contains a 21 amino acid-long signal peptide followed by 8-subunit (36,440 Mr) and a-subunit (20,190 Mr). The translation product of this gene is relatively rich in sulphur amino acids, containing 4 met and 5 Cys residues.

A second legumin gene, denoted *legD*, has been sequenced (Bown et al., 1985). The complete nucleotide sequence of this gene shows that it is a non-functional gene "pseudogene", has two stop codons near the 5' end as well as deletions and frame shift errors when compared with *legA* and no transcripts from this gene were detected in developing seeds. A significant homology was found between *legA* and *legD* in their coding sequences, whereas partial homology was found between them in the intron sequences and the immediate 5' flanking sequences. Other flanking sequences of the two genes showed no significant homology, apart from the presence of polyadenylation signals 3' end to both coding sequences. The intervening sequences (IVS-1 and IVS-3) of *legD* are homologous in location to their correspondents in *legA* but slightly shorter (84 and 92 bp, respectively). IVS-2 can not be compared because of a deletion in the coding sequence 3' end of this intron including its 3'
A fourth legumin gene, denoted $\text{legC}$, has been sequenced (Gatehouse, unpublished results). The complete nucleotide sequence of this gene shows a strong homology with $\text{legA}$ in the coding sequences, the intervening sequences and the 5' and 3' flanking regions. Thirteen base substitutions were found in $\text{legC}$ (compared with $\text{legA}$), three active substitutions in the coding sequences resulted in three different amino acid residues and ten substitutions in the non-coding sequences. The direct repeated sequences found in two cDNA clones coding for legumin protein in *Pisum sativum* (Lycett et al., 1984a) are present in both $\text{legA}$ and $\text{legC}$.

1.6 Ribosomal RNA

The ribosomal RNA coding sequences of the eukaryotic nucleus (rDNA) are well known to be among the most conserved DNA sequences with respect to base-pair substitutions. Eukaryotic nuclear genes encoding the 17S, 5.8S and 25S rRNA are arranged as tandem repeats, each repeat containing a coding region for all 3 rRNAs and a non-transcribed spacer (NTS). The coding region is the most conserved part of the repeat and has much less length variation than the NTS region. The three genes are transcribed as a single ribosomal precursor which is subsequently processed in several steps. The internal transcribed spacer (ITS) region between the 17S and 25S rRNA is less conserved than the NTS region (Jorgensen et al., 1982). The non-transcribed region is the source of the length variation in a large number of plant species so far investigated. In legumes, only part of the non-transcribed region is composed of repeats, while the
rest of the region is also variable in length and nucleotide sequence (Jorgensen et al., 1982). Following a diagram showing pea rDNA repeats.

Diagrammatic representation of pea rDNA repeats. NTS = non-transcribed spacer, ITS = internal transcribed spacer, SR = subrepeat and R = Eco RI restriction site.

The ribosomal repeat length varies in different plant species; 9-10 kbp in barley, rye and wheat (Appels et al., 1980; Gerlach and Bedbrook, 1979); more than 11.0 kbp in carrot (Kato et al., 1982); 8.8 kbp in lupin (Rafalski et al., 1983); 9 kbp in maize (Messing et al., 1984); 9-12 kbp in pea (Ellis et al., 1984); 10 kbp in pumpkin (Siegl and Kolacz, 1983); more than 11.0 kbp in radish (Delseny et al., 1983) and 9.0 kbp in soybean (Friedrich et al., 1979).

Although the repeat units of species all are presumed to contain the same information, they frequently are not exactly alike. Heterogeneity has been observed in length, nucleotide sequence and base modification. Two types of length heterogeneity have been described. The first type that occurs in Drosophila and other insect species results from the absence or presence of a variable length insertion which interrupts the 28S rRNA coding sequence (Glover,
1981). The second type is a consequence of differing numbers of base-pair subrepeat in the non-transcribed spacer region (Krystal and Arnheim, 1978; Gerlach and Bedbrook, 1979; Oono and Sugiura, 1980; Rae et al., 1981; Lamppa et al., 1984). In *Pisum sativum* and *Vicia faba*, the subrepeat (SR) was found to be 180 and 326 bp, respectively (Jorgenson et al., 1982; Yakura et al., 1984). Nucleotide sequence heterogeneity has been described in both coding and non-coding regions (Gourse and Gerbi, 1980; Maggini and Cormona, 1981; Jorgensen et al., 1982; Siegel and Kolacz, 1983). A third kind of heterogeneity, base modification has been described for the rDNA of a number of species and in most cases probably results from variable methylation of cytosine residue (Jorgensen et al., 1982; Siegel and Kolacz, 1983; Messing et al., 1984). However, no repeat length heterogeneity was observed in rDNAs of *Allium cepa* (Maggini and Cormona, 1981), lupin (Rafalski et al., 1983), radish (Delseney et al., 1983) and soybean (Varsanyi-Breiner et al., 1979). The studies of Jorgensen et al. (1982) and Ellis et al. (1984) on pea have shown length variation in rDNA both within and between different lines. Such length heterogeneity was used as marker loci in different genetic studies e.g. determination the mode of the inheritance of ribosomal RNA genes in crosses between different pea lines (Ellis et al., 1984) and in this study in an attempt to estimate the chromosomal location of ribosomal RNA clusters in *Pisum sativum*. Finally, the spacer length variation in barley rDNA was found to be under the control of two co-dominant Mendelian loci. These two loci were located on chromosomes 6 and 7 by Saghai-Maroof et al. (1984).
It has been known for many years that ribosomal RNA genes are multiple and clustered at the secondary chromosome constrictions which are often "nucleolus organizers" (Attardi and Amaldi, 1970; Birnstiel et al., 1971; De-Faria, 1976; Burger and Knälmann, 1983). There is considerable variation between plant genera in the number of rRNA genes per haploid complement (Ingle and Sinclair, 1972), even between plant species which have the same number of chromosomes (Maher and Fox, 1973). *Pisum sativum* was found to have 3,900 rRNA copies per haploid complement (Ingle and Sinclair, 1972). The correlation between the number of rRNA cistrons and the number of nucleolus organizers has been investigated in different plant species; in maize, Phillips et al. (1971) found that the number of rRNA cistrons is approx. doubled in a strain possessing a cytologically-visible "duplication" of the nucleolus organizer, whereas in different wheat genotypes, Mohan and Flavell (1974) found a clear departure from a correlation between the number of rRNA cistrons and the number of nucleolus organizers. In different polyploid *Aegilops* species, Teoh et al. (1983) found that the number of rRNA gene sites revealed by *in situ* hybridization with rRNA radioactive probe are consistently higher than the number of metaphase satellited chromosomes stained by conventional means, whereas in diploid species, the number of satellited chromosomes almost indicate the number of rRNA gene sites. In *Pisum*, Morrison and Lin (1955) identified two pairs of satellited chromosomes at the long arms, which were classified as chromosomes 4 and 7 (Blixt, 1974; Lamm, 1981). Finally, cytological molecular hybridization of radioactive rRNA or rDNA to its complementary DNA *in situ* on metaphase
chromosomes of different plant species has located the rRNA genes to
the nucleolus organizer regions of the chromosomes (Brady and Clutter,

1.7 Objectives, rationale and content of the present research

This thesis has as its aim an investigation of variation between
genes in Pisum sativum, not only at the biochemical phenotypic level,
i.e. in terms of variation in polypeptide gene products, but also at
the genotypic level, i.e. in terms of variation in the genes
themselves.

Variation in biochemical characteristics at the phenotypic level
can be investigated through isoenzyme banding patterns on gel
electrophoresis, or by variation in polypeptide band patterns on gel
electrophoresis, if the variant polypeptides can be clearly assigned
to known protein species. Mapping of these biochemical markers onto
the genome map of pea then gives known genetic loci that can be used
as markers in crossing experiments and/or correlated to useful
phenotypic characteristics. A similar rationale underlies the use of
genotypic variation in terms of restriction fragment polymorphism,
where variant restriction fragments in genomic DNA, as detected by a
specific or semi-specific cloned DNA probe, are correlated with a
phenotypic characteristic and can be mapped to a genetic locus.

Can a correspondence be found between genotypic and phenotypic
variants? In theory, biochemical phenotypic variation in polypeptides
can (usually) be assigned to variation in the primary amino acid
sequence of those polypeptides, and hence to a genotypic variation in
the nucleic acid sequence of their encoding genes. Restriction fragment polymorphism is not usually due to variation in gene sequences, since the majority of DNA does not code for gene products; however, variation at a site near to a coding gene will be very closely linked to variation actually within the gene itself. Even within the coding sequence, not all nucleotide base changes will give rise to variants in amino acid sequence, or to variants that are analysable other than by amino acid sequencing. Variation at the genotypic level, when analysed in molecular terms, will therefore being expected to be much greater than that observed at the phenotypic level.

The value of a "two-pronged" approach nevertheless lies in the possibility of being able to deduce correspondence between regions of the genome, and eventually genes, and phenotypic characteristics where the molecular basis of the phenotype is not known, through the use of known genetic loci which can be cloned via complementary DNAs identified from the known encoded polypeptide. For example, isolation of a genomic clone corresponding to a gene mapping close to the r locus controlling round/wrinkled seeds in pea could lead to isolation of the r gene itself through a "gene walking" strategy. In this thesis the following techniques of examining both phenotypic and genotypic charaters are used in order to explore ways in which such an integration of characters can be achieved.

(i) Three isoenzyme systems (amylase, esterase and glutamate oxaloacetate transaminase) were examined in seeds of pea (Pisum sativum L.) in order to find clear variation in their band patterns on gel electrophoresis between different lines.
The inheritance of these isoenzyme systems was investigated by the analysis of the F2 segregates from crosses between the lines which showed variation in the isoenzyme banding patterns. Linkage between the genes for these isoenzymes and selected morphological markers was investigated to locate the enzyme structural gene locations onto the pea genome.

(ii) Different pea lines were screened for genetic variation in vicilin 50,000 Mr subunit band patterns on SDS-PAGE. Crosses between different lines were carried out and linkage relations between the gene for vicilin 50,000 Mr subunit and morphological and biochemical marker genes were tested to investigate the inheritance and mapping of a putative vicilin structural gene locus (Vc-1).

(iii) Genomic DNA was isolated from leaves of different pea lines and purified twice on caesium chloride gradients. The pure DNA was digested to completion with Eco RI. The fragments were then separated by electrophoresis on agarose gels and transferred to nitrocellulose filter. The genomic blots were hybridized with the radioactive probe of legumin in order to find length variation in the restriction fragments to enable the molecular mapping of legumin genes onto pea genome.

(iv) A third legumin gene from pea nuclear genome, referred to as legB, was sequenced using the M13 sequencing system and "dideoxy chain termination" method of Sanger et al. (1977). The coding sequence, the 3' flanking region were completely sequenced and the homologies between this gene and other legumin genes (legA and legC) were estimated.
(v) Different pea lines were screened for length variation in rDNA fragments of Eco RI digests. The lines which showed length variation in the restriction fragments were crossed together. DNA was then isolated from single F1 and F2 plants and the inheritance of ribosomal RNA genes and its linkage relations with other morphological markers were tested.

(vi) Cytolocalization of rRNA genes to pea chromosomes was carried out by in situ hybridization of rDNA radioactive probe to metaphase chromosomes from root tips of pea (Pisum sativum L.) line 110. Finally, physical mapping was carried out to estimate the locations of rRNA gene sites on the chromosomes in map units.
2. MATERIALS AND METHODS
2.1 Biological materials

Table (1) lists the pea lines and their genotypes which were tested for their isoenzyme, storage protein and nucleic acid patterns, and between which crosses were made. Seeds of these lines were supplied by Dr. S. Blixt, Landskrona, Sweden. Seeds of the var. Feltham First were from Sutton Seeds Ltd., Reading, Berks., UK. In addition, seeds of a number of isogenic pairs of pea lines differing only at the r locus within each pair were supplied by Dr. A.E. Slinkard, University of Saskatchewan, Saskatoon, Canada.

2.2 Chemicals

Reagents, unless otherwise indicated, were obtained from BDH Chemicals Ltd., Poole, Dorset, UK, and were of analytical grade or the best available. The following materials were purchased from the designated sources.

Adenosine triphosphate (ATP); L-aspartic acid, bovine serum albumin, 4-chloro-1-naphthol, Coomassie brilliant blue R250, dithiothreitol (DDT), ethidium bromide, fast blue B salt (o-dianisidine), herring sperm DNA, iodine, α-Ketoglutaric acid, α-naphthyl acetate, poladenylic acid (poly A), potassium iodide, Pyridoxal-5'-phosphate and Tris (Trizma base, reagent grade) were from Sigma Chemical Co., Poole, Dorset, UK.

Sephadex G50 superfine and Ficoll 400 were from Pharmacia Fine Chemicals, Uppsala, Sweden.

Agarose (electrophoresis grade), restriction enzymes, IPTG (Isopropyl-β-D-thio-galactopyranoside) and E. coli DNA polymerase I large fragment (Klenow enzyme) were from Bethesda Research Laboratories (UK) PLC. Cambridge, UK.

Nick translation kits, $^3$H, $^{32}$P and $^{35}$S radio-labelled nucleotides were from Amersham International, White Lion Road, Amersham, UK.

3MM and GF/C papers were from Whatman PLC. Springfield Mill, Maidstone, Kent, UK.
Deoxy- and dideoxynucleoside triphosphates were from Pharmacia P.L. Biochemicals Inc., Pharmacia (Great Britain) Ltd., Milton Keynes, Bucks., UK.

Ilford nuclear research emulsion was from Ilford Limited, Basidon Essex, UK.

Nitrocellulose filters (BA 85, 0.45 µm) were from Schleicher and Schüll, Anderman and Co. Ltd., Kingston-upon-Thames, Surrey, UK.

Caesium chloride and sodium chloride (A.R.) were from Koch-Light Ltd., Haverhill, Suffolk, UK.

Bacto-Trypton, Bacto-Agar and Bacto-Yeast Extract were from Difco Laboratories, Detroit, Michigan, USA.

X-Gal (5-bromo-4-chloro-3-indolyl- β-galactoside) and T4 DNA ligase were from The Boehringer Corporation (London) Ltd., Lewes, East Sussex, UK.

2.3 Recombinant clones

pHA1 : a plasmid clone containing pea ribosomal RNA gene repeat was supplied by Dr. R.E. Cuellar of Plant Breeding Institute, Cambridge, UK.

M13 (pHA1) : a phage clone containing pea ribosomal RNA gene repeat was supplied by Dr. J.A. Gatehouse, Dept. of Botany, Univ. of Durham, Durham, UK.

pRC 6.2 and pRC 5.4 : plasmid clones containing legumin genes, legA and legB, respectively, were supplied by Dr. R.R.D. Croy, Dept. of Botany, Univ. of Durham, Durham, UK.

2.4 M13 vectors and bacterial strains

M13; mp8, mp9, mp18, mp19 and E. coli strains, JM101 and JM105 were supplied by Dr. J. Messing, Dept. of Biochemistry, Univ. of Minnesota, St. Paul, MN, USA.

2.5 Genotypes and nomenclatures

Amy = locus for amylase variant band patterns on PAGE.
Cve = locus for convicilin band patterns on SDS-PAGE.
Est = locus for esterase variant band patterns on PAGE.
Got = locus for GOT variant band patterns on PAGE.
Lg-1 = legumin major α-subunit band patterns on SDS-PAGE.
Vc-1 = vicilin 50,000 Mr polypeptide band patterns on SDS-PAGE.
LegB = legumin gene from genomic clone λLeg2.

2.6 Growth of biological materials

Seeds of pea lines were germinated on Alkathene polyethylene beads in a dark spray room at 20°C with water misting at regular intervals for 4-5 days. Seedlings were then transferred to plastic pots containing compost (Levington : John Innes No.2 = 1 : 3 parts) in the Botanic Gardens, University of Durham. Plants were grown under natural daylight between March and November, 1982 for the parents and F1 plants and between March and June, 1983 for F2 plants. The plants were watered weekly with commercial nutrient feed containing all necessary elements, and at all other times with tap water.

The flowers were emasculated at the closed bud stage by gripping the bud between two fingers; the petals were opened gently and the anthers were removed with forceps leaving behind an emasculated flower which was cross pollinated immediately. The pollinated flowers were kept inside empty gelatin capsules.

2.7 Preparation of single seed extracts

The testa was removed from a part of the seed and a quarter of the seed (half of the cotyledon) was cut off taking care that the radicle and the other parts of the embryo were not injured. The cotyledon was then ground and a known weight of the seed meal (10 mg and 30 mg for total protein and enzymes, respectively) was suspended in 100 µl of extraction buffer (see below). The resuspension was left at 4°C for 12-16 h, then clarified by centrifugation in a microcentrifuge for 5 min. Subsequently 10 µl and 30 µl aliquous of
the supernatant were loaded into the gels for total protein and
enzymes, respectively.

2.8 Extraction buffers

The extraction buffer for total protein consisted of 0.2 M Tris-
HCl pH 6.8, 2% (w/v) sodium dodecyl sulphate (SDS), 10% (w/v) sucrose
and 1% (v/v) 2-mercaptoethanol, whereas that for isoenzymes consisted
of 0.025 M sodium phosphate pH 7.3 with 20% (w/v) sucrose.

2.9 Polyacrylamide gel electrophoresis (PAGE)

2.9.1 Isoenzymes

Polyacrylamide gel electrophoresis under non-dissociating
condition was carried out by the method of Gabriel (1971). 7.5%
acrylamide gels were prepared using the following stock solutions:

<table>
<thead>
<tr>
<th>Stock solution</th>
<th>Separating gel</th>
<th>Stacking gel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acrylamide</td>
<td>4.5 g</td>
<td>0.9 g</td>
</tr>
<tr>
<td>NN-Methylene-bisacrylamide (Bis)</td>
<td>0.1 g</td>
<td>25.0 mg</td>
</tr>
<tr>
<td>1M Tris-HCl pH 8.8</td>
<td>22.5 ml</td>
<td>-</td>
</tr>
<tr>
<td>1M Tris-HCl pH 6.8</td>
<td>-</td>
<td>2.5 ml</td>
</tr>
<tr>
<td>Distilled water</td>
<td>34.0 ml</td>
<td>17.0 ml</td>
</tr>
<tr>
<td>1% Ammonium persulphate</td>
<td>1.5 ml</td>
<td>0.5 ml</td>
</tr>
<tr>
<td>NNNN-Tetramethyl ethylene diamine (TEMED)</td>
<td>20 μl</td>
<td>20 μl</td>
</tr>
</tbody>
</table>

Electrophoresis buffer (10x stock) pH 8.3/litre

Glycine 141.0 g
Tris base 30.0 g

2.9.2 Total Protein

SDS-polyacrylamide gel electrophoresis was carried out as
previously described (Laemmli, 1970; Matta et al., 1981). 12% acryl-
amide gels were prepared using the following stock solutions:
Stock solution  

<table>
<thead>
<tr>
<th>Stock solution</th>
<th>Separation gel</th>
<th>Stacking gel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acrylamide stock</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(containing 30g acrylamide, 135 mg*)</td>
<td>25.8 ml</td>
<td>3.0 ml</td>
</tr>
<tr>
<td>Bis and distilled water to 100 ml</td>
<td>22.5 ml</td>
<td>-</td>
</tr>
<tr>
<td>1M Tris-HCl pH 8.8</td>
<td>-</td>
<td>2.5 ml</td>
</tr>
<tr>
<td>1M Tris-HCl pH 6.8</td>
<td>8.6 ml</td>
<td>13.8 ml</td>
</tr>
<tr>
<td>Distilled water</td>
<td></td>
<td></td>
</tr>
<tr>
<td>degassed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1% Ammonium persulphate</td>
<td>1.5 ml</td>
<td>0.5 ml</td>
</tr>
<tr>
<td>10% SDS</td>
<td>0.6 ml</td>
<td>0.2 ml</td>
</tr>
<tr>
<td>TEMED</td>
<td>20 μl</td>
<td>10 μl</td>
</tr>
</tbody>
</table>

* 435 mg for the stacking gel stock solution

Electrophoresis buffer (10x stock) pH 8.3/litre

- Glycine: 141.0 g
- Tris base: 30.0 g
- SDS: 10.0 g

Electrophoresis was initially carried out at 15 mA for 1 h and then subsequently at 25 mA until the tracking dye (0.1% "w/v" bromophenol blue in ethanol) reached the bottom of the gel.

2.10 General gel assay methods

2.10.1 Amylase

Amylase bands were stained according to the method of Brewbaker et al. (1968). The gels were incubated for 5 min in 0.2 M sodium acetate buffer pH 5.0, followed by incubation for 2 h in 1% (w/v) soluble starch in 0.2 M sodium acetate buffer pH 5.0. The gels were then washed for 5 min in 0.1 M acetic acid followed by staining in potassium iodide/iodine solution (0.1 g potassium iodide, 20 mg iodine, 100 ml of 0.1 M acetic acid).

2.10.2 Esterase

Esterase assay was based on the method of Shaw and Prasad (1970).
Esterase bands were detected after pre-incubation of the gels for 5 min in 0.2 M sodium acetate buffer pH 5.0 containing α-naphthyl acetate (1 mg/ml), and staining for 15 min at 37°C in 0.2 M sodium acetate buffer pH 5.0 containing fast blue B salt (1 mg/ml).

2.10.3 Glutamate oxaloacetate transaminase (GOT)

GOT assay was carried out according to the methods of Shaw and Prasad (1970). GOT bands were visualized by incubation for 10 min at 37°C in 200 ml of a staining solution containing 0.2 M sodium acetate buffer pH 5.0, L-aspartic acid (500 mg), α-Ketoglutaric acid (70 mg), Pyridoxal-5-phosphate (10 mg) and fast blue B salt (100 mg).

2.10.4 Total protein

Protein bands were stained overnight with 0.05% Coomassie brilliant blue in methanol : acetic acid : water (50 : 7 : 43). Developed gels were destained in methanol : acetic acid : water (50 : 7 : 43), photographed and dried between layers of uncoated cellophane.

2.10.5 Immunological detection of vicilin polypeptides after "Western" blotting to nitrocellulose filters

The method used was based on that of Towbin et al. (1979). Vicilin subunit bands were transferred from SDS-polyacrylamide gels to nitrocellulose filters by electroblotting as follows:

The unstained gel was placed on a presoaked (in water) nitrocellulose filter ensuring that no air bubbles were trapped (they would prevent the current from flowing through). Sponges and elastic pads
were used to hold the gel and the filter together. The assembly was placed in a Trans-Blot Cell (Bio. Rad Laboratories) containing transfer buffer (25 mM Tris base, 192 mM glycine, 20% "v/v" methanol, pH 8.3), and the transfer was allowed to proceed for 12 h at 60V. The filter was then removed and incubated in 100 ml of Tris-saline buffer (20 mM Tris-HCl pH 7.4, 0.9% "w/v" NaCl) containing 5% (w/v) BSA for 1 h at 40°C in a heat-sealed plastic bag. The filter was rinsed briefly at 25°C in 100 ml of Tris-saline buffer in a plastic bag. It was then incubated for 2-3 h with 40 µl of affinity-purified, rabbit antivicilin IgG diluted into 30 ml of Tris-saline buffer, 5% BSA in a plastic bag (at 25°C). The filter was washed for 45 min in 4 changes of Tris-saline buffer (100-200 ml per wash) at 25°C. It was then incubated for 1 h at room temperature with 40 µl of swine, peroxidase-conjugated, antirabbit IgG (Orion Diagnostica, Helsinki, Finland) at 25°C. The filter was then washed in Tris-saline buffer as above. 20 ml of a solution of 4-chloro-1-naphthol (3 mg/ml in methanol) were diluted into 100 ml of Tris-saline buffer, and 40 µl of 30% (v/v) hydrogen peroxide solution were added. The nitrocellulose filter was placed in this mixture and kept in the dark. The reaction was stopped after suitable intense staining had been achieved (20-30 min) by washing the filter thoroughly in water. The blot was dried between sheets of 3MM paper and stored in the dark.

2.11 Preparation of pea DNA

The method used to prepare DNA from plant leaves was essentially that of Graham (1978). The green leaves (14 days after germination) were ground in a precooled mortar after freezing in liquid nitrogen.
All steps from this point onwards were done as quickly as possible, in order to keep nuclease activity at a minimum. The ground tissue were homogenized in homogenizing buffer (see below), and the solution was made 1M with sodium perchlorate (5M stock). 0.5 volume redistilled phenol and 0.5 volume chloroform/octanol (1% octanol) were then added, and the mixture was shaken on a rotary shaker at 4°C for 50 min, and then centrifuged, in order to sediment leaf debris. The supernatant was removed, shaken with an equal volume of chloroform/octanol (1% octanol) and centrifuged at 8000 rpm for 1 min at 10°C to separate phases. The aqueous phase was collected and the nucleic acids were precipitated by the addition of 3 volumes of ethanol at -20°C. The nucleic acids were spooled out on a spatula, and redissolved in resuspension buffer (see below) by shaking very gently for 24 h, in order to minimize shearing of the high molecular weight nucleic acids. The nucleic acid solution was digested with pronase (500μg/ml, previously autodigested at 37°C for 2 h) for 3 h at 37°C. The solution was made up to 8.3 ml with resuspension buffer, and 8 g of CsCl were added, along with ethidium bromide (EtBr) to a final concentration of 300 μg/ml. The solution was mixed and centrifuged at 40,000 rpm in a MSE Superspeed Ultracentrifuge using a 10 x 10 ml rotar at 15°C for 36 h. The DNA bands were collected from the tube by side puncture using a wide bore hypodermic needle. The DNA solution was recentrifuged in a fresh CsCl solution as before. The DNA bands were collected, and the ethidium bromide removed by extensive extraction with isoamyl alcohol saturated with resuspension buffer. The DNA solution was then dialyzed against several changes of TE buffer (see below) over a period of 12 h. DNA was then alcohol precipitated, the
pellet vacuum dried, and redissolved in a minimal volume of TE buffer.

<table>
<thead>
<tr>
<th>Homogenizing buffer</th>
<th>Resuspension buffer, pH 8.0</th>
<th>TE buffer, pH 8.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1 M NaCl</td>
<td>50 mM Tris-Hcl</td>
<td>10 mM Tris-HCl</td>
</tr>
<tr>
<td>0.025 M EDTA</td>
<td>10 mM EDTA</td>
<td>1 mM EDTA</td>
</tr>
<tr>
<td>2% SDS</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.12 Restriction endonuclease digestion

Pea genomic DNA was digested with the required restriction enzymes according to the enzyme optimal reaction conditions described by Maniatis et al. (1982). A total volume of 20 µl restriction solution containing 1 µg DNA, 2 µl of the appropriate 10x digestion buffer (see below) and 10 units of the restriction enzyme was incubated for 1 h at 37°C. The restriction solution was heated for 10 min at 70°C and cooled rapidly, in order to separate annealed cohesive ends. 10 µl of the agarose loading beads (see below) were added and the mixture was loaded into the gel.

### Restriction endonuclease digestion buffer (10x)

<table>
<thead>
<tr>
<th>Buffer</th>
<th>NaCl</th>
<th>Tris-HCl</th>
<th>MgCl₂</th>
<th>Dithiothreitol (DDT)</th>
<th>pH 7.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0</td>
<td>1 mM</td>
<td>10 mM</td>
<td>1 mM</td>
<td>1 mM</td>
</tr>
<tr>
<td>Medium</td>
<td>50 mM</td>
<td>10 mM</td>
<td>10 mM</td>
<td>1 mM</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>100 mM</td>
<td>50 mM</td>
<td>10 mM</td>
<td>1 mM</td>
<td></td>
</tr>
</tbody>
</table>

### Agarose loading beads

A mixture containing 10 mM Tris-HCl pH 8.0, 10 mM EDTA, 30% (v/v) glycerol, 0.1% (w/v) bromophenol blue, 0.1% (w/v) xylene cyanol and 0.2% (w/v) agarose was autoclaved and extruded through a syringe and fine needle.

2.13 Agarose gel electrophoresis

Horizontal slab gels used to separate, identify and isolate DNA fragments, were prepared by boiling the agarose in 180 ml distilled water (the weight of the agarose depended upon the agarose gel concen-
The mixture was allowed to cool to 70°C, then 20 ml of 10x Tris-acetate buffer (see below), and 20 μl EtBr (10 mg/ml) were added, the mixture was poured into the gel apparatus and allowed to set for 1 h. Samples to be electrophoresed were mixed with agarose loading beads and loaded into the gel slots using a micropipette. Gels were usually electrophoresed either at 100 V for 4 h or at 30 V for 12-16 h. In order to detect the DNA bands, EtBr was included in both the gel and the electrophoresis buffer (see below). After electrophoresis was completed, the gels were photographed using transmitted Ultraviolet light at 254 nm and Polaroid film ASA 3000 with an exposure time of 9 sec at f 5.6.

<table>
<thead>
<tr>
<th>Tris-acetate buffer (10x)/litre</th>
<th>Electrophoresis buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tris base</td>
<td>Distilled water</td>
</tr>
<tr>
<td>glacial acetic acid</td>
<td>1900 ml</td>
</tr>
<tr>
<td>0.5 M EDTA (pH 8.0)</td>
<td>Tris-acetate buffer (10x)</td>
</tr>
<tr>
<td></td>
<td>210 ml</td>
</tr>
<tr>
<td></td>
<td>EtBr (10 mg/ml)</td>
</tr>
<tr>
<td></td>
<td>200 μl</td>
</tr>
</tbody>
</table>

2.14 Recovery of DNA fragments from agarose gels

Legumin sequences "legA" were excised from the plasmid pRC 6.2 (pDUB 21) for nick translation by restriction with Bam HI restriction enzyme and purified by agarose gel electrophoresis according to the method of Yang et al. (1979). The digestion solution containing legumin sequence (3.1 kbp) and plasmid DNA (2.7 kbp) was electrophoresed on 0.5% agarose gel. The agarose slices containing legumin sequence were cut and transferred to eppendorf tubes, 3 volumes of 8 M sodium perchlorate were added and the tube swirled until the agarose had dissolved. A 6 mm circle of GF/C filter was cut and placed on a layer of 3 MM paper above a folded tissue and a nappy liner. The filter was wetted with 6 M sodium perchlorate (in TE buffer) and then
the DNA/agarose solution was applied in 30 μl aliquotes and allowed to soak through. The filter was then washed with 1600 μl of 6 M sodium perchlorate (in TE buffer) followed by washing with 1600 μl of ethanol (in 30 μl aliquotes). The filter was air dried for about 3 min, placed in a small eppendorf tube with 20 μl of 1 mM Tris-HCl buffer pH 7.5 and heated at 37°C for 30 min. The bottom of the tube was punctured and the solution collected by centrifugation for 5 min into a large eppendorf tube. This was followed by washing with 20 μl of 1 mM Tris-HCl buffer, pH 7.5 and the solution collected by centrifugation for 5 min and stored at -20°C. legB sequences were excised from the plasmid pRC 5.4 (pDUB 23) as above except that the restriction enzymes Hind III and Pst I were used to cut the legumin coding sequences into two small fragments (0.8 and 1.74 kbp) to be identified from the plasmid DNA containing fragment (2.7 kbp).

2.15 Isolation of DNA from low melting point agarose gels

The isolation of legB coding sequence from the plasmid pRC 5.4 (pDUB 23) for cloning purposes was carried out using low melting point agarose (LMP) according to the protocol supplied by BRL. Low melting point agarose was used to prepare 0.5% gel using Tris-acetate buffer (see Section 2.13). The agarose was boiled in water, cooled to 37°C before pouring. The gel was then placed in the tank whilst the comb and the gel surround were still in positions, covered with buffer prior to removing the comb and surround, and run overnight at 30 V. In order to remove DNA, the required fragments were cut and transferred to eppendorf tubes. The agarose was then melted by heating to 65°C and 2 volumes of 50 mM Tris-HCl, 0.5 mM EDTA pH 8.0 were added, mixed and
placed at 37°C for 10 min. An equal volume of phenol was added, mixed by vortexing for 15 sec and centrifuged for 5 min. The upper aqueous layer was removed and re-extracted with phenol. This was then re-extracted with an equal volume of chloroform : isoamyl alcohol solution (24 : 1), centrifuged for 4 min and the upper aqueous layer was removed. 0.1 volume of 3 M sodium acetate and 3 volumes of ethanol were added. This was left at -20°C for 1 h and the precipitate was collected by centrifugation for 7 min, washed with 1 ml of 80% ethanol, dried under vacuum for 10 min and resuspended in sterile water.

2.16 Transfer of DNA from agarose gels to nitrocellulose filters by "Southern" blotting

The technique used was essentially that described by Southern (1979). After gel electrophoresis was completed, the gel was transferred to a glass dish and denatured by soaking in several changes of denaturing solution (see below) for 1 h. It was then neutralized by soaking for another h in neutralizing solution (see below), followed by equilibration for a third h in 20x SSC (see below) and transferred to blotting apparatus in a cold room (4°C). The gel was carefully placed onto the apparatus and a sheet of nitrocellulose (previously soaked for 1 min in water) was soaked for 20 min in 20x SSC and placed on the gel. A piece of 3 M M paper was wetted in 20x SSC and placed over the filter, followed by three more layers of dry 3 M M paper, four layers of disposable nappies and a weight of about 500 g. The transfer of DNA was allowed to proceed for 16-24 h. After the transfer was completed, the filter was baked between two layers of 3 M M paper for
2 h at 80°C under vacuum. The filter can be stored at this stage for several months.

<table>
<thead>
<tr>
<th>20x SSC pH 7.0/litre</th>
<th>Denaturing solution</th>
<th>Neutralizing solution pH 7.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sodium chloride</td>
<td>1.5 M NaCl</td>
<td>3.0 M NaCl</td>
</tr>
<tr>
<td>Sodium citrate</td>
<td>0.5 M NaOH</td>
<td>0.5 M Tris base</td>
</tr>
<tr>
<td></td>
<td>1.0 mM EDTA</td>
<td>1.0 mM EDTA</td>
</tr>
</tbody>
</table>

2.17 \(^{32}\)P-labelling of double-stranded DNA (ds-DNA) by nick translation

The ds-DNA was labelled in vitro to a specific activity of \(>10^7\) cpm/µg using the nick translation kit supplied by Amersham International as described in its instructions. A reaction mixture was made of 20 µl \(\alpha^{(32}P)\) dCTP (200 µCi), 0.25 µg DNA, 10µl nucleotide/buffer solution (100 µM dATP, 100 µM dGTP and 100 µM dTTP in a concentrated nick translation buffer solution containing Tris-HCl pH 7.8, MgCl\(_2\) and 2-mercaptoethanol) and 10 µl enzyme solution (5 units DNA polymerase I and 100 µg DNase I in a buffer solution containing Tris-HCl pH 7.5, MgCl\(_2\), glycerol and bovine serum albumin). The nick translation reaction was allowed to proceed for 2 h at 15°C and terminated by adding 5 µl of 10% SDS. The products were separated from unincorporated label by gel filtration on a 5 cm column of Sephadex G50 superfine. The Sephadex was hydrated before use by boiling for 1 h in an elution buffer containing 150 mM NaCl, 10 mM EDTA, 50 mM Tris-HCl pH 7.5 and 0.1% (w/v) SDS. A sterile, disposable plastic pipette was plugged with sterile, siliconized glass wool, the column poured, and equilibrated with elution buffer. The nick translated DNA sample was carefully layered on top of the column and its progress through the column monitored with a Grieger counter. 0.5 ml fractions were collected, 1 µl from each fraction dispersed in toluene
scintillation fluid (3.37 g PPO/667 ml toluene, 333 ml Triton X-100 per litre), and the radioactivity counted in a Packard Tricarb Scintillation Counter. The appropriate fractions were pooled and used in hybridization reactions. The probe prepared by this method was denatured by heating for 5 min at 90°C before use.

2.18 $^{32}\text{P}$-labelling of single-stranded DNA (ss-DNA) by primer extension on M13 templates

The ss-DNA probe was prepared as previously described by Hu and Messing (1982). An annealing mixture of 1 µl (7 ng) M13 labelling primer, 1 µl (100 ng) M13 single-stranded DNA (encoding ribosomal RNA), 1 µl of 10x Klenow reaction buffer (100 mM Tris-HCl pH 7.5, 66 mM MgCl$_2$ and 600 mM NaCl), 1 µl of 0.1 M dithiothreitol (DDT) and 3 µl H$_2$O. The mixture was incubated at 65°C for 15 min, then removed to room temperature. 1 µl of a mixture of dATP, dGTP, dTTP at 500 µM each, 1 µl (10 µCi) of α-($^{32}$P) dCTP and 1 µl of DNA polymerase (Klenow, about 0.5 unit) were added to the mixture. The DNA synthesis reaction was allowed to proceed at 15°C for 90 min, and then stopped by adding 1 µl of 250 mM EDTA pH 8.3.

2.19 $^{3}$H-labelling of rDNA by nick translation

The cDNA clone "pHAI" (sequences encoding ribosomal RNA) was labelled to a specific activity of $>10^7$ cpm/µg using the nick translation kit supplied by Amersham International as described in its instructions. 50 µl ($^{3}$H)dCTP (50 µCi) were lyophilized to dryness, followed by adding 50 µl H$_2$O, 20 µl nucleotide/buffer solution (see Section 2.17) and 1 µg DNA. The solution was completed to 90 µl with H$_2$O followed by adding 10 µl enzyme solution (see Section 2.17). The
mixture was incubated for 2 h at 15°C, then 5 μl of 10% SDS were added to terminate the reaction. The products were separated from unincorporated label by gel filtration as described in Section 2.17. The labelled DNA was recoverd from the pooled fractions by alcohol precipitation and redissolved in a hybridization solution containing 50% (v/v) formamide, 50% (v/v), 2x SSC, 5 μg/ml poly (A) and 5 μg/ml denatured herring sperm DNA.

2.20 Hybridization of 32P-labelled probes to filter-bound DNA

Labelled DNA (sequences encoding legumirn or ribosomal RNA) was hybridized to the ss-DNA on the filter as described by Hu and Messing (1982). The baked filters (see Section 2.16) were prehybridized at 43°C for 2 h in 50 ml of a prehybridization solution (see below). The prehybridization solution was then replaced by a hybridization solution (see below) and the reaction was allowed to proceed at 43°C for 12-16 h on a shaking water path. The filter was then washed three times (15 min each) with 2x SSCP (standard saline citrate phosphate buffer, 1x SSCP is 120 mM NaCl, 15 mM sodium citrate, 13 mM KH₂PO₄, 1 mM EDTA, adjusted to pH 7.2 with NaOH) and 0.1% sarkosyl (N-lauroyl sarcosine, sodium salt), at room temperature, then four times (15 min each) at 50°C in 0.2x SSCP and 0.1% sarkosyl. The filter was then dried in a 80°C oven for 15 min and prepared for autoradiography.

<table>
<thead>
<tr>
<th>Prehybridization solution</th>
<th>Hybridization solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>5x Denhardt's solution</td>
<td>1x Denhardt's solution</td>
</tr>
<tr>
<td>5x SSC</td>
<td>5x SSC</td>
</tr>
<tr>
<td>50% (v/v) formamide</td>
<td>50% (v/v) formamide</td>
</tr>
<tr>
<td>50 mM phosphate buffer, pH 6.8</td>
<td>20 mM phosphate buffer, pH 6.8</td>
</tr>
<tr>
<td>250 μg/ml denatured herring sperm DNA</td>
<td>100 μg/ml denatured herring sperm DNA</td>
</tr>
<tr>
<td>100 μg/ml poly (A)</td>
<td>100 μg/ml poly (A)</td>
</tr>
<tr>
<td>0.1% (w/v) SDS</td>
<td>0.1% (w/v) SDS</td>
</tr>
<tr>
<td>1.0% (w/v) glycine</td>
<td>The radioactive probe</td>
</tr>
</tbody>
</table>
2.21 Autoradiography

The nitrocellulose filter to be autoradiographed was dried in a vacuum oven for 15 min and mounted on a piece of blotting paper which was then taped to a glass plate. Radioactive marker ink was spotted on the blotting paper to mark the line of the slots and the whole encased in a plastic bag in order to prevent contamination of the intensifying screens and prevent the sample from sticking to the film. In the dark room, a sheet of X-ray film was flashed using a Sunpak flashgun at a distance of 3 feet and placed over the encased nitrocellulose filter. An intensifying screen was placed over the X-ray film, followed by a glass plate. The two plates were then held together with rubber bands, wrapped in several layers of black plastic bags to exclude light, and exposed at -80°C for 1-3 days (exposure times varied depending on the counts present in the sample). At the end of the exposure time, the film was removed from the sample in a dark room, and developed by immersing in Kodak X0 MAT developer for 8 min, followed by washing with water for 1 min and then fixed for 3 min in Kodak fixer. The developed film was then dried at room temperature.

2.22 DNA cloning in M13 vectors

Appropriate DNA fragments were digested with various restriction enzymes and cloned in M13 vectors; mp8, mp9, mp18 and mp19. The clones were propagated in the host E. coli JM 101 or JM 105 and single-stranded DNA (ss-DNA) was purified from the phage according to the
2.22.1 Initial restriction of insert and vector DNAs

In order to digest the vector DNA (the double-stranded replicative form DNA "RF DNA" of the phage M13; mp8, mp9, mp18 and mp19) and the fragment to be cloned (insert DNA), digestion solutions were made of 1 µl of DNA (0.01 and 0.2 µg of vector and insert DNA, respectively), 4 µl of 10x restriction enzyme buffer (Section 2.12), 1 µl (10 units) of the restriction enzyme and 14 µl of H₂O. The reaction mixture was incubated for 1 h at 37°C, an equal volume of phenol was added, mixed by vortexing for 15 sec and centrifuged for 5 min. The upper aqueous layer was taken and re-extracted with an equal volume of chloroform : isoamyl alcohol (24 : 1), this was centrifuged for 3 min and the upper aqueous layer was removed. 10 µg of tRNA and 3 volumes of ethanol were added, this was then left at -20°C for 7 min, washed with ethanol, dried under vacuum for 5 min and resuspended in 10 µl of sterile water.

2.22.2 Ligation

A ligation mixture was made of 5 µl of the digested insert DNA, 10 µl of the digested vector DNA (see Section 2.22.1), 2 µl of 10x ligase buffer (0.66 M Tris-HCl, pH 7.5; 50 mM MgCl₂; 50 mM DTT and 10 mM ATP), 1 µl T4 DNA ligase and 2 µl H₂O. The ligation was allowed to proceed at room temperature for 2 h, 3 volumes of ethanol were added and this was then left at -20°C for 1 h. The precipitate was collected by centrifugation for 7 min, washed with ethanol, vacuum dried for 5 min and resuspended in 10 µl of 0.1 M CaCl₂.
2.22.3 Preparation of competent cells

100 μl of overnight growing *E. coli* strain JM 101 or JM 105 were inoculated into 50 ml of 1x YT (see Section 2.22.6) and incubated at 37°C for 5 h with agitation. This was then cooled on ice and transferred to a 100 ml MSE polyprepelene centrifuge tube and centrifuged at 7000 rpm for 5 min at 4°C. The precipitate was resuspended in 20 ml of 0.1 M CaCl$_2$ and left on ice for 1 h. This was then centrifuged as before, the precipitate was resuspended in 2 ml of 0.1 M CaCl$_2$ and left on ice overnight.

2.22.4 Transformation and plating

The resuspended ligations (see Section 2.22.2) were mixed with 100 μl of competent cells (Section 2.22.3) and left on ice for 40 min. 0.01 μg of RF M13 vectors DNA were suspended in 20 μl of 0.1 M CaCl$_2$ and added to 100 μl competent cells as a control. YT soft agar (see Section 2.22.6) was melted; 3 ml were pipetted into a sterile small test tube and left in a water bath at 42°C. After transformation was completed, the competent cells with the ligations were shocked at 42°C for 2 min and added to 10 μl of 0.1 M IPTG, 50 μl of 2% X-Gal and 200 μl of fresh *E. coli* cells. This was mixed, added to the 3 ml soft agar and poured quickly into prewarmed (37°C) agar plates. The plates were swirled gently to give even coating, left at room temperature to set, inverted and incubated at 37°C overnight.

2.22.5 Preparation of M13 subclones (single-stranded DNA templates)

After overnight growth, transformed cells should have formed plaques which in this case, are areas of retarded growth. Infected
cells from recombinant colourless plaques were grown to produce single-stranded DNA. 200 μl of overnight culture of E. coli were used to inoculate 100 ml of 2x YT (Section 2.22.6) and distributed into 2 ml aliquots in sterile universal bottles. Each bottle was inoculated with a colourless plaque using sterile wooden cocktail sticks and left to grow overnight at 37°C with shaking. This was then transferred to eppendorf tubes, centrifuged for 5 min and the supernatant was recentrifuged for 4 min. To 1 ml of the supernatant, 100 μl of 50% (w/v) polyethylene glycol (in 50 mM Tris-HCl, pH 7.5) were added, mixed and left at room temperature for 20 min to precipitate the phage from the bacteria. This was followed by centrifugation for 3 min; the supernatant was discarded, the precipitate recentrifuged for 10 sec, the excess of polyethylene glycol removed with syringe and the precipitate resuspended in 100 μl of TE buffer. The resuspended phage was deprotenized by phenol and chloroform extractions. The ss-DNA was ethanol precipitated, vacuum dried and resuspended in 20 μl of TE buffer. 2 μl of the resuspended DNA were added to 18 μl of TE buffer and 10 μl of agarose beads and loaded into agarose gel to check that the virus had grown and the template material (ss-DNA) had successfully been recovered. After running the gel, the recombinant DNA should have reduced mobility, but this is only observable with inserts greater than 300 bp.

2.22.6 Media and solutions used in DNA cloning

1x YT medium/litre
8 g bactotrypton, 5 g yeast extract, 5 g NaCl.
2x YT medium/litre
16 g bactotrypton, 10 g yeast extract, 5 g NaCl.

1x YT agar
15 g/litre bactoagar to 1x YT medium. This should give sufficient agar for approx. 40 plates.

1x YT soft agar (top agar)
6 g/litre bactoagar to 1x YT medium

X-Gal (2%)
2% X-Gal in dimethylformamide must be prepared immediately before use.

IPTG (0.1M)
28.3 mg/ml sterile water.

2.23 Screening M13 subclones

The M13 subclones which were prepared earlier (see Section 2.22.5) were screened via filter hybridization by using "dot-blot" assays. LegA or legB sequences were used as radioactive probes (see Sections 2.17 and 2.18). Dot hybridization experiment was carried out as outlined by Hu and Messing (1982). Approx. 0.5 µg samples of M13 subclones (ss-DNA) were spotted onto nitrocellulose filter by means of a Hydrodot apparatus (BRL), according to the protocol supplied with the apparatus. A nitrocellulose filter (previously soaked for 1 min in water) was soaked in 20x SSC for 20 min and placed on the apparatus. DNA samples (0.5 µl ss-DNA mixed with 20 µl of 20x SSC) were loaded and allowed to soak through for 5 min under vacuum. The nitrocellulose filter was then baked between two sheets of 3 M M paper for 2 h at 80°C under vacuum. The baked filter was prehybridized, hybridized and autoradiographed as previously described (Sections 2.20 and 2.21).

2.24 The orientation test of M13 subclones

To test whether two recombinant phages have DNA complementary to
each other, an orientation test was carried out according to BRL ICN manual, 1980. A reaction mixture containing 1 μl of each recombinant phage DNA, 16.5 μl of TE buffer, 3 μl of 2.5 M NaCl and 5 μl of SDS/formamide dye (see below) was incubated at 65°C for 1 h. This was then spun and loaded into a 0.7% agarose gel. If the two viral DNAs hybridize via their inserts they form a figure eight-like structure and migrate slower in the gel than the single viral circle.

**SDS/formamide dye**
3% (w/v) SDS, 0.1% (w/v) bromophenol blue, 60% (v/v) formamide and 25 mM EDTA pH 8.0.

### 2.25 DNA sequencing

"Dideoxy chain termination sequencing" method of Sanger *et al.* (1977) was used to determine the nucleotide sequence of pea legumin gene "legB" according to the protocol described in Amersham International Pamphlet. The ss-DNA of M13 subclones (Section 2.22.5) were used as templates in sequencing reactions.

#### 2.25.1 Annealing primer to template

An annealing mixture was made of 5 μl of ss-DNA template, 1 μl of M13 universal primer (15 bp), 1.5 μl of 10x Klenow reaction buffer (100 mM Tris-HCl pH 7.5, 66 mM MgCl₂ and 600 mM NaCl) and 3.5 μl of double-distilled water. The mixture was heated for 5 min at 85°C and allowed to cool very slowly. The annealed primer/template mix can be stored at -20°C.
2.25.2 Preparation of deoxy and dideoxy NTP mixes

Deoxy NTP working solutions
The samples of the supplied stock solutions were diluted to give 0.5 mM working solutions.


<table>
<thead>
<tr>
<th>NTP</th>
<th>A°</th>
<th>C°</th>
<th>G°</th>
<th>T°</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5 mM</td>
<td>20 μl</td>
<td>1 μl</td>
<td>20 μl</td>
<td>20 μl</td>
</tr>
<tr>
<td>0.5 mM</td>
<td>20 μl</td>
<td>20 μl</td>
<td>1 μl</td>
<td>20 μl</td>
</tr>
<tr>
<td>0.5 mM</td>
<td>20 μl</td>
<td>20 μl</td>
<td>20 μl</td>
<td>1 μl</td>
</tr>
<tr>
<td>1x TE buffer</td>
<td>20 μl</td>
<td>20 μl</td>
<td>20 μl</td>
<td>20 μl</td>
</tr>
</tbody>
</table>

Dideoxy NTP working solutions
The samples of the supplied stock solutions were diluted to give the following working solutions: 0.125 mM ddATP, 0.5 mM ddCTP, 0.5 mM ddGTP and 1.0 mM ddTTP.

dNTP/ddNTP mixes
To each dNTP mix, an equal volume of the corresponding ddNTP working solution was added e.g. 25 μl A°+25 μl ddATP = A°/ddATP mix etc.

2.25.3 Sequencing reaction
To the annealing tube (see Section 2.25.1), 1 μl of α-(35S) dATP and 1 μl of DNA polymerase (Klenow fragment) were added and mixed carefully by pipetting in and out. 2.5 μl of this mixture (template/primer/label/enzyme) were added to the ACGT mixes in their vials (4 vials for each clone, each vial contains 2 μl of the dNTP/ddNTP mixes i.e. A°+ddATP, C°+ddCTP, G°+ddGTP and T°+ddTTP). The components were span briefly to the bottom of the vial and incubated for 20 min at 30°C. 1 μl of chase solution (0.5 mM dATP) was added to each vial, span briefly and incubated for 15 min at 30°C. At the end of the incubation, 5 μl of formamide dye (see below) were added and the mixtures stored at -20°C until loading into the gel.

Formamide dye
80% (v/v) formamide, 10 mM NaOH, 1 mM EDTA pH 8.0, 0.1% (w/v) xylene cyanol, 0.1% (w/v) bromophenol blue.
2.25.4 Preparation, electrophoresis and autoradiography of sequencing gels

Two 6% thin polyacrylamide gels were used for sequencing three different clones. Glass plates were cleaned extensively with decon, water and ethanol until a high polish was obtained. The smaller glass plate was siliconized, dried, cleaned with distilled water and polished with ethanol. The two glass plates were separated by 0.5 mm perspex spacers and taped together. A gel solution was made of 30 g urea, 9 ml of 40% acrylamide/bisacrylamide stock (see below), 6 ml of 10x TBE buffer (see below) and distilled water to 50 ml. It was filtered through a millipore filtration apparatus and degassed. 300 μl of 10% ammonium persulphate and 20 μl of TEMED were added and the gel was poured free of bubbles and allowed to set.

Gels were pre-electrophoresed at 1500 V for 1 h before sequencing samples were loaded. The samples to be loaded were heated at 90°C for 5 min, cooled rapidly in ice, and 5 μl of the nucleotide reactions were loaded into 4 separate tracks in both gels using a syringe. The gels were switched off for as short a time as possible during loading of samples. The first gel was run until the bromophenol blue had migrated to the bottom of the gel (about 2 h) and the second gel was run until 30 min after the xylene cyanol had migrated off the gel (about 5 h). At the end of each run, the tape holding the plates together was cut, and the perspex spacers removed. The plates were then levered apart, the gel was stuck to a sheet of 3 MM paper, and dried for 1 h at 60°C in a vacuum drier. The dried gel was autoradiographed for 2-4 days at -80°C using a preflashed film and intensifying screen as previously described (Section 2.21).
40% acrylamide stock
38 g acrylamide, 2 g bisacrylamide, distilled water to 100 ml, 5 g Amberlite MB1 were added, stirred for 30 min and filtered through a sintered glass filter to remove resin. The fine particles were removed using a millipore filtration apparatus. This stock solution was kept at 4°C.

10x Tris-borate buffer (TBE) pH 8.3/litre
108 g Tris base, 55 g boric acid, 9.3 g EDTA.

2.26 Cytology and in situ hybridization

The main and side roots of pea line 110 were cut off, pretreated for 2 h at room temperature in a concentrated solution of 1, 4-dichlorobenzene, fixed for 24 h at room temperature in either Pienaar's mixture (6 : 3 : 2 methanol, chloroform, propionic acid) or 3 : 1 ethanol, acetic acid and stored in 70% ethanol at 4°C. After fixation in Pienaar's mixture, the root tips were stored in 80% methanol for 24 h at 4°C, then transferred to 70% ethanol for permanent storage.

The root tips were transferred to 3 : 1 ethanol, acetic acid for 10 min, then hydrolyzed in 1 N HCl for 8 min at 60°C, washed in distilled water and stored in 70% ethanol before squashing in either aceto-carmine for conventional staining or 45% acetic acid for in situ hybridization.

The in situ hybridization was carried out according to the methods described by Pardue and Gall (1975) and Hutchinson et al. (1980). The root tips were squashed in a drop of 45% acetic acid on a "subbed" slide i.e., covered with a thin layer of an aqueous mixture of 0.1% gelatin and 0.01% chrome alum (chromium potassium sulphate), and covered with a siliconized cover slip. The cover slip was then removed after freezing in liquid nitrogen and the slide was transferred to 96% ethanol for 10 to 15 min before being air dried.
Prior to hybridization with the ribosomal RNA probe \((^3\text{H-rDNA})\), slides were incubated for 2 h in 100 \(\mu\text{g/ml}\) RNase (prepared from a 10 mg/ml stock which had been incubated at 80°C for 10 min to inactivate DNases) in 2x SSC at 37°C. Slides were then washed three times (5 min each) in 2x SSC at room temperature followed by dehydration through 70% and 96% ethanol. \(^3\text{H-rDNA}\) probe (see Section 2.19) was placed over the preparation and covered with an acid washed cover slip (100,000 cpm in 5 \(\mu\text{l}\) hybridization solution were applied per slide). The slide and probe was then dipped in water at 67°C for 30 sec to denature the DNA of the chromosomes and the hybridization was allowed to proceed for 16 h at 40°C. After hybridization, the cover slip was removed and the slides were washed three times (5 min each) in 2x SSC at 45°C and then incubated for 15 min in 2x SSC at 60°C to dissociate nonspecific hybrids. This was followed by washing three times (5 min each) in 2x SSC at room temperature, incubation for 1 h in 100 \(\mu\text{g/ml}\) RNase at room temperature and dehydration through 70% and 90% ethanol before being air dried. The slides were then covered with a very thin layer of 50% solution of Ilford emulsion in a dark room, air dried and stored over silica gel in dark at 4°C for 10 days. After exposure time, the autoradiographs were developed for 3 min in Kodak X0 MAT developer and fixed for 2 min in Kodak fixer. Finally, the slides were air dried, stained for 3 min in Carbol fuchsin stock stain (Darlington and LaCour, 1969), rinsed with water, air dried and mounted in immersion oil.

2.27 Physical mapping of ribosomal RNA gene sites

The physical mapping of ribosomal RNA gene clusters was based on
the estimation of the nucleolus organizer (satellite constriction) sites in map units on the satellited chromosomes (Dr. R. Flavell, personal communication; Martini and Flavell, 1985). Fifty-four pairs of the satellited chromosomes (4 and 7) were measured accurately from root-tip metaphase preparations of pea (Pisum sativum L.) line 110 (standard for the normal structural type). Each chromosome was measured as 4 separate parts i.e. SA, the short arm; LA, the distance between the centromere and the satellite constriction (the long arm); SC, the satellite constriction; and Sat, the satellite. The length of the long arm (LA) and the satellite constriction (SC) were transformed to percentages of the total length of the chromosome as follows;

\[
\text{Long arm(\%)} = \frac{LA}{SA+LA+Sat} \times 100
\]

\[
\text{Satellite constriction (\%)} = \frac{SC}{SA+LA+Sat} \times 100
\]

The distance between the centromere and the satellite constriction of chromosomes 4 and 7 (the ribosomal RNA gene sites) were calculated in map units using the genetic map of Blixt (1974). The total length of chromosomes 4 and 7 are 247 and 109 map units, respectively.

2.28 Spectrophotometric determination of DNA

To determine the amount of DNA, readings were taken of 1 ml samples in a Pye Unicam SP8-150 spectrophotometer at 260 nm. An O.D. of 1 corresponding to 50 µg/ml of ds-DNA.

2.29 Statistical methods

Chi-square values were calculated following the formula given by Fisher (1925).
\[ X = \frac{d^2}{e} \]

Where "d" stands for deviation of observed value from the expected value, "e" represents the corresponding expected value and "\( \Sigma \)" is the symbol for "summation of". The percentages of crossing over and probable error were calculated by the "product ratio" method of Immer (1930). One of the co-dominant factors in the allelic pair of the parental plants was considered as dominant, so that in a hybrid cross involving dominant and co-dominant factors, those factors were in coupling phase.

2.30 Stains for chromosomes

**Aceto-carmine**
A mixture containing 0.5 g carmine, 45 ml glacial acetic acid and 55 ml distilled water was boiled for 5 min, left to cool and filtered.

**Carbol fuchsin**

A) **Stock stain**
- 10 ml of 3% basic fuchsin (in 70% ethanol)
- 90 ml of 5% phenol (in distilled water)

B) **Staining solution**
- 45 ml of solution A
- 6 ml glacial acetic acid
- 6 ml formalin
**Table (1)** Numbers, genotypes and karyotypes of pea lines (from Blixt, 1981). D, N and T for duplication, normal and translocation, respectively.

<table>
<thead>
<tr>
<th>Line No.</th>
<th>Karyotype</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>102</td>
<td>N</td>
<td>pa, a Br, pre, tl, td, fl, le, mifo, Tra, r</td>
</tr>
<tr>
<td>110</td>
<td>N</td>
<td>a, td, D, n, v, fov, mifo, Tra, F, Fs,</td>
</tr>
<tr>
<td>180</td>
<td>N</td>
<td>Kp, td, D, Np, te, Pl, M, F, Fs</td>
</tr>
<tr>
<td>200</td>
<td>N</td>
<td>a, Br, un, fl, le, bt, mifo, i, r</td>
</tr>
<tr>
<td>360</td>
<td>T(1-4-7)</td>
<td>a, wsp, un, fl, bt, mifo, Tra, i, r</td>
</tr>
<tr>
<td>611</td>
<td>T(3-5-7)</td>
<td>Kp, td, D, Pl, M, F, Fs</td>
</tr>
<tr>
<td>807</td>
<td>N</td>
<td>Kp, D, M</td>
</tr>
<tr>
<td>851</td>
<td>N</td>
<td>pa, vim, b, k, dt, pr, wb, tl, st, td, D, fl, pro, le, di, mifo, s, coh, cor, F, Fs, i, r,</td>
</tr>
<tr>
<td>1238</td>
<td>N</td>
<td>b, k, Bra, dt, pr, wb, tl, d, fl, fr, fru, pro, coh, le, t, cp-1, te, gp, pur, sru, s, u, i, r</td>
</tr>
<tr>
<td>1249</td>
<td>N(2-3-5)+D</td>
<td>b, k, wb, ins, td, D, fl, cp-1, rup, mifo, s, rag, cal, mp, z</td>
</tr>
<tr>
<td>1263</td>
<td>N</td>
<td>a, Br, dt, Tra</td>
</tr>
<tr>
<td>1533</td>
<td>N</td>
<td>b, k, pe, wb, tl, td, D, le, bt, cp-1, sur, lap, mifo, s, F, i, r</td>
</tr>
<tr>
<td>1552</td>
<td>N</td>
<td>r</td>
</tr>
<tr>
<td>1722</td>
<td>T(3-7)</td>
<td>vim, brac, td, D, Pl, v, bt, F, gl, r</td>
</tr>
<tr>
<td>2162</td>
<td>N</td>
<td>a, Br, fn, fl, le, Amp-1, mifo, Tra, i, r</td>
</tr>
<tr>
<td>2747</td>
<td>N</td>
<td>a, Br, le, Tra, i</td>
</tr>
<tr>
<td>3080</td>
<td>N</td>
<td>a, r</td>
</tr>
<tr>
<td>5452</td>
<td>N</td>
<td>a, le, bt, r</td>
</tr>
<tr>
<td>5478</td>
<td>N</td>
<td>a, un, fl, le, mifo, i, r</td>
</tr>
</tbody>
</table>
Table (2) Key to gene symbols and their locations on pea chromosomes (from the *Pisum* -Genebank Gene-Symbols, Blixt, S. Weibullsholm Plant Breeding Institute, Landskrona, Sweden (1981)).

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Location</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>1</td>
<td>Uncoloured flowers, seeds, axils</td>
</tr>
<tr>
<td>b</td>
<td>3</td>
<td>Flowers deep rose pink</td>
</tr>
<tr>
<td>Br</td>
<td>4</td>
<td>Flowers with bracts</td>
</tr>
<tr>
<td>Bra</td>
<td>3</td>
<td>As Br</td>
</tr>
<tr>
<td>bt</td>
<td>7</td>
<td>Pod apex pointed</td>
</tr>
<tr>
<td>coh</td>
<td>5</td>
<td>Reducing internode length</td>
</tr>
<tr>
<td>cp</td>
<td>5</td>
<td>Concavely curved pods</td>
</tr>
<tr>
<td>d</td>
<td>1</td>
<td>Absence of maculum ring colour</td>
</tr>
<tr>
<td>dt</td>
<td>-</td>
<td>Shortening distance from axil to first flower</td>
</tr>
<tr>
<td>F</td>
<td>3</td>
<td>Seed coat with violet spots</td>
</tr>
<tr>
<td>fl</td>
<td>6</td>
<td>Leaves and stipules without air pockets under epidermis</td>
</tr>
<tr>
<td>fov</td>
<td>3</td>
<td>Seed with impression over radicula</td>
</tr>
<tr>
<td>fru</td>
<td>4</td>
<td>Increasing number of stem branches</td>
</tr>
<tr>
<td>Fs</td>
<td>5</td>
<td>As F</td>
</tr>
<tr>
<td>gp</td>
<td>5</td>
<td>Pods yellow</td>
</tr>
<tr>
<td>i</td>
<td>1</td>
<td>Green cotyledons</td>
</tr>
<tr>
<td>k</td>
<td>2</td>
<td>Wings reduced, keel-like</td>
</tr>
<tr>
<td>Kp</td>
<td>4</td>
<td>Keel coloured with anthocyanin</td>
</tr>
<tr>
<td>le</td>
<td>4</td>
<td>Shortening internodes</td>
</tr>
<tr>
<td>M</td>
<td>3</td>
<td>Brown marbling of seed testa</td>
</tr>
<tr>
<td>mifo</td>
<td>2</td>
<td>Close-set small and shallow impressions on testa</td>
</tr>
<tr>
<td>n</td>
<td>4</td>
<td>Thick and fleshy pod wall</td>
</tr>
<tr>
<td>pa</td>
<td>7</td>
<td>Dark green foliage colour</td>
</tr>
</tbody>
</table>
Table (2) continued

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Location</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Pl</em></td>
<td>6</td>
<td>Black hilum colour on seeds</td>
</tr>
<tr>
<td><em>pr</em></td>
<td>-</td>
<td>Shortens inflorescences</td>
</tr>
<tr>
<td><em>pre</em></td>
<td>-</td>
<td>As <em>pr</em></td>
</tr>
<tr>
<td><em>pro</em></td>
<td>4</td>
<td>Stem branches growing out horizontally, then rising at 45° angle</td>
</tr>
<tr>
<td><em>pur</em></td>
<td>1</td>
<td>Pod colour purple</td>
</tr>
<tr>
<td><em>r</em></td>
<td>7</td>
<td>Cotyledons wrinkled; starch phenotypically compound</td>
</tr>
<tr>
<td><em>s</em></td>
<td>2</td>
<td>Tragacanth excretion on seed coat outside; seeds stick together</td>
</tr>
<tr>
<td><em>sru</em></td>
<td>1</td>
<td>Pod along upper suture with anthocyanin stripe</td>
</tr>
<tr>
<td><em>t</em></td>
<td>1</td>
<td>Thicker stem</td>
</tr>
<tr>
<td><em>td</em></td>
<td>4</td>
<td>No dentation of foliage</td>
</tr>
<tr>
<td><em>te</em></td>
<td>5</td>
<td>Narrow pod</td>
</tr>
<tr>
<td><em>tl</em></td>
<td>7</td>
<td>Leaflets in the place of tendrils</td>
</tr>
<tr>
<td><em>Tra</em></td>
<td>4</td>
<td>Seed coat with &quot;oily&quot; spots from tragacanth excretion on inside</td>
</tr>
<tr>
<td><em>U</em></td>
<td>5</td>
<td>Seed coat violet coloured</td>
</tr>
<tr>
<td><em>v</em></td>
<td>4</td>
<td>Pod-wall inside with patches of sclerenchyma</td>
</tr>
<tr>
<td><em>wb</em></td>
<td>2</td>
<td>Upper surface of stipules and lower of leaflets, stems and pods with very reduced wax layer</td>
</tr>
<tr>
<td><em>wsp</em></td>
<td>7</td>
<td>Upper surface of leaflets normally waxy, other parts waxless</td>
</tr>
</tbody>
</table>

Gene symbols with first letter capitals indicate phenotypes dominant or co-dominant over the normal phenotype, others recessive.
3. RESULTS
3.1 Genetic variation at isoenzyme loci

Different pea lines (Table 1) were screened by electrophoresis of seed extracts on polyacrylamide gels in order to find genetic variation in the banding patterns of the following isoenzyme systems: alkaline phosphatase, alcohol dehydrogenase, amylase, catalase, DNase, esterase, glutamate oxaloacetate transaminase, lipase, phosphorylase, polynucleotide phosphorylase and RNase. 7.5% polyacrylamide gels were used and the isoenzyme bands were detected in the gels according to the assay methods (Section 2.10). The following three systems were found to give clear variation in seed isoenzyme band patterns to allow a study of the inheritance and mapping of the structural genes for these isoenzymes in pea; amylase, esterase and glutamate oxaloacetate transaminase (GOT). Suitable lines were chosen from the tested lines to set up three separate crosses. The choice of the lines depended upon the variation in the isoenzyme banding patterns as well as the morphological markers (Table 2).

3.1.1 Variation at the amylase (Amy) locus

Reciprocal crosses were made between line 110 which has amylase type "A" and line 1238 which has type "B" (Fig.1a). Type "A" has a major band with faster migration, whereas type "B" has a major band with slower migration (Rf values of 0.57 and 0.55, respectively, with respect to the dye front). Line 1238 was double recessive for the characters; dwarfness, procumbent branches, absence of maculum ring, absence of tendrils, pink flowers, keel-like wings, yellow pods, sticky seeds, absence of air pockets below the leaf epiderms, green cotyledons, wrinkled seeds and reduced wax on the upper surface of
stipules, lower of leaflets, stems and pods, thus of the genotype le, p, d, t, b, k, gp; s, f, i, r and wb, respectively. On the other hand, line 110 was characterized by the absence of anthocyanin pigment, thus of genotype a.

3.1.2 Variation at the esterase (Est) locus

Reciprocal crosses were made between line 102 which has esterase type "A" and line 110 which has type "B" (Fig. 2a). Both types have two different zones of esterase bands, a nonvariant zone with one band of slow migration (Rf 0.55) and a variant zone of two bands in each type with Rf values of 0.69 and 0.65 in type "A" and 0.65 and 0.62 in type "B". Line 102 was double recessive for the characters; dwarfness, dark green foliage, absence of anthocyanin pigment, short inflorescences, leaflets in the place of tendrils, no dentation of foliage, leaves and stipules without air pockets under epidermis, close-set small and shallow impressions on testa and wrinkled seeds, thus of genotype le, pa, a, p, t, td, f, mifo and r respectively. In addition, line 102 was also characterized by the presence of bracteoles, thus of the genotype Br. The genotype and phenotype of line 110 were described earlier in cross "110 x 1238" (Section 3.1.1).

3.1.3 Variation at the glutamate oxaloacetate transaminase (Got) locus

Reciprocal crosses were made between line 1238 which has GOT type "A" and line 1263 which has type "B" (Fig. 3a). Both types have two different zones of GOT bands, a nonvariant zone with two bands of Rf 0.31 and 0.28 for the fast and slow bands, respectively, and a variant
zone which has one band in each type with faster migration in type "A" and slower migration in type "B" (Rf values 0.41 and 0.38, respectively). Line 1263 was characterized by the absence of anthocyanin pigment thus of the genotype a. The genotype and phenotype of line 1238 were described earlier in cross "110 x 1238" (Section 3.1.1)

3.2 Inheritance and mapping of isoenzyme genes

3.2.1. Amylase

F1 plants of the straight and reciprocal crosses between line 110 and line 1238 showed the expected dominant expression of the genes A, D, F1, Gp, K, Le, Pro, Te, TL and Wb whereas F1 seeds showed the expected dominant expression of the genes I and R and no segregation was observed. The band pattern of amylase in the F1 seeds on 7.5% polyacrylamide gels was found to be the additive pattern "AB" i.e. contained bands from both the parental lines "A" and "B". The F1 patterns were the same irrespective of which parent was the female parent and no segregation was observed (Fig.1a). These results indicate that there is no maternal effect on the inheritance of the amylase gene (Amy) and the parental lines were homozygous for this locus.

Statistical analysis of the F2 plants showed the monogenic inheritance and dominance of the genes A, F1, Gp, K, Le, Pro, Te, TL, Wb and the complementary case between A and D, whereas F2 seed showed the monogenic inheritance and dominance of the genes I and R (Table 3). The segregation of the band patterns of amylase on 7.5% polyacrylamide gels in the F2 seeds is shown in Fig.1b.
observed frequencies of the parental and hybrid band patterns "A, AB, B" are given in Table 3 and $X^2$ values of 1.4473 for the segregation ratio 1 : 2 : 1 at 2 degrees of freedom support the hypothesis that the amylase band pattern on PAGE is controlled by a single pair of co-dominant alleles.

In order to establish whether or not the genetic locus for the amylase subunit band pattern (Amy) was linked to the selected morphological markers, the F2 data were used to calculate Chi-square values of the segregation patterns. The calculated $X^2$ values (Table 4) between amylase type A, AB, B and the genes a, d, fl, gp, i, k, le, pro, r, te, tl and wb, showed that the Amy locus segregates independently of the loci a, d, fl, gp, i, le, pro, r, te, tl and dependently with k and wb. These results indicate that the gene Amy for amylase subunits and the genes k for keel-like wings and wb for waxless stipules are located on the same chromosome. On the other hand the genes k and wb have showed the expected dependent segregation with a cross-over value of 9.00±1.76 between them (Table 4). Cross-over values were calculated between the linked loci (Table 4). The values of 25.50±3.02 and 33.50±3.50 were found between the Amy gene and the genes k and wb, respectively.

3.2.2. Esterase

F1 plants of the straight and reciprocal crosses between line 102 and line 110 showed the expected dominant expression of the genes Br, Fl, Le, N, Te and Tl, whereas F1 seeds showed the expected dominant expression of the gene R and no segregation was observed. The band pattern of esterase in the F1 seeds on 7.5% polyacrylamide gels
*Only the variant bands (arrowed) were analysed and that the esterase gene (Est) refers to these bands. Another variant band was observed but the segregation of this band was difficult to investigate.
was found to be the additive pattern "AB" i.e. contained bands from both the parental lines "A" and "B". The F1 patterns were the same irrespective of which parent was the female parent and no segregation was observed (Fig. 2a). These results indicate that there is no maternal effect on the inheritance of the esterase gene (Est) and the parental lines were homozygous for this locus.

Statistical analysis of the F2 plants showed the monogenic inheritance and dominance of the genes Fl, Le, N, Te, Tl and the complementation between Br and Bra whereas F2 seeds showed the monogenic inheritance and dominance of the gene R (Table 5). The segregation of the band patterns of esterase on 7.5% polyacrylamide gels in the F2 seeds is shown in Fig. 2b. The observed frequencies of the parental and hybrid band patterns "A, AB, B" are given in Table 5, and $X^2$ value of 0.1345 for the segregation ratio 1 : 2 : 1 at 2 degrees of freedom support the hypothesis that the esterase band pattern on PAGE is controlled by a single pair of co-dominant alleles.

To investigate the linkage between the gene for esterase subunit band pattern (Est) and the genes for the selected morphological markers from the F2 data, Chi-square values of the segregation patterns for esterase subunits and the morphological markers were calculated. The calculated $X^2$ values (Table 6) between esterase type A, AB, B and the genes Br/Bra, fl, le, n, r, te and tl showed that the Est locus segregates independently of the loci fl, le, n, r, te, and tl but is linked to the locus Br for bracteoles with cross-over value of 28.43±4.52 between them (Table 6).
3.2.3 Glutamate oxaloacetate transaminase (GOT)

F1 plants of the straight and reciprocal crosses between line 1238 and line 1263 showed the expected dominant expression of the genes A, D, Fl, Gp, K, Le, Pro, Te, Tl and Wb whereas F1 seeds showed the expected dominant expression of the genes I and R and no segregation was observed. The band pattern of GOT in the F1 seeds on 7.5% polyacrylamide gels was found to be the additive pattern "AB" i.e. contained bands from both the parental lines "A" and "B". The additive pattern "AB" of GOT bands of the F1 seeds was found to contain not only the parental bands "A" and "B" but also an additional band of intermediate migration. The F1 patterns were the same irrespective of which parent was the female parent and no segregation was observed (Fig. 3a). These results indicate that there is no maternal effect on the inheritance of the GOT gene (Got) and the parental lines were homozygous for this locus.

Statistical analysis of the F2 plants showed the monogenic Mendelian inheritance and dominance of the genes A, Fl, Gp, K, Le, Pro, Te, Tl, Wb and the complementation between A and D whereas the F2 seeds showed the monogenic inheritance and dominance of the genes I and R (Table 7). The segregation of the band patterns of GOT on 7.5% polyacrylamide gels in the F2 seeds is shown in Fig. 3b. The observed frequencies of the parental and hybrid band patterns "A, AB, B" are given in Table 7, and \( \chi^2 \) value of 2.5715 for the segregation ratio 1:2:1 at 2 degrees of freedom support the hypothesis that GOT band pattern on PAGE is controlled by a single pair of co-dominant alleles.

In order to investigate the linkage between the gene for GOT
subunit band patterns (Got) and the genes for the selected morphological markers from the F2 data, Chi-square values of the segregation patterns for GOT subunits and the morphological markers were calculated. $X^2$ values (Table 8) between GOT type A, AB, B and the genes $a$, $d$, $fl$, $gp$, $i$, $k$, $le$, $pro$, $r$, $te$, $tl$ and $wb$ indicate that the gene for GOT subunits (Got) segregates independently of the loci $fl$, $gp$, $i$, $k$, $le$, $pro$, $r$, $te$, $tl$, $wb$ and dependently with the loci $a$ and $d$. The linkage between the gene for GOT subunits (Got) and the loci $a$ and $d$ gave cross-over values of 23.78±3.00 and 40.28±5.88 between Got locus and the loci $a$ and $d$, respectively. These results indicate that the gene for GOT isoenzyme subunits (Got) and the genes $a$ for absence of the anthocyanin pigment and $d$ for absence of maculum ring are located on the same chromosome.
Fig. 1 PAGE (7.5% acrylamide gel) of amylase isoenzyme banding patterns of pea seeds.

a) The parental lines 110 (amylase type A) and 1238 (type B) and their F1 in both the straight (F1/S) and reciprocal (F1/R) crosses (type AB).

b) The F2 segregation (type; A, AB, B).
Fig. 1

a. [Image of gel electrophoresis with bands labeled L.1238, F1/S, F1/R, and L.110.]

b. [Image of gel electrophoresis with bands labeled A, B, AB, AB, AB, A, AB, AB, A, AB, AB, AB.]
Fig. 2 PAGE (7.5% acrylamide gel) of esterase isoenzyme banding patterns of pea seeds.

a) The parental lines 102 (esterase type A) and 110 (type B) and the F1 in both the straight (F1/S) and reciprocal (F1/R) crosses (type AB).

b) The F2 segregation (type; A, AB, B). Variant bands identified by >.
Fig. 2

a. L.102  F1/S  F1/R  L.110
   A  AB  AB  B
   +  +  +

b. F2
   AB  B  AB  AB  AB  B  AB  B  AB  A  A  AB
   +  +  +
Fig. 3 PAGE (7.5% acrylamide gel) of GOT isoenzyme banding patterns of pea seeds.

a) The parental lines 1238 (GOT type A) and 1263 (type B) and their F1 in both the straight (F1/S) and reciprocal (F1/R) crosses (type AB).

b) The F2 segregation (type; A, AB, B). Variant bands identified by ••.
Fig. 3

a. L1263  F1/S  AB  F1/R  AB  L1238

b. F2

AB  AB  A  AB  A  B  A  B  B  AB  AB  A
Table (3) Analysis of the F2 segregation of amylase band patterns on PAGE (Amy) and the selected marker genes studied in a cross between two pea lines (110 x 1238).

<table>
<thead>
<tr>
<th>Gene</th>
<th>F2 phenotypic classes and observed frequency</th>
<th>Segregation ratio</th>
<th>$\chi^2$</th>
<th>Probability range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amy ; A, AB, B</td>
<td>28 AA</td>
<td>73 Aa</td>
<td>33 aa</td>
<td>1 : 2 : 1</td>
</tr>
<tr>
<td>A, a</td>
<td>98 AA</td>
<td>36 Aa</td>
<td></td>
<td>3 : 1</td>
</tr>
<tr>
<td>D, d</td>
<td>72 AA</td>
<td>36 Aa</td>
<td></td>
<td>9 : 7</td>
</tr>
<tr>
<td>Fl, fl</td>
<td>96 AA</td>
<td>36 Aa</td>
<td></td>
<td>3 : 1</td>
</tr>
<tr>
<td>Gp, gp</td>
<td>101 AA</td>
<td>33 Aa</td>
<td></td>
<td>3 : 1</td>
</tr>
<tr>
<td>I, i</td>
<td>100 AA</td>
<td>34 Aa</td>
<td></td>
<td>3 : 1</td>
</tr>
<tr>
<td>K, k</td>
<td>104 AA</td>
<td>30 Aa</td>
<td></td>
<td>3 : 1</td>
</tr>
<tr>
<td>Le, le</td>
<td>103 AA</td>
<td>31 Aa</td>
<td></td>
<td>3 : 1</td>
</tr>
<tr>
<td>Pro, pro</td>
<td>98 AA</td>
<td>36 Aa</td>
<td></td>
<td>3 : 1</td>
</tr>
<tr>
<td>R, r</td>
<td>98 AA</td>
<td>36 Aa</td>
<td></td>
<td>3 : 1</td>
</tr>
<tr>
<td>Te, te</td>
<td>100 AA</td>
<td>34 Aa</td>
<td></td>
<td>3 : 1</td>
</tr>
<tr>
<td>Tl, tl</td>
<td>96 AA</td>
<td>38 Aa</td>
<td></td>
<td>3 : 1</td>
</tr>
<tr>
<td>Wb, wb</td>
<td>105 AA</td>
<td>29 Aa</td>
<td></td>
<td>3 : 1</td>
</tr>
</tbody>
</table>

*Amy, A (line 110)*
*Amy, AB (F1)*
*Amy, B (line 1238)*
Table (4) Segregation of amylase band patterns on PAGE (*Amy*) in the F2 generation of a cross between two pea lines (110 x 1238) with respect to the selected marker genes. Segregation ratio is 9 : 3 : 3 : 1. Cp for coupling and Rp for repulsion.

<table>
<thead>
<tr>
<th>Gene pair</th>
<th>F2 phenotypic classes and observed frequency</th>
<th>( \chi^2 )</th>
<th>Probability range</th>
<th>Cross-over ± probable error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amy - a (Rp)</td>
<td>AB 77, Ab 24, aB 21, ab 12</td>
<td>2.3316</td>
<td>0.60 - 0.50</td>
<td></td>
</tr>
<tr>
<td>Amy - d (Cp)*</td>
<td>AB 61, Ab 40, aB 13, ab 20</td>
<td>4.4720</td>
<td>0.20 - 0.10</td>
<td></td>
</tr>
<tr>
<td>Amy - fl (Cp)</td>
<td>AB 72, Ab 29, aB 24, ab 9</td>
<td>0.8456</td>
<td>0.90 - 0.80</td>
<td></td>
</tr>
<tr>
<td>Amy - gp (Cp)</td>
<td>AB 76, Ab 25, aB 25, ab 8</td>
<td>0.0232</td>
<td>&gt; 0.99</td>
<td></td>
</tr>
<tr>
<td>Amy - i (Cp)</td>
<td>AB 77, Ab 24, aB 23, ab 10</td>
<td>0.5804</td>
<td>0.90 - 0.80</td>
<td></td>
</tr>
<tr>
<td>Amy - k (Cp)</td>
<td>AB 88, Ab 13, aB 16, ab 17</td>
<td>20.1626</td>
<td>&lt; 0.01</td>
<td>25.50 ± 3.02%</td>
</tr>
<tr>
<td>Amy - le (Cp)</td>
<td>AB 77, Ab 24, aB 26, ab 7</td>
<td>0.3415</td>
<td>0.99 - 0.95</td>
<td></td>
</tr>
<tr>
<td>Amy - pro (Cp)</td>
<td>AB 72, Ab 29, aB 26, ab 7</td>
<td>1.0049</td>
<td>0.90 - 0.80</td>
<td></td>
</tr>
<tr>
<td>Amy - r (Cp)</td>
<td>AB 75, Ab 26, aB 23, ab 10</td>
<td>0.5274</td>
<td>0.95 - 0.90</td>
<td></td>
</tr>
<tr>
<td>Amy - te (Cp)</td>
<td>AB 75, Ab 26, aB 25, ab 8</td>
<td>0.0498</td>
<td>&gt; 0.99</td>
<td></td>
</tr>
<tr>
<td>Amy - tl (Cp)</td>
<td>AB 81, Ab 20, aB 23, ab 10</td>
<td>1.9603</td>
<td>0.60 - 0.50</td>
<td></td>
</tr>
<tr>
<td>Amy - wb (Cp)</td>
<td>AB 85, Ab 16, aB 20, ab 13</td>
<td>8.1427</td>
<td>0.05 - 0.02</td>
<td>33.50 ± 3.50%</td>
</tr>
<tr>
<td>k - wb (Cp)</td>
<td>AB 99, Ab 5, aB 6, ab 24</td>
<td>67.2338</td>
<td>&lt; 0.01</td>
<td>9.00 ± 1.76%</td>
</tr>
</tbody>
</table>

* The segregation ratio is 27 : 21 : 9 : 7
Table (5) Analysis of the F2 segregation of esterase band patterns on PAGE (Est) and the selected marker genes studied in a cross between two pea lines (102 x 110).

<table>
<thead>
<tr>
<th>Gene</th>
<th>F2 phenotypic classes and observed frequency</th>
<th>Segregation ratio</th>
<th>$\chi^2$</th>
<th>Probability range</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AA</td>
<td>Aa</td>
<td>aa</td>
<td></td>
</tr>
<tr>
<td>Est ; A, AB, B</td>
<td>32</td>
<td>69</td>
<td>33</td>
<td>1 : 2 : 1</td>
</tr>
<tr>
<td>Br / Bra, br / bra</td>
<td>71</td>
<td>63</td>
<td>39</td>
<td>9 : 7</td>
</tr>
<tr>
<td>Fl, fl</td>
<td>95</td>
<td>39</td>
<td>39</td>
<td>3 : 1</td>
</tr>
<tr>
<td>Le, le</td>
<td>100</td>
<td>34</td>
<td>34</td>
<td>3 : 1</td>
</tr>
<tr>
<td>N, n</td>
<td>102</td>
<td>32</td>
<td>32</td>
<td>3 : 1</td>
</tr>
<tr>
<td>R, r</td>
<td>107</td>
<td>27</td>
<td>27</td>
<td>3 : 1</td>
</tr>
<tr>
<td>Te, te</td>
<td>101</td>
<td>33</td>
<td>33</td>
<td>3 : 1</td>
</tr>
<tr>
<td>Tl, tl</td>
<td>105</td>
<td>29</td>
<td>29</td>
<td>3 : 1</td>
</tr>
</tbody>
</table>

$Est$, A (line 102)
$Est$, AB (F1)
$Est$, B (line 110)
Table (6) Segregation of esterase band patterns on PAGE (Est) in the F2 generation of a cross between two pea lines (102 x 110) with respect to the selected marker genes. Segregation ratio is 9 : 3 : 3 : 1. Cp for coupling and Rp for repulsion.

<table>
<thead>
<tr>
<th>Gene pair</th>
<th>F2 phenotypic classes and observed frequency</th>
<th>$\chi^2$</th>
<th>Probability range</th>
<th>Cross-over ( \pm ) probable error</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>Ab</td>
<td>aB</td>
<td>ab</td>
<td></td>
</tr>
<tr>
<td>Est-Br (Cp)*</td>
<td>60</td>
<td>41</td>
<td>11</td>
<td>22</td>
</tr>
<tr>
<td>Est-f (Rp)</td>
<td>69</td>
<td>32</td>
<td>26</td>
<td>7</td>
</tr>
<tr>
<td>Est-le (Rp)</td>
<td>76</td>
<td>25</td>
<td>24</td>
<td>9</td>
</tr>
<tr>
<td>Est-n (Cp)</td>
<td>75</td>
<td>26</td>
<td>27</td>
<td>6</td>
</tr>
<tr>
<td>Est-r (Rp)</td>
<td>78</td>
<td>23</td>
<td>29</td>
<td>4</td>
</tr>
<tr>
<td>Est-te (Rp)</td>
<td>73</td>
<td>28</td>
<td>28</td>
<td>5</td>
</tr>
<tr>
<td>Est-tl (Rp)</td>
<td>78</td>
<td>23</td>
<td>27</td>
<td>6</td>
</tr>
</tbody>
</table>

* The segregation ratio is 27 : 21 : 9 : 7
Table (7) Analysis of the F2 segregation of GOT band patterns on PAGE (Got) and the selected marker genes studied in a cross between two pea lines (1238 x 1236).

<table>
<thead>
<tr>
<th>Gene</th>
<th>F2 phenotypic classes and observed frequency</th>
<th>Segregation ratio</th>
<th>$X^2$</th>
<th>Probability range</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Got$; A, AB, B</td>
<td>AA  24</td>
<td>Aa  66</td>
<td>aa  36</td>
<td>1:2:1</td>
</tr>
<tr>
<td>$A$, $a$</td>
<td>A   90</td>
<td>a   36</td>
<td></td>
<td>3:1</td>
</tr>
<tr>
<td>$D$, $d$</td>
<td>Fl  71</td>
<td>fl  55</td>
<td></td>
<td>9:7</td>
</tr>
<tr>
<td>$Fl$, $fl$</td>
<td>102</td>
<td>24</td>
<td></td>
<td>3:1</td>
</tr>
<tr>
<td>$Gp$, $gp$</td>
<td>92</td>
<td>34</td>
<td></td>
<td>3:1</td>
</tr>
<tr>
<td>$I$, $i$</td>
<td>Le  89</td>
<td>le  37</td>
<td></td>
<td>3:1</td>
</tr>
<tr>
<td>$K$, $k$</td>
<td>97</td>
<td>29</td>
<td></td>
<td>3:1</td>
</tr>
<tr>
<td>$Le$, $le$</td>
<td>98</td>
<td>28</td>
<td></td>
<td>3:1</td>
</tr>
<tr>
<td>$Pro$, $pro$</td>
<td>93</td>
<td>33</td>
<td></td>
<td>3:1</td>
</tr>
<tr>
<td>$R$, $r$</td>
<td>95</td>
<td>32</td>
<td></td>
<td>3:1</td>
</tr>
<tr>
<td>$Te$, $te$</td>
<td>94</td>
<td>32</td>
<td></td>
<td>3:1</td>
</tr>
<tr>
<td>$Tl$, $tl$</td>
<td>98</td>
<td>28</td>
<td></td>
<td>3:1</td>
</tr>
<tr>
<td>$Wb$, $wb$</td>
<td>96</td>
<td>30</td>
<td></td>
<td>3:1</td>
</tr>
</tbody>
</table>

$Got$, A (line 1238)
$Got$, AB (F1)
$Got$, B (line 1263)
Table (8) Segregation of GOT band patterns on PAGE (Got) in the F2 generation of a cross between two pea lines (1238 x 1263) with respect to the selected marker genes. Segregation ratio is 9 : 3 : 3 : 1. Cp for coupling and Rp for repulsion.

<table>
<thead>
<tr>
<th>Gene pair</th>
<th>F2 phenotypic classes and observed frequency</th>
<th>( X^2 )</th>
<th>Probability range</th>
<th>Cross-over range ± probable error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Got - a (Cp)</td>
<td>76 14 14 22</td>
<td>33.5485</td>
<td>&lt; 0.01</td>
<td>23.78±3.00%</td>
</tr>
<tr>
<td>Got - d (Rp)</td>
<td>57 19 12 12</td>
<td>4.8148</td>
<td>0.20 - 0.10</td>
<td>40.28±5.88%</td>
</tr>
<tr>
<td>Got - fl (Rp)</td>
<td>75 15 27 9</td>
<td>4.0317</td>
<td>0.30 - 0.20</td>
<td></td>
</tr>
<tr>
<td>Got - gp (Rp)</td>
<td>65 25 27 9</td>
<td>1.2098</td>
<td>0.80 - 0.70</td>
<td></td>
</tr>
<tr>
<td>Got - i (Rp)</td>
<td>64 26 25 11</td>
<td>2.2258</td>
<td>0.60 - 0.50</td>
<td></td>
</tr>
<tr>
<td>Got - k (Rp)</td>
<td>71 19 26 10</td>
<td>1.7178</td>
<td>0.70 - 0.60</td>
<td></td>
</tr>
<tr>
<td>Got - le (Rp)</td>
<td>69 21 29 7</td>
<td>1.6614</td>
<td>0.70 - 0.60</td>
<td></td>
</tr>
<tr>
<td>Got - pro (Rp)</td>
<td>67 23 26 10</td>
<td>1.0406</td>
<td>0.80 - 0.70</td>
<td></td>
</tr>
<tr>
<td>Got - r (Rp)</td>
<td>69 21 26 10</td>
<td>1.1535</td>
<td>0.80 - 0.70</td>
<td></td>
</tr>
<tr>
<td>Got - te (Rp)</td>
<td>70 20 24 12</td>
<td>2.7337</td>
<td>0.50 - 0.40</td>
<td></td>
</tr>
<tr>
<td>Got - tl (Rp)</td>
<td>71 19 27 9</td>
<td>1.5484</td>
<td>0.70 - 0.60</td>
<td></td>
</tr>
<tr>
<td>Got - wb (Rp)</td>
<td>67 23 29 7</td>
<td>1.5485</td>
<td>0.70 - 0.60</td>
<td></td>
</tr>
</tbody>
</table>
3.3 Genetic variation at storage protein loci

3.3.1 Variation at the vicilin (Ve) locus

In order to investigate the inheritance and chromosomal location of vicilin 50,000 Mr subunit in pea (Pisum sativum L.), different pea lines were screened using 12% SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and the total protein subunit band patterns were shown in Fig.4. In general, little variation in band patterns of vicilin subunits at approx. 50,000 Mr or 33,000 Mr was observed, in contrast to the pronounced variation in legumin a-subunits (4 different patterns in 8 lines tested) and frequent, although less pronounced variations in the convicilin band (3 different patterns in 8 lines tested). However, line 3080 showed a double band at 50,000 Mr whereas all other lines showed a single thicker band; the latter pattern was also given by a further 8 lines screened. Line 611 gave a double banded pattern similar to line 3080. No variation in the vicilin 33,000 Mr subunits was observed in this study, and variation in other vicilin polypeptides was not examined. A cross of lines 1263 and 3080 was selected as suitable for testing the inheritance of vicilin 50,000 Mr subunit and the segregation between the variation in vicilin band pattern and that in legumin a-subunits; in addition, the seed surface round/wrinkled gene r differed between these lines (genotypes and phenotypes of lines 1263 and 3080 were presented in Tables 1 and 2).

3.3.2 Variation at the legumin (Lg) locus

3.3.2.1 Restriction fragment length polymorphism

DNA was isolated from leaves of different pea lines and purified by banding twice in caesium chloride gradients (Section 2.11). The
pure DNA was digested to completion with Eco RI restriction enzyme, the digested fragments were separated by electrophoresis on agarose gels and transferred to nitrocellulose filter by "Southern" blotting as described in Sections 2.12, 2.13 and 2.16. The genomic blot was hybridized with radioactive legumin probe (32P-legumin sequences) as described in Sections 2.17 and 2.20, in order to find length variation in the restriction fragments of legumin genes. Fig. 5 shows Eco RI digests of total DNA from leaves of different pea lines after "Southern" transfer to nitrocellulose filter and hybridization with legumin genomic probe. Three different patterns were found i.e. pattern "A" which was found in the majority of the investigated lines (102, 200, 1238, 3080, 5452, 5478 and Feltham First) and patterns "B" and "C" which were found in lines 110 and 1263, respectively (Fig. 5). The observed restriction fragments detected by legumin probes in pea genomic DNA are known to represent single copy genes (Croy et al., 1982), and the present result confirms the presence of several copies of legumin coding sequences. There are no Eco RI sites either within the coding sequence or the internal introns of legumin genes previously sequenced (Lycett et al., 1984b; Bown et al., 1985). The variation between different pea lines can be used in locating the genes onto the genome.

3.3.2.2 The nucleotide sequence of legumin gene" legB "

The genomic clone "Leg2" (Fig. 6a) was characterized by Croy et al. (paper in preparation) and restriction fragments of this clone were subcloned in the plasmid vector pUC8. A subclone containing legB sequences, designated pRC 5.4, was used for further cloning and
sequencing operations. *LegB* sequences were isolated from the plasmid and cloned in M13 vectors; mp8, mp9, mp18 and mp19 as previously described (Section 2.15 and 2.22). The ss-DNA templates for sequencing reactions were prepared (Section 2.22.5) and screened for *legB* sequences by dot hybridization (Section 2.23). The *LegB* containing fragments were tested for their orientation as previously described (Section 2.24) and sequenced by the "dideoxy chain termination" method of Sanger *et al.* (1977) using α-(35S)dATP as a radioactive label (Section 2.25). The sequencing map and autoradiographs of the nucleotide sequence are shown in Fig. 6b and 7, respectively. Fig. 8 shows the nucleotide sequence of the coding regions and 3' flanking region of *legB*, the nucleotide sequence of *legA*, the predicted amino acid sequence of *legA* and the amino acid residues differing in *legB*. The comparison between *legA* and *LegE* shows that *LegE* has a general form typical of an eukaryotic gene with three intervening sequences (IVS-1, IVS-2 and IVS-3) of lengths 88, 88 and 85 bp, respectively. The boundary sequences of the three introns obey the CT/AG boundary rules of Breathnach *et al.* (1978). The sequences of IVS-1 and IVS-2 show complete homology in location and nucleotide sequence with their correspondents in *legA*, whereas IVS-3 sequence shows two base substitutions over its 5' end and a deletion of 14 bp near the 3' end.

The coding sequence of *legB* shows very strong homology with *legA* with substitution at only 8 amino acid residues (see Fig. 8) as a consequences of substitution of 8 active bases. The three direct repeated sequences of pea legumin gene family at the C-terminal of the α-polypeptide chain (Lycett *et al.*, 1984a) are present in *legB*
sequence without significant change.

The 3' flanking sequence of legB shows strong homology with that of legA with 10 base substitutions and deletion of two bases 5' of poly A site 3, whereas it shows complete homology with that of legC (Gatehouse, unpublished results). No change was observed in the three polyadenylation signals either in their locations or their nucleotide sequences. These signals downstream of both legA and legB are, AATAAGAAAA, AATAAATAAA and AATAAATAAA, respectively (Fig.8).

The predicted amino acid composition of legB gene product was compared with that of legA (Lycett et al., 1984b) in Table 9. The results show that both legA and legB have 5 cysteine residues (1% of total amino acid residues) and legB has 5 methionine residues (1% of total), compared with 4 residues (0.8% of total) in legA. Similar variations were found between the two genes in some of the amino acids (± 1 residue) except that legA has 2 extra isoleucine residues.
Fig. 4 SDS-PAGE (12% acrylamide gel) of total protein extracts of seeds of different pea lines (numbered), and of vicilin from var. Feltham First (VcFF) and from lines 360 and 611 (Vc360, Vc611). Variation in vicilin 50,000 Mr (Vc-1) and legumin α-subunit (Lg-1α) band patterns is indicated.
Fig. 4
Fig. 5  *Eco RI* digests of total DNA from leaves of different pea lines (numbered) after fractionation on 0.7% agarose gels, autoradiographs of hybridization of $^{32}$P-labelled probe (legumin coding sequence, labelled to a high specific activity of $3 \times 10^8$ cpm/μg) to total genomic DNA after "Southern" blotting to nitrocellulose filter. Arrow indicates band due to linearized pDUB 21 used as internal standard.
Fig. 6  a) Restriction map of \(\lambda\)Leg2. \(\lambda\) represents the legumin coding sequence, \(\lambda L\) and \(\lambda R\) are the left and right arms of bacteriophage \(\lambda\), respectively.

b) Sequencing map of legB gene. The figure shows the restriction sites and the sequence strategy. The coding regions are presented as heavy lines and the symbols for restriction enzyme cleavage sites are: A=AccI, B=Bam HI, E=Eco RI, G=BgI II, H=Hind III, N=Bst NI, X=Xho I, P=Pst I, S=Sau 3A.
Fig. 6

(a) E H B B B G H E H E B

\( \lambda L \) \( \lambda R \)

5' \[ \text{Leg B} \] 3'

(b) 200 bp

HX S S A NS P X B SH
Fig. 7 Autoradiographs of the nucleotide sequence of DNA using the "dideoxy chain termination" method with M13 sequencing system. The reading ladder is GATC (C=guanine, A=adenine, T=thymine, C=cytosine).
Fig. 7
Fig. 8 Nucleotide sequence of the legumin gene legB of pea (Pisum sativum L.). The sequence of legA is also given. The predicted amino acid sequence of legA is shown underneath the DNA sequence and the residues differing in legB are given above the sequence. Introns (IVS-1, IVS-2, IVS-3) and the other features discussed in the text are shown. Base substitutions and deletions in the non-coding regions and the 3' flanking region are indicated by (v) above the sequence. The three direct repeated sequences are underlined and the termini of the mature α- and β-subunits marked by brackets.
### Table (9) Comparison between the predicted amino acid composition of **LegA** and **LegE** gene products. **LegA** data are from Lycett et al. (1984b).

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>α-subunit no. of residues</th>
<th>% of total</th>
<th>β-subunit no. of residues</th>
<th>% of total</th>
<th>α+β-subunits no. of residues</th>
<th>% of total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>legA</td>
<td>legB</td>
<td>% of total</td>
<td>legA</td>
<td>legB</td>
<td>% of total</td>
</tr>
<tr>
<td>Ala (A)</td>
<td>12</td>
<td>12</td>
<td>3.9</td>
<td>3.9</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>Arg (R)</td>
<td>35</td>
<td>36</td>
<td>11.3</td>
<td>11.3</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Asn (N)</td>
<td>20</td>
<td>20</td>
<td>6.4</td>
<td>6.4</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td>Asp (D)</td>
<td>18</td>
<td>17</td>
<td>5.8</td>
<td>5.8</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Cys (C)</td>
<td>3</td>
<td>3</td>
<td>1.0</td>
<td>1.0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Gln (Q)</td>
<td>33</td>
<td>32</td>
<td>10.6</td>
<td>10.3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Glu (E)</td>
<td>43</td>
<td>43</td>
<td>13.8</td>
<td>13.8</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>Gly (G)</td>
<td>27</td>
<td>26</td>
<td>8.7</td>
<td>8.4</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>His (H)</td>
<td>7</td>
<td>7</td>
<td>2.3</td>
<td>2.3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Ile (I)</td>
<td>14</td>
<td>12</td>
<td>4.5</td>
<td>3.9</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Leu (L)</td>
<td>15</td>
<td>15</td>
<td>4.8</td>
<td>4.8</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Lys (K)</td>
<td>13</td>
<td>13</td>
<td>4.2</td>
<td>4.2</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Met (M)</td>
<td>3</td>
<td>3</td>
<td>1.0</td>
<td>1.0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Phe (F)</td>
<td>12</td>
<td>12</td>
<td>3.9</td>
<td>3.9</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Pro (P)</td>
<td>15</td>
<td>15</td>
<td>4.8</td>
<td>4.8</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Ser (S)</td>
<td>14</td>
<td>15</td>
<td>4.5</td>
<td>4.8</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Thr (R)</td>
<td>6</td>
<td>7</td>
<td>1.9</td>
<td>2.3</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Trp (W)</td>
<td>2</td>
<td>2</td>
<td>0.6</td>
<td>0.6</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Tyr (Y)</td>
<td>8</td>
<td>8</td>
<td>2.6</td>
<td>2.6</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Val (V)</td>
<td>11</td>
<td>13</td>
<td>3.5</td>
<td>4.2</td>
<td>14</td>
<td>13</td>
</tr>
</tbody>
</table>

The one and three letters symbols for amino acids are given in Methods in Enzymology (1984), 106 page 4.
3.4. Inheritance and mapping of storage protein genes

3.4.1 Inheritance of vicilin (50,000 Mr) and legumin (α-subunit) genes

Ten F1 seeds from each reciprocal cross of lines 1263 which has vicilin and legumin type "A" and 3080 which has type "B" were examined, and showed additive inheritance patterns for the vicilin 50,000 Mr and legumin α-subunit band patterns "AB". The F1 patterns were the same irrespective of which parent was the female parent and no segregation was observed (Fig. 9). These results indicate that there is no maternal effect on the inheritance of the major vicilin subunit (50,000 Mr) and legumin α-subunits and the parental lines were homozygous for both loci. 308 F2 seeds were examined, and showed the expected monogenic Mendelian inheritance and dominance of the r gene (Table 10). The segregation of the band patterns for the vicilin 50,000 Mr subunit and the α-subunit of legumin in the F2 seeds on 12% polyacrylamide gels is shown in Fig. 9. The observed frequencies of the parental and hybrid band patterns of vicilin 50,000 Mr and the legumin α-subunit are given in Table 10, and support the hypothesis that the vicilin 50,00 Mr and legumin α-subunit band patterns are controlled by a single pair of co-dominant alleles. Like that (Lg-1) controlling the legumin α-subunit band pattern, the locus which controls the major vicilin subunit (50,000 Mr) is subsequently referred to as Ve-1.

The observed variation in band patterns was confirmed to be due to vicilin polypeptides by the transfer of the protein subunits from an SDS-PAGE analysis of total extracts to nitrocellulose filters and specific detection of vicilin polypeptides by antivicilin IgG ("Western" blotting) as described in Section 2.10.5. Fig. 10a shows vicilin subunits from seeds of the parental lines 1263 and 3080 and
their F1 and F2 as detected by antivicilin IgG reaction on nitrocellulose filter after "Western" blotting of the total protein extracts from SDS-PAGE (12% polyacrylamide gels). A strong reaction was observed with the variant 50,000 Mr subunits whereas other vicilin polypeptides which reacted with the antivicilin IgG were invariant. Although the blot shown could not be scored for variant bands, the correlation was confirmed on blot/gel pairs not shown.

3.4.2 Chromosomal location of vicilin (50,000 Mr) gene locus

In order to establish whether or not the genetic locus for the vicilin 50,000 Mr subunit band pattern was linked to the markers \( r \) and \( Lg-1 \), the F2 data were used to calculate cross-over values (Table 11). No recombinants were observed between the \( Ve-1 \) pattern of line 1263 and wrinkled seed surface (\( r \)) leading to a cross-over value of 0 between the \( Ve-1 \) and \( r \) loci, and suggesting that these genes are closely linked to the extent of being within less than 2 cross-over units of each other. In confirmation of earlier results, the major legumin gene "\( Lg-1 \)" was found to be linked to the gene \( r \) on chromosome 7 (cross-over value 36.2±4.9). A cross-over value of 29.5±3.3 between \( Lg-1 \) and \( Ve-1 \) was also determined; in this case the genes are in coupling phase (Table 11) so this figure is likely to be more accurate. No linkage to other genes segregating in this cross was observed. Examination of 181 F2 seeds from the cross of lines 360 and 611 (these data were taken from Dr. J.A. Gatehouse) also showed no recombinants between the wrinkled seed surface \( r \) phenotype and line 611 (\( Ve-1 \)) pattern, and gave a cross-over value of 16.2±2.1 between \( Lg-1 \) and \( Ve-1 \) (coupling phase), in reasonable agreement with data from the other cross (1263 x 3080). Although these lines are known to contain transpositions involving chromosome 7 (Blixt, 1977)
it is most unlikely that this would affect the very close linkage between Va-1 and r.

The close linkage of the vicilin locus" Va-1 " with the r locus was further supported by the examination of seeds from five pairs of isogenic lines (TRT4/TRT3, TRT6/TRT5, TRT8/TRT7, TRT12/TRT11 and TRT18/TRT17) differing at the r locus. SDS-polyacrylamide gel electrophoresis of total protein extracts of these seeds showed that within each pair band patterns were identical, except that in 3 out of the 5 isogenic pairs examined one member of the pair had an extra band in the vicilin 50,000 Mr region, at approx. 48,000 Mr (Fig.10b). This band reacted with antivicilin IgG after the transfer of total protein extracts from polyacrylamide gel to nitrocellulose filter using a "Western" blotting procedure (see Section 2.10.5). Since these seeds were F8 selections, Va-1 and r must be closely linked for the different phenotypes at the two loci to have segregated together. Although, an exact figure can not be calculated from the small sample analysed, Va-1 and r were suggested to be within 10 cross-over units of each other, since this distance implies co-segregation in approx. 50% of F8 selections assuming the parental lines differed in both their Va-1 and r phenotypes.
Fig. 9 SDS-PAGE (12% acrylamide gel) of total protein extracts of seeds of pea lines 1263 and 3080, and F1 and F2 seeds from their cross. Vc-1 indicates variation in vicilin 50,000 Mr subunit band pattern, Lg-1α variation in legumin α-subunit band pattern. The band patterns of the parental lines 1263 (vicilin and legumin type A), 3080 (type B), the F1 in both the straight (F1/S) and reciprocal (F1/R) crosses (type AB) and the F2 segregation (type; A, AB, B) of vicilin 50,000 Mr subunit and legumin α-subunit are shown at the top and bottom of the plate, respectively.
Fig. 9

<table>
<thead>
<tr>
<th>L1263</th>
<th>F1/S</th>
<th>F1/R</th>
<th>F2</th>
<th>L3080</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>AB</td>
<td>AB</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Vc) 71
(Vc) 50
(Vc) 33
(Lg - β) 22
Mr × 10^3

Lg-1α
Vc-1

Legumin
Vicilin
Fig. 10 SDS-PAGE (12% acrylamide gel) of total protein extracts of pea seeds.

a) The parental lines 1263 and 3080, and F1 and F2 seeds from their cross and of vicilin from var. Feltham First (VcFF) after "Western" blotting to nitrocellulose filter.

b) Isogenic round/wrinkled lines (T), compared with corresponding "Western" blot of vicilin subunit (Vc). R indicates round seed phenotype, r wrinkled seed phenotype. > indicates an extra band in vicilin 50,000 Mr region, at approx. 48,000 Mr.
Fig. 10

a. 

![Image](image_url)

b. 

![Image](image_url)
Table (10) Analysis of the F2 segregation of vicilin 50,000 Mr subunit (Ve-1) and legumin \( \alpha \)-subunit (Lg-1) band patterns on SDS-PAGE and round/wrinkled seed surface (R/r) in a cross between two pea lines (1263 x 3080).

<table>
<thead>
<tr>
<th>Gene</th>
<th>F2 phenotypic classes and observed frequency</th>
<th>Segregation ratio</th>
<th>( \chi^2 )</th>
<th>Probability range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ve-1; A, AB, B</td>
<td>AA 68, Aa 160, aa 80</td>
<td>1 : 2 : 1</td>
<td>1.4026</td>
<td>0.50 - 0.40</td>
</tr>
<tr>
<td>Lg-1; A, AB, B</td>
<td>AA 84, Aa 158, aa 66</td>
<td>1 : 2 : 1</td>
<td>2.3117</td>
<td>0.40 - 0.30</td>
</tr>
<tr>
<td>R, r</td>
<td>AA 238, Aa 70</td>
<td>3 : 1</td>
<td>0.8485</td>
<td>0.50 - 0.40</td>
</tr>
</tbody>
</table>

Ve-1, A (line 1263)  
Ve-1, AB (F1)  
Ve-1, B (line 3080)

Leg-1, A (line 1263)  
Leg-1, AB (F1)  
Leg-1, B (line 3080)
Table (11) Segregation of vicilin 50,000 Mr subunit (Ve-1) and legumin α-subunit (Lg-1) band patterns on SDS-PAGE in the F2 generation of a cross between two pea lines (1263 x 3080) with respect to the selected marker gene (R/r) for round/wrinkled seed surface. Segregation ratio is 9 : 3 : 1 : 1. Cp for coupling and Rp for repulsion.

<table>
<thead>
<tr>
<th>Gene pair</th>
<th>F2 phenotypic classes and observed frequency</th>
<th>$x^2$</th>
<th>Probability range</th>
<th>Cross-over ± probable error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AB</td>
<td>Ab</td>
<td>aB</td>
<td>ab</td>
</tr>
<tr>
<td>Lg-1 - r</td>
<td>164</td>
<td>74</td>
<td>60</td>
<td>10</td>
</tr>
<tr>
<td>Ve-1 - r</td>
<td>170</td>
<td>68</td>
<td>70</td>
<td>0</td>
</tr>
<tr>
<td>Ve-1 - Lg-1 (Cp)</td>
<td>193</td>
<td>47</td>
<td>31</td>
<td>37</td>
</tr>
</tbody>
</table>

Line 3080 was taken as dominant for Ve-1 and Lg-1 band patterns; thus (3080) includes seeds showing F1 protein patterns.
3.5 Genetic variation at the ribosomal RNA gene sites

The purified DNA of pea lines which were tested for genetic variation in legumin restriction fragments (see Section 3.3.2.1) was used to investigate the inheritance and chromosomal location of rRNA gene sites in pea (*Pisum sativum* L.). The total genomic DNA was digested with different restriction enzymes i.e. *Bam* HI, *Bgl* II, *Eco* RI and *Hind* III. Double digestions with *Hind* III + *Bam* HI, *Hind* III + *Bgl* II and *Hind* III + *Eco* RI were also carried out (see Section 2.12). The restriction fragments were electrophoretically separated on agarose gels and transferred to nitrocellulose filter by "Southern" blotting as described in Sections 2.13 and 2.16. The genomic blot was hybridized with ribosomal RNA radioactive probe 

\[ ^{32}P\text{-rDNA} \]

as described in Sections 2.17, 2.18 and 2.20, in order to find length variation in the restriction fragments of ribosomal RNA genes. The restriction enzyme *Eco* RI was found to give the clearest distinction between three different band patterns (Fig.11a); pattern "A" has four fragments of size 6.5, 6.1, 5.1 and 3.7 kbp (line 110), pattern "B" has three fragments of size 5.5, 5.1 and 3.7 kbp (lines 102, 200, 1238, 3080, 5452 and 5478) and pattern "C" has three fragments of size 5.8, 5.1 and 3.7 kbp (line 1263 and var. Feltham First). The band corresponding to the fragment of 3.7 kbp was found in all the lines investigated either in this study (Fig.11a) or in others (Jorgensen et al., 1982; Ellis et al., 1984). This fragment contains most of the coding sequence for 25S rRNA and none of the external spacer region (Ellis et al., 1984). The variant bands in all lines studied were found to correspond to fragments of size 5.1 - 6.5 kbp. These fragments contain the non-transcribed spacers.
A cross of line 110 and line 1238 was selected as suitable for testing the inheritance of ribosomal RNA genes; in addition, the genes \( le \), \( r \) and \( tl \) for dwarfness, round/wrinkled seed surface and absence of tendrils, respectively, differed between these lines which allow the linkage between these genes and the genes for ribosomal RNA to be investigated.

### 3.6 Inheritance and mapping of ribosomal RNA genes

#### 3.6.1 Inheritance of ribosomal RNA genes

DNA was isolated from leaves of F1 plants of the reciprocal crosses between lines 110 which has rDNA fragment pattern "A" and 1238 which has pattern "B" (Fig.11a). The DNA was purified by banding in caesium chloride gradients and digested with \( \text{Eco RI} \) (see Sections 2.11 and 2.12). The digested fragments were separated by agarose gel electrophoresis and transferred to nitrocellulose filters using "Southern" blotting. The genomic blot was then hybridized with the radioactive probe of ribosomal RNA (\(^{32}\text{P}-\text{rDNA}\)). The banding pattern of \( \text{Eco RI} \) restriction fragments of rDNA in the F1 plants was found to be the additive pattern "AB" i.e. contained bands from both parents "A" and "B". The additive pattern "AB" of the F1 was found to be the same in the reciprocal crosses irrespective of which parent was the female parent and no segregation was observed (Fig.11b). These results indicate that there is no maternal effect on the inheritance of ribosomal RNA genes and the two parental lines were homozygous.

Twenty eight F2 plants were examined by following the same methods which used in the analysis of the parental and F1 plants. The results obtained showed the expected mongenic inheritance and dominance of the
genes \( Le, R \) and \( Tl \) (Table 12). The segregation of \( Eco \) RI restriction fragments (pattern; A, AB, B) in the F2 plants is shown in Fig.11b. The observed frequencies of the parental and hybrid patterns of rDNA (A, AB, B) are given in Table 12, and support the hypothesis that ribosomal RNA genes were controlled by a single pair of co-dominant alleles.

3.6.2 Molecular mapping of ribosomal RNA genes

In order to establish whether or not the genes for ribosomal RNA were linked to the selected morphological markers; \( Le/le \) for plant height and \( R/r \) for round/wrinkled seed surface, linkage relations between \( Eco \) RI restriction fragments of rDNA (pattern; A, AB, B) of the twenty eight F2 plants which were used in investigating the inheritance of rRNA genes and the selected morphological markers (\( le \) and \( r \)) were tested. The data presented in Table 13, show that no linkage was detected between ribosomal RNA genes and the morphological markers (\( le \) and \( r \)). Linkage was not tested between ribosomal RNA genes and the gene \( tl \) for absence of tendrils because of the close linkage between the genes \( r \) and \( tl \) (5 map units, Blixt, 1974); the number of F2 plants was not large enough to show recombination between them. Since the genes \( le \) and \( r \) were located on the short arm of chromosomes 4 and 7, respectively (Blixt, 1974), the genes for ribosomal RNA are located either in the long arm of these chromosomes (the nucleolus organizer sites, Morrison and Lin, 1955; Blixt, 1974; Lamm, 1981; Davies and Cullis, 1982) or on different chromosomes.
### 3.6.3 Cytolocalization of ribosomal RNA genes

Fig. 12a shows the metaphase chromosomes from root tips of pea line 110 (*Pisum sativum* L.) after aceto-carmine staining (see Section 2.26). In most metaphase plates only two pairs of submetacentric chromosomes carry distinct satellites attached to the telomere of the long arms. These two pairs of homologous chromosomes were identified by Morrison and Lin (1955) and classified as chromosomes 4 and 7 by Blixt (1974) and Lamm (1981). Since the satellite constrictions are often "nucleolus organizers" that is the genetic sites of ribosomal RNA genes, rRNA radioactive probe (\(^3\)H-rDNA) was used to locate the transcriptionally active rRNA genes to positions onto pea chromosomes as described in Sections 2.19 and 2.26. After *in situ* hybridization, some of the metaphase plates were shown to have silver grains at four sites (Fig. 12b), apparently corresponding to the satellite constrictions. However, identification of the chromosomes was difficult because of their sensitivity to the lengthy procedures of *in situ* hybridization.

### 3.6.4 Physical mapping of ribosomal RNA genes

In order to locate the exact position of ribosomal RNA gene sites onto the telomere of the long arm of chromosomes 4 and 7, the distance between the centromere and the satellite constriction (rRNA gene sites) was measured in metaphase chromosomes after aceto-carmine staining and transformed to a percentage of the total length of the chromosome as previously described (Section 2.27). This distance, referred to as the long arm "LA" (Fig. 13), was found to constitute 55.76% and 55.45% of the total length of chromosomes 4 and 7,
respectively (Table 14). The genetic map of Blixt, 1974 was used to calculate the distance between the centromere and the satellite constriction in map units. The results obtained showed that ribosomal RNA gene clusters lie on the telomere of the long arm at 138 and 60 map units from the centromere of chromosomes 4 and 7, respectively. Table 14 also shows that the satellite constriction (SC) of chromosomes 4 and 7 were 22.54% and 10.76% of the chromosome length (SA+LA+Sat), respectively. This indicates that the nucleolus organizer (satellite constriction) of chromosome 4 is about twice the size of that of chromosome 7 which suggests that the volume of the nucleolus on chromosome 4 is nearly twice the volume of that on chromosome 7 (Martini and Flavell, 1985).
Fig. 11 Eco RI digests of total DNA from pea leaves after fractionation on 0.7% agarose gels, autoradiographs of hybridization of $^{32}$P-rDNA to total genomic DNA after "Southern" blotting to nitrocellulose filter.

a) Different lines (numbered).

b) The parental lines 110 (rDNA type A) and 1238 (type B), their F1 in both the straight (F1/S) and reciprocal (F1/R) crosses (type AB) and the F2 segregation (type; A, AB, B).
Fig. 11

a.

![Image of gel electrophoresis with molecular weights and bands labeled with kbp.

b.

![Image of gel electrophoresis with bands labeled with kbp.]}
Fig. 12 Metaphase mitotic chromosomes from root tips of pea line 110 (Pisum sativum L.), 2n=14.

a) After aceto-carmine staining. Arrows indicate chromosomes 4 and 7.

b) After in situ hybridization to $^3$H-rDNA from cloned pea ribosomal gene repeat (pHAI). Arrows indicate rRNA gene sites.
Fig. 13 The physical map of chromosomes 4 and 7 from root-tip metaphase preparations of pea line 110 (Pisum sativum L.), 2n=14. Different parts of the chromosome are shown (SA = the short arm, LA = the long arm, SC = the satellite constriction and Sat = the satellite).
Table (12) Analysis of the F2 segregation of Eco RI restriction fragments of rDNA (rRNA) and the selected marker genes studied in a cross between two pea lines (110 x 1238).

<table>
<thead>
<tr>
<th>Gene</th>
<th>F2 phenotypic classes and observed frequency</th>
<th>Segregation ratio</th>
<th>$\chi^2$</th>
<th>Probability range</th>
</tr>
</thead>
<tbody>
<tr>
<td>rRNA, A, AB, B</td>
<td>AA 7 Aa 16 aa 5</td>
<td>1 : 2 : 1</td>
<td>0.8571</td>
<td>0.07 - 0.60</td>
</tr>
<tr>
<td>rRNA, A</td>
<td>Le 21 le 7</td>
<td>3 : 1</td>
<td>0.0000</td>
<td>0</td>
</tr>
<tr>
<td>rRNA, A</td>
<td>R 22 r 6</td>
<td>3 : 1</td>
<td>0.1429</td>
<td>0.90 - 0.80</td>
</tr>
<tr>
<td>rRNA, A</td>
<td>Tl 22 tl 6</td>
<td>3 : 1</td>
<td>0.1429</td>
<td>0.90 - 0.80</td>
</tr>
</tbody>
</table>
Table (13) Segregation of Eco RI restriction fragments of rDNA (rRNA) in the F2 generation of a cross between two pea lines (110 x 1238) with respect to the selected marker genes. Segregation ratio is 9 : 3 : 3 : 1. Cp for coupling and Rp for repulsion.

<table>
<thead>
<tr>
<th>Gene pair</th>
<th>F2 phenotypic classes and observed frequency</th>
<th>$\chi^2$</th>
<th>Probability range</th>
<th>Cross-over ± probable error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AB</td>
<td>Ab</td>
<td>aB</td>
<td>ab</td>
</tr>
<tr>
<td>rRNA - r (Cp)</td>
<td>17</td>
<td>5</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>rRNA - R (Rp)</td>
<td>15</td>
<td>7</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>rRNA - le (Cp)</td>
<td>18</td>
<td>3</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>rRNA - Le (Rp)</td>
<td>15</td>
<td>6</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>r - le (Cp)</td>
<td>16</td>
<td>5</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>
Table (14) Average length (in microns) of the different parts of chromosomes 4 and 7 from root-tip metaphase preparations of pea (*Pisum sativum* L.) line 110. SA = the short arm, La = the long arm, SC = the satellite constriction, Sat = the satellite. LA and SC were transformed to percentages of the total length of the chromosomes as described in section 2.27.

<table>
<thead>
<tr>
<th>Chromosome No.</th>
<th>length in microns</th>
<th>% of total (LA)</th>
<th>% of total (SC)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SA</td>
<td>LA</td>
<td>SC</td>
</tr>
<tr>
<td>4</td>
<td>± 1.42</td>
<td>± 2.45</td>
<td>± 1.00</td>
</tr>
<tr>
<td></td>
<td>± 0.26</td>
<td>± 0.42</td>
<td>± 0.14</td>
</tr>
<tr>
<td>7</td>
<td>± 1.54</td>
<td>± 2.78</td>
<td>± 0.55</td>
</tr>
<tr>
<td></td>
<td>± 0.24</td>
<td>± 0.48</td>
<td>± 0.12</td>
</tr>
</tbody>
</table>
4. DISCUSSION
4.1 Isoenzymes

Enzymes are proteins which catalyze chemical reactions. Each enzyme may exist in the same organism in more than one molecular form, and such multiple molecular forms of an enzyme in a single organism have been designated isoenzymes (Markert and Moller, 1959). An enzyme coded at a specific gene locus very often occurs in structurally different forms in different individuals of a plant species. Such allelic variants presumably arose by single gene mutations in earlier generations. In most cases they probably represent the substitution of one amino acid for another in the polypeptide chain as a result of single base substitution in the DNA coding sequences. Thus, the differences between the allelic forms generally reflect only a slight change in the overall structures of the protein. The extensive degree of allelic enzyme variation that occurs in most natural populations was largely discovered by the techniques of gel electrophoresis.

In many instances, the enzyme families are made up of heteromeric isoenzymes containing polypeptides encoded by different loci as well as of homomeric isoenzymes in which the polypeptides are derived from single loci. In other cases heteromers are not found because the enzymes are monomeric. There are two different examples involved in this study; first, amylase and esterase are monomeric molecules, each contains single polypeptide chain (A or B). Second, GOT is a homodimeric molecule contains two identical polypeptides (AA or BB).
Studies dealing with genetic aspects of isoenzymes in higher plants started on the multiple esterases in maize endosperm (Schwartz, 1960). Since that time most of the studies of genetic control of isoenzymes in higher plants have reported monogenic inheritance with co-dominance relations between alleles for several isoenzyme systems in different plant species. Gates (1978) found that the intensity of staining of esterase bands from cotyledons of legumes could vary in reciprocal crosses. This may indicate that the expression of the alleles depends on the cytoplasmic environment in which they are expressed, and that some paternal alleles are suppressed in the maternal tissue. This explanation was also given by McDonald and Brewbaker (1972) in other isoenzyme systems and Davies (1973) for the maternal inheritance of globulin patterns in pea.

Amylase affords a good opportunity for the study of multiple molecular forms of enzymes. Both α-amylase and β-amylase are known to occur in higher plants and their chemical and physiological activities are fairly well understood. The first genetic studies of amylase isoenzymes in plants were reported by Frydenberg and Nielsen (1965) in barley, suggesting simple Mendelian control of the amylases in higher plants. This was followed by the study of Scandalios (1966) in inbred lines of Zea mays. He found electrophoretic variants of multiple amylases and the hybrid patterns obtained from F1 crosses generally consisted of the maximum number of bands present in either parents. In both electrophoretic (Juliano and Varner, 1969; Yomo and Varner, 1973; Al-Helal, 1985) and biochemical (Swain and Dekker, 1966) assays, α-amylase has been distinguished from β-amylase and phosphorylase by
i) α-amylase was stable to heating to 70°C for 15 min in the presence of Ca++, while the other enzymes were inactivated; ii) α-amylase was inactivated by EDTA while the other enzymes retained their activity; iii) α-amylase and debranching enzyme degraded soluble starch completely, leaving no colour when stained with iodine, while β-amylase left a residual colour level due to α-limit dextrins; and iv) phosphorylase degraded no starch in the absence of phosphate. None of these tests was carried out in this study, since the main amylase enzyme in pea cotyledons has been shown to be β-amylase by Al-Helal (1985). Further, the colour of the bands on acrylamide gel after assaying for amylase was "pink" which supports the conclusion that the investigated enzyme is β-amylase.

Schwartz (1960) separated esterases from crude extracts of immature maize seeds by starch gel electrophoresis and detected the sites of enzyme activity by a histochemical methods. The pattern of sites of enzyme activity obtained by electrophoresis has been termed a zymogram (Hunter and Markert, 1957). Genetic variation in esterase zymogram from germinating seeds of 12 pea varieties was found by Frankel and Garber (1965). This was followed by the detection of electrophoretic variants of esterase and leucine aminopeptidase activity in crude extracts from cotyledons of germinating seedlings of Phaseolus vulgaris and P. coccineus by West and Garber (1967a). Zymograms of Phaseolus vulgaris and P. coccineus, the interspecific hybrids, their progeny, backcrosses of the interspecific hybrids and the parental species, and their progeny were compared by West and Garber (1967b) to determine the inheritance of two sites of esterase activity. The genetic data indicated monogenic Mendelian inheritance
of two sites of investigated esterases. Brown and Allard (1969) investigated several enzyme systems in two populations of *Zea mays* and demonstrated that each of these systems was governed by a single locus. One of these systems was esterase which was found to be controlled by a single Mendelian locus with three alleles. Similar results were obtained by Wall and Whitaker (1971) after the electrophoretic analysis of crude seed extracts of *Cucurbita ecuadorensis*, *C. maxima*, their F1 and F2, and three of the four possible interspecific backcrosses.

Transaminases interconvert amino acids and keto acids and play an important role in maintaining amino acid balance during the development and differentiation of plant tissue. The most common transaminase reactions involve glutamic acid and its keto acid analogue, ketoglutaric. The observation that the F1 band pattern of glutamate oxaloacetate transaminase (GOT) has an additional band of intermediate mobility with respect to the variant parental bands suggests that this enzyme is a dimeric in structure and that the intermediate isoenzyme represents an allodimer or hybrid enzyme containing the two parental type subunits. This phenomenon of dimeric GOT was observed in *Zea mays* (MacDonald and Brewbaker, 1972; Goodman *et al.*, 1980), *Pinus sylvestris* (Rudin, 1975), *Pinus rigida* (Guries and Ledig, 1978), and *Vicia faba* (Suso and Moreno, 1982). In all these cases, GOT banding pattern was found to be controlled by co-dominant alleles at single loci.

The results obtained from this study indicate that the genetic variation in the isoenzyme systems previously indicated in higher plants can be used in pea to map the controlling genes onto the
Each of the three isoenzyme systems investigated, amylase, esterase and glutamate oxaloacetate transaminase was found to be controlled by co-dominant alleles at a single locus. These results are in agreement with the results obtained by Frydenberg and Nielsen (1965), Scandalios (1966) for amylase; Frankel and Garber (1965), West and Garber (1967b), Brown and Allard (1969), Wall and Whitaker (1971), Weeden and Marx (1984) for esterase; MacDonald and Brewbaker (1972), Rudin (1975), Guries and Ledig (1978), Goodman et al. (1980), Suso and Moreno (1982) for GOT.

The F2 segregation results from the cross involved amylase gene (Amy) showed that this gene is on chromosome 2 and linked to the loci k and wb for keel-like wings and waxless stipules, respectively. The deduced cross-over values (Table 4), predict a position of Amy in Pisum; wb...9...k...25...Amy (Fig. 14).

Analysis of the data from the cross involved the gene for esterase isoenzyme (Est) was complicated due to the polymeric genes affecting the bracteole phenotype. The early studies of Lamprecht and Mrkos (1950) indicated that the gene Br causes (both in dominant and homozygous states) the development of bracteoles, but later Lamprecht (1953) found that the development of bracteoles was dependent upon two polymeric genes in dominant states. Apart from the earlier known gene Br, dominance is necessary for the other gene; which is signified by the symbol Brα. From the gene map (Blixt, 1974), the genes Br and Brα are on chromosomes 4 and 3, respectively. Esterase isoenzyme patterns were shown to be linked to the bracteole phenotype (Table 6), and since the lines in this cross differ in their genotype with respect to Br, linkage must be between Br and Est. The failure to
No attempt was made to correlate any of these genes to the gene mapped in this study.
detect linkage between \( le \) and \( Est \) suggests that \( Est \) is on the opposite side of \( Br \) to \( le \) i.e. the tentative map position is \( Est...28...Br...60...le \) (Fig.14). Since some esterase bands were found to be invariant in this study, this suggests the presence of several esterase loci in the pea genome. Weeden and Marx (1984) showed that pea leaf esterase zymogram contains four esterase i.e. \( EST-1, EST-2, EST-3 \) and \( EST-4 \); the \( EST-4 \) locus was mapped to position on chromosome 7.

The gene for glutamate oxaloacetate transaminase in \( Pisum \) was found to be on chromosome 1 by linkage with the loci \( a \) and \( d \) for the absence of anthocyanin pigment and the absence of maculum ring, respectively. The cross-over values of 24 and 41 map units from the genes \( a \) and \( d \), respectively (Table 8) fix its position as lying between these loci; \( a...24...Got...41...d \) (Fig.14). The predicted value for the \( a...d \) distance, 65 map units, agrees well with that in the standard \( Pisum \) gene map (76 units; Blixt, 1974). This \( Got \) locus does not correspond to the major isoenzyme bands, but to faster migrating minor bands; it appears to correspond to the \( GOT-1 \) locus of Przybylska et al. (1982). It seems likely that this is only one of two or more \( GOT \) loci, since another zone of invariant band pattern was found. This zone was found to contain two bands of slow migration (Rf 0.31 and 0.28) which was previously reported as \( GOT-2 \) by Przybylska et al. (1982) in \( Pisum \) and "zone A" by Suso and Moreno (1982) in \( Vicia \). Weeden and Marx (1984) found that the mitochondrial aspartate aminotransferase (AAT-m) from pea leaves was also controlled by co-dominant alleles at a single locus and mapped the controlling gene to position on chromosome 2.

The mapping of these isoenzyme loci to defined positions on the
pea genome offers the opportunity to isolate genes which will be assignable to those loci. Of the three loci studied, esterase and GOT are clearly complex and it is likely that a number of genes encoding these proteins are present, distributed over the genome. In the case of β-amylose, a single locus is clearly possible, and isolation of the gene encoding this protein would give a sequence that would be located with some confidence to chromosome 2. Isolation of such a gene could be carried out via a cDNA isolated from a pea cotyledon cDNA library, if the enzyme itself were purified and used to raise antibodies to allow identification of cDNA clones by conventional methods.

4.2 Storage proteins

A number of proteins from seeds of *Pisum sativum* L. were investigated regarding the inheritance and mapping the controlling genes onto the genome. Apart from the early report by Davies (1973) that the female parent has some effect on the inheritance of pea storage proteins, all the proteins so far investigated have been reported to be under the control of a single pair of co-dominant genes i.e. vicilin (Hynes, 1968; Thomson and Schroeder, 1978), acidic subunits of legumin (Thomson and Schroeder, 1978; Casey 1979b; Matta and Gatehouse, 1982), and albumin (Blixt et al., 1980).

Legume plants synthesize large amounts of proteins, which subsequently accumulated and stored in membrane-bounded organelles "protein bodies". Vicilin-like protein i.e. vicilin from pea (*Pisum sativum*) and broad bean (*Vicia faba*), β-conglycinin from soybean (*Glycine max*) and phaseolin from bean (*Phaseolus vulgaris*) are very complicated proteins of molecular weight range from 140,000 to
about 225,000 and sedimentation constant 6-7S. This protein is a trimer of polypeptides of Mr 50,000 and many also have smaller polypeptides between 12,000 and 33,000 Mr (Croy et al., 1980b; Gatehouse et al. 1981). The vicilin locus (Ve-l) identified in the present study is likely to represent a structural gene for vicilin 50,000 Mr because the observed difference in 50,000 Mr subunit band patterns corresponds most nearly of all vicilin polypeptides to the initial translation products, and thus to genomic coding sequences. It is possible, however, that the differences in 50,000 Mr subunit band patterns observed on SDS-PAGE are the result of some kind of post-translational modification not determined by structural genes, e.g. differences in carbohydrate side chains, but this is unlikely since the carbohydrate moiety of vicilin is known to be associated almost exclusively with the 16,000 Mr polypeptide (Davey et al., 1981). This locus relates to a polypeptide different from those associated with the loci identified by Thomson and Schroeder (1978), although because of the precursor-product relationships of vicilin polypeptides of approx. 50,000 Mr and less than 35,000 Mr respectively, it is not necessarily different from their loci "Veb" or "Vcc". The symbol Ve-l is used in accordance with the recommendations of the Pisum Genetic Association (Blixt, 1977) to allow for further loci, if shown to be distinct when mapped, to be designated Ve-2, etc.

Further vicilin loci are possible, since vicilin is known to be encoded by a small family (<10 copies) of genes on the pea chromosomes (Gatehouse et al., 1983; Lycett et al., 1983b). However, some or all of the vicilin structural genes may be arranged in one or more closely linked arrays which would result in the number of "classical" vicilin loci being considerably fewer than the number of genes determined by
DNA hybridization. The present vicilin locus is not closely associated with either legumin or convicilin loci and thus vicilin gene expression must be controlled independently of the other storage proteins.

Close linkage of the vicilin locus with the r locus (Fig. 14) is interesting in view of reports that legumin: vicilin ratio in peas is affected by the round/wrinkled genotype (Davies, 1980; Schroeder, 1982), although the actual amounts of legumin in mg per seed does not differ significantly between round and wrinkled lines (Schroeder, 1982). It is thus perhaps tempting to postulate a direct effect of the r locus genotype on vicilin synthesis, as was suggested for legumin synthesis (Davies, 1980); however, even if this were the case, the pleiotropic effects of the wrinkled phenotype (Shia and Slinkard, 1977) would make the existence of such a direct control very difficult to demonstrate. Since vicilin genomic clones are now available (R. Sawyer, unpublished results) it is likely that at least some of these represent the Ve-1 locus mapped in this study. However, it will be difficult to correlate the genes to protein subunits since the protein sequence data are not sufficiently accurate to allow precise matching to genomic sequences.

Legumin-like protein such as legumin from pea (Pisum sativum) or broad bean (Vicia faba) and glycinin from soybean (Glycine max) have a sedimentation constant of 11-12S and molecular weight of 300,000-360,000 and consists of six subunits each of 60,000-65,000 Mr. Each subunit contains two polypeptides of Mr 38,000-42,000 (α- or acidic subunits) and Mr 17,000-20,000 (β- or basic subunits) held together by a single disulphide bond (Derbyshire et al., 1976, Croy et al., 1979; Matta et al., 1981). The genetic variations obtained by
SDS-PAGE analysis were used to examine the inheritance of legumin in pea by Thomson and Schroeder (1978), Casey (1979b), and Matta and Gatehouse (1982). The results obtained from this study showed that α-subunits behave as the product of a single Mendelian gene, which is in agreement with the previous reports. The controlling genes were located near the r locus on chromosome 7 by Davies (1980) and Matta and Gatehouse (1982). The latter authors located the genes for acidic subunits of the major legumin subunit pairs to chromosome 7 at 15 map units distant from the tL locus on the side away from the r locus. The present study showed that legumin genes for acidic subunits are 36 map units from the r locus on chromosome 7. Moreover, this locus showed some linkage with the locus for vicilin storage protein with 29 map units between them (Table 11). The inheritance of β-subunits has not been investigated because of a lack of suitable variants.

The legumin locus Lg-1 must correspond to at least one of the legumin genes legA, legB or legC on the following grounds. The polypeptides encoding major legumin α-subunits are always inherited as a unit, and no segregation has been observed between them (Casey, 1979; Matta and Gatehouse, 1982). LegA corresponds in sequence to major legumin polypeptides, and is identical to a sequenced cDNA species, and thus is expressed (Lycett et al., 1984a). Amino acid sequence variants in major legumin polypeptide corresponding to several of the amino acid substitutions encoded by legB and legC have been detected (Gilroy and Gatehouse, unpublished results) and in one case (threonine at position 2) could be correlated with different major legumin polypeptides. Thus legA must map to the Lg-1 locus, and either or both of legB and legC must be at this locus also. Since the genes
do not overlap on the existing genomic clones, isolation of further genomic sequences will be necessary to map the overall organisation of this locus.

Legumin subunits which are encoded by a small gene family (~7 members; Shirsat, 1984) map to more than one locus (Casey, 1979b; Matta and Gatehouse, 1982). The independent segregation of some legumin genes observed by Matta and Gatehouse (1982) indicates the presence of other locations for legumin genes either on chromosome 7 and are not linked to the one already located or in any part of the genome. Moreover, a different kind of legumin (80,000 Mr) was observed by Chrispeels et al. (1982) and showed some kind of structural similarity with the conventional 60,000 Mr legumin precursor (Domoney and Casey, 1984). Domoney and Casey (1985) have shown that these minor legumin polypeptides are encoded by 3-4 genes, and that these genes form a separate subfamily to the 4 genes encoding the major legumin polypeptides. The sequence of a gene from the 'minor legumin' subfamily shows considerable sequence divergence from the legA, legB and legC genes (Gatehouse et al., in preparation). Since the major and minor legumin polypeptide genes belong to separate subfamilies, different genomic locations for the two subfamilies is not unexpected.

Protein electrophoresis has been widely used in different genetical studies, but the electrophoretic variants in some complicated proteins such as legumin and vicilin were not enough for mapping all the controlling genes. Thus, the method mentioned by Polans and Thompson (1983) which based on the analysis of DNA restriction fragment length polymorphisms was used in this study to investigate the possibility of obtaining genetic variation in legumin fragment
Fig. 14 Relative map positions on the seven chromosomes of *Pisum sativum* of the loci examined in this study which are shown to the right of the chromosome and the marker loci which are shown to the left of the chromosome.

The chromosome lengths and the map positions of the marker genes were taken from the genetic map of Blixt, 1974.
banding patterns between different pea lines after digestion with restriction enzymes. The results obtained showed a wide range of variation in Eco RI restriction fragments. In only 7 different pea lines digested with one enzyme, four different kinds of length polymorphisms were observed (see Fig. 5). This indicates the potential power of this technique for future analysis of different genotypes.

The complete nucleotide sequence of legB gene shows typical eukaryotic gene features. The coding sequence starts with the methionine initiation codon "ATG" at the 5' end and ends with the termination codon "TAG" at the 3' end. The nucleotide sequence codes for 517 amino acid residues which are the components of the signal peptide, α-subunit and β-subunit (21, 311 and 185 amino acid residues, respectively) of legumin protein. Lycett et al. (1984b) have shown the predicted amino acid composition of legA gene product as well as molecular weights of 36,441 and 20,185 for the α- and β-subunits, respectively. When this was compared with the predicted amino acid composition of legB gene product, the results show no significant variation between the two genes in most of the amino acid residues (Table 9). The sulphur amino acid cysteine shows no variation between the two genes whereas the number of methionine residues increased by 25% in legB (4 and 5 residues in legA and legB, respectively).

There are three intervening sequences present in legB gene, two within the sequence encoding the legumin α-subunit and one within the sequence encoding the β-subunit. Apart from a short deletion of 14 base pairs in the 3' region of the IVS-3 of legB, the three intervening sequences show strong homology in nucleotide sequence, location
and boundary sequences with the previously published legA (Lycett et al., 1984b) and legC (Gatehouse, unpublished results) and in agreement with the plant consensus sequences (Slightom et al., 1983). Meanwhile it was difficult to make a comparison with legumin gene legD, which is a "pseudogene" in the pea legumin gene family because of long deletions in the coding sequences. However the IVS-1 and IVS-3 of legD gene are homologous in location to legA introns but slightly shorter; 84 and 92 bp, respectively, in legD compared with 88 and 99 bp, respectively, in legA but the nucleotide sequences differ significantly. The IVS-2 of legD can not be compared because of a long deletion in its 5' end.

The nucleotide sequence of the 3' flanking region of legB gene shows a strong homology with that of legA with the same three polyadenylation signals. These signals downstream of the two genes are respectively AATAAGAAAA, AATAAAATAAAA, and AATAAATAAAA. In contrast, legD gene shows a complete divergence immediately after the stop codon with three polyadenylation signals in its immediate 3' flanking sequences, none of these correspond to the polyadenylation signals in legA or legB and only the first is of the multiple overlapping type found in legA and legB. The fourth sequenced gene in this subfamily, legC, shows greater homology to legB than legA (Gatehouse, unpublished results) in that its 3' noncoding sequence is identical, and most of the nucleotide substitutions in the introns and exons are common to legB and legC. However, legC is identical to legA rather than legB at some positions (see Fig.15), and this shows that legB and legC are not simply a duplicated gene but have undergone some evolution away from each other.
It may be concluded that the functional legumin genes (\textit{legA, legB} and \textit{lege}) so far investigated (Fig.15) show a strong homology, even the non-functional gene (\textit{legD}) sequence shows a partial homology with the members of the legumin gene family. This suggests that these genes may evolved from the same ancestor through the evolution processes by means of mutations or crossing-overs. However, the nucleotide sequences of these genes show the variation between them in the coding regions, the non-coding regions and the 5' and 3' flanking regions as nucleotide base substitutions. This may prove that DNA sequencing is the best technique for comparing gene structures, establishing variations between genes and providing the essential information for the genetic manipulations of plant genes. It is interesting to note that sequence divergences between \textit{legA, legB} and \textit{lege} do not show a strong trend towards changes at silent codon positions, or indeed towards changes in non-coding compared to coding sequence. A region from approx. -300 bases relative to the start of coding sequence, to approx. 700 bp after the stop codons is virtually identical in all three genes (Lycett \textit{et al.}, 1985). The \textit{legA, legB} and \textit{lege} genes may thus represent genes that have duplicated very recently on an evolutionary time scale.

The third large protein in legume seeds and the richest in sulphur amino acids is albumin. This protein accounts for 20-35\% of the extractable protein of pea cotyledons (Schroeder, 1982). Albumin fraction which was isolated and characterized by Croy \textit{et al.} (1984) was found to contain two major proteins which designated as PMA-L "pea major albumin, large" and PMA-S "pea major albumin, small". PMA-L has a Mr \~{}53,000 and consists of two 25,000 Mr subunits, whereas PMA-S has
Fig. 15 Comparison of the legumin genes; legA, legB and legC of pea (Pisum sativum L.). The coding regions are presented as heavy lines, intervening sequences (IVS) and other features discussed in the text are shown.
Fig. 15

Polyadenylation sites

TATA Box
CAAT Box

End
End

Inverted repeats

I VS-1 IVS-2

αβ site
I VS-3

Leg A
Leg B
Leg C

Strong homology

Intermediate homology (>70%)

Weak or no homology

Deletions

Inverted repeats
a Mr ~ 48,000 and consists of two 24,000 Mr subunits. Genetic variations have been found in albumin banding pattern form pea cotyledons by Przybylska et al. (1973, 1977). Such variations were used by Blixt et al. (1980) to investigate the inheritance of albumin banding pattern on PAGE. Their results showed that albumin inheritance is determined by two alleles at one locus which appear to be weakly linked with the gene pl on chromosome 6.

A cross between two different pea lines, one with high legumin and low albumin and the other with low legumin and medium to high albumin has been constructed by Schroeder and Brown (1984). The result showed that legumin and albumin were negatively correlated. Since these two fractions are rich in sulphur amino acid contents, the quality of pea protein in terms of sulphur amino acid content can not be improved by raising the content of one of these two proteins, but the first step is to break the negative correlation between them and to combine high levels of legumin with high level of albumin. Schroeder and Brown (1984) found that out of a total of 141 F2 seeds from the cross mentioned above, only 4 round seeds contain high levels of legumin and albumin. However, no seeds with such combinations were found between the F2 wrinkled population which indicate the effect of the r locus on seed protein composition. The effect of the r locus on legumin of pea seeds was previously reported by Davies (1980). According to the data of the total sulphur amino acids from the cross mentioned above, Schroeder and Brown (1984) suggested that sulphur amino acid content of pea can be improved by breeding, but that the required selection regime must take both legumin and albumin contents into account. Improving the quality of proteins in terms of sulphur
amino acids was discussed in details by Boulter (1982). He suggested two ways for improving the sulphur amino acid content of legume seeds. First, increasing the amount of albumins as a whole by increasing the relative proportion of embryo axis to endosperm since the proportion of albumin to total protein is higher in the axis. Second, improvement could be made in different legumes by changing the ratio legumin : vicilin in favour of legumin since legumin is the richer in sulphur amino acids. Finally, a new era of research has been started aiming at the use of genetic engineering techniques for improving the quantity and quality of food crops and different prospects have been discussed by Larkins, 1983; Comai and Stalker, 1984; Fraser, 1984; Hess, 1984; Miflin and Lea, 1984; Shaw, 1984. However, plant storage proteins are the main target for such improvement for the reasons mentioned earlier (Section 1.1).

4.3 Ribosomal RNA

In all plant species so far investigated for the distribution of rRNA cistrons throughout the plant genome, it was demonstrated that the nucleolus was the site where the precursor of rRNA was synthesized and the location of the rRNA genes at the nucleolus organizer region (satellite constriction) where the nucleolus is built was established. Nucleolus size is correlated with the size of nucleolus organizer (NOR) constriction in metaphase chromosomes and poorly correlated with the number of ribosomal RNA genes in the NOR region. Martini and Flavell (1985) found that the nucleolus on chromosome 1B of the wheat variety Chinese Spring is twice the volume of that on chromosome 6B even though only half as many ribosomal RNA genes reside at the 1B locus compared with the number at the 6B locus. In the present study
the secondary constrictions of chromosomes 4 and 7 of Pisum sativum were found to be 22.54 and 10.76% of the total length of the chromosomes, respectively (see Table 14). This indicates that the NOR constriction of chromosome 4 is about twice the size of that of chromosome 7 which suggests that the volume of the nucleolus on chromosome 4 is nearly twice the size of that on chromosome 7. However, this provides no information about the number of rRNA genes at each of the two NORs. In the majority of plant species so far investigated, rRNA gene clusters were located to the satellite constriction on the short arm of the chromosome (De-Faria, 1972, 1976) and in few species on the long arm e.g. Vicia (Burger and Knälmann, 1980) and Pisum (Davies and Cullis, 1982).

The congruity between the number of rRNA gene sites and the satellited chromosomes found in the present study indicates that the metaphase chromosomes stained by conventional means to show the satellite constrictions reveal the total number of rRNA gene sites. Similar results were found by Teoh et al. (1983) in diploid Aegilops species, whereas a clear departure from a correlation between the number of NOR regions and the number of rRNA gene clusters was found by Mohan and Flavell (1974) in different wheat genotypes and Teoh et al. (1983) in polyploid Aegilops species.

In the present study a cross between two different pea lines was set up in an attempt to investigate the inheritance of rRNA genes and locate the controlling genes onto the pea genome by the combined use of the classical linkage methods and the analysis of the fragment length polymorphisms after digestion with restriction enzymes. The results obtained from this cross (see Tables 12 and 13) showed the
simple Mendelian inheritance of Eco RI digests of rDNA in the F2 plants and no linkage was detected with the morphological markers $Le/le$ for plant height and $R/r$ for round/wrinkled seed surface (chromosomes 4 and 7, respectively). This may confirm the mapping of rRNA gene sites to the satellite constrictions on the long arm of chromosomes 4 and 7, since the morphological markers ($le$ and $r$) involved in this cross are known to lie on the short arms of the chromosomes (Blixt, 1974). Davies and Cullis (1982) cultured pea cotyledons in the presence of 2-4D to stimulate some cells with polytene chromosomes to enter the prophase and used these chromosomes for *in situ* localization of rRNA gene sites. Their results showed the presence of rRNA genes in the NOR region of chromosomes 4 and 7.

In order to estimate the chromosomal locations of rRNA genes in map units, an alternative method has been employed. In this method, the satellited chromosomes were measured in the metaphase mitotic preparations and the distance between the centromere and the satellite constriction (the nucleolus organizer region) on the telomere of the long arm of chromosomes 4 and 7 was calculated as a percentage of the total length of the chromosome as described in Section 2.27. The physical mapping procedure was based on three points. First, the accurate measurement of the satellited chromosomes form root-tip metaphase preparations of pea (*Pisum sativum* L.) line 110 (standard for the normal structural type). Second, the satellite constrictions were not included in the total length of the chromosomes because they seem to contain no DNA and therefore, they are not involved in recombination processes. Third, the pea genetic map of Blixt (1974) shows that the total length of chromosomes 4 and 7 are 247 and 109 map
Table (14) shows that the long arm (the distance between the centromere and the satellite constriction) of chromosomes 4 and 7 constitutes 55.76 and 55.45% of the total length of the chromosome, respectively. When these percentages were transformed to map units, the satellite constrictions were found to be at 138 and 60 map units from the centromeres of the two chromosomes, respectively. The conclusion drawn from these results was that the rRNA gene sites are not located at random on the long arm, but the location is controlled by a relationship between the centromere and the telomere of the chromosome. Finally, it can be suggested that not only the centromere and the telomere of the short arm are the controlling factors of the rRNA gene location as has been established in several plant species but also the telomere of the long arm and the total length of the chromosome have equal effects.

4.4 Analysis of the genetic variation and mapping of biochemical marker genes

In this study three different approaches for analysing the genetic variation between genes were investigated. First, testing the variation between genes at the protein level through the analysis of banding patterns of isoenzymes or storage protein genes on PAGE. The variations detected were used to locate the controlling genes to positions onto the genome. Four different nuclear genes in the pea genome; Amy, Est, Got and Ve-1, encoding the amylase, esterase, glutamate oxaloacetate transaminase isoenzymes and vicilin storage protein, respectively, were investigated to see whether the phenotypic
variations observed between different lines of peas are corresponding to genotypic variations at the controlling gene loci. Analysis of the data obtained from different crosses between different pea lines has shown that these genes were inherited additively and their structural genes have been located to positions on chromosomes 2, 4, 1 and 7 respectively (see Fig. 14). However, some proteins are very complicated e.g. pea storage proteins legumin and vicilin have several subunits of different molecular weights and the variation between these subunits needs screening of a very large number of lines. Second, testing the variation at the DNA level through the analysis of the restriction fragment length polymorphisms was used to investigate the inheritance and chromosomal locations of rRNA genes in Pisum sativum L. The results showed that rRNA genes are under simple Mendelian control, but mapping the genes to positions onto the chromosomes was difficult because of a lack of morphological markers. However, this technique seems to be very useful for investigating genetic variations in a wide range of genes in different plant species. Third, DNA sequencing of the genes seems to be the most powerful tool for future investigations. From the nucleotide sequence, a complete picture of the gene can be drawn and variation between genes can be detected at the level of single base changes which is essential for engineering the genes. In this study a legumin gene (legB) was completely sequenced and compared with other members of the legumin gene family (see Fig. 8 and 15).

Mapping of biochemical marker genes onto plant genomes can be of benefit to plant breeding and molecular biology e.g. i) Marker genes, when mapped to positions onto the chromosome, can be used in further constructions of the genetic map of plant species and as chromosome
markers in various molecular and cytological studies as previously discussed (Section 1.2). ii) Marker genes which are tightly linked to desirable genes may be discovered through constructing the linkage map of a given species. Such marker genes could then be used to easily detect the desirable genes in segregating populations. In tomato an acid phosphatase isoenzyme which is tightly linked to a gene for nematode resistance has been discovered, allowing selection for nematode resistance to be done at an earlier stage and more efficiently than can be done by exposure to the parasite. This linkage relationship is now being extensively used to transfer nematode resistance from one tomato cultivar to another (Tanksley and Rick, 1980). In barley, the close linkage of Hor1 and Hor2 (the genes for hordein storage protein) with MLa and MLk (the mildew resistance loci) was detected by Jensen et al. (1980) and Shewry et al. (1980), allowing the protein loci to be used as markers in screening barley plants for the disease resistance. iii) Biochemical marker genes which are closely linked to physiological or morphological genes can be used in the isolation of these genes by a partial digestion of the DNA and isolation of the overlapping fragments along the chromosome. For example, the close linkage between the gene Ve-1 for vicilin 50,000 Mr subunit and the gene R/r for round/wrinkled seed surface in the pea genome, may enable the isolation of the gene r itself which is a pleiotropic gene i.e. has an effect on various characteristics in pea plants especially storage protein content (Davies, 1980; Schroeder, 1982; Schroeder and Brown 1984).
REFERENCES


Brewbaker J.L., Upadhyya M.D., Makinen Y. and MacDonald T. (1968) Isoenzyme polymorphism in flowering plants III. Gel electrophoresis methods and applications. Physiol Plant. 21, 930-940.


Casey R. (1979a) Immuonaffinity chromatography as a means of purifying legumin from Pisum (pea) seeds. Biochem J. 177, 509-520.


Martini G. and Flavell R. (1985) The control of nucleolus volume in wheat, a genetic study at three developmental stages Heredity 54(1), 111-120.


Rudin D. (1975) Inheritance of glutamate oxaloacetate transaminase (GOT) from needles and endosperm of Pinus sylvestris. Hereditas 80; 296-300.


Vasil V. and Hildebrandt A.C. (1965a) Growth and tissue formation from single isolated cells in microculture. Science 147, 1454-1455.


