# MAPPING GENES ESSENTIAL FOR EMBRYO DEVELOPMENT IN ARABIDOPSIS THALIANA

By

DAVID ANDREW PATTON

Bachelor of Science in Arts and Sciences
Oklahoma State University
Stillwater, Oklahoma
1984

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1986

Submitted to the Faculty of
the Graduate College of
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 1991

TABLE OF CONTENTS


Chapter — Page

I. LITERATURE REVIEW

The Study of Plant Embryo Development....1
Variant Cell Lines of Carrot Defective in Somatic Embryogenesis....2
Defective-Kernel Mutants of Maize....4
Embryo-Lethal Mutants of *Arabidopsis*....9

II. MAPPING EMBRYO-LETHAL MUTATIONS WITH VISIBLE VISIBLE MARKERS AND TELOTRISOMICS....23

Introduction....23
*Arabidopsis* as a Model System for the Study of Plant Embryo Development....23
General Features....23
Genetic Features....27
Molecular Characteristics....28
Why Map Embryo-Lethal Mutants of *Arabidopsis*?...30
Materials and Methods....34
Growth of Plant Material....34
Isolation of Embryo-Lethal Mutants....34
The *bio1* Mutant....36
Mutant *emb22*....38
Mutant *emb30*....39
Wild-Type Ecotypes....39
Landsberg "Erecta" Ecotype....39
Columbia Ecotype....40
Multiple Marker Lines....40
General Information on Multiple Markers....40
W2 Multiple Marker....41
W4 Multiple Marker....44
W8 Multiple Marker....48
W10 Multiple Marker....48
W11 Multiple Marker....51
W13 Multiple Marker....52
Single Marker Lines....54
The *dis2* Mutation....54
The *clv2* Mutation....54
Telotrisomic Marker Lines....56
General Information....56
Telotrisomic Trla....56
Telotrisomic Trlb....58

Chapter Page


Linkage Analysis....58
Mapping With Visible Markers....58
Backcrosses....62
Constructing the Genetic Map....64
Telotrisomic Analysis....66
Results....68
Linkage Analysis....68
Backcross Analysis....75
Telotrisomic Analysis....78
Discussion....82

III. CONSTRUCTION OF NEW MULTIPLE MARKER LINES TO FACILITATE MAPPING OF EMBRYONIC LETHALS....85

Introduction....85
Materials and Methods....87
Plant Material....87
Multiple Marker Construction....90
The DP23 Multiple Marker Line....90
The DP28 Multiple Marker Line....92
The DP27 Multiple Marker Line....92
Results and Discussion....93

IV. MAPPING WITH RFLP MARKERS....97

Introduction....97
Materials and Methods....103
Results and Discussion....107
High-Resolution RFLP Mapping Strategy....107
RFLP Mapping of the EMB30 Locus....109
RFLP Mapping of the BIO1 Locus....111
Concluding Remarks....116

LITERATURE CITED....118

APPENDICES....131

APPENDIX A - PRINTOUT OF F2 SEGREGATION DATA FROM THE BUILD AND SUMMARY PROGRAMS....131
APPENDIX B - EXAMPLE OF HOW TO USE THE CHI PROGRAM TO DETECT LINKAGE....138
APPENDIX C - USING THE RECF2 COMPUTER PROGRAM....143
APPENDIX D - HOW TO USE GENMAP....150
APPENDIX E - USING CHROMAP AND PRINTAPLOT COMPUTER PROGRAMS TO PRINT THE LINKAGE MAP....155
APPENDIX F - CONVERSION OF RECOMBINATION FREQUENCY (P) INTO CENTIMORGANS USING THE KOSAMBI (1944) MAPPING FUNCTION....161
APPENDIX G - DERIVATION OF FORMULAE USED TO CALCULATE MINIMUM NUMBER OF F2 PLANTS REQUIRED TO DETECT LINKAGE....163

Chapter Page

APPENDIX H - CHI-SQUARE CALCULATIONS FOR THE DISTRIBUTION OF F2 PROGENY IN CLASSES A-D FOR THE CROSS AN X EMB22.....166
APPENDIX I - CHI-SQUARE CALCULATION FOR THE DISTRIBUTION OF F2 PROGENY IN CLASSES E AND F FOR THE CROSS CHL1 X EMB22 .....168
APPENDIX J - CHI-SQUARE CALCULATION FOR THE DISTRIBUTION OF F2 PROGENY IN CLASSES E AND F FOR THE CROSS TTG X BIO1.....170
APPENDIX K - CHI-SQUARE CALCULATION FOR THE DISTRIBUTION OF F2 PROGENY IN CLASSES I AND II FOR THE BACKCROSS W2 X EMB30.....172
APPENDIX L - MOLECULAR BIOLOGY PROTOCOLS USED IN RFLP MAPPING.....174

LIST OF TABLES

Table Page

1. OVERVIEW OF VISIBLE MARKERS USED IN THIS STUDY.....43

2. EXPECTED FREQUENCY OF $F_2$ PLANTS OBTAINED FOLLOWING CROSSES WITH VISIBLE MARKERS.....69

3. MINIMUM NUMBER OF $F_2$ PLANTS REQUIRED FOR LINKAGE DETECTION.....70

4. RECOMBINATION ESTIMATES BETWEEN EMBRYONIC LETHALS AND VISIBLE MARKERS.....73

5. RESULTS FROM BACKCROSSES BETWEEN emb30-1 AND LINKED VISIBLE MARKERS.....77

6. LINKAGE DETECTION WITH TELOTRISOMICS.....80

7. EXPECTED FREQUENCIES OF $F_2$ PLANTS WITH DISTINCT RFLP PATTERNS OBTAINED FOLLOWING CROSSES WITH LANDSBERG OR NIEDERZENZ ECOTYPES.....108

LIST OF FIGURES


Figure — Page

1. Phenotypes of Defective Kernel Mutants of Maize Impaired in the Formation of Shoot Primordia ..... 10

2. Lethal Phase of Embryos Found in Aborted Seeds From 60 Lethal Mutants of Arabidopsis thaliana ..... 13

3. Lethal Phase of Arrested Embryos Found in Aborted Seeds From Six Embryo-Lethal Mutants of Arabidopsis thaliana ..... 15

4. Illustration of a Silique From a Plant That is Heterozygous for an Embryo-Lethal Mutation ..... 16

5. Lethal Phases of 32 Embryo-Lethal Mutants of Arabidopsis thaliana ..... 18

6. Drawing of Arabidopsis thaliana (L.) Heynh ..... 24

7. Stages of Normal Embryo Development in Arabidopsis thaliana ..... 25

8. Stages of Early Embryo Development in Capsella bursa-pastoris ..... 26

9. A General Strategy for Producing Genetic Mosaics ..... 31

10. Schematic Drawing of a Chimeric M1 plant ..... 35

11. Embryos From Selected Lethal Mutants of Arabidopsis ..... 37

12. Linkage Map of Arabidopsis Showing Locations of Visible Markers Used in This Study ..... 42

13. Photograph of Rossete Leaves Showing Wild-Type Trichomes ..... 45

14. Photograph of Rossete Leaves From dis1 Plant Showing Distorted Trichomes ..... 46

15. Comparison of Stems From Wild-Type (top) and cer2 (bottom) Plants ..... 49

Figure Page

16. Comparison of Seeds From Wild-Type (left) and tt3 (right) Plants....53

17. Comparison of Siliques From clv2 (top) and Wild-Type (bottom) Plants....55

18. Comparison of 4 Week Old Tr1a (left) and Wild-Type (right) Plants....57

19. Comparison of 4 Week Old Tr1b (left) and Wild-Type (right) Plants....59

20. Drawing of Strategy Used to Map Embryonic Lethals With Visible Markers....60

21. Drawing of Strategy Used for Backcrosses....63

22. Drawing of Strategy Used for Mapping Embryonic Lethals With Telotrisomics....67

23. Linkage Map of Arabidopsis thaliana With Embryonic Lethals in Bold....76

24. Linkage Map With Visible Markers Used in Multple Marker Construction....88

25. Drawing of Strategy Used to Create Multiple Markers From Existing Visible Markers....89

26. Strategy Used to Isolate cer2 as a Monogenic Line From W10....91

27. Schematic Drawing That Illustrates the Theory Behind RFLP Mapping....100

28. Drawing of Strategy Used in High-Resolution RFLP Mapping of Embryonic Lethals....104

29. RFLP Linkage Map of Arabidopsis thaliana....106

30. Southern Blot of DNA From 4 Pools of biol x Niederzenz F2 Plants Probed With λbAt331....112

31. Southern Blot of DNA From Genetic Reconstruction and 5 pools of biol x Landsberg F2 Plants Probed With λbAt558....113

32. Southern Blot of DNA From 5 Pools of biol x Landsberg F2 Plants Probed With λbAt233....114

## LIST OF SYMBOLS

| | |
|---|---|
| °C | degrees Celsius |
| cM | centi Morgans, a genetic unit defined as number of crossovers per 100 progeny |
| cm | centimeter |
| EMS | ethyl methanesulfonate, a potent alkylating agent and mutagen |
| $F_1$ | first generation following a cross |
| $F_2$ | second generation following a cross |
| h | hour |
| HSP | heat shock protein |
| IAA | indole-3-acetic acid |
| Kb | Kilobases |
| $M_1$ | mutagenized individual, first generation |
| $M_2$ | second generation produced following mutagenesis |
| ml | milliliter |
| mm | millimeter |
| NAA | naphthaleneacetic acid |
| RFLP | restriction fragment length polymorphism |
| SDS-PAGE | sodium dodecylsulfate - polyacrylamide gel electrophoresis |
| T-DNA | transfer DNA, fragment of DNA tranferred during Agrobacterium infection |
| µl | microliter |
| µm | micrometer |
| YAC | yeast artificial chromosome |

# CHAPTER I

## LITERATURE REVIEW

### The Study of Plant Embryo Development

Embryogenesis in plants is a relatively simple morphological process. However, at the molecular level a complex series of interactions regulate the temporal and spatial expression of genes which ultimately control development of the embryo. Plant hormones and other regulatory molecules must act in concert to control development in a predictable and reproducible way. Much of what is known about these processes comes from descriptive, experimental, and biochemical studies (Maheshwari 1950; Raghavan 1976; Bewley and Black 1985). These studies have addressed many features of embryogenesis, ranging from the structure and organization of the egg, zygote, and embryo, to hormonal, nutritional, and metabolic aspects that play a key role during seed development.

Regulation of embryogenesis has also been studied at the molecular level (Higgins 1984; Dure 1985; Goldberg et al. 1989). These studies suggest that many of the important events that affect both morphogenesis and differentiation occur during the early stages of embryo development and that as many as 20,000 genes may be expressed during plant embryogenesis.

Genetic analysis of embryo development in plants has been

facilitated by the isolation of mutants defective in embryogenesis. The best examples of this type of mutant include: cell culture lines of carrot which fail to complete somatic embryogenesis *in vitro*, defective kernel mutants of maize, and embryo-lethal mutants of *Arabidopsis*.

## Variant Cell Lines of Carrot Defective in Somatic Embryogenesis

A number of different plant species can produce somatic embryos when grown in liquid suspension culture (Ammirato 1983). Many of these somatic embryos can then be transferred to a solid medium where they can complete development and form normal plants. Somatic embryogenesis appears to mimic zygotic embryogenesis even in the absence of endosperm and surrounding maternal tissue. One genetic approach to understanding somatic embryogenesis in carrot has been to isolate cell lines that do not form somatic embryos at a high temperature (Breton and Sung 1982; Giuliano et al. 1984; Schnall et al. 1988). The aim in these studies has been to identify genes or gene products that play a key role in embryogenesis.

The variant cell lines described by Breton and Sung (1982) were isolated from haploid cell cultures that were not treated with mutagens. These temperature-sensitive lines were identified by their inability to produce embryos at the restrictive temperature (32° C) while retaining the ability to produce somatic embryos at the permissive temperature (24° C). Three types of variant cell lines were isolated using this approach; the cell lines either stopped growing (ts-growth), formed callus only, or formed a mixture of callus and somatic embryos that became arrested during embryo development (ts-emb$^-$). This final class of variants was a group of mutants potentially defective in processes

specific to embryo development. Further characterization of one mutant line (ts59) has shown that the altered gene product is required twice during development: 5 days after the initiation of embryogenesis and during the transition from the heart to torpedo stage of development (Lo Schiavo et al. 1988). Further analysis of this mutant, using two-dimensional gel electrophoresis, showed that protein patterns were significantly different from wild-type. The main difference in ts59 appeared to be due to a lack of secondary modification in one subclass of heat-shock proteins (HSPs). It appears in this case that a developmental mutant isolated at high temperature was defective in a gene that does not play a direct role in embryogenesis, but rather is important for protecting the cell from damage due to high temperature.

Variant cell lines with a temperature-sensitive phenotype are potentially useful for identifying genes that play a critical role in embryogenesis. The main drawback to this group of mutants is that they have been induced in a haploid cell line that cannot be regenerated to give fertile plants. This limitation makes it difficult to perform many genetic manipulations that are commonly used to characterize developmental mutants.

A second group of ts-growth and ts-emb$^-$ variants of carrot has been described by Giuliano et al. (1984). In this study, diploid suspension culture cells were mutagenized with EMS, transferred to fresh embryogenic media, and then tested for their ability to produce somatic embryos at 24° and 31° C. One line, ts2, consistently lacked the ability to produce somatic embryos at 31° C. This phenotype was maintained in cell lines started from variant ts2 plants that were regenerated at the permissive temperature. Unfortunately, genetic

studies on the inheritance pattern and developmental studies describing zygotic embryo development in regenerated ts2 plants have not been reported.

Schnall et al. (1988) described the isolation and initial characterization of 21 new variant cell lines that were defective in somatic embryogenesis at 33° C. These variants were somaclonal in nature and isolated from a predominantly diploid cell culture. Six classes of temperature-sensitive variants were identified. Three of the variant cell lines were classified as ts-emb$^-$. Based on temperature-shift experiments, one of the ts-emb$^-$ variants was found to be defective in a factor that was constitutively required for somatic embryogenesis. Two other mutants appeared to blocked at the globular stage of development.

Defective-Kernel Mutants of Maize

The earliest reports of genetic factors causing defective kernels in maize were made by Jones (1920), Demerec (1923), Mangelsdorf (1923, 1926), Brink (1927), Wentz (1930), and Emerson (1932). These recessive lethal factors were first identified in corn improvement programs where part of the improvement strategy included inbreeding of heterozygous lines. The inbreeding of heterozygotes in this case uncovered recessive traits that had a negative effect on seed development. Several different types of defective kernels were described in these early works. The major differences in the phenotypes were the amount of endosperm and the size of the embryo present within the seed. The phenotypes that resulted were given a number of names such as zygotic lethal, germless, and defective seed.

Of these earliest reports, Mangelsdorf (1926) described the most comprehensive group of mutants with defective kernels. Fourteen different defective seed (de) lines were identified following selfing, because they produced ears with a high percentage of aborted seeds. All of these mutations segregated as single Mendelian recessive factors. Complementation crosses were made among these 14 mutants and only two mutants (de-5 and de-11) appeared to be defective in the same gene. Mangelsdorf estimated the stage of developmental arrest by expressing the weight of defective kernels as a percentage of the normal kernel weight. Defective kernels had weights that ranged from 2% (de-14) to 49% (de-2) of the normal kernel weight. This method of estimating the stage of developmental arrest from average kernel weight was generally good for determining the extent of endosperm and embryo development. However, it was not possible in every case to accurately determine the stage of the arrested embryo by simply weighing the seed. For example, de-4 seeds, which usually lacked a visible embryo, were often heavier than seeds from other mutants that did contain a visible embryo. The mutant gene in most of these de lines did not appear to be expressed exclusively in the seed because homozygous mutant plants grown from defective kernels were usually abnormal and died after only a few weeks of growth.

Defective kernel mutants of maize have subsequently been divided into three classes: (1) defective endosperm mutants that produce altered endosperm but have an embryo capable of growing into normal plants; (2) germless mutants with a defective embryo but a normal endosperm; and (3) defective seed mutants where development of both the embryo and the endosperm are defective. The third class of mutants was not always

lethal; in some cases mutant embryos of this type were capable of growing into normal plants.

Defective endosperm mutants include the miniature seed mutant (Lowe and Nelson 1946), mutant de-17 (Brink and Cooper 1947), several mutants described by Mangelsdorf (1926), the sugary, floury, shrunken, and waxy mutants described by Coe and Neuffer (1977) and a series of endosperm mutants which fail to accumulate dry matter (Manzocchi et al. 1980a, 1980b). Microscopic analysis of sections through defective seeds produced by many of these mutants has shown that the flow of nutrients from maternal tissues to the developing embryo and endosperm is often disturbed by changes in the chalazal region of the embryo sac and is probably the cause of reduced endosperm development (Lowe and Nelson 1946). In a more recent paper, Torti et al. (1984) describe a specific defect in de*-B18, one of the defective endosperm mutants originally described by Manzocchi et al. (1980a). This particular mutant accumulates 15-fold less of the endogenous auxin indole-acetic acid (IAA) in the endosperm of mature kernels relative to wild-type. Furthermore, exogenous applications of the synthetic auxin naphthalene-acetic acid (NAA) to developing seeds were found to normalize the weight of de*-B18 mutant kernels. These results suggest that the altered gene in de*-B18 is involved in IAA metabolism in the endosperm.

Germless mutants of maize produce kernels with a normal endosperm but no visible embryo a maturity. Germless mutants were initially described by Demerec (1923) and subsequently by Wentz (1930) and Sass and Sprague (1950). Some germless mutants were found that formed visible mutant embryos during the intermediate stages of seed development; these embryos then decayed during the final stages of seed

development and were occasionally not found in mature seeds (Sass and Sprague 1950). Relatively few of these defective kernel mutants were truly germless because most kernels contained arrested embryos.

An extensive collection of defective kernel mutants has been described more recently by Neuffer and Sheridan (1980). In this study, pollen was mutagenized with EMS and then used to pollinate 72 ears. A total of 3919 $M_1$ kernels were produced, 3461 of these grew into mature $M_1$ plants which were selfed to give 3172 ears with $M_2$ kernels. Screening these ears revealed that 855 $M_1$ plants segregated for recessive kernel mutants, 432 of these were classified as embryo-lethal because defective kernels from these mutants did not germinate. Neuffer and Sheridan used the term "defective kernel" in a generalized way to describe the 855 recessive kernel mutants that produced 25% defective kernels when selfed. A list of 26 descriptive terms, each with a two or three letter symbol, was presented in order to more accurately define each kernel phenotype. Neuffer and Sheridan (1980) also defined four types of defective kernel mutants: (1) those with defective endosperm and a non-viable embryo; (2) those with defective endosperm but a viable embryo that could germinate and produce an abnormal plant; (3) those with defective endosperm but a normal, viable embryo; and (4) those with normal endosperm and a defective embryo or "germless" kernels. The original 855 recessive kernel mutants included 432 type 1, 59 type 2, 147 type 3, and 3 type 4. The remaining 214 were not classified.

Type 1 mutants were of most interest because they produced kernels with embryos that were arrested between zygotic and mature stages of development. A subset of 194 type 1 mutants were chosen for further analysis. Mapping studies were performed using a series of B-A

translocations and the method developed by Roman and Ullstrup (1952). This technique was used to locate 89 mutations to chromosome arms. The mapping results indicated that these mutations were scattered randomly throughout the genome. Another set of experiments included the construction of discordant kernels to test whether mutant embryos could be rescued by normal endosperm. The results of these studies on embryo-endosperm interaction indicated that in most cases normal endosperm could not rescue mutant embryos. In the few cases where the embryo was rescued, the resulting kernels germinated but died soon after depletion of endosperm reserves. These results support the idea that the endosperm does play a nurturing role during the development of the embryo.

In a companion paper, Sheridan and Neuffer (1980) described the lethal phases of embryos in defective kernels and also characterized the response of mutant embryos in culture in an attempt to identify auxotrophic mutants. Most mutants had immature kernels that were lighter in color and smaller than adjacent normal kernels. In most cases the mutant embryos were no more than one-half to two-thirds the size of comparable normal, mature embryos. One-hundred and two mutants were examined in culture on basal and enriched media; 21 simply enlarged or failed to respond on either medium, while 81 produced roots and shoots on at least one medium. Most mutants responded in a similar way on the two different media, while 16 grew better on basal and 23 showed superior growth on enriched medium. Among this latter group, 10 were classified as potential auxotrophs, one of these was auxotrophic but turned out to be allelic with pro-1, a previously identified proline mutant described by Gavazzi et al. (1975).

In a subsequent study, Sheridan and Neuffer (1982) further characterized 14 additional defective kernel mutants that formed defective shoot primordia. Five mutants were blocked at the proembryo stage and were allelic, the remaining 9 mutants were blocked at later stages of development (Figure 1). Eleven mutants representing 7 distinct loci were analyzed in discordant kernels to test the relationship of the embryo and endosperm in determining the mutant phenotype. In all cases, the presence of normal endosperm failed to rescue the mutant embryo. Reciprocal experiments also revealed that mutant endosperm had little effect on normal embryos. These results indicate that the mutant phenotype in this group was determined largely by the genotype of the embryo.

Embryo-Lethal Mutants of Arabidopsis

The first comprehensive study of embryo-lethal mutants in Arabidopsis was reported by Müller (1963). Approximately 20,000 seeds from the geographical race "Dijon" were imbibed for 42 h and then treated with X-irradiation. These $M_1$ seeds were then grown in soil and the resulting plants were screened for the presence of abnormal embryos. Abnormal embryos were found in a total of 3080 $M_1$ plants. Dry $M_2$ seeds were saved from 875 of these $M_1$ plants and used in subsequent studies. Chlorophyll mutations were identified by segregation in 26 of these families. Seventy-two families with interesting phenotypes were examined more closely with respect to the phenotype of the aborted seeds and arrested embryos produced by heterozygous plants. A different group of 60 randomly-chosen families was used for determining the diversity of lethal phases and segregation patterns of the mutant alleles.

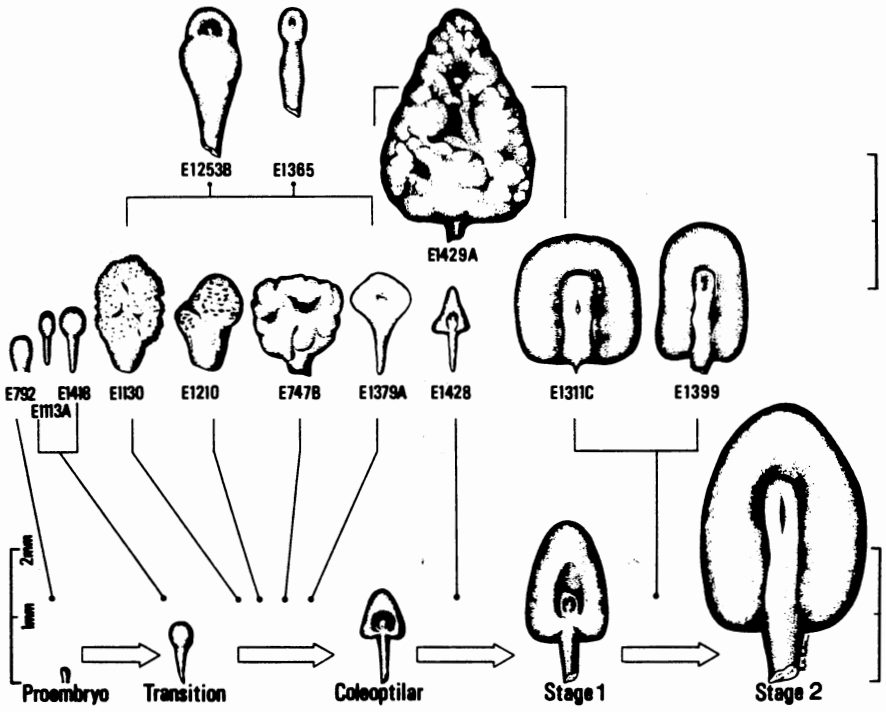



Figure 1. Phenotypes of Defective Kernel Mutants of Maize Impaired in the Formation of Shoot Primordia. Stages of normal development are shown across the bottom according to Abbe and Stein (1952). Arrested embryos from defective kernels are shown above the corresponding normal stage of development. Figure reprinted from Sheridan and Neuffer (1982).

The following nomenclature was devised to more accurately describe the different classes of seeds found in heterozygous fruits (siliques). Class I, the embryo fails to develop beyond the heart stage. Four types of class I phenotypes were defined: sicca, brevis, vana, and diffusa. A more complete description of these 4 types of class I mutants was as follows: sicca, maximal seed length of 0.3 mm, aborted seeds turn brown while normal seeds in the same silique are at the linear to curled cotyledon stage; brevis, maximal seed length is 0.4 mm, aborted seeds turn brown at the time when normal seeds have mature green embryos; vana, maximal seed length is 0.5 mm, aborted seeds remain colorless, then turn brown when normal seeds have mature green embryos; diffusa, like vana except that aborted seeds are pale green, then turn brown when normal seeds have mature green embryos.

Class II, the embryo clearly has differentiated hypocotyl and cotyledons, but does not reach normal size. Seeds reach normal length of 0.5 mm. Two types of class II mutants were defined; murca, with green embryos, and parva with embryos that are not green.

Class III, the embryo is normal in size and structure. This class has 4 types: fusca, albina, xantha, and chlorina. Fusca mutant embryos have dark, wine-red regions of variable size, usually in the cotyledons. Albina embryos are white, even after the removal of the seed coat. Xantha embryos are cream colored or yellow, and turn light or bright yellow after removal of the seed coat. Chlorina embryos are light greenish-yellow, and turn intensely greenish-yellow after removal of the seed coat.

Most of the $M_1$ families examined by Müller segregated in subsequent generations for a single recessive lethal factor. Four families were

found that segregated for two independent and phenotypically distinguishable lethal factors in subsequent generations. The segregation ratios and lethal phases of a group of 60 randomly-chosen mutants were also determined. The lethal phases of this group are summarized in Figure 2. A relatively large number of mutants were blocked at the transition between the globular and heart stages of development. The segregation ratio in heterozygous plants of each mutant was compared with the expected 25% aborted seeds using the chi-square test. In the random sample, 29 mutants had segregation ratios that were not significantly different from 25%, 27 mutants had slightly reduced segregation ratios with a median ratio of 21.1%, and 4 mutants had severely reduced segregation ratios (less than 10% aborted seeds). Müller concluded that approximately 50% of the mutations being studied had a somewhat negative effect on the germination or growth of pollen tubes.

One mutant, brevis 1420 showed a non-random distribution of aborted seeds within heterozygous siliques. A total of 51 siliques from three heterozygous plants were examined with respect to the relative position of normal and aborted seeds in the siliques. These siliques contained a total of 3058 seeds, approximately 85% normal seeds and 15% aborted seeds. Over 95% of the aborted seeds were located in the distal-half of the siliques. This observation led to the conclusion that the brevis 1420 mutant allele inhibits aspects of gamete competition that take place during self-fertilization.

The work described by Müller (1963) was pioneering in several respects. He devised a system that could be used for testing the mutagenic potential of various treatments with chemicals or irradiation.

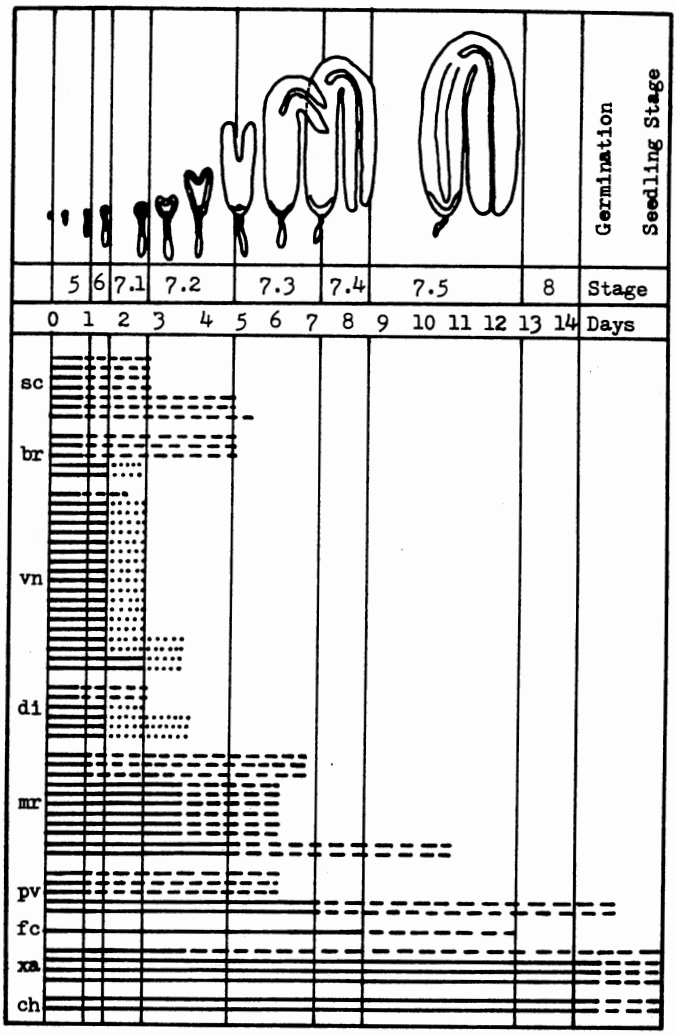



Figure 2. Lethal Phase of Embryos Found in Aborted Seeds From 60 Lethal Mutants of Arabidopsis thaliana. Stages of normal development are shown across the top and the mutant classes are listed along the left side. The solid horizontal bars represent the extent of normal development for each mutant. The dashed lines following the bars represent the lethal phase of each mutant. Figure reprinted from Müller (1963).

This work also served as an inspiration to others to pursue Arabidopsis as a model system for the genetic analysis of plant embryo development (Meinke and Sussex 1979a).

Meinke and Sussex (1979b) described the isolation and characterization of six new EMS-induced embryo-lethal mutants of Arabidopsis with early lethal phases (see Figure 3). These mutants were non-allelic and were shown to segregate as single Mendelian recessive traits at 18°, 25° and 32° C. Meinke (1982) subsequently demonstrated that two of these mutants showed a non-random distribution of aborted seeds in siliques produced by heterozygous plants; which was evidence for gametophytic expression of the mutant gene. Figure 4 illustrates the idea of a non-random distribution of aborted seeds within a heterozygous silique. These data support the idea that there is overlap between sporophytic and gametophytic gene expression. Another of the six mutants originally described by Meinke and Sussex (1979b) had a particularly interesting phenotype: the embryo proper arrested at an early globular stage, but embryonic arrest was followed by abnormal growth of the suspensor. Serial sections through aborted seeds from this mutant, called 50B, demonstrated that the suspensor in this mutant contained between 15 and 150 cells (Marsden and Meinke 1985), while the suspensor of normal embryos is made up of only 6 to 8 cells. Two of the six original mutants described by Meinke and Sussex (1979b), 50B and 71E, have not been maintained because seed stocks lost viability.

Since 1980, over 90 additional embryo-lethal (or embryo-defective) mutants have been isolated and characterized by the Meinke laboratory in an attempt to further our understanding of the genetic control of plant embryo development. A number of different approaches have been taken in

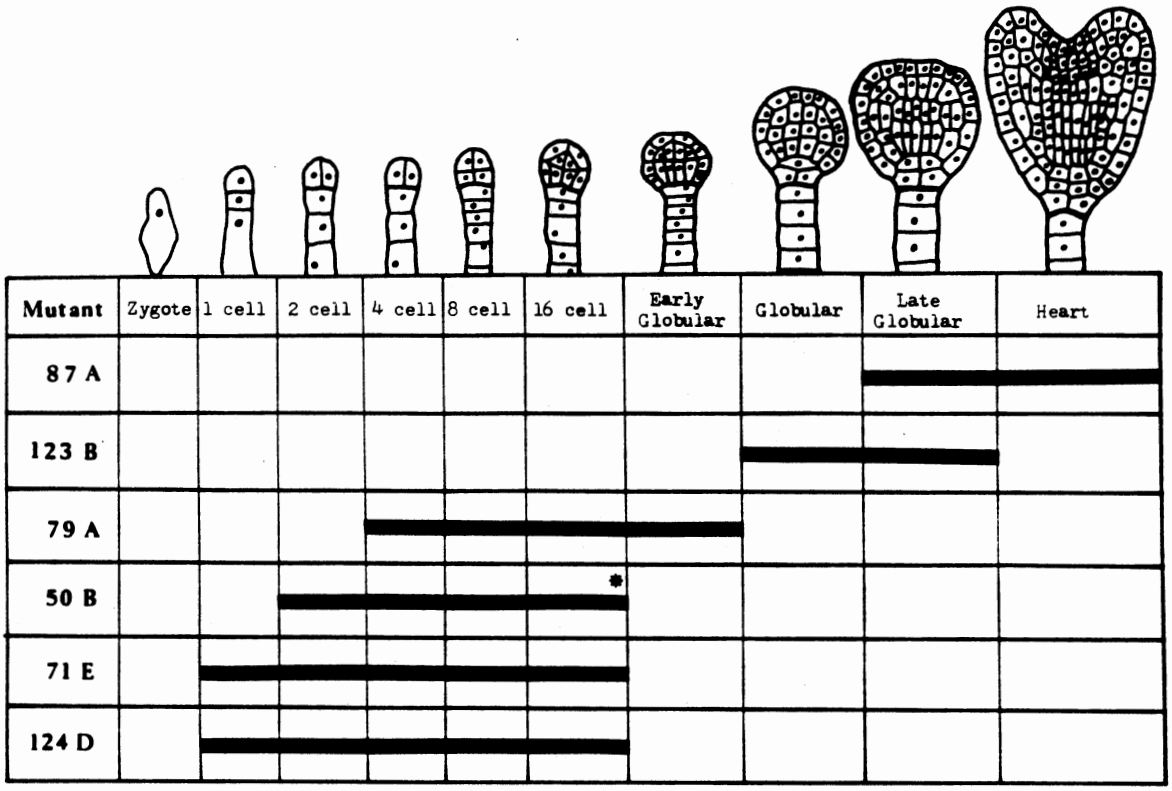



Figure 3. Lethal Phase of Arrested Embryos Found in Aborted Seeds From Six Lethal Mutants of Arabidopsis thaliana. Stages of normal development are shown across the top. The stage of developmental arrest for each mutant is represented by the horizontal black bar. In mutant 50B (*), developmental arrest of the embryo proper is followed abnormal growth of the suspensor. Figure reprinted from Meinke (1982).

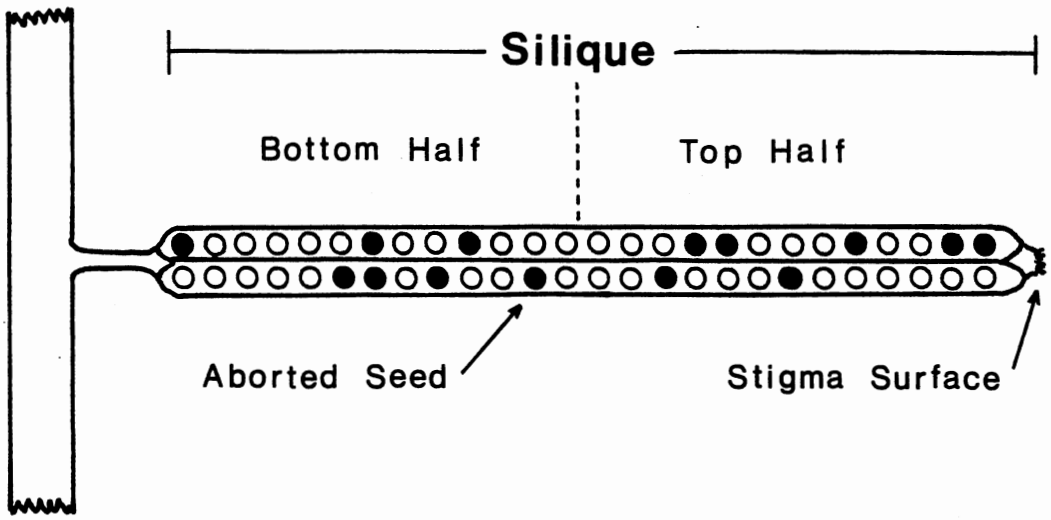



Figure 4. Illustration of a Silique From a Plant That is Heterozygous for an Embryo-Lethal Mutation. The open circles represent normal seeds and the colored circles represent aborted seeds. The top and bottom halves of this silique contain equal numbers of aborted seeds, while the top and bottom halves of siliques from some mutants contain unequal numbers of aborted seeds. Reprinted from Meinke (1982).

the analysis of these mutants. Initial studies summarized the segregation ratios and lethal phases of 32 EMS-induced embryo-lethal mutants (Meinke 1985). Figure 5 demonstrates the range of lethal phases obtained in this group. Further studies with 17 of these mutants showed that arrested embryos could be grown in culture (Baus et al. 1986). Two mutants showed particularly interesting responses; one mutant (112A-2A) formed rootless plants in culture and another mutant (122G-E) appeared to be auxotrophic when tested on basal and enriched media. The latter mutant was subsequently shown to require biotin for normal growth (Schneider et al. 1989). Shellhammer and Meinke (1990) then showed that arrested embryos from heterozygous 122G-E (biol) plants contained 5-fold less biotin than corresponding wild-type embryos. Recent work on biol has dealt with determining the biochemical defect in this mutant.

Using SDS-PAGE of total seed protein, Heath et al. (1986) examined the levels of storage proteins present in normal and aborted seeds. The aim of that study was to find a mutant that blocked at an early stage of development but continued cellular maturation and accumulated seed storage proteins. These studies showed that levels of two major seed storage proteins were reduced in aborted seeds from eight of the nine mutants tested. Dry seeds from homozygous mutant 130B-A-2 plants accumulated nearly normal levels of storage proteins, while green seeds from this mutant had reduced levels of storage proteins. In a related study, Patton and Meinke (1990) described the ultrastructure of embryos from normal and aborted seeds from four mutants. The goal was to determine the degree of cellular differentiation in arrested embryos by using protein bodies as developmental markers. Mature protein bodies accumulated in the hypocotyl and cotyledons of normal embryos only

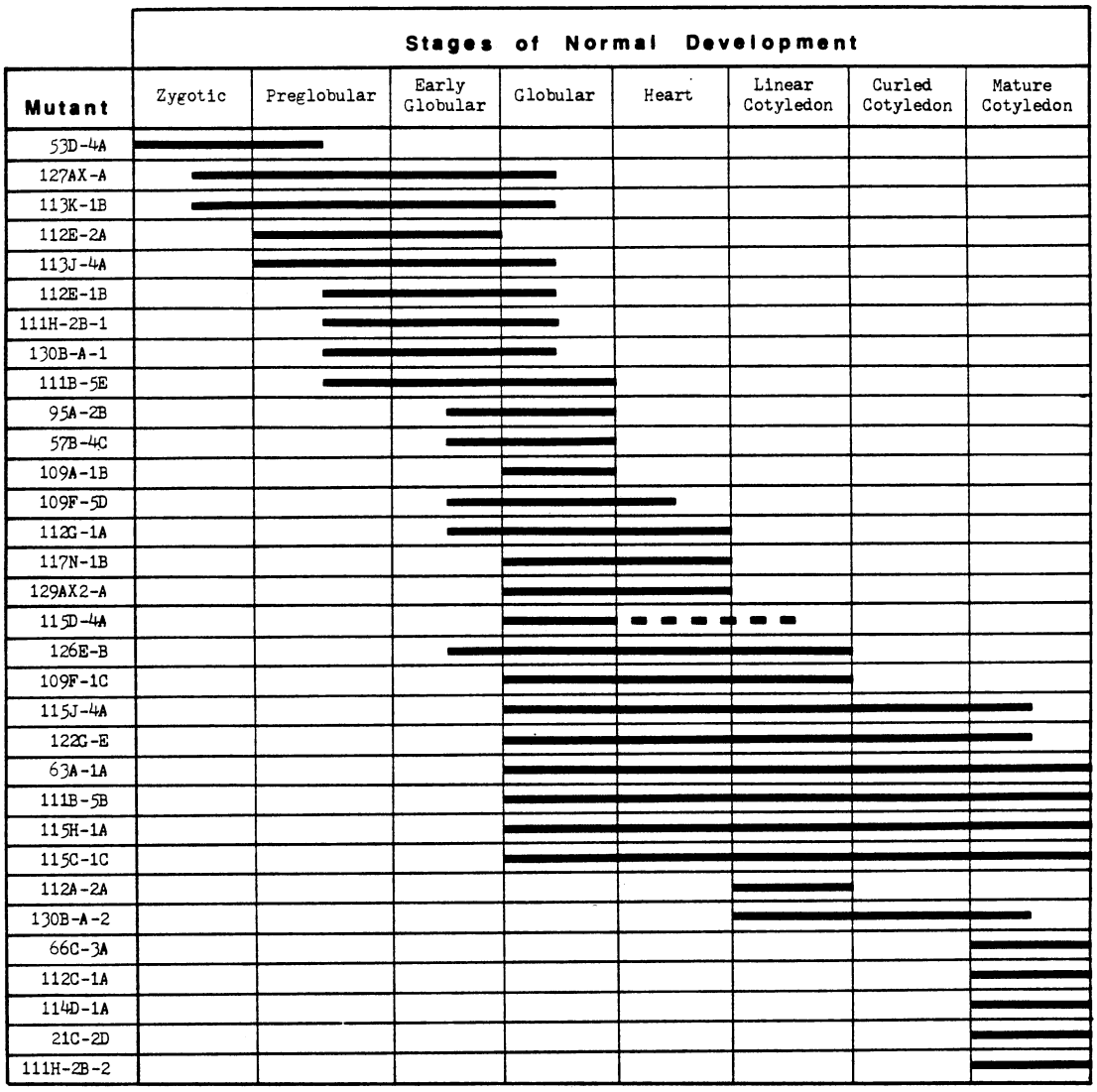



Figure 5. Lethal Phases of 32 Embryo-Lethal Mutants of Arabidopsis thaliana. Normal stages of development are listed across the top. The solid black bar represents the lethal phase of each mutant listed. The dashed line for mutant 115D-4A approximates the extent of the lethal phase. Figure reprinted from Meinke (1985).

during the final stages of development. One mutant (emb22) consistently lacked mature protein bodies, while another mutant (emb30) accumulated mature protein bodies in the hypocotyl and cotyledons. Other mutants showed variable patterns of protein accumulation; the albino mutant 114D-4A had protein bodies that were only partially filled, while emb31 had mature protein bodies in the hypocotyl and immature protein bodies in the cotyledons. The pattern of maturation in emb31 was consistent with the phenotype of this mutant because the cotyledons developed more slowly than the hypocotyl.

Additional embryo-lethal mutants with a wide range of lethal phases have now been tested for their response in culture (Franzmann et al. 1989). Conclusions from this study were that mutants with an early lethal phase failed to turn green in culture and produced only limited callus; mutants with a later lethal phase responded better in culture, many of these mutants produced abnormal leaves, rosettes, and flowers. Two mutants formed phenotypically normal homozygous mutant plants with the exception of producing 100% aborted seeds.

Over 50 new mutants have been recently isolated following EMS seed mutagenesis, X-irradiation, and Agrobacterium-mediated insertional mutagenesis (Meinke et al. 1989; Feldmann et al. 1990). One recent study with some of the T-DNA tagged lines describes the patterns of abnormal development found in some these embryo-lethal mutants (Errampalli et al. in press). This paper also describes an extensive study on the cosegregation of T-DNA markers (kanamycin resistance; nopaline synthesis) and the mutant phenotype in attempt to resolve whether the mutant genes are truly tagged with a T-DNA element. The T-DNA tagged mutants have been examined in preparation for molecular

isolation of the defective genes in these embryonic mutants.

Most of the data presented in Chapter 2 and Chapter 4 of this dissertation were included in a recently submitted manuscript that described the mapping of embryo-lethal mutations with visible markers, telotrisomics, and molecular (RFLP) markers (Patton et al. submitted). The results presented in the mapping manuscript included the addition of 16 new genes to the linkage map. Additionally, a high-resolution RFLP mapping strategy was presented and used to map the bio1 gene to within 0.5 cM of an RFLP marker. My role in this study was to provide our laboratory the methods used to map embryonic lethals, and test these methods on the first three mutations (bio1, emb22, and emb30-1). These procedures were based on the methods originally described by Servítovà and Cetl (1984). I first mapped these three genes with visible markers and telotrisomics, then started working with RFLP markers. Other members of the laboratory (Linda Franzmann, with the help of Karl Hansen and Leigh Mickelson) subsequently mapped an additional 13 genes with visible markers using procedures that I refined during my first year of study towards the Ph.D. degree. All of the mapping data from our laboratory were then combined and used to construct an updated linkage map of Arabidopsis which was presented in the mapping manuscript described above.

Lethal mutants of Arabidopsis have been used to address key questions concerning the genetic, biochemical, and physiological basis of embryo development. A number of diverse approaches have been used to document that many genes expressed during embryo development also play a role during vegetative growth. Many of these genes are likely to code for factors involved in basic cellular metabolism, like bio1, while

others may play a more critical role in the control of early development. The challenge in the future is to determine the function of these genes and how they interact to control a relatively simple developmental process. This is an exciting time because we are at the threshold of being able to meet this challenge using Arabidopsis and gene isolation techniques that include T-DNA insertional mutagenesis, RFLP mapping and chromosome walking.

Results presented in this dissertation represent the first efforts to define the chromosomal location of three mutant genes (bio1, emb22, and emb30-1) being studied in our laboratory. Three different methods were explored for mapping embryo-lethal mutations: mapping with visible markers, telotrisomic analysis, and mapping with RFLP markers. Different visible markers were used to map these three embryo-lethal mutations of Arabidopsis. Backcrosses were used to demonstrate the correct position of emb30 relative to flanking visible markers. Telotrisomic analysis was also explored as a method for mapping embryo-lethal mutations to chromosome arm by screening for a distortion in the expected percentage of aborted seeds produced by telotrisomic $F_1$ plants. Finally, a method of high-resolution RFLP mapping of embryonic lethals was presented and then tested with the bio1 gene in preparation for gene isolation through chromosome walking.

Mapping of embryonic lethals will clearly complement the current analysis of the Arabidopsis genome. These studies will not only help to define the number of complementation groups that are represented among these mutants, but should also facilitate the analysis of the Arabidopsis genome by defining new genes. The addition of embryonic lethals to the standard linkage map will increase the number of useful

genetic markers, improve the resolution of the map, and be a first step toward the construction of balanced-lethal chromosomes. The high-resolution RFLP mapping strategy should facilitate molecular isolation of genes through chromosome walking and also help to integrate the standard linkage and RFLP maps.

# CHAPTER II

# MAPPING EMBRYO-LETHAL MUTATIONS WITH VISIBLE MARKERS AND TELOTRISOMICS

## Introduction

### Arabidopsis as a Model System for Studying Plant Embryo Development

General Features. Arabidopsis thaliana (L.) Heynh. (Figure 6) is an inconspicuous member of the brassica family that grows well in sandy soils in many parts of the temperate world. Arabidopsis plants are easily grown in the laboratory or in a greenhouse with minimal maintenance. Mature plants have a fibrous root system, a compact basal rosette, and main and minor stems with indeterminant inflorescences. Arabidopsis usually self pollinates, but may be crossed in the laboratory with the aid of a dissecting microscope and fine-tipped forceps. The fruits, called siliques, have 2 locules, each containing up to 30 seeds at maturity. The typical life cycle of Arabidopsis ranges from 4 to 6 weeks.

Embryo development in Arabidopsis follows a predictable pattern of development (Figure 7) that takes from 10 to 14 days depending upon environmental conditions. Early embryonic development begins with the zygote, which divides via mitosis giving rise to the two-celled embryo. One of these cells gives rise to the embryo proper while the other gives rise to the suspensor as outlined in Figure 8. The embryo turns green


Figure 6. Drawing of Arabidopsis thaliana (L.) Heynh. (A) Mature plant showing rosette leaves, main stem and lateral braches with terminal inflorescence; (B) Epidermal hairs (trichomes) on the surface of rosette leaves; (C) Flower at the time of pollination; (D) Floral petal (E,F) Flower with sepals and petals removed, showing the position of the stamens; (G) Mature pistil, shown with sepals, petals and stamens removed; (H) Mature silique with valves split to reveal two rows of seeds; (I) Mature seed. Drawing from Ross-Craig (1948).

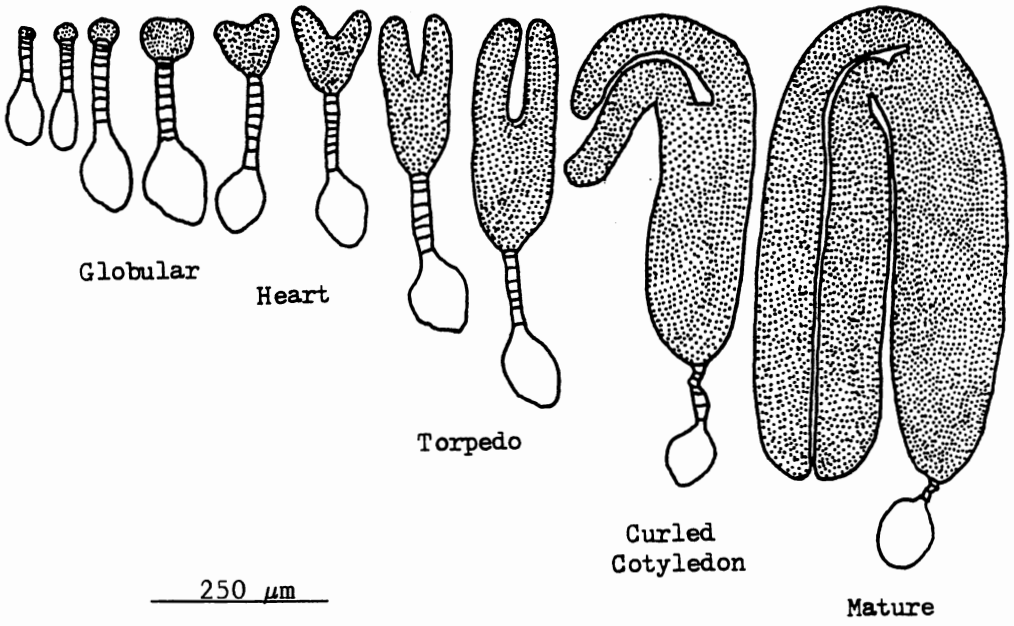



Figure 7. Stages of Normal Embryo Development in Arabidopsis thaliana. The embryo is composed of two major parts, the embryo proper (stippled region) and the suspensor (clear cells). The embryo proper develops into the mature embryo, while the suspensor degenerates during the final stages of development. Drawing from Meinke (1986).

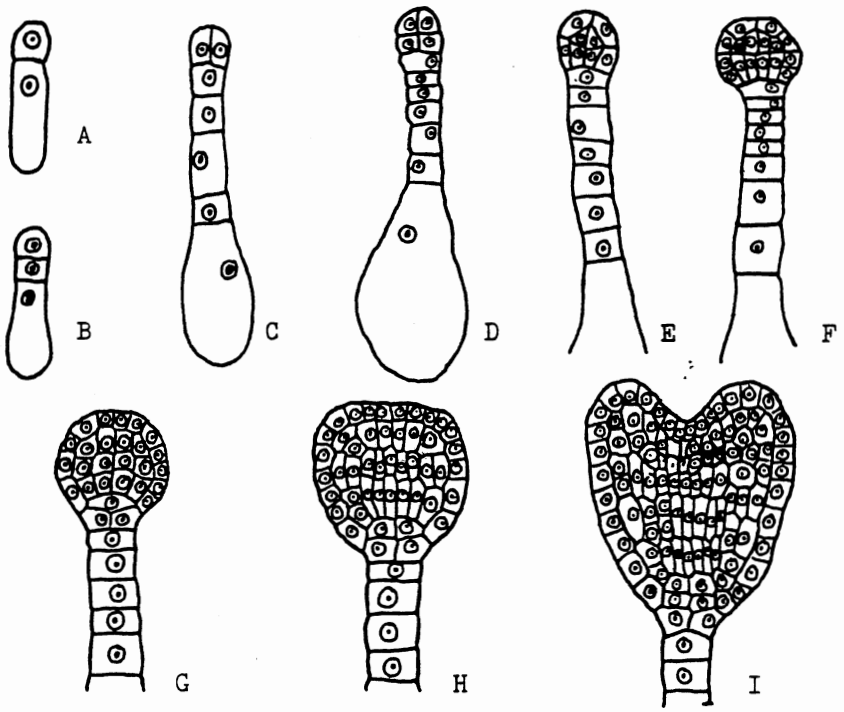



Figure 8. Stages of Early Embryo Development in Capsella bursa-pastoris. (A, B, C) Development of the proembryo; (D) octant stage with four of the eight cells of the embryo proper visible; (E) 16-cell stage, the time of initiation of the epidermal layer; (F) early globular; (G) globular; (H) late-globular; (I) heart stage. Early development in Arabidopsis follows a similar pattern as shown here. Drawing from Meinke (1979).

at the heart stage and remains green until desiccation when the embryo turns pale inside the dark brown seed coat. The suspensor is composed of 6 to 8 cells and is believed to play a role in delivering maternal nutrients to the developing embryo proper (Yeung 1980). The cells of the suspensor degenerate during the final stages of embryo development and are generally not found in mature seeds.

Genetic Features. Arabidopsis is particularly amenable to the isolation of mutants. In Arabidopsis, mutations have been identified in genes for nearly every aspect of plant growth and development. Mutations that affect embryo development are among the most commonly seen following mutagenesis with either X-irradiation (Müller 1963), EMS (Meinke 1986), or T-DNA insertion (Errampalli et al. 1991). The linkage map of Arabidopsis contains over 80 visible mutations with a wide range of morphological aberrations (Koornneef 1990). Considering the number of mutations that have been isolated, this is a relatively small number of genes on the genetic map. Trisomic and telotrisomic lines have also been identified in Arabidopsis which can be used for linkage studies (Koornneef 1983). Some of these lines however, are difficult to screen and maintain.

Complementation analysis in Arabidopsis is feasible, but only when working with a limited number of genetic loci. Performing complementation tests among embryo-lethal mutants is facilitated by the fact that the $F_1$ progeny are screened as mature green seeds, not as mature $F_1$ plants. However, cross-pollination in Arabidopsis is tedious because the young flower of the female parent must be emasculated prior to self fertilization, then hand-pollinated with a detached dehiscent

anther from the donor plant. Each cross typically takes 5 minutes and two to three successful pollinations are usually required to get meaningful results from a complementation test. Thus,it would take at least 300 hours just to perform the crosses for allelism tests among 50 mutants.

A number of different tissues of Arabidopsis including roots, leaves, stems, and seeds are susceptible to Agrobacterium-mediated T-DNA transformation, (Lloyd et al. 1986; Schmidt and Willmitzer 1988; Valvekens et al. 1988; Feldmann and Marks 1987). The frequency of transformation in these procedures has not yet reached the level required for shotgun transformation; a method in which random DNA fragments from wild-type plants are transformed into homozygous mutant tissue in an attempt to identify a piece of DNA that can complement (or rescue) mutant phenotypes.

Many of the tools that are required for basic genetic studies are presently available in Arabidopsis, but many of the genetic stocks that are found in other model systems are not available. For example, there are not many good stocks available for initial linkage analysis. Other tools like translocations, well-characterized deletions, duplications, inversions and balanced lethal chromosomes simply do not exist in Arabidopsis at this time. As pointed out by Fink (1988), these shortcomings will need to be addressed if Arabidopsis is to become a model genetic system like yeast, Drosophila, or Caenhorhabditis.

Molecular Characteristics. One of the main reasons Arabidopsis has received so much attention lately has to do with the molecular organization of the genome. Arabidopsis has very little repetitive DNA

and one of the smallest genomes (7 x $10^4$ Kb) among flowering plants (Leutwiler et al. 1984; Meyerowitz and Pruitt 1985). Other features include 2 separate RFLP maps (Chang et al. 1988; Nam et al. 1989), a yeast artificial chromosome (YAC) library with average inserts of 150 kb (Ward and Jen 1990), and a partially-completed physical map of the Arabidopsis genome (Goodman et al. 1989) consisting of overlapping cosmid clones prepared by the method described by Coulsen et al. (1986). In addition, a large international effort is currently being organized with the following goals: sequencing the entire genome of Arabidopsis (Palca 1989) and identifying by mutation and cloning as many genes as possible in order to examine in detail the relationship between mutant phenotype and gene function in higher plants. As part of this effort, resource and stock centers will be set up to maintain and distribute clones, mutants and seed stocks. A computer network with well over 100 subscribers called "Electronic Arabidopsis" has already been in operation at Michigan State University that allows researchers to exchange ideas and obtain up-to-date information from colleagues on matters concerning Arabidopsis (Sommerville 1989).

In summary, Arabidopsis has one of the best combinations of morphological, genetic, and molecular characteristics for the study of plant embryo development. Other model systems like maize have the problem of many linkage groups and a large haploid genome (5 x $10^6$ Kb) which may hinder the molecular isolation of genes. The main limitations in using carrot as a system for studying embryogenesis include the lack of genetic information available and the fact that many of the embryonic mutants affect somatic rather than zygotic embryogenesis.

Why Map Embryo-Lethal Mutants of Arabidopsis? As just pointed out, embryo-lethal mutants of Arabidopsis are an ideal system for the genetic analysis of plant embryo development. Because these mutants are so abundant, they can help strengthen the genetic characteristics of Arabidopsis. Once these mutations are placed on the standard genetic map they may be used as visible markers to help map mutations isolated in other labs. Embryonic lethals are particularly useful as genetic markers because (1) they are readily recovered following mutagenesis; (2) compared with other visible markers, only a small number of $F_2$ plants are required for linkage detection; (3) plants heterozygous for embryonic lethals can be identified by dissecting mature siliques following self-pollination; and (4) the absence of pleiotropic effects allows rapid identification of other visible markers in plants segregating for embryonic lethals.

Mapping these lethal mutations is also the first step in creating balanced lethal chromosomes, which have been invaluable to genetic analysis of Drosophila and Caenhorhabditis. Balanced lethal chromosomes usually contain inversions to suppress crossing over and flanking recessive lethal genes that enforce heterozygosity in subsequent generations. Mapping will also be required to construct genetic mosaics which could be used to perform clonal analysis studies and examine cell-autonomy of the lethal gene. These studies would require that the lethal mutation map near the end of a chromosome and distal of a linked cell-autonomous marker. See Figure 9 for a description of the strategy used to construct genetic mosaics.

Complementation analysis among embryonic lethals can be simplified by first mapping mutations to linkage groups, then limiting allelism

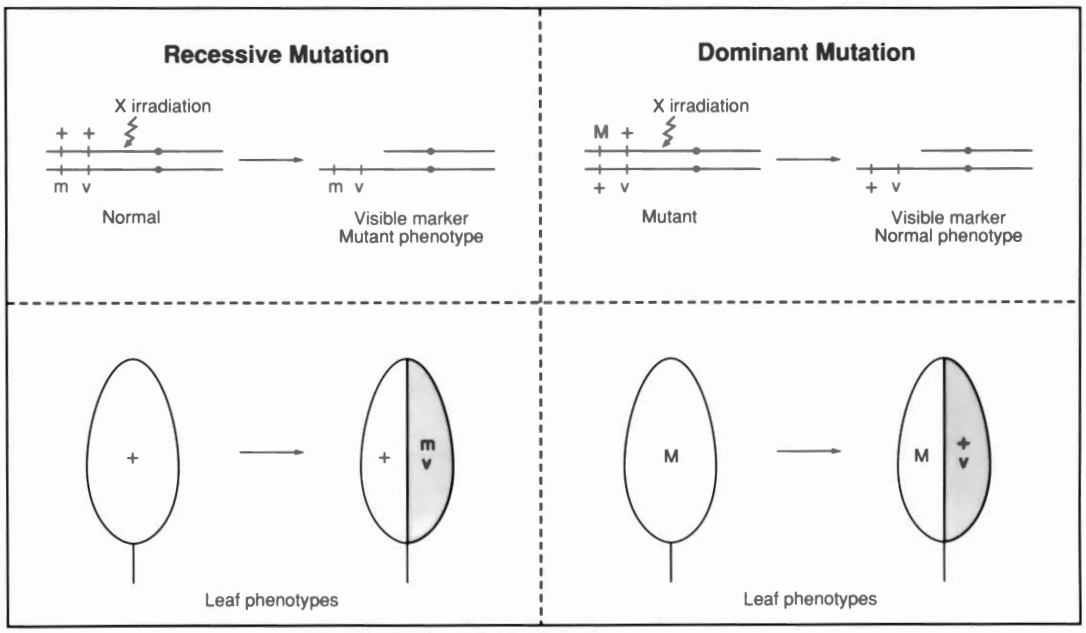



Figure 9. A General Strategy for Producing Genetic Mosaics. The left half of the box shows the strategy for analyzing a recessive mutation (m) in mosaic plants and the strategy on the right is for analyzing a dominant mutation in mosaic plants (M). In both strategies, (v) represents the recessive allele of a visible marker gene with a leaf phenotype and (+) represents the wild-type allele of the corresponding genes. The lower half of the box shows the phenotype of the leaves before and after X-irradiation. The shaded portion of the leaf represents the mosaic sector, identified by the visible marker phenotype. Figure reprinted from Meinke (1991).

tests to mutations that map to the same chromosome. This strategy may be used in the near future to help identify as many complementation groups as possible that are represented among embryonic lethals. This approach may help to reach one of the goals of the International Arabidopsis Genome Project which is to identify by mutation as many genes as possible that play a role in the life cycle of the plant.

A final reason for mapping embryonic mutations in Arabidopsis is that it is the first step in isolating genes through chromosome walking. This strategy involves determining which linkage group the gene is on, then locating the closest flanking molecular (RFLP) marker using RFLP mapping techniques (a method for high-resolution RFLP mapping of embryo-lethal mutations is presented in Chapter 4). Tightly-linked RFLP markers can then be used as starting material to isolate large (150 Kb) DNA fragments from a genomic YAC library. These YACs should on the average represent 1 cM of the genome. New fragments are made from the ends of the starting YAC and used to isolate adjacent overlapping YACs. This process continues until the gene of interest is covered by a YAC. Verification of gene isolation will require transformation of the putative gene back into plants and complementation of the mutant phenotype. Mapping additional embryo-lethal mutants of Arabidopsis should therefore complement the genetic analysis of embryogenesis in plants, strengthen the genetic characteristics of Arabidopsis, and facilitate the isolation of genes through chromosome walking.

The remainder of this chapter will focus on the procedures used in the linkage analysis of three embryo-lethal mutations: bio1, emb22, and emb30-1. Visible markers and telotrisomic lines are described along with the methods used to make crosses, collect and analyze data, and

construct an updated linkage map. The efficiency of two different methods for mapping embryonic lethals with visible markers is also compared. The third chapter describes the construction of new multiple marker lines that will be useful in the rapid detection of linkage between new mutations and visible markers. The fourth chapter describes how RFLP markers can be used to map embryo-lethal mutations at high resolution. Also, an example using this strategy is presented where the bio1 gene of Arabidopsis is mapped to within 0.5 cM of an existing RFLP marker.

# Materials and Methods

## Growth of Plant Material

All plants used in this study were grown at 23 ± 3° C in either 3 inch pots or 100 mm tissue culture plates under 40 W fluorescent shop lights set for 16h light/ 8h dark cycles. The soil mixture contained 12 parts coarse vermiculite (Terra-Lite, W.R. Grace Co., Cambridge, MA), 3 parts potting soil (Redi-Earth Peat-Lite, W.R. Grace) and 1 part sterile sand. Plants grown in soil were watered daily with a solution containing 1.6 g/l of 7-16-19 All Purpose Hyponex (Hyponex Co., Fort Wayne, IN) and 0.1 g/l of 15-16-17 Peat Lite Special (Peters Co., Allentown, PA). Seeds placed in culture were first surface sterilized for 30 sec in 95% ethanol, 6 min in 50% Clorox and rinsed with at least three changes of sterile distilled water.

## Isolation of Embryo-Lethal Mutants

Embryo-lethal mutants used in this study were isolated and characterized previously as described by Meinke (1985). The mutagenic procedure consisted of treating mature dry seeds of the Columbia ecotype with 0.2% EMS in water for 7h with gentle stirring; the seeds were then washed with water for 12h prior to planting. The resulting chimeric $M_1$ plants (see Figure 10) were grown in soil and allowed to self-pollinate. At least 4-6 successive siliques were screened in each $M_1$ plant for the presence of 25% aborted seeds to ensure sampling from each sector. Normal seeds present in the siliques containing 25% aborted seeds should be a mixture of heterozygous and wild-type with respect to the lethal

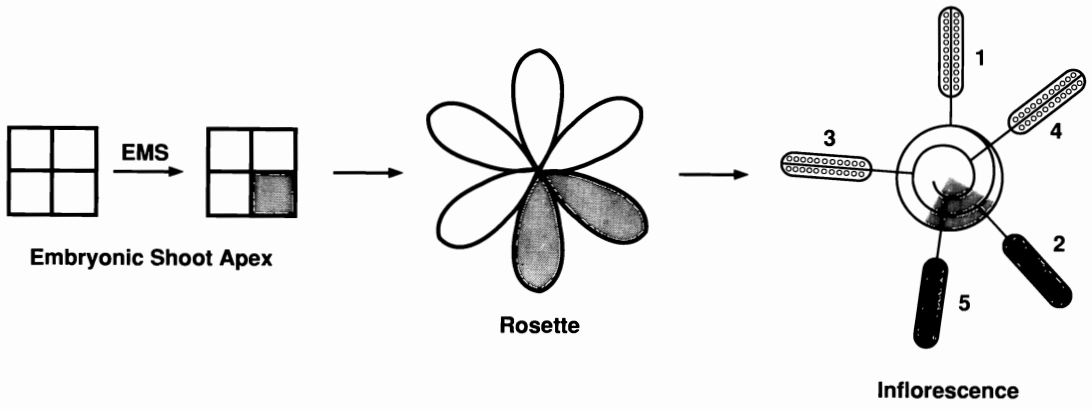



Figure 10. Schematic Drawing of Chimeric $M_1$ Plant. The four squares represent the embryonic shoot apex. One cell (shaded) becomes mutagenized at an emb locus during EMS treatment and subsequently gives rise to the mutant sector. Siliques that grow out of this mutant sector will contain 25% aborted seeds (dark) and 75% normal seeds (light) following self-pollination. Figure reprinted from Meinke (1991).

gene. Heterozygotes were identified among these normal seeds by planting in soil and screening siliques again for 25% aborted seeds.

Figure 11 demonstrates the diversity of phenotypes obtained from this group of mutants. Three embryo-lethal mutants with distinctive phenotypes were chosen for this study: bio1, emb22, and emb30. The eventual goal is to isolate the genes defined by these mutations and determine their function during embryo development.

The bio1 Mutant. This mutant, originally called 122G-E, is a biotin auxotroph and has a broad lethal phase ranging from the globular to mature cotyledon stages of development (see Figure 5, p.18). A typical arrested bio1 embryo is shown in Figure 11. The bio1 mutant was chosen for this study for several reasons: (1) very few auxotrophs have been identified in plants; (2) virtually nothing is known about biotin biosynthesis in plants; (3) this is the first example of an embryonic lethal of Arabidopsis that has a known biochemical lesion; and (4) studying bio1 should lead to a better understanding of mutations that affect housekeeping functions essential for embryogenesis.

Further analysis of bio1 supports the model which predicts that arrested embryos from lethal mutants defective in housekeeping functions should be white in color and have a broad lethal phase (Meinke 1991). The logic behind the first part of this model is as follows: housekeeping genes are likely to be required for normal chloroplast maintenance, so mutants defective in these genes will likely have defective chloroplasts and therefore lack the normal green pigmentation. The second part of the model predicts that auxotrophic mutants will have a broad lethal phase; this prediction stems from the fact that the

Figure 11. Embryos From Selected Lethal Mutants of Arabidopsis. Top row from left to right the embryos are: mature wild-type; emb16; emb22; and biol. Bottom row from left to right emb30-1, emb18, emb31; and mature wild-type. The maximal length of the mature wild-type embryos is approximately 500 $\mu$m.

homozygous mutant embryo receives nutrients from surrounding heterozygous maternal tissue which may provide variable levels of the required nutrient or gene product which in turn leads to the variable lethal phase. This model has not been adequately tested but is instead a working model that will likely require modification once a number of genes have been cloned and analyzed. Two other nutritional mutants have been identified in Arabidopsis, but neither becomes lethal during embryo development. The thiamine (th) mutants described by Li and Redéi (1968) appear to get sufficient levels of thiamine from maternal sources to survive embryo development, but mutant plants never survive past the seedling stage. The trp1 mutant of Arabidopsis is defective in the synthesis of tryptophan, but does not become lethal during embryo development, probably because the trp1 gene is duplicated in the genome (Last and Fink 1988). The second gene appears to produce enough gene product for the completion of embryo development but not enough for the completion of the life cycle.

The best way to verify this model would be to isolate genes from a number of lethal mutants and determine their function. Mutations in other genes which appear to play a more direct role in the regulation of embryogenesis are currently being analyzed at the molecular level in preparation for gene isolation.

Mutant emb22. In previous papers this mutant was referred to as 115D-4A or the "green blimp" mutant. Arrested embryos from emb22 are green and fairly large in size but they consistently lack a differentiated hypocotyl and cotyledons (Figure 11, p.37). Light microscopic examination of serial sections revealed that some internal

differentiation does occur in arrested emb22 embryos but no defined root or shoot meristems were seen (Patton and Meinke 1990). Arrested emb22 embryos grow in culture to produce abnormal shoots and leaves, but never produce flowers or seeds (Franzmann et al. 1989). The emb22 mutant therefore appears to be blocked in morphogenesis of both developing embryos and young shoots. This mutant gene also appears to show gametophytic as well as sporophytic expression (Meinke 1985). The phenotype of mutant emb22 suggests that the defective gene may have a developmental rather than housekeeping function. Until the emb22 gene is cloned and analyzed, this question remains unresolved.

Mutant emb30. Arrested embryos from emb30, originally called 112A-2A, are blocked at the linear cotyledon stage of development (Figure 5, p.18), are green in color, and typically have fused cotyledons and a reduced hypocotyl (Figure 11, p.37). As described earlier, this mutant in culture forms rootless plants with thick abnormal leaves. Serial sections taken through arrested embryos have shown that the root apex appears to be missing in emb30 (Patton and Meinke 1990). These observations are consistent with the idea that the wild-type allele (EMB30) is responsible in part for controlling development of the root apex. Again, the only definitive way to determine the function of EMB30 will to be clone and analyze this gene at the molecular level.

## Wild-Type Ecotypes

Landsberg "Erecta" Ecotype. The Landsberg ecotype of Arabidopsis has been used for many years as a common laboratory strain. All of the visible marker lines used in this study were isolated previously and

maintained in this genetic background. Although Landsberg is often considered a wild-type ecotype, many laboratories maintain the homozygous recessive er (erecta) mutation. Typical plants from this ecotype have short main stems (15-20 cm) that are erect in stature, have 2-4 lateral branches, and short floral internodes (distance between adjacent siliques on the inflorescence is often less than 0.5 cm). Mature siliques produced by Landsberg plants are usually 0.7-1.0 cm in length, 0.2-0.3 cm wide and typically contain 50 seeds.

Columbia Ecotype. When Columbia and Landsberg "erecta" plants are grown together, the most striking differences are the growth habit and the length of the inflorescence. The Columbia ecotype, in contrast to Landsberg, has long (30-40 cm) main stems, which bend under the weight of the inflorescence, and the distance between adjacent siliques is typically 1-2 cm. Mature siliques produced by Columbia plants are noticeably longer and more slender than Landsberg siliques. Typical Columbia siliques are 1-2 cm long, 0.1-0.2 cm wide and contain 50 seeds.

Features that are common to these Arabidopsis ecotypes but are abnormal in visible marker lines include the appearance and distribution of trichomes, presence of epidermal wax on stems, floral morphology, and color of leaves, inflorescences or seeds. Each of these characters will be discussed in the following section.

## Multiple Marker Lines

General Information on Multiple Markers. Multiple markers used in this study were generated previously by Maarten Koornneef and were maintained in the Landsberg "erecta" background. Each multiple marker

used in this study contained two or three recessive mutations on one of the five linkage groups plus the er mutation on chromosome 2. Refer to the genetic map shown in Figure 12 for the location, and Table 1 for a brief description of each of the visible markers used in this study. The paragraphs below summarize the unique features of the multiple markers W2, W4, W8, W10, W11 and W13 and describe the recessive mutations contained in each multiple marker. The letter "W" in the multiple marker designation stands for Wageningen, the location of the Agricultural University in the Netherlands where Maarten Koornneef and his students constructed these multiple marker lines. It should be noted that Maarten Koornneef has recently changed the designation of some these marker lines. In order to be consistent with my laboratory records, the old nomenclature will be used in this document.

W2 Multiple Marker. This multiple marker contains two recessive mutations, an and dis1, that map to the top arm of chromosome 1. The first mutation, an is at 0 cM and exhibits two distinct phenotypes, narrow leaves (both rosette and cauline) and twisted siliques. The narrow leaf phenotype is most striking when comparing wild-type and an cauline leaves. This phenotype can easily be identified in young seedlings grown either in soil or in culture plates. In 7-10 day old seedlings the an leaves are 2 to 3 times longer than they are wide, while wild-type leaves are just about as long as they are wide. The twisted silique phenotype is also easy to screen, the siliques are twisted for 1-3 complete 360° rotations. This phenotype is more pronounced in Columbia plants because the siliques are typically longer and therefore contain more twists than Landsberg siliques. This

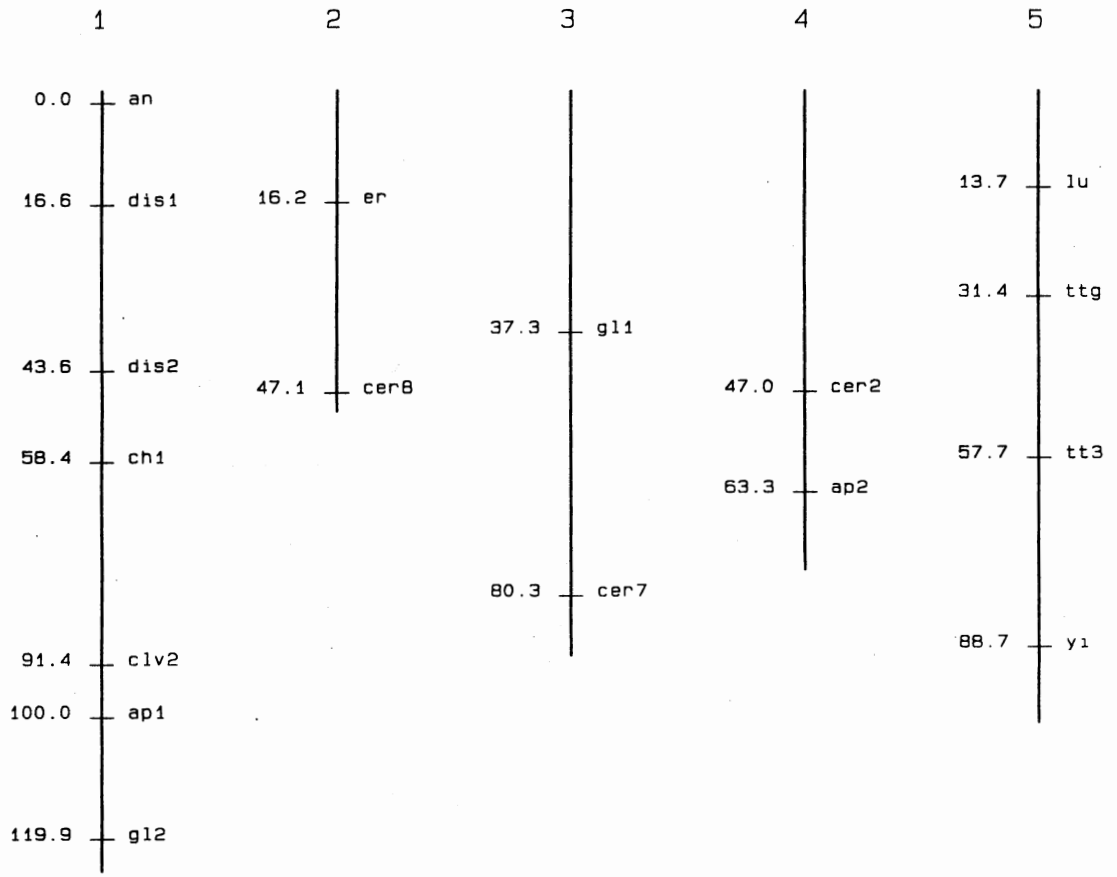



Figure 12. Linkage Map of Arabidopsis Showing Locations of Visible Markers Used in This Study. Marker positions from Patton et al. (submitted).

TABLE 1

OVERVIEW OF VISIBLE MARKERS USED IN THIS STUDY

| Marker | Location | Brief Description of Phenotype |
|---|---|---|
| an | 1: 0 | narrow leaves; twisted siliques |
| dis1 | 1: 16.6 | linear, singed, and distorted trichomes |
| dis2 | 1: 43.6 | similar to dis1 but longer trichomes |
| ch1 | 1: 58.4 | pale-green plants |
| clv2 | 1: 91.4 | fat siliques with extra carpels |
| ap1 | 1: 100.0 | apetalous flowers |
| gl2 | 1: 119.9 | reduced number of trichomes |
| er | 2: 16.2 | erect stems; blunt siliques |
| gl1 | 3: 37.3 | trichomes absent |
| cer7 | 3: 80.3 | shiny stems; reduced epidermal wax |
| cer2 | 4: 47.0 | like cer2 |
| ap2 | 4: 63.3 | like ap1 |
| lu | 5: 13.7 | pale green plants; not seen in culture |
| ttg | 5: 31.4 | trichomes absent; yellow seeds |
| tt3 | 5: 57.7 | yellow seeds |
| yi | 5: 88.7 | yellow inflorescence |

phenotype makes it slightly more difficult to open the valves and examine the contents of the silique but it is not a major hinderance when screening mature siliques for aborted seeds.

The second mutation in the W2 multiple marker is dis1, which maps to 16.6 cM on chromosome 1. This mutation disrupts the formation of trichomes, which are normally found on the upper leaf surface and the lower portion of stems. In wild-type Arabidopsis, the leaf trichomes are usually three-branched (Figure 13), while the stem trichomes are branched either two or three times. Leaf and stem trichomes of dis1 plants are unbranched and singed or twisted in appearance (Figure 14). A 10X hand lens can be used to screen for the dis1 phenotype in young seedlings grown in soil or in culture. The best time to screen seedlings for dis1 is when they have formed 2-3 pairs of true leaves. It should be noted that cotyledons do not have trichomes and should therefore not be confused with true leaves when screening for trichome phenotypes. The dis1 phenotype is essentially the same in both Landsberg and Columbia backgrounds. The combination of an and dis1 in the W2 multiple marker line does not seem to have an adverse effect on general plant growth or seed set.

W4 Multiple Marker. This multiple marker contains three mutations on the lower arm of chromosome 1: ch1, ap1, and gl2. The ch1 mutation is located at 58.4 cM on chromosome 1 and causes plants to develop a lime-green color rather than the normal dark green color. This mutation affects leaves, stems and sepals, or all of the green organs in the Arabidopsis plant. One of the pleiotropic effects of this mutation is an overall reduction in growth rate and ultimately results in smaller

Figure 13. Photograph of Rosette Leaves Showing Wild-Type Trichomes. The smallest rosette leaf here is about 0.5 cm in length.

Figure 14. Photograph of Rosette Leaves From dis1 Plant Showing Distorted Trichomes. The smallest rosette leaf here is about 0.5 cm in length.

plants. This phenotype can be screened in plants grown in soil or in culture. Care must be taken when growing a population of plants that are segregating for ch1 because surrounding green plants tend to quickly overgrow adjacent mutant plants. In this case, the normal plants can be selectively trimmed to allow adequate light for the mutant plants.

The ap1 mutation is located at 100.0 cM on chromosome 1 and causes plants to produce flowers that are missing petals. The flowers of ap1 are normal in every other way. The ap1 phenotype can be screened only in mature plants that are flowering. Seed set is slightly reduced in ap1 plants.

The gl2 mutation is located at 119.9 cM which is near the end of chromosome 1. The gl2 mutation affects the shape and distribution of trichomes on the plant. The leaves of gl2 plants have very few trichomes. Those that are seen are usually linear or branched twice and are found mainly near the leaf margin. This phenotype can be screened with a 10X hand lens in plants grown in soil or in culture. The gl2 phenotype is most easily seen in plants with a number of fully expanded rosette leaves. Screening for gl2 can be tricky because some leaves in wild-type plants show an abnormal distribution of trichomes. The key to identifying plants that are homozygous recessive for gl2 is to make sure that all of the leaves examined are fully expanded and have the phenotypes described above.

Plants from the W4 line tend to grow more slowly than wild-type, flower later, and have slightly reduced seed set. Despite these features and the fact that the gl2 mutation is sometimes hard to screen in segregating populations, the W4 marker line can still be used successfully to map mutations that affect seed development. Each of the

three mutant phenotypes exhibited in W4 look essentially the same in both Landsberg and Columbia backgrounds.

W8 Multiple Marker. The W8 multiple marker has two markers on chromosome 3: gl1 and cer7. The first marker, gl1, maps to the top arm of chromosome 3 at 37.3 cM. Plants that are homozygous recessive for gl1 completely lack leaf trichomes. This phenotype is seen when plants are grown either in soil or in culture. As with most other trichome mutants, the phenotype is most easily seen in young seedlings with 2-3 pairs of true leaves.

The cer7 mutation maps to 80.3 cM near the lower end of chromosome 3. This mutation blocks the accumulation of epidermal wax on the stem. Nine distinct cer loci have been identified and mapped in Arabidopsis. The stems of most of these mutants resemble cer2 which is shown in Figure 15 along with a wild-type stem. The cer7 phenotype can easily be screened with the unaided eye by gently rubbing a small portion of the stem with a finger, and then examining that region under bright light. Plants that have normal epidermal wax will have an overall powdery or dull appearance with a small shiny green region where the wax was removed by rubbing. Mutant plants that lack epidermal wax have stems that are shiny along their entire length, regardless of whether the stems are rubbed with a finger or not. The phenotypes of the W8 marker are similar in the Landsberg and Columbia backgrounds.

W10 Multiple Marker. The two mutations found in W10, cer2 and ap2, map to chromosome 4. The cer2 mutation is located at 47.0 cM near the center of the chromosome. Like cer7 described above, cer2 affects the accumulation of epidermal wax on the stems. A quick comparison of cer2

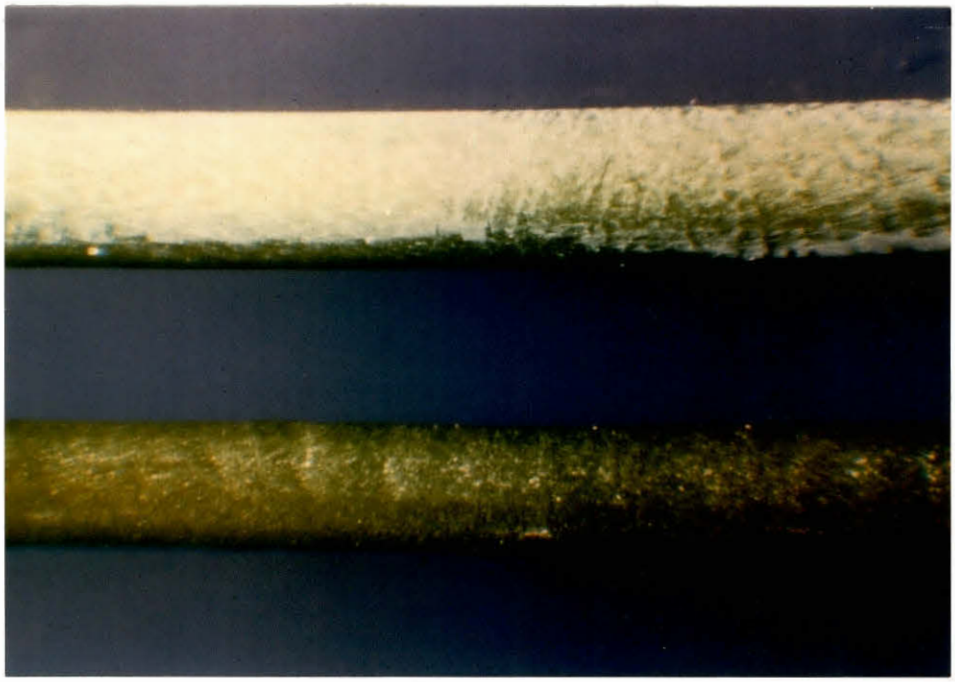


Figure 15. Comparison of Stems From Wild-Type (top) and cer2 (bottom) Plants. Note the absence of wax on the cer2 stem. These stems are approximately 0.2 mm in diameter.

and cer7 shows that the phenotypes of these two mutations are indistinguishable. Figure 15 (p.49) compares the main stems from wild-type and cer2 plants. Note the dull appearance of the wild-type stem when compared to cer2. Screening for the cer2 phenotype is the same as described above for cer7.

The second mutation in W10, ap2, maps to 63.3 cM near the lower end of chromosome 4. The ap2 mutation disrupts the formation of petals in the flower. Plants that are homozygous recessive for ap2 will occasionally produce flowers with small, pale-green petals, but most flowers completely lack petals. This phenotype can quickly be identified in a segregating population because the distal portion of ap2 inflorescences lack the white color provided by normal petals. In our laboratory, most siliques produced by ap2 plants were very short (<0.5 cm) and did not contain seeds, however, occasional siliques did produce 2-4 seeds. These plants produce normal pollen but the dehiscent anthers are not in close proximity to the stigma when pollen is shed. The filaments seem to be pointed at an angle away from the pistil rather than growing parallel to it. The pistil also seems to elongate past the ends of stamens prior to the time of dehiscence, so even if the filaments did grow at the proper angle, the pollen would probably not be deposited onto the stigma. These ap2 flowers can be hand-pollinated if the stamens are removed with fine-tipped forceps and the dehiscent anthers are touched onto the stigma surface. Siliques of ap2 plants produced following hand-pollination typically contain 20 seeds.

The W10 multiple marker is not an ideal marker for mapping mutations that affect seed development because one-fourth of the $F_2$ progeny must be hand-pollinated in order to get enough seeds to screen

for the seed phenotype. The cer2 and ap2 phenotypes do not appear to be affected by genetic background. An additional multiple marker for chromosome 4 has been constructed and is discussed in chapter 3.

W11 Multiple Marker. The W11 multiple marker contains two mutations on chromosome 5, lu and tt3. The lu mutation is at 13.7 cM on chromosome 5 and disrupts timing of flowering and the pigmentation of green tissues. Plants that are homozygous recessive for lu are very pale green and exhibit an extended vegetative growth phase. The pale green phenotype can be seen in young seedlings and mature plants when grown in soil. In contrast, the pale-green phenotype is not seen when plants are grown in culture on a basal medium supplemented with glucose. Homozygous recessive lu plants grow vegetatively for 8-10 weeks producing large rosettes (up to 10 cm in diameter) before they flower, while wild-type plants grown under the same conditions begin to flower after only 4-5 weeks of vegetative growth (rosettes are typically 5 cm in diameter). When these W11 plants finally do flower they seem to produce an unusually large number of opened (mature) flowers at the same time. The overall effect is an inflorescence with 15-20 mature flowers bunched up at the tip, while wild-type plants usually have 5-10 mature flowers at any one time. These mutant plants with a large number of mature flowers also produce a very distinct and sweet-smelling aroma especially if a number of plants are grown together in the same pot. All of the characteristics described above have been observed in both Landsberg and Columbia backgrounds.

The tt3 mutation is located at 57.7 cM near the middle of chromosome 5. This mutation disrupts the accumulation of anthocyanin-

derived pigments in the seed coat. The testa (seed coat) of normal seeds accumulates a brown pigment during the final stages of seed maturation, while mutant tt3 seeds have transparent seed coats which give these seeds an overall yellow color (Figure 16). The yellow color in this case is due to the normal color of the mature dry embryo which is visible through the transparent seed coat rather than yellow pigments in the seed coat. The yellow seed phenotype can be seen with the naked eye by breaking open mature dry siliques on a piece of white paper. Wild-type plants produce dark seeds which are easily seen on the light background, while the seeds produced by mutant tt3 plants are yellow and more difficult to see on the light background. This phenotype can be seen only in mature plants that have produced dry seeds. The tt3 phenotype is the same whether in the Landsberg or Columbia background. The W11 multiple marker does not affect seed set.

W13 Multiple Marker. Like W11, this multiple marker has two markers on chromosome 5, ttg and yi. The ttg mutation is at 34.4 cM and affects both trichome formation and seed coat pigmentation. This mutation causes plants to produce yellow seeds (like tt3) and glabrous stems and leaves (no trichomes, like gl1). These two characteristics are inseparable during segregation and appear to be the result of a single mutational event. The glabrous phenotype can be screened like the other trichome mutations, with a 10X hand lens. In this case, all seedlings with the glabrous phenotype will grow and produce seeds with a transparent seed coat (yellow seeds as in Figure 16).

The yi mutation is located at 88.7 cM near the end of the lower arm of chromosome 5. This mutation causes the center of the rosette and

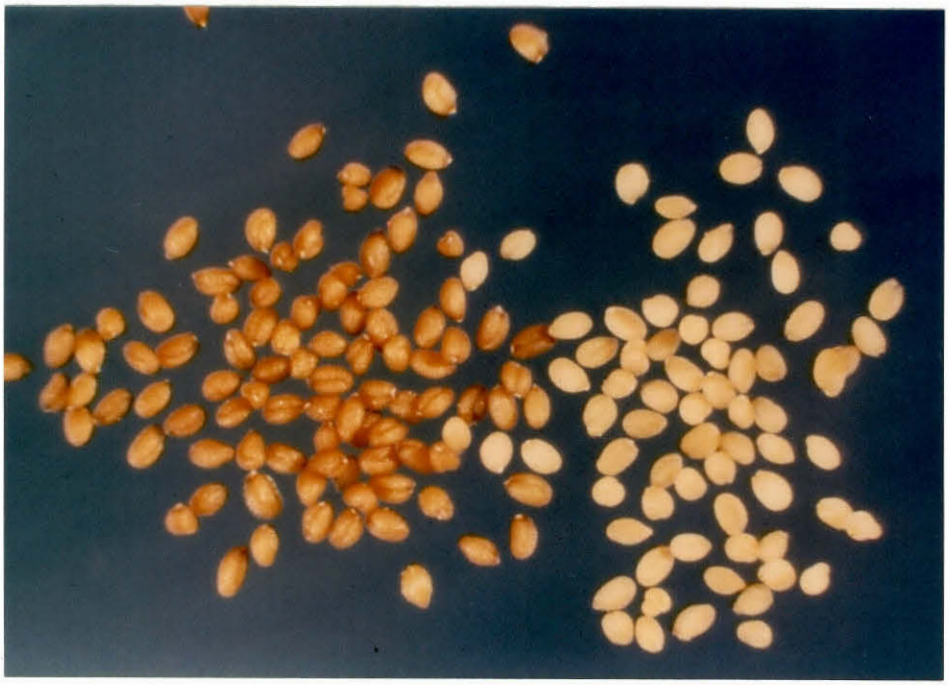


Figure 16. Comparison of Seeds From Wild-Type (left) and tt3 (right) Plants. Note the absence of pigmentation in the seed coat of tt3 seeds. Seeds are approximately 0.5 mm in length.

tips of inflorescences to accumulate a yellow pigment. The bright yellow color can be seen with the unaided eye by examining either the last few centimeters of mature inflorescences or the center of the rosette of young seedlings grown in culture or in soil. The W13 markers, ttg and yi, do not affect seed set and are essentially the same in the Landsberg and Columbia backgrounds.

Single Marker Lines

The dis2 Mutation. The dis2 mutation is located at 43.6 cM near the center of chromosome 1. This mutation causes the plants to form distorted linear trichomes that are much like dis1 trichomes. The main difference between dis1 and dis2 is that dis2 trichomes are noticeably longer and are occasionally branched. This mutation does not affect seed set and looks similar in both the Landsberg and Columbia backgrounds.

The clv2 Mutation. The clv2 mutation maps to 91.4 cM on the lower arm of chromosome 1. This mutation affects the formation of the carpels within the ovary. Siliques produced by wild-type plants have two carpels or valves, while siliques produced by homozygous recessive clv2 plants typically have 3-4 carpels (Figure 17). These siliques are usually shorter and fatter in mutant plants and contain essentially the same number of seeds as normal wild-type siliques. This phenotype seems to be consistent whether in the Landsberg or Columbia background.

Figure 17. Comparison of Siliques From clv2 (top) and Wild-Type (bottom) Plants. Note the extra valve present on the clv2 silique. The siliques in this photograph are approximately 2 cm long.

Telotrisomic Marker Lines

General Information. Telotrisomics are aneuploids that have an extra chromosome that is missing one of the chromosome arms. The extra chromosome arm is attached to the centromere and segregates as an independent genetic unit. Transmission of telocentric chromosomes in telotrisomics is often reduced in both pollen and ovules. A selfed telotrisomic will therefore produce a mixture of normal diploid and telotrisomic progeny. A number of different trisomic and telotrisomic sets have been identified in Arabidopsis (Steinitz-Sears 1963; Röbblen and Kribben 1966, and Koornneef and Van der Veen 1978). The telotrisomic lines used in this study were originally identified as morphological variants in a trisomic line which was obtained by crossing a stable colchicine-induced tetraploid with a normal diploid (Koornneef and Van der Veen 1983). Putative telotrisomics were characterized in detail and shown to distort segregation over the region covered by the extra chromosome arm in crosses with visible markers. The telotrisomics described below were induced and maintained in the Landsberg "erecta" background and were obtained from Maarten Koornneef. Telotrisomic lines were maintained by saving selfed seed from telotrisomic plants then identifying them again in the next generation based on plant morphology.

Telotrisomic Trla. The phenotype of Trla is fairly consistent and easy to see in a segregating population. Trla telotrisomic plants are shorter than wild-type and have dark green leaves and irregular petals (Figure 18). These plants are 5-10 cm shorter than diploids from the same population of seeds. Trla plants are a very dark shade of green which sometimes is not easy to see. The flowers have 2 to 4 long

Figure 18. Comparison of 4 Week Old Trla (left) and Wild-Type (right) Plants. Note the abnormal flowers and dark green leaves on the Trla plants.

slender petals which are often irregularly spaced around the circumference of the flower. Tr1a plants are fertile but produce only one-half to two-thirds of the normal number of seeds when compared to wild-type.

Telotrisomic Tr1b. These plants have long slender stems and small necrotic rosette leaves (Figure 19). The rosette of Tr1b plants also contains a reduced number of leaves. This telotrisomic has normal fertile flowers and shows good seed set. Tr1b plants are easily spotted among a population of plants that are descendants from a telotrisomic parent.

## Linkage Analysis

Mapping With Visible Markers. The strategy used for mapping embryonic lethals with visible markers is shown in Figure 20. Flowers from visible marker lines (vis/vis) were emasculated prior to self-pollination and dusted with pollen from plants heterozygous for an embryo-lethal mutation (emb/EMB). The resulting $F_1$ progeny were screened for the lethal by examining the contents of mature green siliques under a dissecting microscope. $F_1$ plants that contained 25% aborted seeds were double heterozygotes with the recessive alleles in trans with respect to the marker and the lethal (VIS emb/vis EMB). These double heterozygotes were allowed to self and produce $F_2$ seeds. Aborted seeds were removed from seed stocks prior to planting. Usually 27 pots, each containing 9 $F_2$ plants in known positions, were grown from each cross. Mature $F_2$ plants were then screened for both the visible marker and the lethal (AD method). Alternatively, $F_2$ seeds were surface

Figure 19. Comparison of 4 Week Old Trlb (left) and Wild-Type (right) Plants. Note the slender stem and necrotic rosette of the Trlb plant.

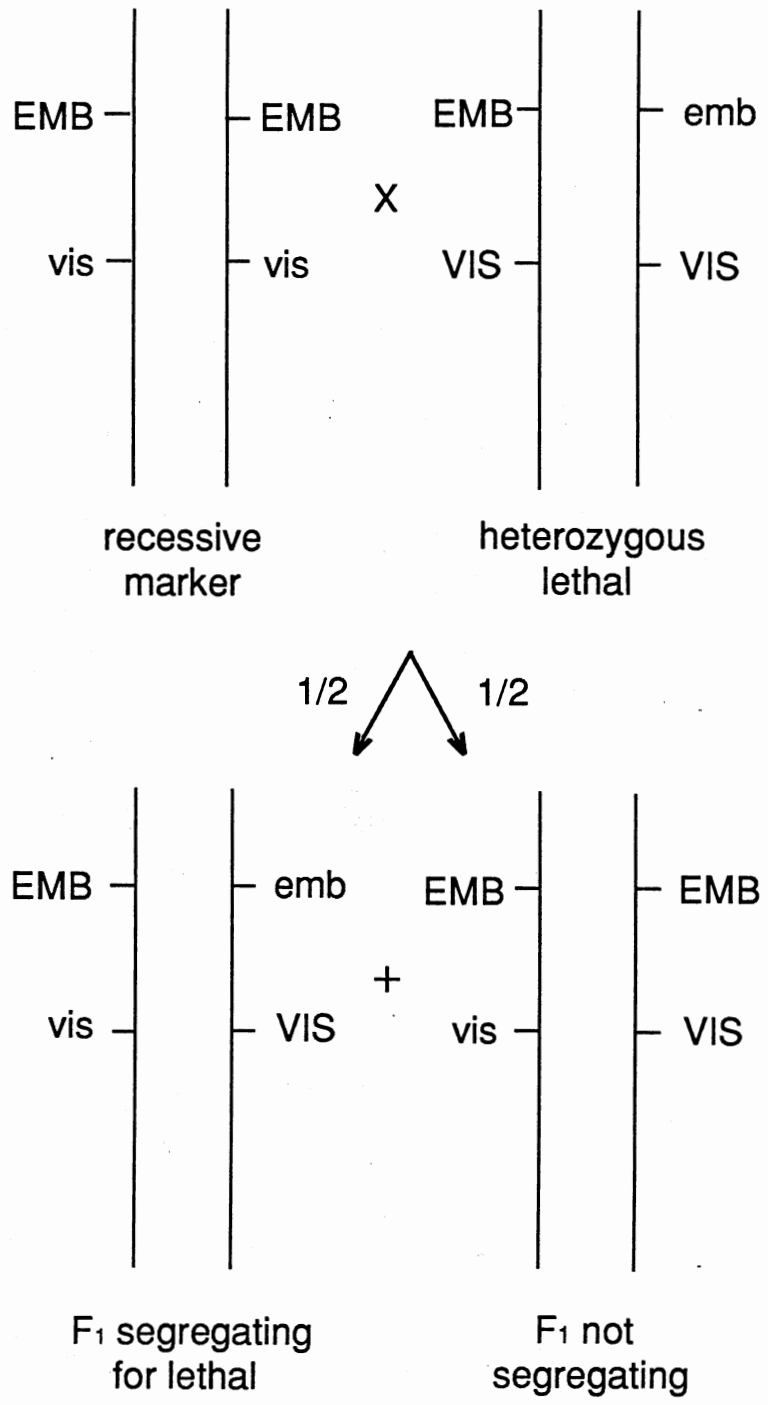



Figure 20. Drawing of Strategy Used to Map Embryonic Lethals With Visible Markers. The visible marker is crossed with pollen from plants heterozygous for the lethal. One-half of the resulting $F_1$ plants will be segregating for the lethal and one-half will not be segregating for the lethal. $F_2$ seeds are saved from $F_1$ plants segregating for the lethal, then planted and screened for the appropriate phenotypes.

sterilized and placed in 100mm tissue culture plates containing an agar medium then screened after germination for the visible marker only (EF method).

Four computer programs (BUILD, SUMMARY, CHI, RECF2) were then used in succession to summarize and analyze data collected using the AD screening method, while only the CHI program was needed to analyze the EF data. The BUILD and SUMMARY programs were written by Patricia Alexander at the OSU computer center (according to my specifications) during the summer of 1988. The CHI program was written by John Garnett, a student at Oklahoma State University (again, according to my specifications), during the spring of 1988. The RECF2 program was written by Piet Stam of the Wageningen Agricultural University (WAU), the Netherlands, and obtained from Maarten Koornneef during a visit to his lab at the WAU in January of 1988. The RECF2 program was described by Koornneef and Stam (1988) in a paper that discusses different methods for mapping genes in Arabidopsis.

Data collected using the AD screening method were entered onto screening sheets that had a separate row for each plant and columns representing each phenotype screened. Segregation data were then entered via keyboard into the BUILD computer program which generated a text file that was essentially a duplicate of the screening sheet. The SUMMARY computer program, which analyzes the BUILD text file, was then used to sum the total number plants in each of the four phenotypic classes. See appendix A (p.131) for sample printouts from the BUILD and SUMMARY programs for a cross between W2, containing visible markers an, dis1, and er and the embryonic lethal emb30. In this example, the visible markers an and dis1 are linked to emb30. Note that the data

shown in Appendix A were actual $F_2$ segregation data.

A computerized version of the chi-square test (CHI) was then used to determine whether the segregation ratios were significantly different from that expected for unlinked genes. The CHI program was written for the sole purpose of analyzing $F_2$ segregation data collected by either the AD or the EF screening method, following crosses between an embryonic lethal and visible markers. Appendix B (p.138) demonstrates how the CHI program was used to determine whether the segregation data presented in Appendix A was significantly different from what was expected for unlinked genes.

Finally, an estimate of the percent recombination between the embryonic lethal and the visible marker was made using AD data and RECF2. This program is a computerized version of the maximum likelihood method (Fisher 1958) based on the method of Servítovà and Cetl (1984). See Appendix C (p.143) for an example of how recombination estimates were made using AD segregation data. In this example, a recombination estimate was made between dis1 and emb30 from the cross described above.

Backcrosses. Backcrosses are required to definitively show the gene order when two genes are tightly linked. Figure 21 demonstrates the crossing strategy used to resolve the gene order in this case. The steps included are: (1) cross a maker line that has two visible markers (EMB vis1 vis2/EMB vis1 vis2) with pollen from a linked heterozygous lethal (EMB VIS1 VIS2/emb VIS1 VIS2); (2) identify heterozygous lethal $F_1$ plants (EMB vis1 vis2/emb VIS1 VIS2); (3) backcross visible marker parent with pollen from heterozygous lethal $F_1$ plant described above; (4) collect 100 to 200 $F_2$ backcross progeny and then plant and screen

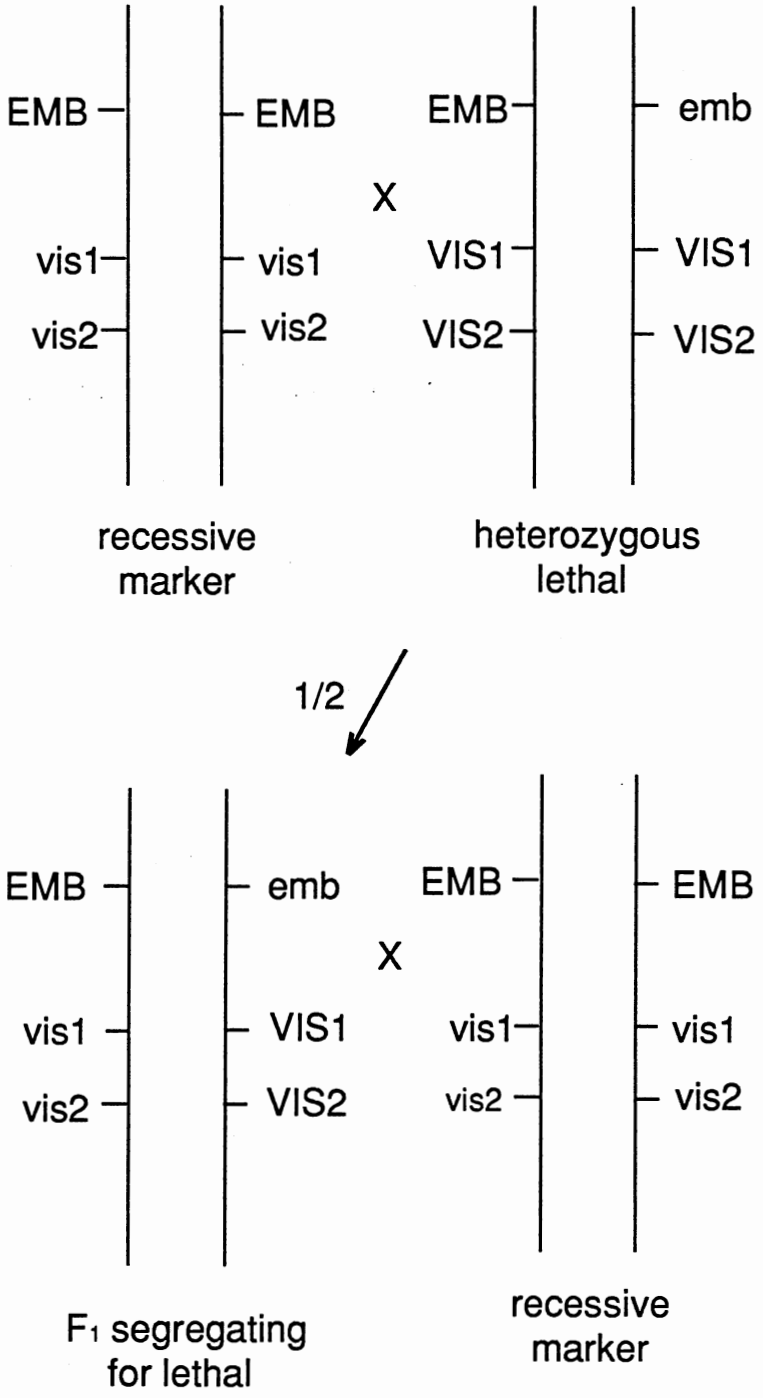



Figure 21. Drawing of Strategy Used for Backcrosses. Plants with two visible markers are crossed with pollen from a heterozygous lethal. One-half of the resulting $F_1$ plants will be segregating for the lethal and also be heterozygous for both of the visible markers; these $F_1$ plants are then backcrossed with the visible marker parent. The backcross progeny are then planted and screened for both visible markers and the lethal.

for both visible markers and the lethal; and (5) resolve the order of the three genes by determining the number of backcross progeny in each of the eight possible genotypic classes. Only one gene order will be consistent with the data because the number of gametes with 2 recombination events over the interval in question will be much less than the number of gametes with one or no crossovers.

Constructing the Genetic Map. The genetic map was constructed with the aid of three computer programs, GENMAP, CHROMAP, and PRINTAPLOT. Detailed instructions on how to use these programs are included in Appendix D (p.150), and Appendix E (p.155). The GENMAP program was written by Piet Stam at the Wageningen Agricultural University and obtained from Maarten Koornneef in January 1988. GENMAP requires data files that contain all available recombination data for each of the chromosomes (shown in Appendix D, p.150). The data files that were used to construct the most recent linkage map of Arabidopsis (Koornneef 1990) were obtained from Maarten Koornneef. In order to use GENMAP, the order of genes on each chromosome must be known. GENMAP first converts recombination estimates from each data set into cM using the Kosambi (1944) mapping function. For short distances (eg. $P < 10\%$), recombination estimates are approximately equal to cM. As the distance gets greater between two loci, the cM value increases exponentially (see Appendix F; p.161). GENMAP then estimates the position of each gene using a best-fit procedure. Chi-square values are calculated for each data set by comparing the original estimated distance between two genes obtained experimentally with the final estimated distance established by GENMAP. This process continues for each pair of genes until a minimum

cumulative chi-square value for all data sets is reached. Suspect data sets with unusually high chi-square values are noted and may either be removed or maintained. The orders of some genes on the map have not been resolved; backcrosses will therefore be required to definitely determine the gene order in these regions. Until these regions are resolved, the gene order suggested by GENMAP will be maintained.

The CHROMAP program was written by Mike Palmer, an assistant professor in the Department of Botany at Oklahoma State University. This program is essentially a mathematical algorithm that reads a data file that contains information on the length of each chromosome and the number and position of each gene. CHROMAP then draws lines corresponding to length of each chromosome and places the gene symbols and appropriate map positions on each of the chromosomes. When two genes are very close and the gene symbols overlap, the program moves the gene symbols apart and then draws a line from the gene symbol to the appropriate position on the chromosome. This program then saves the resulting map in a plotting language that can then be read by plotting programs to generate a hard copy of the map. The commercially available plotting program PRINTAPLOT (Insight Development Corporation) was used to plot the updated linkage maps. These programs were essential for presenting the genetic map in a standard way. Original versions of the updated map were reproduced by a professional graphic artist at considerable expense; with CHROMAP and PRINTAPLOT, new genetic maps can be generated easily with an IBM-compatible computer and a Hewlett Packard printer in 15 minutes. Detailed instructions for how to use CHROMAP and PRINTAPLOT are shown in Appendix E (p.161).

Telotrisomic Analysis. The strategy used for mapping with telotrisomics is shown in Figure 22. Telotrisomic plants (EMB/EMB/EMB) were identified based on their characteristic morphology, then crossed with pollen from plants heterozygous for the lethal (EMB/emb). In this case, $F_1$ plants were scored as being diploid or telotrisomic and also for the presence of aborted seeds following self-pollination. The percentage of aborted seeds produced by telotrisomic $F_1$ plants was expected to be significantly less than 25% if the extra chromosome arm covered the gene of interest. The percentage of aborted seeds produced by diploid (EMB/emb) siblings from the same cross was determined as a control. Chi-square tests were used to determine whether the percentage of aborted seeds produced by telotrisomic $F_1$ plants was significantly less than that produced by diploid $F_1$ plants.

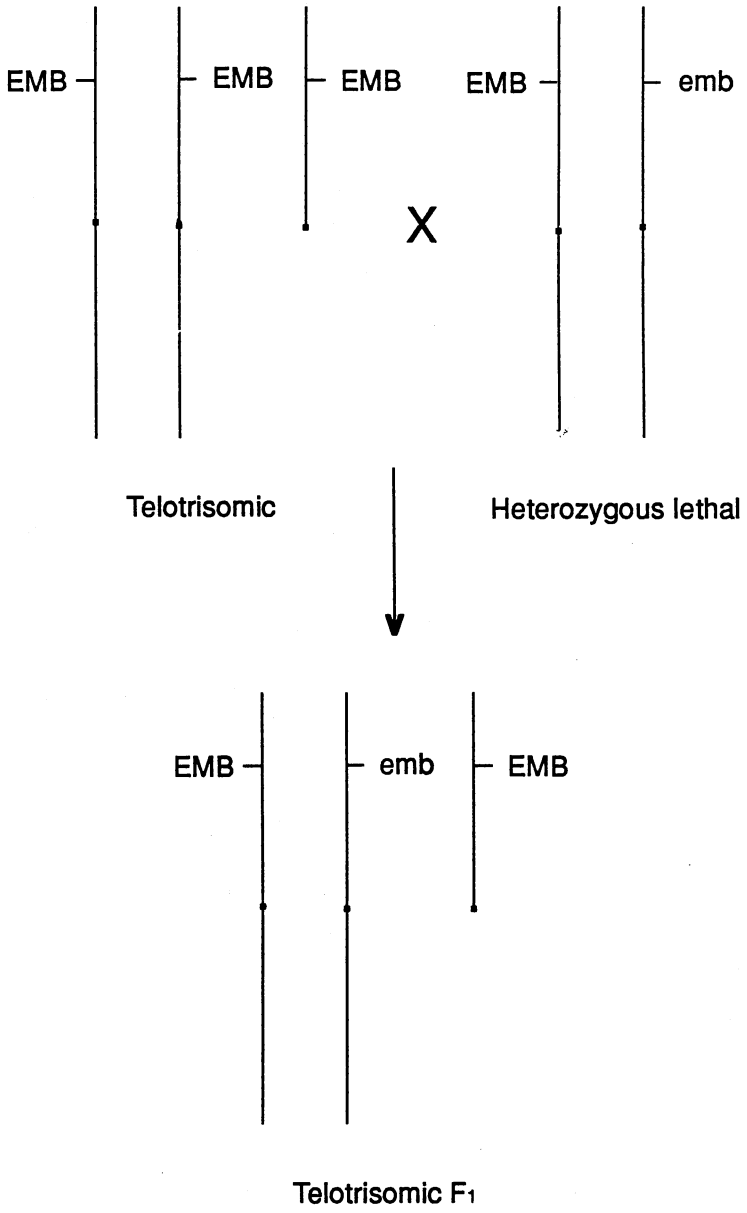



Figure 22. Drawing of Strategy Used for Mapping Embryonic Lethals With Telotrisomics. Telotrisomic plants are crossed with pollen from plants heterozygous for the lethal. $F_1$ plants are screened as diploid or telotrisomic based on plant morphology. The percentage of aborted seeds is calculated for telotrisomic $F_1$ plants that are segregating for the lethal and compared to the pecentage produced by sibling diploid plants that are segregating.

# Results

## Linkage Analysis

Three embryo-lethal mutants with interesting phenotypes (bio1, emb22, and emb30) were chosen for this initial mapping study in preparation for gene isolation. Visible markers used to map these genes had phenotypes that were easy to screen and did not significantly disrupt seed set. Four classes of progeny (A, B, C, D) were observed when $F_2$ plants were screened for both the visible marker and the lethal (Table 2). The AD method of linkage detection was used to assign $F_2$ plants to these four classes. The lethal and visible markers were in repulsion in the $F_1$ generation and linkage was detected by non-random segregation of the marker and lethal genes. Table 2 illustrates how recombinant progeny (A and D) become rare as the visible marker and the lethal get more tightly linked. Only two classes of progeny (E and F) were observed when $F_2$ plants were screened for just the visible marker. Table 2 shows that the ratio of phenotypically normal (E) to mutant (F) progeny should be 3:1 if the visible marker and the lethal are unlinked and 2:1 if the two genes are adjacent. Linkage can therefore be detected using the EF method if the ratio of normal to mutant progeny is significantly less than 3:1.

The minimum number of $F_2$ plants required to detect linkage for each screening method is summarized in Table 3. Values were calculated by substituting the derivation formulae, shown in Table 2, into the chi-square formula as outlined in Appendix G (p.163). Equations were solved for N (number of plants) by entering values of p (recombination percentage) and critical values from the chi-square table for each level

TABLE 2

EXPECTED FREQUENCY OF $F_2$ PLANTS OBTAINED FOLLOWING CROSSES WITH VISIBLE MARKERS

| $F_2$ class | $F_2$ genotype | | Recombination frequency | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Marker | Lethal | 0.0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | Derivation[a] |
| A | +/- | +/+ | 0.000 | 0.063 | 0.120 | 0.170 | 0.213 | 0.250 | $1/3(P)(2P)$ |
| B | +/- | +/m | 0.667 | 0.607 | 0.560 | 0.530 | 0.507 | 0.500 | $2/3(1-P+P^2)$ |
| C | m/m | +/+ | 0.333 | 0.270 | 0.210 | 0.160 | 0.120 | 0.083 | $1/3(1-P)^2$ |
| D | m/m | +/m | 0.000 | 0.060 | 0.110 | 0.140 | 0.160 | 0.167 | $2/3(P)(1-P)$ |
| E | +/- | -- | 0.667 | 0.670 | 0.680 | 0.700 | 0.720 | 0.750 | A + B |
| F | m/m | -- | 0.333 | 0.330 | 0.320 | 0.300 | 0.280 | 0.250 | C + D |

[a] Method for determining frequency of $F_2$ plants in each class from Servítová and Cetl (1984), where P = recombination frequency.

TABLE 3

MINIMUM NUMBER OF $F_2$ PLANTS REQUIRED FOR LINKAGE DETECTION[a]

| Recombination frequency[b] | Screening method[c] | Level of significance | | |
|---|---|---|---|---|
| | | P = 0.05 | P = 0.01 | P = 0.005 |
| 0.1 | AD | 13 | 18 | 20 |
| | EF | 113 | 195 | 232 |
| 0.2 | AD | 27 | 38 | 43 |
| | EF | 147 | 255 | 303 |
| 0.3 | AD | 73 | 105 | 119 |
| | EF | 254 | 438 | 521 |
| 0.4 | AD | 358 | 518 | 587 |
| | EF | 800 | 1384 | 1646 |

[a] Numbers in the table were calculated using formulae derived in Appendix 1.

[b] Between visible marker and embryonic lethal.

[c] Refer to Table 2 and text.

of significance. One striking feature of Table 3 is the small number of $F_2$ plants required for linkage detection with the AD method. For example, fewer than 150 $F_2$ plants are needed to detect genes separated by 30% recombination. If additional plants are screened, linkage can be detected between genes separated by even greater distances. For example, linkage was detected between an and emb22, two genes separated by over 40% recombination on chromosome 1, after screening only 467 $F_2$ plants with the AD method. In this case the number of plants in classes A-D was significantly different from that expected for unlinked genes, $\chi^2 = 16.1$; $P < 0.005$ (see Appendix H, p.166 for calculation).

It appears from Table 3 that the EF method could also be used to detect linkage between genes separated up to 40% recombination, if a sufficient number of $F_2$ plants were screened. The advantage to this approach of linkage detection would be to eliminate the effort required to screen $F_2$ plants for the embryonic lethal (dissecting siliques). In practice, the EF method was not always reliable for linkage detection. In one case 453 $F_2$ seedlings were screened following crosses between chl and emb22 and linkage was not detected ($\chi^2 = 1.6$) even though these genes are separated by less than 5% recombination on chromosome 1 (see Appendix I, p.168 for calculation). In a cross between ttg and bio1, two genes separated by more than 30% recombination on chromosome 5, linkage was detected despite the distance after screening 1364 $F_2$ seedlings ($\chi^2 = 32.4$; $P < 0.005$; Appendix J, p.170). In other cases, the EF method gave misleading results and suggested that genes were linked when in fact the genes were unlinked. Effects of mutant alleles on gametogenesis (certation), reduced viability of homozygous mutant plants, and relatively small sample sizes appear to be largely

responsible for inconsistent results obtained with the EF method. Linkage was readily detected with the AD method by screening 250 $F_2$ plants for both the visible marker and the embryonic lethal. Numbers obtained with the AD method were not only used to detect linkage with chi-square tests but were also used to estimate the percent recombination with the RECF2 computer program. Recombination estimates obtained with this program were accompanied by a standard deviation and an internal chi-square value that compared the observed frequencies of plants in classes A-D with those expected for the estimated level of recombination.

Table 4 summarizes estimates of recombination between embryonic lethals and linked visible markers generated by the RECF2 computer program. Internal chi-square values were generally less than 5.0 with three degrees of freedom, indicating that results obtained were close to those expected for the estimated level of recombination. The consistency of this mapping method is illustrated by results obtained from crosses between W2 (an, dis1, er) and emb30. After screening only 219 $F_2$ plants, the estimated percentages of recombination with emb30-1 were 2.4% (dis1), 10.2% (an), and 45.0% (er). An additional 203 $F_2$ plants were then screened and the final combined estimates were 2.8% (dis1), 10.5% (an), and 49.6% (er). A second allele at this locus (emb30-2) was then mapped with W2 by Linda Franzmann in our laboratory; she screened 230 $F_2$ plants and found recombination estimates that were 1.4% (dis1), 11.3% (an), and 50.6% (er). We are therefore confident that recombination estimates obtained with this method are reproducible and that linkage can routinely be detected between an embryo-lethal mutation and a visible marker separated by up to 35% recombination.

TABLE 4

RECOMBINATION ESTIMATES BETWEEN EMBRYONIC LETHALS AND VISIBLE MARKERS

| Lethal | Marker | Chromosome | Percent recombination[a] | $F_2$ plants screened |
|---|---|---|---|---|
| bio1 | yi | 5 | 11.2 ± 1.5 | 423 |
| | tt3 | 5 | 18.5 ± 2.7 | 221 |
| | ttg | 5 | 35.3 ± 3.0 | 423 |
| | an | 1 | 49.2 ± 5.0 | 222 |
| | dis1 | 1 | 43.4 ± 4.7 | 224 |
| | ch1 | 1 | 63.8 ± 5.3 | 201 |
| | gl2 | 1 | 57.0 ± 5.4 | 201 |
| | ap1 | 1 | 71.5 ± 4.8 | 201 |
| | er | 2 | 50.9 ± 3.6 | 428 |
| | gl1 | 3 | 43.3 ± 3.7 | 361 |
| | cer7 | 3 | 58.1 ± 0.4 | 361 |
| | cer2 | 4 | 62.9 ± 5.0 | 226 |
| | ap2 | 4 | 58.5 ± 5.1 | 226 |
| emb22 | ch1 | 1 | 4.0 ± 0.8 | 450 |
| | dis2 | 1 | 19.1 ± 2.7 | 236 |
| | clv2 | 1 | 24.7 ± 3.6 | 185 |
| | dis1 | 1 | 31.3 ± 2.6 | 466 |
| | an | 1 | 40.0 ± 3.1 | 466 |
| | er | 2 | 51.6 ± 3.5 | 466 |
| | cer2 | 4 | 51.2 ± 5.1 | 221 |
| | ap2 | 4 | 51.3 ± 5.1 | 221 |

TABLE 4 (Continued)

| Lethal | Marker | Chromosome | Percent recombination[a] | $F_2$ plants screened |
|---|---|---|---|---|
| emb30-1 | dis1 | 1 | 2.8 ± 0.7 | 422 |
| | an | 1 | 10.5 ± 1.4 | 422 |
| | er | 2 | 53.0 ± 3.1 | 621 |
| | cer2 | 4 | 42.6 ± 4.9 | 199 |
| | ap2 | 4 | 39.8 ± 4.9 | 199 |
| emb30-2 | dis1 | 1 | 1.4 ± 0.6 | 230 |
| | an | 1 | 11.3 ± 2.0 | 230 |

[a] From the RECF2 computer program ± standard deviation.

The data shown in Table 4 were then combined with recombination estimates obtained from Maarten Koornneef and Linda Franzmann (our laboratory) and entered into GENMAP to determine chromosomal locations of mutant genes and produce the updated linkage map shown in Figure 23. It should be noted that Linda Franzmann (with the help of Karl Hansen and Leigh Mickelson) mapped the additional 13 embryonic lethals using procedures worked out during the course of this study and outlined in this chapter. Suspect data sets (GENMAP $\chi^2 > 5.0$) were maintained in both this map and the previously published map (Koornneef 1990). GENMAP estimates map positions without definitively establishing gene order. The precise locations of some genes on the linkage map therefore remains to be determined. The positions of emb genes were established through backcrosses as described below for emb30 or, estimated with GENMAP by identifying the gene order that gave the lowest cumulative chi-square value.

Backcross Analysis

Backcrosses were used to resolve the position of emb30 relative to linked markers on chromosome 1 (Table 5). Recombination data with an suggested that emb30 was located between an and dis1, but GENMAP placed emb30 below dis1. A further complication was the fact that many of the recombination estimates for genes in this region of chromosome 1 were considered suspect by GENMAP. Backcrosses were performed between W2 (an, dis1) and emb30-1. Eight progeny were identified in classes V and VI (single recombinants if the gene order is an-dis1-emb30) whereas no progeny were identified in classes VII and VIII (double recombinants if the gene order is an-dis1-emb30). The correct gene order was therefore

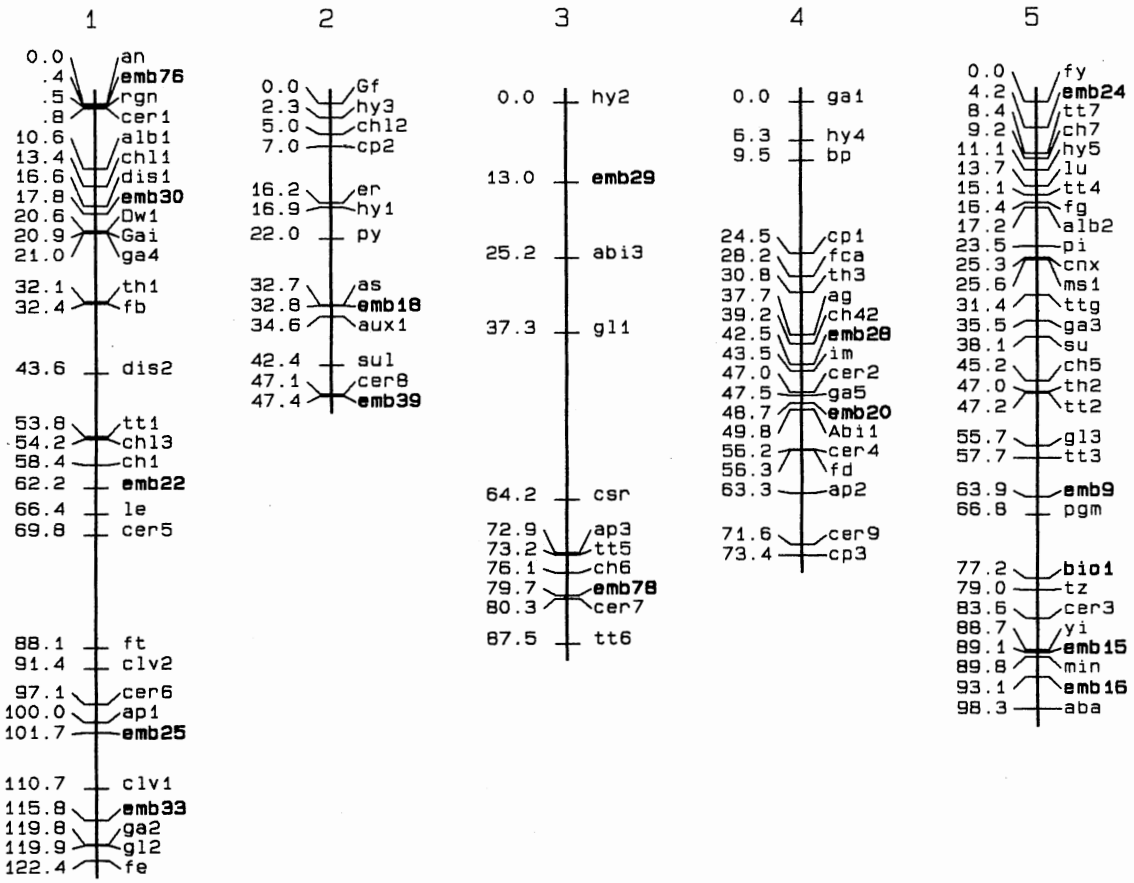



Figure 23. Linkage Map of Arabidopsis thaliana With Embryonic Lethals in Bold. Reprinted from Patton et al. (submitted).

TABLE 5

RESULTS FROM BACKCROSSES BETWEEN emb30-1 AND LINKED VISIBLE MARKERS

| Progeny class[a] | Progeny genotype | | | Crossovers required[b] | Number of progeny |
|---|---|---|---|---|---|
| | an | dis1 | emb30 | | |
| I | m/m | m/m | +/+ | 0 | 59 |
| II | +/- | +/- | +/m | 0 | 101 |
| III | m/m | +/- | +/m | 1 | 20 |
| IV | +/- | m/m | +/+ | 1 | 3 |
| V | m/m | m/m | +/m | 1 | 2 |
| VI | +/- | +/- | +/+ | 1 | 6 |
| VII | m/m | +/- | +/+ | 2 | 0 |
| VIII | +/- | m/m | +/m | 2 | 0 |

[a] Classes VII and VIII should be rare if the gene order is an-dis1-emb30 as shown. Classes V and VI should be rare if the gene order is an-emb30-dis1.

[b] Minimal number of crossovers required in $F_1$ parent to generate observed progeny.

determined to be an-dis1-emb30.

One surprising feature of the backcross results was the unequal frequency of plants obtained in reciprocal classes. For example, classes I and II were not represented equally among the backcross progeny (Table 5, p.77). Only 59 backcross progeny were identified in class I while the remaining 101 were in class II ($\chi^2 = 11.0$; $P < 0.005$; Appendix K, p.172). Classes III and IV were also not represented equally among the backcross progeny. Similar results were also observed by Linda Franzmann following backcrosses with other visible markers and embryonic lethals. Certation and differences in genetic background are two factors that may disrupt the expected distribution of progeny in backcross experiments. Additional backcrosses will nevertheless be required to definitively establish gene order in certain regions of the linkage map.

Telotrisomic Analysis

Telotrisomics were also used to assign genes to linkage groups. This method was explored because aneuploids are particularly useful for linkage detection. In this method, results are obtained in the $F_1$ generation without the need to grow large numbers of $F_2$ plants. Telotrisomics Tr1a and Tr1b were chosen for this study because their distinctive phenotypes facilitated rapid identification in segregating populations. If the embryonic lethal is covered by the extra arm in a segregating $F_1$ plant (EMB/emb/EMB), then siliques produced by that plant should contain significantly less than 25% aborted seeds following self-pollination. Results from crosses between two embryonic lethals and their corresponding telotrisomic (Tr1a x emb30; Tr1b x emb22)

demonstrated that segregation ratios were indeed reduced when the duplicated arm covered the embryonic lethal (Table 6). The percent aborted seeds produced by segregating telotrisomics (EMB/emb/EMB) was determined to be 14.5% for Tr1b x emb22 and 15.9 to 16.6% for Tr1a x emb30 (Table 6). Two types of controls were included in these experiments. The first involved diploid plants that were identified in the $F_1$ generation. Segregation ratios were calculated for EMB/emb heterozygotes identified as diploid (D) in column 3 of Table 6, to demonstrate that approximately 25% aborted seeds were produced in the absence of the duplicated chromosome arm. The second control involved crosses between embryonic lethals and telotrisomics with duplicated arms not covering the embryonic lethal. Results from these crosses (Tr1a x emb22; Tr1b x emb30) demonstrated that segregation ratios were not reduced when the extra arm did not cover the gene of interest (Table 6). Results obtained with Tr1a x emb22 and Tr1b x emb22 also demonstrated that this locus (emb22), known from recombination data to be near the centromere, was located on the lower arm of chromosome 1. The bio1 locus was not mapped with this technique because the telotrisomics for chromosome 5 could not be identified in segregating populations.

The ratio of aborted seeds produced by telotrisomic $F_1$ plants was expected to be determined by transmission rates of monosomic and disomic gametes (Koornneef and Van der Veen 1983) and distances between the embryonic lethal and the centromere (Koornneef 1983). In reciprocal crosses with wild-type Columbia plants, the transmission rate for Tr1b disomic female gametes was 33% (34 telotrisomic and 69 diploid $F_1$ plants). The transmission rate of male disomic gametes was 10% (3 telotrisomic and 27 diploid $F_1$ plants). These values are similar to

TABLE 6

LINKAGE DETECTION WITH TELOTRISOMICS

| Cross | EMB location | $F_1$ plant[a] | Percent aborted seeds | Number of seeds screened | Chi-square[b] |
|---|---|---|---|---|---|
| Tr1b x emb22 | 1b | T | 14.5 | 1091 | 32.2*** |
| Tr1b x emb22[c] | " | D | 21.5 | 995 | -- |
| Tr1a x emb22[c] | 1b | T | 20.1 | 634 | 0.9 |
| Tr1a x emb22[c] | " | D | 21.6 | 713 | -- |
| Tr1a x emb30-1 | 1a | T | 15.9 | 1189 | 52.4*** |
| Tr1a x emb30-1 | " | T | 16.6 | 823 | 30.6*** |
| Tr1b x emb30-1[c] | 1a | T | 22.6 | 836 | 0.7 |
| Tr1b x emb30-1[c] | " | T | 23.2 | 757 | 0.1 |
| Tr1b x emb30-1[c] | " | D | 24.5 | 1117 | -- |
| Tr1b x emb30-1[c] | " | D | 23.2 | 1028 | -- |

[a] Screened as diploid (D) or telotrisomic (T) based on plant morphology. Individual $F_1$ plants are listed separately to demonstrate the consistency of results obtained.

[b] Calculated using non-telotrisomic controls as standards for expected values. *** Significantly different from expected at $P \leq 0.005$.

[c] Negative controls involving either non-telotrisomic $F_1$ plants or telotrisomics with the duplicated arm not covering the gene of interest.

those reported previously (27% female and 7% male) by Koornneef and Van der Veen (1983). It should be noted that Koornneef and Van der Veen (1983) obtained their values following reciprocal crosses with wild-type plants in the Landsberg background. The transmission rates for crosses with Columbia plants was therefore not expected to be identical to those obtained following crosses with Landsberg.

The transmission rate for Trla gametes was more difficult to calculate in reciprocal crosses with wild-type Columbia, because the Trla seed stock segregated for an apetalous phenotype not previously reported. It was not determined whether this class of progeny represented tetrasomic individuals or simply plants segregating for a recessive mutation. Following crosses with pollen from either emb22 or emb30-1, 33.8% female transmission of the Trla disomic gametes was observed in $F_1$ plants (22 telotrisomic and 43 diploid), which appeared to be consistent with the published transmission rate of 28% for female Trla gametes (Koornneef and Van der Veen 1983).

## Discussion

The value of saturating genetic maps has been demonstrated in a variety of organisms, ranging from viruses to eukaryotes. The linkage map of Arabidopsis now contains over 100 visible markers (Koornneef 1990; Figure 23, p.76) but is still far from saturation. Now that Arabidopsis is considered a model system for plant biology, advances must be made in both classical genetics as well as the molecular genetics, in order to complement the current focus on the analysis of the Arabidopsis genome. The standard linkage map of Arabidopsis has been constructed largely by the examination of $F_2$ progeny from selfed $F_1$ plants (Koornneef et al. 1983; Koornneef and Stam 1988; Patton et al. submitted). This mapping method minimizes the number of backcrosses required for mapping, but often fails to establish gene order. Consequently there is some uncertainty associated with both the precise gene order and location of some genes on the updated linkage map shown in Figure 23 (p.76). Questions also remain concerning the effect of environmental factors and sex differences on observed recombination frequencies in Arabidopsis. Zhuchenko et al. (1988) conclude that in Arabidopsis, there are up to 5-fold differences in the observed recombination frequencies, over certain intervals, between the male and female gametes. Additional studies will therefore be required to improve both the density and the resolution of the linkage map.

The efficiency of mapping genes in Arabidopsis with chlorophyll-deficient lethals was noted previously by Servítovà and Cetl (1984). A similar approach has now been used to map embryonic lethals, now the most common class of marker on the genetic map. Linkage can easily be

detected between visible markers and embryonic lethals that are separated by 35% recombination, which translates into about 45 cM with the Kosambi (1944) mapping function (Appendix F, p. 161). This is a conservative estimate of the distance over which visible markers are useful. In the future it should be possible to use new linkage tester lines (described in Chapter 3), that contain visible markers near the center of each chromosome, to detect linkage with over 85% of the new embryonic lethals.

Many additional mutants defective in embryo development have been identified in Arabidopsis but remain to be placed on the linkage map. It should therefore be possible to improve the resolution and utility of the linkage map by mapping additional embryonic lethals. The efficiency of mapping embryonic lethals with visible and RFLP markers (discussed in Chapter 4), and the wide distribution of emb genes throughout the genome should also facilitate the integration of the standard genetic and molecular maps of Arabidopsis.

Considerable effort will be required to saturate for the embryo-lethal phenotype and assign these mutants to complementation groups. Large-scale allelism tests in Arabidopsis are feasible only when working with mutant phenotypes determined by relatively few target genes. The availability of multiple alleles has nevertheless played an important role in the genetic and molecular analysis of developmental pathways in other organisms and would clearly facilitate the analysis of embryonic lethals in Arabidopsis. The most efficient way to identify multiple alleles of embryonic lethals in Arabidopsis may be to assign new mutants to linkage groups, and then perform complementation tests with mutants in the same region. This approach has already led to the identification

of a new allele of emb20 (Linda Franzmann, our lab) and might simplify the task of cataloguing mutants identified in other laboratories.

The addition of embryonic lethals to the Arabidopsis linkage map should also facilitate the construction of balanced lethal chromosomes. Balancers with defined chromosomal inversions and flanking lethal mutations have been used extensively in Drosophila and Caenhorhabditis to enforce heterozygosity, maintain mutagenized chromosomes, and isolate new alleles at a locus of interest. Now that a number of embryonic lethals have been mapped (Patton et al. submitted), the main obstacle to constructing balancers is the identification of appropriate inversions. The cytogenetic approach to isolating inversions has proven successful in maize (Carlson 1988) but is not likely to be feasible in Arabidopsis because of the small size of meiotic chromosomes (Schweizer et al. 1988). Inversions in Arabidopsis will likely have to be identified by screening for segregation distortion following mutagenesis with X-rays.

Previous studies with maize have demonstrated the value of constructing genetic mosaics to examine the role of specific genes in plant development (Hake and Freeling 1986; Poethig 1988). The approach used in these studies was to map the gene of interest, identify a closely-linked visible marker with a cell-autonomous phenotype, then use X-irradiation to generate marked sectors that lacked the dominant allele. A similar approach could be used to determine whether recessive embryonic lethals of Arabidopsis are defective in a diffusible product required for later stages of development. Mapping embryonic lethals of Arabidopsis should therefore not only strengthen the classical genetics of this model system, but also our understanding of the molecular basis of plant growth and development.

# CHAPTER III

# CONSTRUCTION OF NEW MULTIPLE MARKER LINES TO FACILITATE MAPPING OF EMBRYONIC LETHALS

## Introduction

Mapping embryonic lethals of Arabidopsis has been facilitated by the use of marker lines that contain two or more visible mutations. These multiple marker lines (described in Chapter 2) are especially useful when performing backcrosses because they usually contain two linked mutations which can be used to identify the correct placement of linked mutations. These marker lines however, are not efficient for initial linkage detection because they cover at most two different chromosomes. Each multiple marker line carries er on chromosome 2 in addition to the other visible markers on different chromosomes.

A supertester line, called W100, has been constructed in Arabidopsis that contains two visible markers on each of the five chromosomes (Koornneef et al. 1987) that should, in theory, greatly reduce the amount of time required to map new genes. There are a number of reasons why this multiple line has not been used widely among members of the Arabidopsis community. The most common complaints about W100 are the difficulty in maintaining W100 and in screening $F_2$ plants for each of the marker phenotypes. The plants from the W100 line also have a male sterile mutation (ms-1), maintained in the heterozygous state, which obviously disrupts seed set. The ms-1 mutation would have to be

segregated out of W100 before it could be used to map embryonic lethals because the mapping method requires screening mature seeds of $F_2$ for the lethal phenotype. Another problem with the W100 line is that one of the mutations, py, results in plants that require pyrimidine for normal growth. The problem with this marker is that $F_2$ plants must be grown in the absence of pyrimidine in order to screen for this marker. Those $F_2$ plants that are homozygous recessive for py will become pale and die as seedlings; these plants must then be rescued by watering with a pyrimidine-containing solution before they can be screened for the remaining markers. The problem with this kind of marker is the time involved in, and difficulty associated with screening for this phenotype. The only way to be sure that $F_2$ plants are screened correctly is to test progeny from each $F_2$ plant for their ability to grow in the absence of pyrimidine; the ratio of $F_3$ progeny that grow to those that do not grow, indicates the genotype of the $F_2$ parent.

In order to circumvent the problems associated with W100, three new multiple marker lines were constructed that could be used to efficiciently detect linkage between embryonic lethals and linked visible markers. One of the multiple marker lines (DP23) contains 5 mutations, each near the center of one of the chromosomes. With this line it should be possible to detect linkage with most embryonic lethals after screening only one set of $F_2$ plants. This assumes that embryo-lethal mutations will be distributed randomly throughout the genome, which they appear to be (see Figure 23, p.76). The other two marker lines, DP27 and DP28, contain mutations that are located near the ends of chromosomes 1, 3, 4, and 5 which should complement areas not covered by DP23.

## Materials and Methods

### Plant Material

Plants were grown and maintained as described in chapter 2. Visible markers used in the construction of multiple markers were obtained from Maarten Koornneef during December of 1987 and maintained in the Landsberg (er) background. The locations of these markers are shown in Figure 24. The phenotypes of these mutations, with the exception of bp and tt5, are described in the Materials and Methods section of Chapter 2. The bp mutation, on chromosome 4 at 9.5 cM, reduces the distance between siliques on the stem and, more prominently, causes mature siliques to point downward. Plants with the bp mutation are easy to screen at maturity whether in the Landsberg or Columbia background. The tt5 mutation is located at 73.3 cM on chromosome 3 and has a phenotype identical to that of tt3 (yellow seeds at maturity).

The strategy used to create multiple marker lines from single markers is outlined in Figure 25. In general, the strategy was to cross one marker line with pollen from the other marker line. The $F_1$ progeny were screened for the maternal phenotype to make sure that accidental selfs were no longer maintained. The desired $F_1$ plants were allowed to self and the dry $F_2$ seeds were collected. $F_2$ plants were screened for all appropriate phenotypes and dry seeds were saved from those plants that carried all of the recessive marker mutations. The integrity of each new multiple marker line was verified by planting and screening the next generation of seeds.

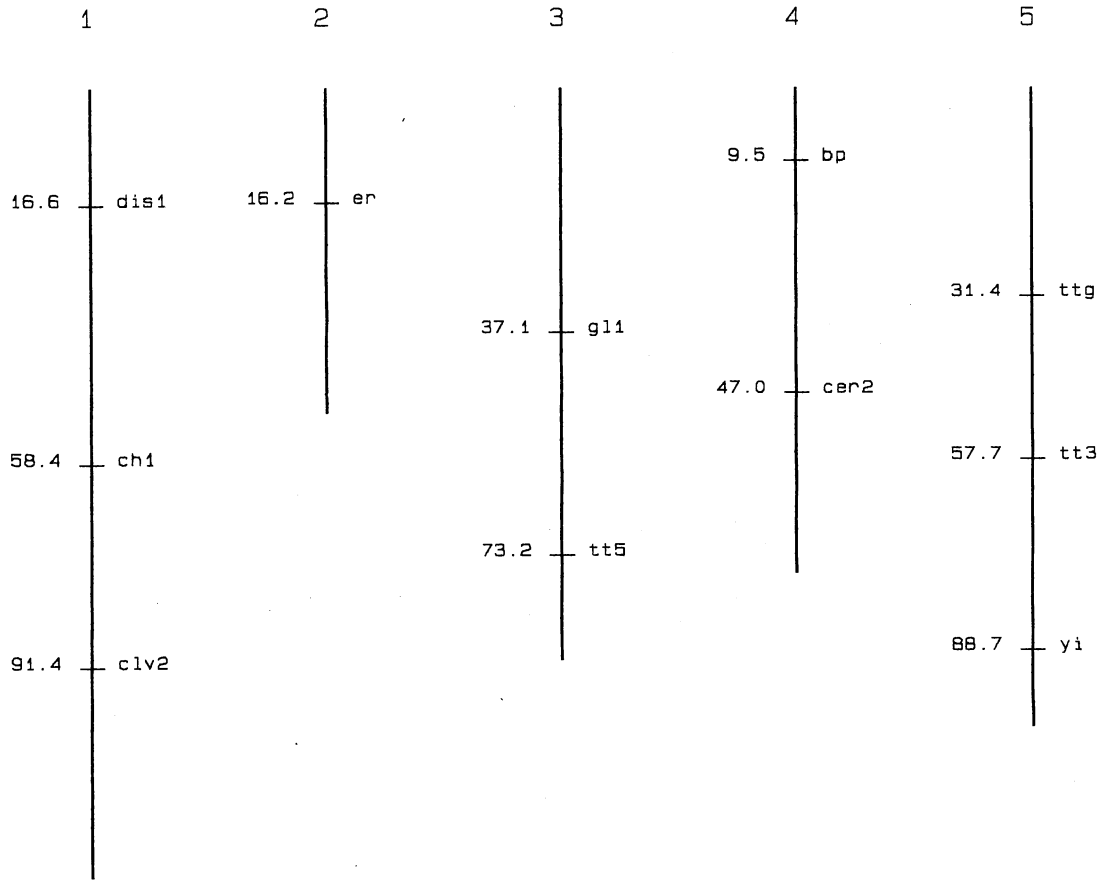



Figure 24. Linkage Map With Visible Markers Used in Multiple Marker Construction. DP23 contains ch1, er, gl1, cer2, and tt3. DP27 contains er, bp, yi, and ttg. DP28 contains dis1, clv2, er, and tt5. Map positions from Patton et al. (submitted)

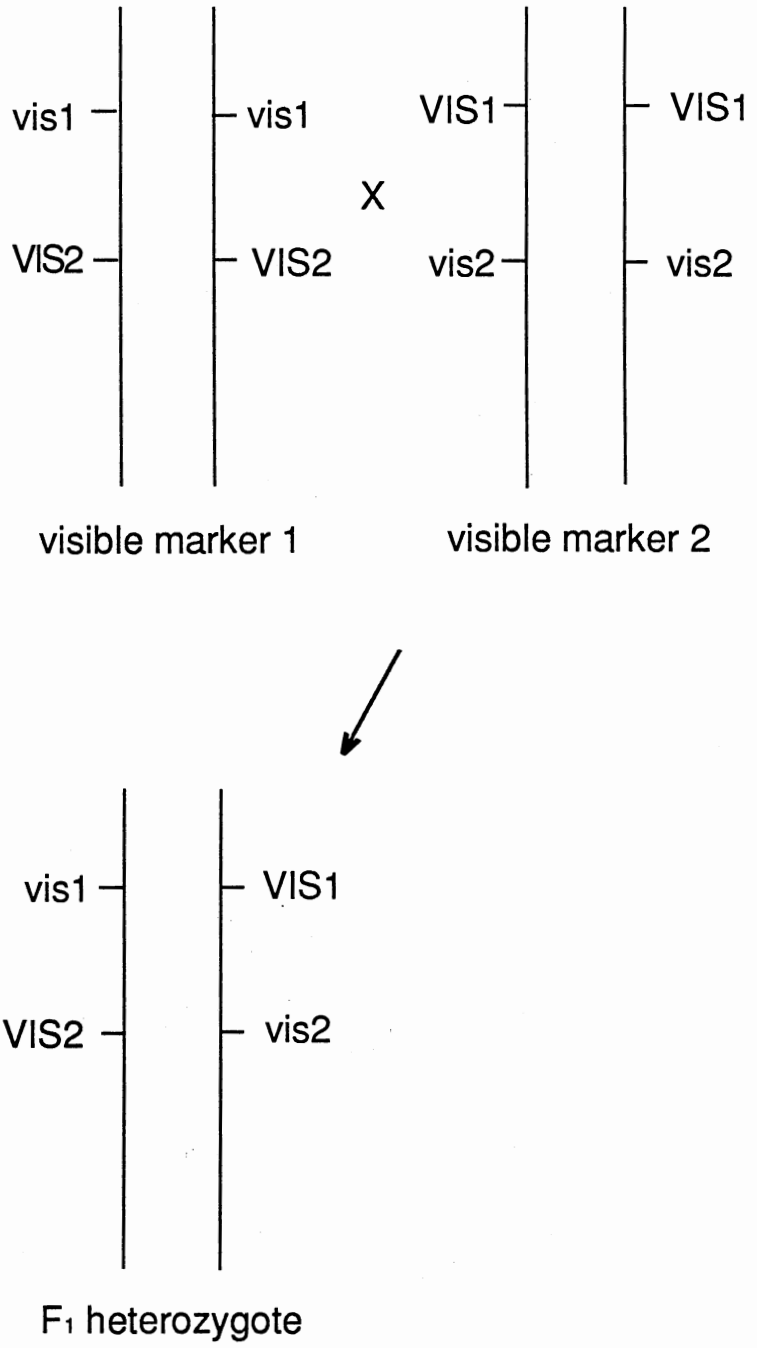



Figure 25. Drawing of Strategy Used to Create Multiple Markers From Existing Visible Markers. Visible marker 1 is crossed with visible marker 2. The resulting $F_1$ double-heterozygote is grown to maturity and allowed to self-pollinate. $F_2$ seeds are then planted and screened for both recessive traits.

Multiple Marker Construction

The DP23 Multiple Marker Line. This multiple marker line contains a total of 5 recessive mutations; chl, er, gl1, cer2, and tt3. The construction of DP23 was slowed because cer2 was not available as a monogenic line, but was part of W10, a multiple marker that also carried ap2. The ap2 mutation was considered undesirable for DP23 because the apetalous flowers often required manual pollination to get adequate seed set for mapping embryonic lethals. The strategy used to obtain cer2 as a monogenic line (shown in Figure 26) was to cross W10 (cer2 ap2/cer2 ap2) with Landsberg erecta wild-type plants (CER2 AP2/ CER2 AP2), save $F_2$ seeds from $F_1$ plants (CER2 AP2/cer2 ap2), and then to screen the $F_2$ progeny for plants that were homozygous recessive for only the cer2 mutation (cer2 AP2/cer2 -). $F_3$ seeds were saved from each of the cer2/cer2 $F_2$ plants, and then tested for the presence of a recessive ap2 allele (AP2/ap2). The majority of the cer2/cer2 $F_2$ plants were heterozygous (AP2/ap2), however several clean cer2/cer2 lines were found and subsequently used in the construction of DP23.

The strategy used to construct DP23 was to make crosses between chl and gl1, and also to make crosses between cer2 and tt3. $F_2$ seeds were saved from $F_1$ plants from each of the crosses (either CH1 gl1/chl GL1 or CER2 tt3/ cer2 TT3). $F_2$ plants that carried both mutations (either chl gl1/chl gl1 or cer2 tt3/cer2 tt3) were identified and seeds were saved. These new marker lines were then crossed to each other and again $F_2$ seeds were saved from $F_1$ plants. A large number of these $F_2$ seeds were planted and screened before a single plant was found that was homozygous for each of the five mutations (chl er gl1 cer2 tt3/chl er gl1 cer2 tt3).

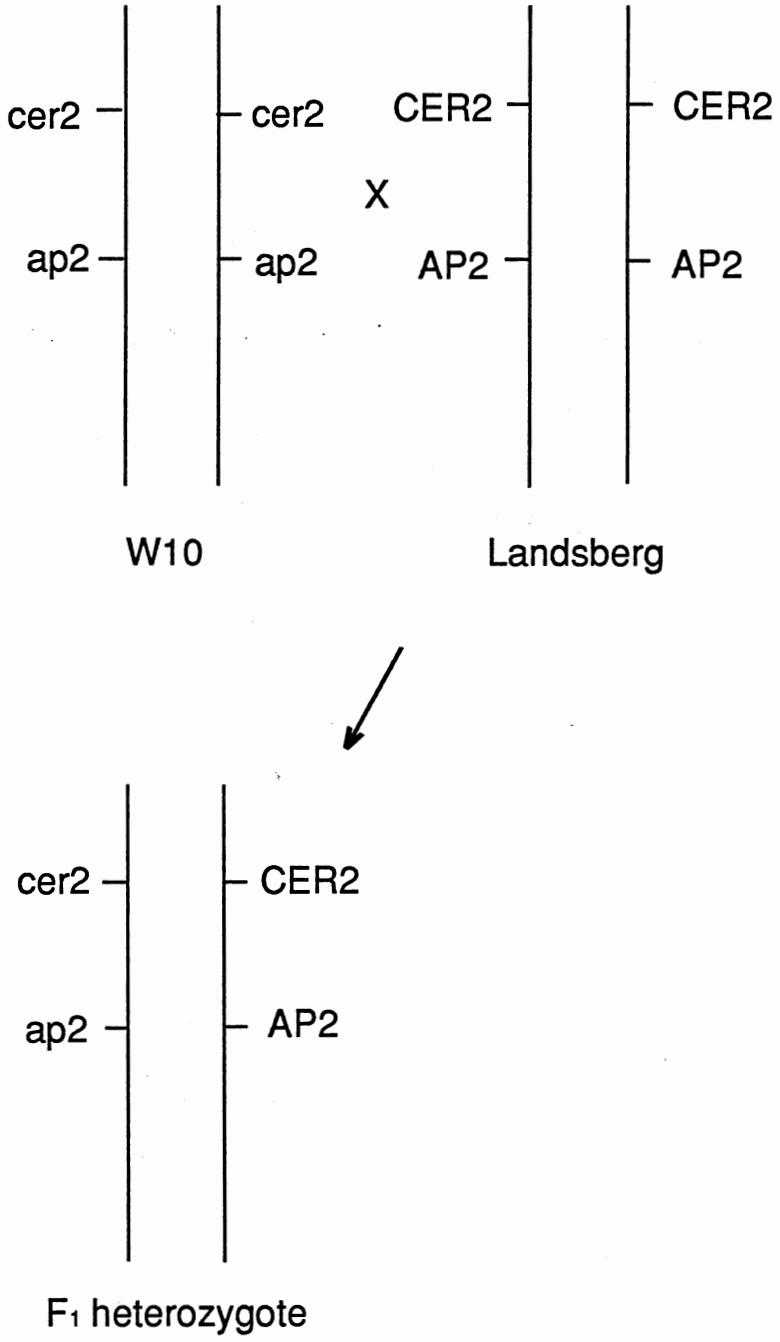



Figure 26. Strategy Used to Isolate cer2 as a Monogenic Line From W10. W10 (cer2 ap2/cer2 ap2) was crossed with pollen from Landsberg wild-type (CER2 AP2/CER2 AP2). The resulting $F_1$ double-heterozygote in this case contained the recessive alleles in the cis configuration and was allowed to self-pollinate. $F_2$ plants that were homozygous recessive cer2/cer2 were progeny then tested in the $F_3$ generation for the presence of the recessive ap2 allele.

The DP 28 Multiple Marker Line. This multiple marker carries 4 mutations: dis1 and clv2 on chromosome 1, er on chromosome 2, and tt5 on chromosome 3. This multiple marker should be able to detect linkage with mutations on the distal portions of chromosome 1, all of chromosome 2, and the lower half of chromosome 3. One of the markers, dis1, was not available as a monogenic line, but was easily segregated away from an in the W2 line using the same strategy employed to recover cer2. The strategy used in the construction of DP28 was to cross dis1/dis1 with clv2/clv2 first and to recover dis1 clv2/dis1 clv2 $F_2$ plants, then cross these plants with tt5/tt5 and finally recover dis1 clv2 tt5/dis1 clv2 tt5 plants.

The DP27 Multiple Marker Line. This multiple marker contains four visible mutations; er on chromosome 2, bp on chromosome 4, and yi and ttg on chromosome 5. In this strategy, bp/bp was crossed with W13 which already contained yi and ttg. $F_2$ plants were identified that were homozygous recessive for all of the mutations. DP27 contains markers near the top of chromosome 4 and near the ends of chromosome 5. These positions were selected to complement the coverage missed by the DP23 and DP28.

## Results and Discussion

The new multiple marker lines DP23, DP27, and DP28 are summarized in Figure 24 (p.88). These multiple marker lines should decrease the amount of time and effort involved in mapping new genes. The major drawback in making new multiple marker lines is the time involved in their construction. For example, it takes at least 4 months just to get dry $F_2$ seeds from one set of crosses. In cases where visible markers had to be separated away from a multiple marker line (eg. cer2 and dis1), 4 months were again required just to get the $F_2$ seeds following crosses with wild-type. The next step was to screen $F_2$ plants for the desired phenotype and save dry $F_3$ seeds (2 months) and test these progeny to determine which plants were homozygous dominant for the other visible marker. These progeny testing experiments took an additional 2-3 months to finally identify plants that contained the desired genotype. In all, it took at least 8 months just to separate one visible marker from another in crosses between W2 and wild-type, and W10 and wild-type.

One other aspect to note was the number of $F_2$ plants that had to be screened before the monogenic lines were identified. In the cross between W2 (an dis1/an dis1) and wild-type (AN DIS1/AN DIS1), a total of 280 $F_2$ seeds were planted and only 7 had the desired phenotype (AN dis1/- dis1). Two of these plants were later shown to be wild-type (AN/AN) for the other marker, while the other 5 were heterozygous (AN/an) at this locus. This translates into 0.7% (2/280) $F_2$ plants with the desired genotype. This result was not too surprising because these two genes are separated by approximately 15% recombination. The expected frequency of the desired $F_2$ class was therefore the product of

the expected frequency of recombination between these two genes for the male and female gametes (0.075♀ x 0.075♂) or 0.56% which is very close to the observed value of 0.7%.

The isolation of cer2 as a monogenic line (described in Figure 26, p.91) was easier than expected because the other recessive mutation (ap2) segregated away from cer2 in the $F_2$ generation at a higher than expected frequency. In this case, 12 out 120 $F_2$ plants were recessive (cer2 AP2/cer2 -) for the first marker and five of those (4% of the total $F_2$ planted) were the desired genotype (cer2 AP2/cer2 AP2). This was surprising because these cer2 and ap2 are separated by exactly the same distance (16.6 cM) as an and dis1 on the linkage map. One would therefore expect approximately the same frequency of $F_2$ plants with a monogenic genotype. The difference observed in these segregating populations could be due to small sample size, peculiarities in the recombination frequency in that region, a negative effect of the recessive ap2 allele on gametogenesis (certation) or uncertainties in the genetic map. These differences were first thought to be due to small sample size, but a chi-square test suggests that the observed difference in this sample was significant ($\chi^2 = 27.9$; $P < 0.005$). A second explanation for these differences could be the phenomenon observed by Zhuchencko et al. (1988) where extreme differences were measured in the recombination frequencies between the male and female gametes over certain intervals. It is interesting to note that over the interval between cer and ap2, four times more recombination was measured in male gametes than in female gametes, which was one of the largest differences for all intervals tested. Another reason for having more plants than expected in the desired $F_2$ class could be that the ap2

allele exhibits certation. This hypothesis however is not supported by the frequency of ap2/ap2 plants observed in 870 $F_2$ plants; if anything, the ap2 allele contributed to fertilization more frequently than the AP2 allele in these populations. This phenomenon then becomes hard to explain unless one considers that cer2 and ap2 may actually be separated by a greater distance on chromosome 4 than is reflected in the genetic map, which is a possibility given the uncertainties associated with genetic maps at this level of resolution.

In each case during the construction of DP27 and DP28, the frequency of obtaining double recessive $F_2$ plants following crosses between two linked visible markers (dis1/dis1 x clv2/clv2 and cer2/cer2 x bp/bp) was not significantly different from expected ($P > 0.05$). This was probably due to the fact that these sets of genes were separated by such great distances on their respective chromosomes that recombination occurred freely between them.

In the construction of DP23, the strategy relied on pure random assortment rather than recombination because the recessive mutations involved were on different chromosomes. In the final cross during DP23 construction (ch1 gl1/ch1 gl1 x cer2 tt3/cer2 tt3) approximately 1 out of 250 $F_2$ plants or $(0.25)^4$ was expected to have all of the recessive markers. In reality, approximately 500 $F_2$ plants were screened before one plant was identified as having all of the recessive traits, which is not significantly different from expected.

If we assume that linkage can be detected between loci separated by 35% recombination (Chapter 2), which translates into approximately 45 cM by the Kosambi (1944) mapping function (Appendix F, p.161), the DP23 multiple marker should cover about 88% (278 cM covered/430 cM total) of

the genome with respect to detecting linkage between embryonic lethals and the visible markers. Most of the genome not covered by DP23, about 33 cM on distal portions on chromosome 1 and 13 cM near the top end of chromosome 5, is covered by the other two markers DP27 and DP28. This means that if embryo-lethal mutations occur randomly throughout the genome, then 88% of the new mutations should be assigned to a linkage group with one cross. This marker should therefore greatly facilitate future complementation analysis among embryo-defective mutants. As already stated, this strategy requires that mutations be mapped to linkage group and then complementation tests will be limited only to linked mutations. This strategy is one step toward the goal of defining as many genes as possible in Arabidopsis that play a role during plant growth and development.

# CHAPTER IV

# MAPPING WITH RFLP MARKERS

## Introduction

In the 1960's, a group of bacterial enzymes was discovered that could digest incoming bacteriophage DNA during the process of infection (Boyer 1971). These enzymes were said to "restrict" the establishment of phage within the cell. These enzymes, called restriction enzymes (Lederberg and Meselson 1964), have been instrumental in the growth of modern molecular biology. The first restriction enzymes that were described (type I) bind at specific sites but require a number of cofactors for in vitro activity and appear to cleave DNA molecules at more or less random sites that are over 1000 base pairs away. Type III restriction enzymes also bind to DNA at specific sites but cleavage occurs over 20 bases away from the recognition site. Type I and type III restriction enzymes have therefore not been used extensively in molecular cloning procedures because the cleavage occurs at non-specific sequences. A third class of restriction enzymes (type II) requires only $Mg^{++}$ as a cofactor and binds and cleaves DNA at specific sequences. The sequence-specificity of type II restriction enzymes allows complex nuclear genomes to be reduced into a population of DNA fragments with discrete sizes with similar ends. At least 475 distinct restriction endonucleases have been isolated from bacteria; approximately 200 are type II restriction enzymes (Roberts 1988), most of which are

commercially available.

The restriction enzymes that are commonly used for restriction fragment length polymorphism (RFLP) analysis usually have a 4-6 base pair (bp) recognition sequence. The frequency of cleavage can be estimated by making the assumption that each of the four different nucleotides occurs randomly and at equal frequency in the genome. A 6 bp recognition enzyme would therefore be expected to cut once every 4096 or ($4^6$) bp within the genome, while a 4 bp recognition enzyme would cut once every 256 or ($4^4$) bp. In plants however, genomic restriction fragments range in size from a few bp to more than 20 Kb. There is also substantial variation in the frequency of cleavage for enzymes with recognition sequences that are the same size. For example, in tomato, the restriction enzyme Dra I (recognition sequence 5′ TTTAAA 3′) cuts nuclear DNA much more frequently than Sst I (GAGCTC) (Bernatzky 1988). In addition, restriction enzymes that are sensitive to methylated DNA do not cut well in plants because of the high degree of methylation.

When plant nuclear DNA is digested with a restriction enzyme, hundreds of thousands of fragments are generated. In order to study the restriction pattern from a given chromosomal locus, these fragments must be fractionated and the fragments of interest distinguished from all other similar-sized fragments. DNA fragments are generally separated by gel electrophoresis, and the individual fragments are visualized using Southern (1975) blotting and subsequent hybridization to a cloned and labeled homologous sequence.

At a given pH, all DNA fragments have roughly the same charge-to-mass ratio. The DNA molecules are negatively charged due to the sugar-phosphate backbone, and move at a constant rate through an electric

field regardless of size. However, when DNA is electrophoretically separated through a medium such as agarose, the fragments move according to size. The fragments can be thought of as "snaking" their way through small pores in the agarose matrix. Small DNA fragments will therefore move faster than larger fragments through agarose during electrophoresis. There is a good inverse relationship between the migration rate of DNA in agarose and the log of its molecular weight (Aaij and Borst 1972).

After electrophoresis, DNA fragments are transferred in their original pattern from the agarose to a membrane or filter. Because large DNA fragments do not transfer efficiently, the gel is first treated with dilute acid to partially depurinate the DNA molecules, then treated with a strong base to denature the double-stranded molecules. The single stands break at the depurinated sites thus liberating smaller fragments which can be efficiently transferred. The DNA is then covalently bound to the membrane in the single-stranded state. This allows labeled complementary strands of DNA (probes) to anneal to their specific homolog among the separated fragments. DNA probes can be labeled using radioactivity in the form of $^{32}P$, or with a nucleotide analog such as digoxigenin deoxyuridine-triphosphate. Fragments can then be identified by autoradiography or immunological detection.

An illustration of the theory behind RFLP analysis is shown in Figure 27. The difference (polymorphism) in the size of restriction fragments is due to variations in the distribution of restriction sites. These variations are caused by insertions, deletions, or the gain or loss of restriction sites by base substitutions. It should be noted that variation in restriction sites at a given locus does not

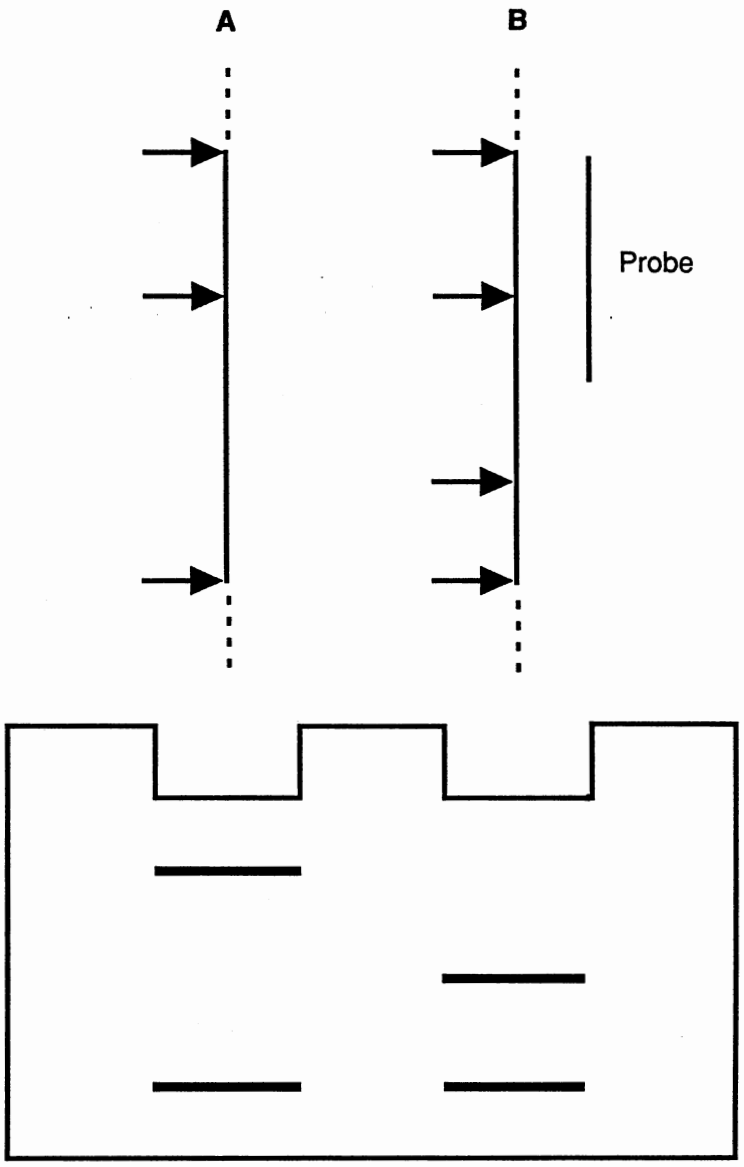



Figure 27. Schematic Drawing That Illustrates the Theory Behind RFLP Mapping. In this example, a small portion of a chromosome is shown for individuals A and B. The restriction sites (arrows) are shown for the locus of interest. Note that the probe shown at right detects the polymorphism (due to additional restriction site in B) between these two individuals on a Southern blot shown below.

necessarily mean that there is functional difference in the sequences examined. Since restriction sites are actual DNA sequences, the variation in these sites have been used to study genetic divergence between individuals (Engels 1981; Templeton 1983; Palmer et al. 1985).

Sequences that hybridize to a given probe come from discrete chromosomal loci. Different alleles at these loci can be detected by variation in restriction fragments. Restriction fragments are therefore ideally suited to be genetic markers. If the individuals in Figure 27 (p.100) were crossed, then the $F_1$ progeny would contain both polymorphic alleles and be recognized as a heterozygote on a Southern blot. This illustrates the codominant nature of RFLP markers. These alleles would segregate in a Mendelian fashion in backrossses or in the $F_2$ generation. Linkage can therefore be detected using a simple $\chi^2$ test when segregation is significantly different from expected for unlinked loci.

Grodziker et al. (1974) first demonstrated that restriction fragments could be used as genetic markers. In this study, a temperature-sensitive mutation was mapped in adenovirus by using a restriction fragment marker. Since then, extensive studies have resulted in the construction of RFLP linkage maps to help identify genetic diseases in man (Botstein et al. 1980; Bhattacharya et al. 1984; Stephans et al. 1990). This technique, in conjunction with chromosome walking, has already led to the discovery of the defective gene in over 70% of cystic fibrosis patients (Riordan et al. 1989). In plants, linkage analysis with DNA markers has been accomplished in maize (Evola et al. 1986), tomato (Vallejos et al. 1986), pea (Polans 1985), Arabidopsis (Chang et al. 1988; Nam et al. 1989), and many other species. The utility of RFLP markers in plant genetics has been

discussed previously (Burr et al. 1983; Tanksley 1983). Some of the applications include: estimating levels of variation in germplasm collections, monitoring the purity of hybrid seeds, selecting (via linkage) agronomically important traits, studying the components of quantitative traits, facilitating gene isolation through chromosome walking, and identifying the products of cell fusion experiments.

The major limitation to RFLP analysis is the cost. Restriction enzymes are expensive and large amounts of these enzymes are needed to analyze many individuals in a segregating population. The complexity of the genome being studied is reflected in the amount of DNA, and restriction enzyme required, for RFLP analysis. Plants with large genomes such as onion or cereals require up to 10 $\mu$g of DNA per individual for RFLP analysis, while in Arabidopsis only 1 $\mu$g of DNA per individual is required to detect single-copy sequences. Traditional Southern analysis requires the use of $^{32}P$ in the production of labeled probes. This isotope is short-lived, hazardous, and expensive. Non-radioactive techniques have been recently produced that allow the production of stable probes and detection of single copy sequences.

A safe and cost-effective method for high-resolution RFLP mapping of embryonic mutations of Arabidopsis is presented in this chapter. This method will facilitate the molecular isolation of genes that play an essential role during embryo development and complement the current analysis of the genetic control of plant embryo development. An example of this method is presented where the biol locus of Arabidopsis is mapped to within 0.5 cM of an existing RFLP marker.

## Materials and Methods

The wild-type ecotypes used for RFLP mapping were Landsberg (er), described in Chapter 2, and Niederzenz which was obtained from both Elliot Meyerowitz (California Institute of Technology) and Mark Estelle (Indiana University, Bloomington). Embryo-lethal mutants in the Columbia background were also described in Chapter 2. The strategy used to map embryonic lethals with RFLP markers is shown in Figure 28. Embryo-lethal mutants in the Columbia background were crossed with either Landsberg (er) or Niederzenz ecotypes to produce polymorphic $F_1$ plants. $F_2$ seeds were collected from heterozygous (EMB/emb) $F_1$ plants following self pollination. $F_2$ plants were then screened for the lethal as soon as the first siliques became mature. Batches of 5 wild-type (EMB/EMB) $F_2$ plants were frozen on dry ice and stored at -80° C in sealed plastic bags.

Total plant DNA was isolated from each pool of 5 plants using the procedure described by Richards (1990) which is outlined in Appendix L (p.174). DNA yields were typically 50-80 $\mu$g per pool as measured by fluorescence in the presence of Hoescht dye #33258 using a mini-fluorometer (Hoefer). Plant DNA was cut with 5-fold excess of restriction enzyme (Bethesda Research Labs), separated on 0.75% agarose gels, and then transferred to Nytran (Schleicher & Schuell) membranes with a vacuum blotter and crosslinked with UV irradiation (Stratagene) as described in Appendix L.

Molecular markers used in RFLP mapping were lambda clones from the RFLP map of Chang et al. (1988) and obtained from Elliot Meyerowitz (California Institute of Technology, Pasadena). Details concerning the

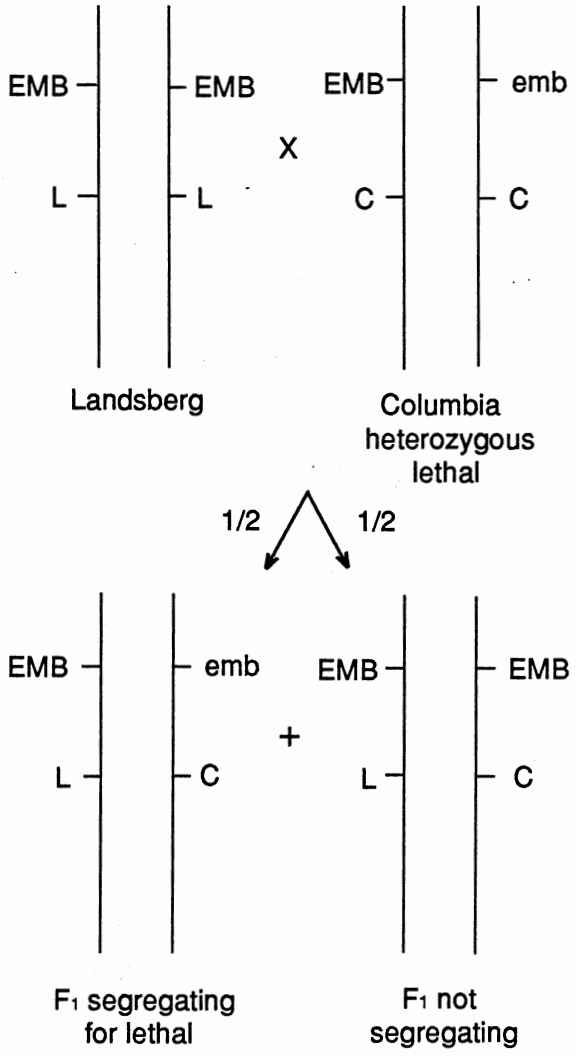



Figure 28. Drawing of Strategy Used in High-Resolution RFLP Mapping of Embryonic Lethals. The letters (L) and (C) represent different RFLP alleles of the same locus. Wild-type (EMB L/EMB L) Landsberg plants are crossed with a Columbia plant heterozygous (EMB C/emb C) for the lethal. Plants which are segragating for the lethal represent one-half of the $F_1$ generation as shown and are allowed to self-pollinate. $F_2$ plants are screened for the lethal and batches of 5 wild-type (EMB/EMB) are saved for DNA isolation and Southern blotting.

construction and mapping of these clones are described by Chang et al. (1988). In general, low copy number DNA sequences from Arabidopsis strain Columbia were cloned into the phage replacement vector λsep6-lac5 (Davis et al. 1980). The phage were then labeled and used as probes on Southern blots that contained genomic restriction digests from Columbia, Landsberg, and Niederzenz ecotypes. Clones that recognized polymorphisms were then mapped in $F_2$ and $F_3$ generations following crosses between either Niederzenz and Columbia, Columbia and Landsberg, or Niederzenz and Landsberg. The resulting RFLP linkage map which contains these markers is shown in Figure 29.

In our laboratory, phage clones were propagated in the E. coli host strain C600, obtained from Elliot Meyerowitz (along with the phage), using the procedures outlined in Maniatis et al. (1982). Phage DNA was isolated from plate lysates using Prep-eze columns (5prime → 3prime, Inc.) according to the procedure provided with the columns. Purified phage DNA was cut with Eco RI and labeled with the reagents provided in the Genius non-radioactive DNA labeling kit (Boehringer Mannheim) using the protocol described in Appendix L (p.174). Hybridization and immunological detection were also carried out using reagents provided with the Genius kit according to the protocols described in Appendix L (p.174).

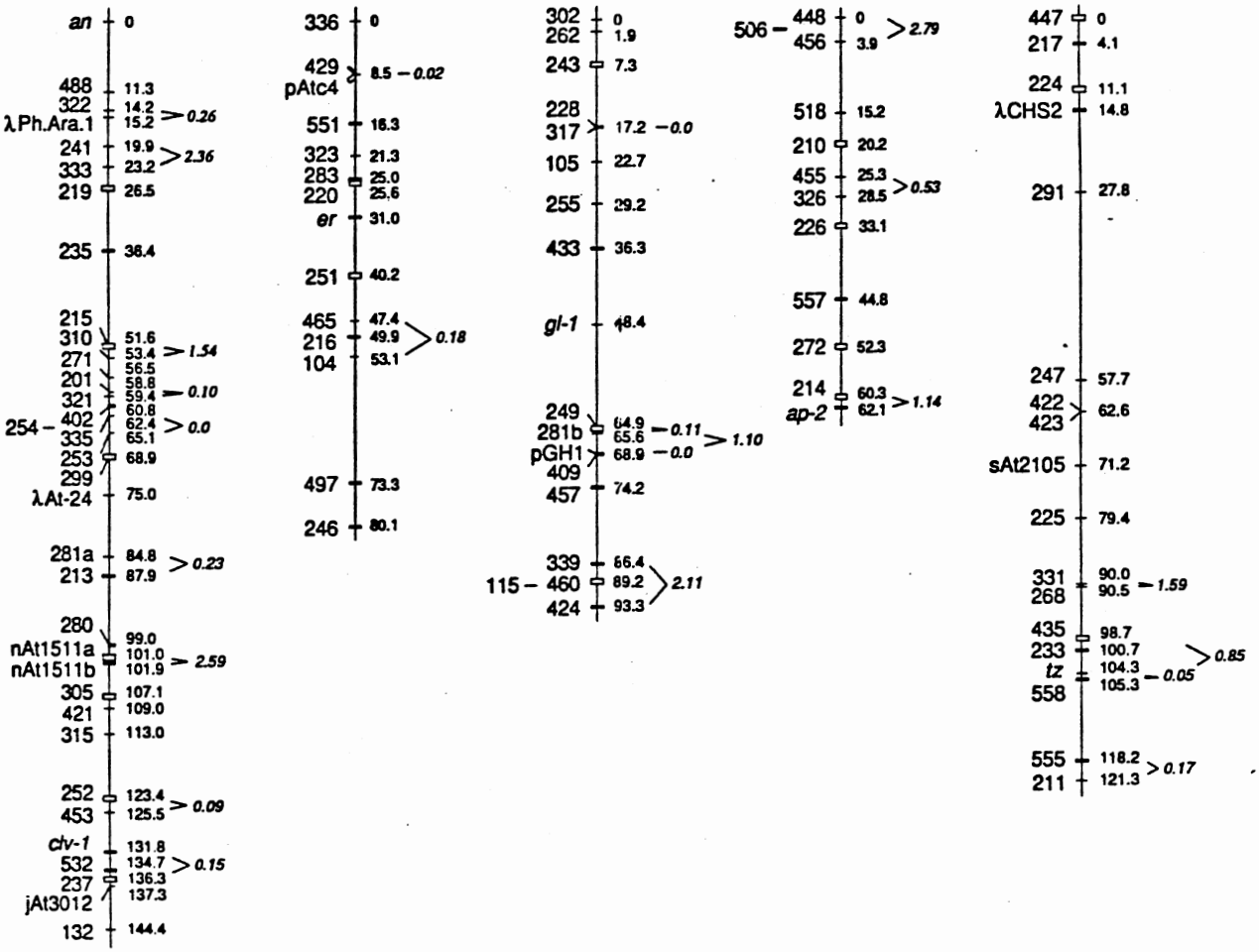



Figure 29. RFLP Linkage Map of Arabidopsis thaliana. From Chang et al. (1988).

## Results and Discussion

### High-Resolution RFLP Mapping Strategy

Although mapping genes with visible markers is important for genetic studies, localizing genes relative to molecular markers is required for gene isolation through chromosome walking. RFLP mapping of embryonic lethals is facilitated by the ability to identify heterozygous (EMB/emb) plants in the $F_2$ generation. Mapping other types of recessive mutations requires progeny testing in the $F_3$ to determine the genotype of $F_2$ plants. The ability to distinguish heterozygous (EMB/emb) from homozygous wild-type (EMB/EMB) plants in the $F_2$ generation was instrumental in devising the high-resolution RFLP mapping strategy.

When co-dominant RFLP markers are used to map embryonic lethals, three RFLP patterns and six classes of plants should be present in the $F_2$ generation following self-pollination of heterozygous (EMB/emb) $F_1$ plants. Table 7 summarizes the expected frequency of $F_2$ plants in each RFLP class at different levels of recombination between the RFLP marker and the embryonic lethal. Frequencies are listed separately for EMB/EMB and EMB/emb plants because only EMB/EMB $F_2$ plants were used for RFLP mapping. Note that EMB/EMB plants with Columbia-specific bands (classes II and III) are rare when the lethal and the RFLP marker are closely linked. For example, only 2% of the EMB/EMB $F_2$ plants are in classes II and III when the lethal and RFLP marker are separated by 1% recombination. This frequency increases to 75% if the RFLP marker and lethal are unlinked. The presence of Columbia-specific bands on Southern blots is therefore a direct measure of recombination between the RFLP marker and the embryonic lethal. In contrast, EMB/emb $F_2$

TABLE 7

EXPECTED FREQUENCIES OF $F_2$ PLANTS WITH DISTINCT RFLP PATTERNS OBTAINED FOLLOWING CROSSES WITH LANDSBERG OR NIEDERZENZ ECOTYPES

| $F_2$ genotype | RFLP class[a] | Recombination frequency between RFLP site and EMB locus | | | | | |
|---|---|---|---|---|---|---|---|
| | | 0.01 | 0.05 | 0.10 | 0.25 | 0.35 | 0.50 |
| EMB/emb | I | 0.010 | 0.0475 | 0.090 | 0.1875 | 0.2275 | 0.250 |
| | II | 0.980 | 0.9050 | 0.820 | 0.6250 | 0.5450 | 0.500 |
| | III | 0.010 | 0.0475 | 0.090 | 0.1875 | 0.2275 | 0.250 |
| EMB/EMB | I | 0.9800 | 0.902 | 0.810 | 0.560 | 0.423 | 0.250 |
| | II | 0.0199 | 0.095 | 0.180 | 0.380 | 0.455 | 0.500 |
| | III | 0.0001 | 0.003 | 0.010 | 0.060 | 0.122 | 0.250 |

[a] I = RFLP pattern produced by Landsberg or Niederzenz parent (EMB/EMB); II = pattern produced by $F_1$ heterozygote; III = pattern produced by Columbia (EMB/emb) parent. With co-dominant RFLP patterns, rare recombinants (classes II and III) are readily detected among pooled EMB/EMB plants.

plants produce bands on Southern blots that are characteristic of both parents, regardless of the percent recombination. Following discussions with Eric Ward and George Jen at Ciba-Geigy Corporation, I reasoned that the efficiency of RFLP mapping with embryonic lethals could be increased by analyzing DNA from pools of EMB/EMB $F_2$ plants and determining the frequency of rare recombinant plants through the presence of Columbia-specific bands on Southern blots. This pooling strategy should reduce the number of DNA preparations required for high-resolution RFLP mapping and facilitate the isolation of genes through chromosome walking.

RFLP Mapping of the EMB30 Locus

Initial RFLP mapping studies were carried out with the hope of identifying molecular markers that were tightly linked to the emb30 locus on chromosome 1. The standard genetic map and RFLP maps of Arabidopsis have not yet been fully integrated, but there are a few markers in common between these maps. Markers in the emb30 region on the RFLP map of Chang et al. (1988), shown in Figure 29 (p.106), were selected because of their ability to recognize polymorphisms between the Columbia ecotype and one of the other two ecotypes used to construct the map. Markers λbAt219 and λbAt254 were both used in initial studies as probes on blots that contained DNA from Columbia, Niederzenz, and Columbia/Niederzenz $F_1$ plants. In each case, the DNA was digested with an enzyme that was supposed to give polymorphic bands. However, these markers consistently failed to recognize differences between the Columbia and Niederzenz ecotypes.

With respect to the emb30 markers, 7 months were spent trying to resolve the problem of identifying the expected polymorphisms. During

this time, over 20 different plant DNA preparations were made and tested on a total of 8 Southern blots with either the λbAt219 or λbAt254 probe. There are a number of different reasons that could explain why these probes did not work. One explanation would be that the lambda phage used to make the DNA probes were not the appropriate strains. Over 20 different RFLP markers, in the form of lambda phage strains, were obtained from Elliot Meyerowitz as high-titer lysates and it is possible that some of these stocks were contaminated or mixed either before they were sent or after they were received. Nevertheless, every precaution was taken when growing these phage not to mix the different strains. A second possibility is that the polymorphic bands were actually present but not detected with the Genius nonradioactive DNA detection system. This explanation does not seem to be likely because later experiments have shown that the Genius system is more than sensitive enough to detect polymorphisms in single-copy sequences of Arabidopsis.

Another possibility is that the Niederzenz ecotype obtained from Elliot Meyerowitz was actually the Columbia ecotype. This was not fully tested, but in later experiments there appeared to be differences in the flowering time between individual plants derived from this seed stock. Since that time, a new seed stock of the Niederzenz ecotype was obtained from Mark Estelle (Indiana University). The Estelle-derived Niederzenz was not tested against the Meyerowitz-derived Niederzenz on Southern blots because by this time all of the DNA was used up from the Niederzenz (Meyerowitz) plants and no further plants were grown from this seed stock.

RFLP Mapping of the BIO1 Locus

The next set of experiments used the RFLP marker λbAt331 from the bio1 region on chromosome 5. Plant DNA from Columbia, Niederzenz (Estelle) and Landsberg ecotypes was digested with the appropriate restriction enzyme and tested for polymorphisms using labeled λbAt331 as a probe on Southern blots. In these experiments polymorphisms were detected between all three ecotypes as expected. Once the appropriate polymorphisms were detected, the next experiment was to show that recombinants could be detected among pooled $F_2$ plants. Again λbAt331 was chosen as a probe because this marker appeared to be roughly 15 cM from the bio1 locus which should give approximately 1 recombinant per 5 BIO1/BIO1 $F_2$ plants or 1 recombinant per pool. Figure 30 shows the resulting blot where 3 of the 4 pools tested had recombinants.

Comparison of the linkage and RFLP maps revealed that two markers λbAt233 and λbAt558 should be within 5 cM of bio1 on chromosome 5. These two markers were labeled and used as probes in a series of Southern blots in order to determine which marker was most closely linked to bio1. In these blots each pool contained 5 $F_2$ plants to ensure detection of a single recombinant homologue. As a control, the equivalent of a single recombinant homologue was detected in a sample containing 10 μg of genomic DNA (Figure 31, lane 3). The sensitivity of this method was also confirmed by the detection of recombinants in several pools (e.g. Figure 32, lane 5).

The presence of recombinants in 3 of 4 pools of bio1 x Niederzenz $F_2$ plants probed with λbAt331 demonstrated that this marker was not tightly linked to the bio1 locus. Two recombinants were detected out of

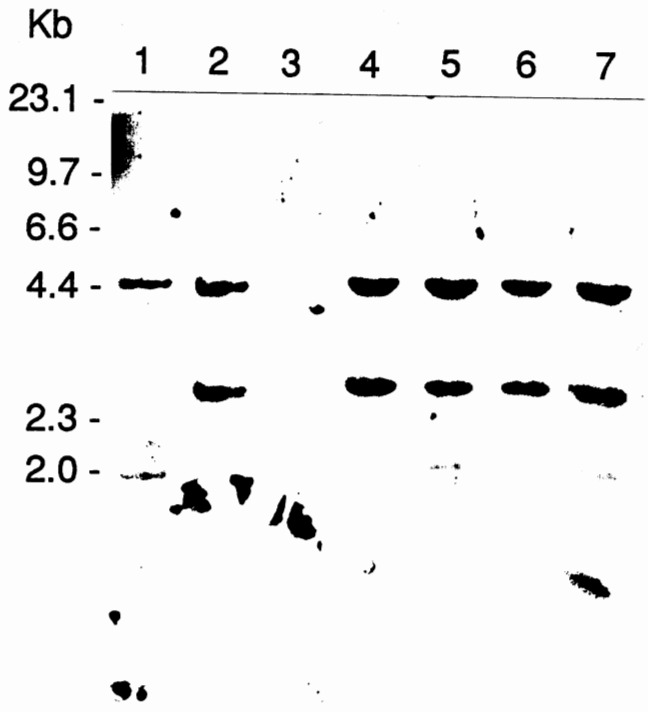



Figure 30. Southern Blot of DNA From 4 Pools of bio1 x Niederzenz $F_2$ Plants Probed With λbAt331. All lanes contained genomic DNA digested with the restriction enzyme Eco RI. Lane 1: 2 μg Columbia DNA; lane 2: 2 μg Niederzenz DNA; lane 3: blank; lanes 4-7: 10 μg each from different pools of 5 BIO1/BIO1 $F_2$ plants from the bio1 x Niederzenz cross. Note the presence of Columbia-specific bands in lanes 6,7, and 8.

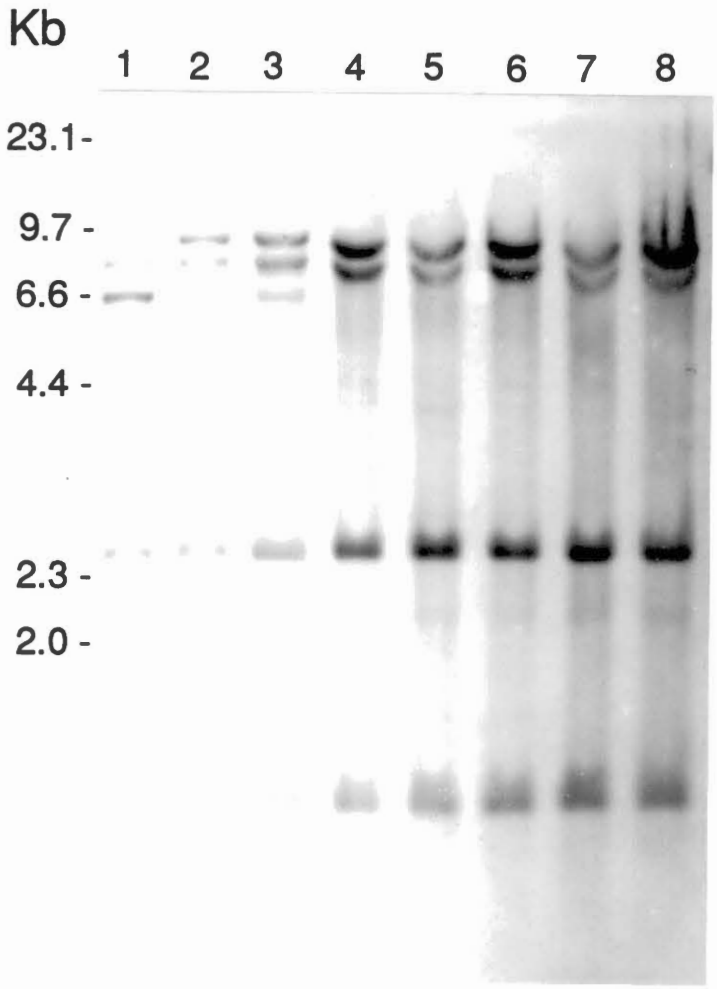



Figure 31. Southern Blot of DNA From Genetic Reconstruction and 5 Pools of bio1 x Landsberg $F_2$ Plants Probed With λbAt558. All lanes contained genomic DNA digested with Xba I. Lane 1: 2 μg of Columbia DNA; lane 2: 2 μg of Landsberg DNA; lane 3: 9 μg of Landsberg and 1 μg of Columbia DNA; lanes 4-8: 10 μg of DNA from different pools of 5 BIO1/BIO1 F2 plants. Note the absence of Columbia-specific bands in these pools.

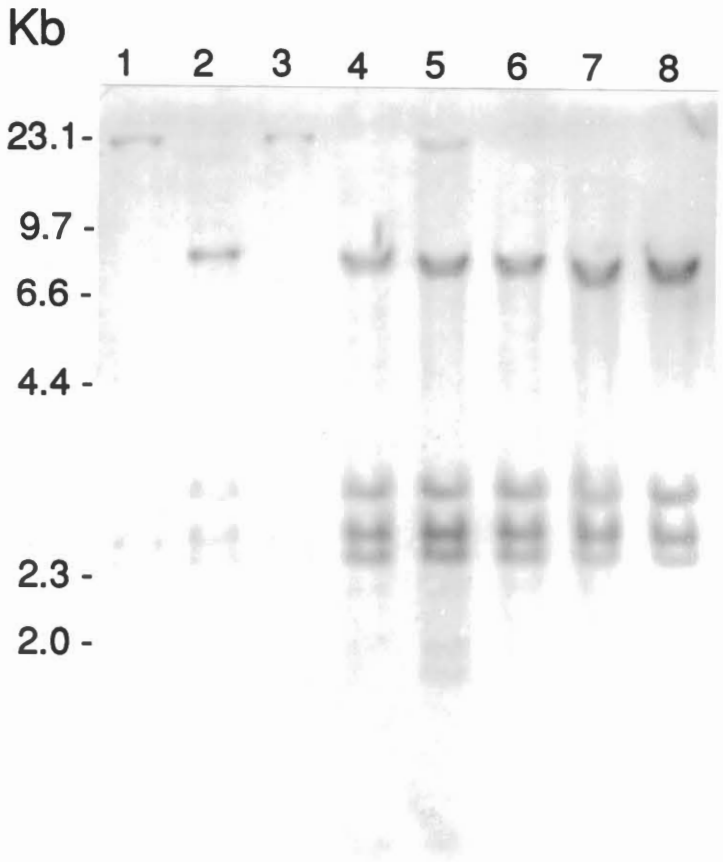



Figure 32. Southern Blot of DNA From 5 Pools of biol x Landsberg $F_2$ Plants Probed With λbAt233. All lanes contained genomic DNA digested with Xba I. Lane 1: 2 μg of Columbia DNA; lane 2: 2 μg of Landsberg DNA; lane 3: 1 μg of Columbia DNA; lanes 4-8: 10 μg of DNA from the same 5 pools shown in Figure 31. Note that lane 5 shows evidence of at least 1 recombinant plant.

14 pools of bio1 x Landsberg $F_2$ plants probed with λbAt233. This corresponds to approximately 3 cM between λbAt233 and bio1. In contrast, no recombinants were detected among 19 pools of bio1 x Landsberg $F_2$ plants probed with λbAt558. It therefore appears that λbAt558 is within 0.5 cM or 70 Kb (Chang et al. 1988) of the bio1 locus. The recent identification of a 220 Kb YAC that hybridizes to λbAt233 (Ward and Jen 1990) and a 160 Kb YAC that hybridizes to λbAt558 (Eric Ward, personal communication) should aid molecular analysis of this region and facilitate the isolation of BIO1 through chromosome walking.

In conclusion, it appears from these preliminary experiments that the pooling strategy described above will be useful for future RFLP mapping studies. It should also be noted that this strategy could be used to map mutations other than embryonic lethals as long as the homozygous recessive individuals are viable and can be readily identified in the $F_2$ generation. This mapping method should not only save time by reducing the number of DNA preparations required but should also result in a higher resolution estimate of recombination per Southern blot. In conjunction with advances in the resolution of the RFLP map and large-DNA technology (YAC,s and pulse-field gel electrophoresis), this RFLP mapping method should facilitate molecular isolation of genes of Arabidopsis through chromosome walking. These techniques should then be useful in answering important questions that remain about how genes control developmental processes such as plant embryogenesis.

## Concluding Remarks

The goals of this dissertation were to document as clearly as possible a review of the literature, results from experiments, and the relevance of this work in relation to the field of the genetic analysis of embryogenesis. Chapter 1 included a review of work on the genetic control of plant embryogenesis. Most of the early work was done on defective-kernel mutants of maize. Maize continues to be a promising system, especially for understanding the role of the endosperm in embryogenesis, and as a model system for studying monocot embryogenesis.

A second experimental system, mutants of carrot that are unable to complete somatic embryogenesis, was also reviewed. At this time, there are some major limitations in studying embryogenesis in carrot. The most important drawbacks are the lack of genetic information, and the fact that many of the embryonic mutants have been induced in haploid cell culture lines that cannot be regenerated into fertile plants.

Finally, a review of embryo-lethal mutants of Arabidopsis was given. An extensive study by Müller (1963) was reviewed where over 800 embryo-defective mutants were isolated and 60 were characterized in some detail. In the latter part of the 1970's, this paper inspired David Meinke, then a graduate student at Yale University, to pursue the study of plant embryogenesis through the isolation and characterization of embryo-lethal mutants of Arabidopsis. In his thesis, Meinke (1979) described six mutants which were defective in embryo development. A series of papers then followed, which described in detail, different aspects of the work presented in his thesis. After taking a position at Oklahoma State University, Meinke and colleagues have isolated and

characterized over 100 new embryo-lethal mutants. Highlights from this work include the isolation of developmentally interesting mutants (emb22 and emb30), the isolation of a biotin auxotroph (bio1), and numerous interesting conclusions regarding the genetic control of plant embryo development. Results from these studies have been summarized in no less than 15 publications and numerous conferences.

The field of Arabidopsis genetics and molecular biology has experienced a tremendous infusion of scientists over the past 5 years. Many types of mutants are being isolated and characterized by members of laboratories from all over the world. Most of these studies have the long-term goal of cloning genes responsible for their respective processes. The purpose of the work summarized in this dissertation was four-fold: (1) to demonstrate that embryonic lethals of Arabidopsis could be readily mapped with visible markers, telotrisomics, and RFLP's; (2) to enhance the genetic characteristics of Arabidopsis by adding new markers to the genetic map and by providing material which could be used in the future to construct balanced-lethal chromosomes; (3) to create new genetic stocks which could be useful in the future to map embryonic lethals and other mutations; and (4) to supply our lab with the most up to date techniques for RFLP mapping in preparation for gene isolation through chromosome walking. I feel that this dissertation has met these goals and should serve as a reference, in the future, to others interested in mapping not only embryo-defective, but other types of mutants in Arabidopsis.

LITERATURE CITED

Aaij, C. and P. Borst, 1972 The gel electrophoresis of DNA. Biochim. Biophys. Acta. **269:** 192-200.

Abbe, E. C. and O. L. Stein, 1952 The growth of the shoot apex in maize: embryogeny. Amer. J. Bot. **41:** 285-293.

Ammirato, P. V., 1983 Embryogenesis. pp. 82-123 in Handbook of Plant Cell Culture, Volume I, edited by D. A. Evans, W. R. Sharp, W. R. Ammirito, and Y. Yamada. Macmillan Publishing, New York.

Baus, A. D., L. Franzmann and D. W. Meinke, 1986 Growth in vitro of arrested embryos from lethal mutants of Arabidopsis thaliana. Theor. Appl. Genet. **72:** 577-586.

Beckett, J. B., 1978 B-A translocations in maize. I. Use in locating genes by chromosome arm. J. Hered. **69:** 27-36.

Bernatzky, R., 1988 Restriction Fragment Length Polymorphism. pp. C2: 1-18 in Plant Molecular Biology Manual, edited by S. B. Gelvin and R. A. Schliperoort. Kluwer Academic, Boston.

Bewley, D. J. and M. Black, 1985 Seeds: Physiology of development and germination. Plenum Press.

Bhattacharya, S. S., A. F. Wright, J. F. Clayton, W. H. Price, C. I. Phillips, C. M. E. McKeown, M. Jay, A.C. Bird, J. Pearson, 1984 Close genetic linkage between X-linked Retinitis pigmentosa and a restriction fragment length polymorphism identified by recombinant DNA probe L1.28. Nature **309:** 253-255.

Botstein, D., R. L. White, M. Skolnick, R. Davis, 1980 Construction of a genetic linkage map in man using using restriction fragment length polymorphisms. Am. J. Hum. Genet. **32:** 314-331.

Boyer, H. W., 1971 DNA restriction and modification mechanisms in bacteria. Ann. Rev. Microbiol. **25:** 153-176.

Breton, A. M. and Z. R. Sung, 1982 Temperature-sensitive carrot variants impaired in somatic embryogenesis. Dev. Biol. **90:** 58-66.

Brink, R. A., 1927 A lethal mutation in maize affecting the seed. Amer. Nat. **61:** 520-530.

Brink R. A. and D. C. Cooper, 1947 Effect of the de-17 allele on development of the maize caryopsis. Genetics **32:** 350-368.

Burr, B., S. V. Evola, F. A. Burr, J. S. Beckmann, 1983 The application of restriction fragment length polymorphisms to plant breeding. pp. 45-59 in Genetic Engineering Principles and Methods, Vol. 5., edited by Setlow, J. K. and A. Hollander. Plenum, New York.

Carlson, W. R., 1988 The cytogenetics of corn. pp. 259-343 in Corn and Corn Improvement, Ed. 3, edited by G. F. Sprague and J. W. Dudley. Amer. Soc. Agron., Madison, Wisconsin.

Chang, C., J. L. Bowman, A. W. DeJohn, E. S. Lander and E. M. Meyerowitz, 1988 Restriction fragment length polymorphism linkage map for Arabidopsis thaliana. Proc. Natl. Acad. Sci. USA **85:** 6856-6860.

Coe E. H., Jr. and M. G. Neuffer, 1977 The genetics of corn. pp.111-223 in Corn and Corn Improvement, Ed. 2, edited by G. F. Sprague. Amer. Soc. Agron., Madison, Wisconsin.

Coe, E. H., Jr., M. G. Neuffer and D. A. Hoisington, 1988 The genetics of corn. pp. 81-258 in Corn and Corn Improvement, Ed. 3, edited by G. F. Sprague and J. W. Dudley. Amer. Soc. Agron., Madison, Wisconsin.

Coulson, A., J. Sulston, S. Brenner and J. Karn, 1986 Toward a physical map of the genome of the nematode Caenorhabditis elegans. Proc. Natl. Acad. Sci. USA **83:** 7821-7825.

Davis, R. W., D. Botstein, and J. R. Roth, 1980 Advanced bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Demerec, M., 1923 Heritable characters in maize. XV. Germless seeds. J. Hered. **14:** 297-300.

Dure, L., 1985 Embryogenesis and gene expression during seed formation. Oxford Surveys of Plant Molecular and Cell Biology. **2:** 179-197.

Emerson, R. A., 1932 A recessive zygotic lethality resulting in 2:1 ratios for normal vs male sterile and colored vs colorless pericarp in $F_2$ of certain maize hybrids. Science. **75:** 566.

Engels, W. R., 1981 Estimating genetic divergence and genetic variability with restriction endonucleases. Proc. Natl. Acad. Sci. USA **78:** 6329-6333.

Errampalli, D., D. Patton, L. Castle, K. Hansen, L. Mickelson, J. Schnall, K. Feldmann and D. Meinke, 1991 Embryonic lethals and T-DNA insertional mutagenesis in Arabidopsis thaliana. Plant Cell.

Estelle, M. A., and C. R. Somerville, 1986 The mutants of Arabidopsis. Trends Genet. **2:** 89-93.

Evola, S. V., F. A. Burr, B. Burr, 1986 The suitability of restriction fragment length polymorphisms as genetic markers in maize. Theor. Appl. Genet. **71**: 765-771.

Feldmann, K. A., and M. D. Marks, 1987 Agrobacterium-mediated transformation of germinating seeds of Arabidopsis thaliana: A non-tissue culture approach. Mol. Gen. Genet. **208**: 1-9.

Feldmann, K. A., T. J. Carlson, S. A. Coomber, C. E. Farrance, M. A. Mandel and A. M. Wierzbicki, 1990 T-DNA insertional mutagenesis in Arabidopsis thaliana. pp.109-120 in Horticultural Biotechnology, edited by A. B. Bennett and S. D. O'Neill. Alan R. Liss, New York.

Fink, G. R., 1988 Notes of a bigamous biologist. Genetics **118**: 549-550.

Fisher, R. A., 1958 The Genetical Theory of Natural Selection. Ed. 2. Dover Press, New York.

Franzmann, L., D. A. Patton and D. W. Meinke, 1989 In vitro morphogenesis of arrested embryos from lethal mutants of Arabidopsis thaliana. Theor. Appl. Genet. **77**: 609-616.

Gavazzi, G., Nava-Rachi, M., and C. Tonelli, 1975 A mutation causing proline requirement in maize. Theor. Appl. Genet. **46**: 339-346.

Giuliano, G., Lo Schavio, F., and M. Terzi, 1984 Isolation and developmental characterization of temperature-sensitive carrot cell variants. Theor. Appl. Genet. **67**: 179-183.

Goldberg, R. B., S. J. Barker and L. Perez-Grau, 1989 Regulation of gene expression during plant embryogenesis. Cell **56**: 149-160.

Goodman, H. M., B. Hauge, B. den Boer, J. Giraudat and H. G. Nam, 1989 Isolation of developmentally important genes of Arabidopsis. J. Cell. Biochem. Suppl. **13D**: 233.

Grodziker, T., J. Williams, P. Sharp, J. Sambrook, 1974 Physical mapping of temperature-sensitive mutations of adenovirus. Cold Spring Harbor Symp. Quant. Biol. **39**: 439-446.

Gusella, J. F., R. E. Tanzi, M. A. Anderson, W. Hobbs, K. Gibbons, R. Raschtchian, T. C. Gilliam, M. R. Wallace, N. S. Wexler, P.M. Conneally, 1984 DNA markers for nervous system diseases. Science **225**: 1320-1326.

Hake, S., and M. Freeling, 1986 Analysis of genetic mosaics shows that the extra epidermal cell divisions in Knotted mutant maize plants are induced by adjacent mesophyll cells. Nature **320**: 621-623.

Heath, J. D., R. Weldon, C. Monnot and D. W. Meinke, 1986 Analysis of storage proteins in normal and aborted seeds from embryo-lethal mutants of Arabidopsis thaliana. Planta **169**: 304-312.

Higgins, T. J. V., 1984 Synthesis and regulation of major proteins in seeds. Ann. Rev. Plant Physiol. **35**: 191-221.

Jones, D. F., 1920 Heritable characters in maize. IV. A lethal factor: defective seeds. J. Hered. **11**: 161-167.

Koornneef, M., 1983 The use of telotrisomics for centromere mapping in Arabidopsis thaliana (L.) Heynh. Genetica **62**: 33-40.

Koornneef, M., 1990 Linkage map of Arabidopsis thaliana (2n = 10), pp. 95-99 in Genetic Maps. Ed. 5, Book VI, edited by S. J. O'Brien. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Koornneef, M., and P. Stam, 1988 Procedures for mapping by using $F_2$ and $F_3$ populations. Arabidopsis Inform. Serv. **25**: 35-40.

Koornneef, M. and J. H. Van der Veen, 1978 Gene localization with trisomic Arabidopsis thaliana. Arabidopsis Inform. Serv. **15:** 38-43.

Koornneef, M., and J. H. Van der Veen, 1983 Trisomics in Arabidopsis thaliana and the location of linkage groups. Genetica **61:** 41-46.

Koornneef, M., J. van Eden, C. J. Hanhart, P. Stam, F. J. Braaksma and W. J. Feenstra, 1983 Linkage map of Arabidopsis thaliana. J. Hered. **74:** 265-272.

Koornneef, M., C. J. Hanhart, E. P. van Loenen-Martinet and J. H. Van der Veen, 1987 A marker line, that allows the detection of linkage on all Arabidopsis chromosomes. Arabidopsis Inform. Serv. **23:** 46-50.

Kosambi, D. D., 1944 The estimation of map distances from recombination values. Ann. Eugen. **12:** 172-175.

Last, R. L., and G. R. Fink, 1988 Tryptophan requiring mutants of Arabidopsis thaliana. Science **240:** 305-310.

Lederberg, S. and M. Meselson, 1964 Degradation of non-replicating bacteriophage DNA in non-accepting cells. J. Mol. Biol. **8:** 623-628.

Leutwiler, L. S., B. R. Hough-Evans, and E. M. Meyerowitz, 1984 The DNA of Arabidopsis thaliana. Mol. Gen. Genet. **194:** 15-23.

Li, S. L., and G. P. Redéi, 1968 Thiamine mutants of the crucifer, Arabidopsis. Biochem. Genet. **3:** 163-170.

Lloyd, A. M., A. R. Barnason, S. G. Rogers, M. C. Byrne, R. T. Fraley, and R. B. Horsch, 1986 Transformation of Arabidopsis thaliana with Agrobacterium tumefaciens. Science **234:** 464-466.

Lo Schiavo, F., Giuliano, G., and Z. R. Sung, 1988 Characterization of a temperature-sensitive carrot cell mutant impaired in somatic embryogenesis. Plant Sci. **54:** 157-164.

Lowe, J., and O. E. Nelson, 1946 Miniature seed: A study in the development of a defective caryopsis in maize. Genetics. **31:** 525-533.

Maheshwari, P., 1950 An Introduction to the Embryology of Angiosperms. McGraw-Hill, New York.

Mangelsdorf, P. C., 1923 The inheritance of defective seeds in maize. J. Hered. **14:** 119-125.

Mangelsdorf, P. C., 1926 The genetics and morphology of some endosperm characters in maize. Conn. Agric. Exp. Station Bull. **279:** 513-620.

Maniatis, T., Fritch, E. F., and J. Sambrook, 1982 Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Manzocchi, L. A., Daminati, M. G., Gentinetta, E., and F. Salamini, 1980a Viable defective endosperm mutants in maize. 1. Kernel weight, protein fraction, and zein subunits in mature endosperms. Maydica. **25:** 105-116.

Manzocchi, L. A., Daminati, M. G., and E. Gentinetta, 1980b Viable defective endosperm mutants in maize. 2. Kernel weight, nitrogen and zein accumulation during endosperm development. Maydica. **25:** 199-210.

Marsden, M. P. F., and D. W. Meinke, 1985 Abnormal development of the suspensor in an embryo-lethal mutant of Arabidopsis thaliana. Amer. J. Bot. **72:** 1801-1812.

Meinke, D. W., 1979 Isolation and characterization of embryo-lethal mutants of Arabidopsis thaliana. Ph.D. Dissertation, Yale University, New Haven, Connecticut.

Meinke, D. W., 1982 Embryo-lethal mutants of Arabidopsis thaliana: Evidence for gametophytic expression of the mutant genes. Theor. Appl. Genet. **63:** 381-386.

Meinke, D. W., 1986 Embryo-lethal mutants and the study of plant embryo development. Oxford Surv. Plant Mol. Cell. Biol. **3:** 122-165.

Meinke, D. W., 1991 Genetic analysis of plant development. (in press) in Plant Physiology, Volume X, Plant Growth and Development, edited by F. C. Steward and R. G. S. Bidwell. Academic Press, New York.

Meinke, D. W., and A. D. Baus, 1985 Gametophytic gene expression in embryo-lethal mutants of Arabidopsis thaliana. pp. 15-20 in Biotechnology and Ecology of Pollen, edited by D. L. Mulcahy, G. B. Mulcahy and E. Ottaviano. Springer-Verlag, New York.

Meinke, D. W., and I. M. Sussex, 1979a Embryo-lethal mutants of Arabidopsis thaliana: A model system for genetic analysis of plant embryo development. Dev. Biol. **72:** 50-61.

Meinke, D. W., and I. M. Sussex, 1979b Isolation and characterization of six embryo-lethal mutants of Arabidopsis thaliana. Dev. Biol. **72:** 62-72.

Meinke, D., D. Patton, J. Shellhammer, A. Reynolds-Duffer, L. Franzmann, T. Schneider and K. Robinson, 1989 Developmental and molecular genetics of embryogenesis in Arabidopsis thaliana. pp. 121-132 in The Molecular Basis of Plant Development, edited by R. Goldberg. Alan Liss, New York.

Meyerowitz, E. M., and R. E. Pruitt, 1985 Arabidopsis thaliana and plant molecular genetics. Science **229:** 1214-1218.

Müller, A. J., 1963 Embryonentest zum Nachweis rezessiver Letalfaktoren bei Arabidopsis thaliana. Biol. Zentralbl. **82:** 133-163.

Nam, H. G., J. Giraudat, B. den Boer, F. Moonan, W. D. B. Loos, B. M. Hauge and H. M. Goodman, 1989 Restriction fragment length polymorphism linkage map of Arabidopsis thaliana. Plant Cell **1:** 699-705.

Neuffer, M. G., and W. F. Sheridan, 1980 Defective kernel mutants of maize. I. Genetic and lethality studies. Genetics. **95:** 929-944.

Palca, J., 1989 Genome projects are growing like weeds. Science **245:** 131.

Palmer, J. D., R. A. Jorgensen, W. F. Thompson, 1985 Chloroplast DNA variation and evolution in Pisum: Patterns of change and phylogenetic analysis. Genetics **109:** 195-213.

Patton, D. A., and D. W. Meinke, 1990 Ultrastructure of arrested embryos from lethal mutants of Arabidopsis thaliana. Amer. J. Bot. **77:** 653-661.

Patton, D. A., Franzmann, L. H., and D. W. Meinke, Submitted. Mapping genes essential for embryo development in Arabidopsis thaliana. Mol. Gen. Genet.

Poethig, S., 1988 A non-cell-autonomous mutation regulating juvenility in maize. Nature **336:** 82-83.

Polans, N. O., N. F. Weeden, W. F. Thompson, 1985 Inheritance, organization and mapping of the rbcs and cab multigene families in pea. Proc. Natl. Acad. Sci. USA **32:** 5083-5087.

Raghavan, V., 1976 Experimental Embryogenesis in Vascular Plants. Academic Press, New York.

Richards, E., 1990 Preparation of genomic DNA from plant tissue. pp. 2.3.1-2.3.3 in Current Protocols in Molecular Biology, edited by F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith and K. Struhl. Greene Publishing and Wiley-Interscience, New York.

Riordan, J. R., J. M. Rommens, B. Kerem. N. Alon, R. Rozmahel, Z. Grzelczak, J. Zielinski, S. Lok, N. Plavsic, J.-L. Chou, M. L. Drumm, M. C. Iannuzzi, F. S. Collins, and L.-C. Tsui, 1989 Identification of the cystic fibrosis gene: Cloning and characterization of complementary DNA. Science **245:** 1066-1073.

Röbbelen, G. and F. J. Kribben, 1966 Erfahrungen bei der Auslese von Trisomen. Arabidopsis Inform. Serv. **3:** 16-17.

Roberts, R. J., 1988 Restriction enzymes and their isoschizomers. Nucleic Acids Res. (suppl.) **16:** r271.

Roman, H., and J. Ullstrup, 1952 The use of A-B translocations to locate genes in maize. Agron. J. **43:** 450-454.

Ross-Craig, S. 1948 Drawings of British Plants. III. Cruciferae, plate 37. G. Bell, London.

Sass, J. E., and G. F. Sprague, 1950 The embryology of 'germless' maize. Iowa State Coll. J. Sci. **24:** 209-218.

Schmidt, R., and L. Willmitzer, 1988 High efficiency Agrobacterium tumefaciens-mediated transformation of Arabidopsis thaliana leaf and cotyledon explants. Plant Cell Rep. **7:** 583-586.

Schnall, J. A., Cooke, T. J., and D. E. Cress, 1988 Genetic analysis of embryogenesis in carrot cell culture: Initial characterization of six classes of temperature-snesitive variants. Dev. Genet. **9:** 49-67.

Schneider, T., R. Dinkins, K. Robinson, J. Shellhammer and D. W. Meinke, 1989 An embryo-lethal mutant of Arabidopsis thaliana is a biotin auxotroph. Dev. Biol. **131:** 161-167.

Schweizer, D., P. Ambros, P. Gründler and F. Varga, 1988 Attempts to relate cytological and molecular chromosome data of Arabidopsis thaliana to its genetic linkage map. Arabidopsis Inform. Serv. **25:** 27-34.

Servítová, J., and I. Cetl, 1984 The use of recessive lethal chlorophyll mutants for linkage mapping of Arabidopsis thaliana (L.) Heynh. Arabidopsis Inform. Serv. **21:** 59-64.

Shellhammer, J., and D. Meinke, 1990 Arrested embryos from the biol auxotroph of Arabidopsis thaliana contain reduced levels of biotin. Plant Physiol. **93:** 1162-1167.

Sheridan, W. F., and M. G. Neuffer, 1980 Defective kernel mutants of maize. II. Morphological and embryo culture studies. Genetics **95:** 945-960.

Sheridan, W. F., and M. G. Neuffer, 1982 Maize developmental mutants. J. Hered. **73:** 318-329.

Somerville, C., 1989 Arabidopsis blooms. Plant Cell **1:** 1131-1135.

Southern, E. M., 1975 Detection of specific sequences among DNA fragments separated by gel electrophoresis. Methods Enzymol. **68:** 152-176.

Steinitz-Sears, L. M., 1963 Chromosome studies in Arabidopsis thaliana. Genetics **48**: 483-490.

Stephans, J. C., M. L. Cavanaugh, M. I Gradie, M. L. Mador, K. K. Kidd, 1990 Mapping the human genome: Current status. Science **250**: 237-244.

Tanksley, S. D., 1983 Molecular markers in plant breeding. Plant Mol. Biol. Rep. **1**: 3-8.

Templeton, A. R., 1983 Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and apes. Evolution **37**: 221-244.

Torti, G., Lombardi, L., and L. A. Manzocchi, 1984 Indole-3-acetic acid content in viable defective endosperm mutants of maize. Maydica. **29**: 335-343.

Vallejos, C. E., Tanksley, S. D., Bernatzky, R., 1986 Localization in the tomato genome of DNA restriction fragments containing sequences homologous to the rRNA (45S), the major chlorophyll a/b binding polypeptide and the ribulose bisphosphate carboxylase genes. Genetics **112**: 93-105.

Valvekens, D., M. Van Montagu and M. Van Lijsebettens, 1988 Agrobacterium tumefaciens-mediated transformation of Arabidopsis thaliana root explants by using kanamycin selection. Proc. Natl. Acad. Sci. USA **85**: 5536-5540.

Ward, E. R., and G. C. Jen, 1990 Isolation of single-copy-sequence clones from a yeast artificial chromosome library of randomly-sheared Arabidopsis thaliana DNA. Plant Mol. Biol. **14**: 561-568.

Wentz, J. B., 1930 The inheritance of germless seeds in maize. Iowa Agric. Exp. Station Res. Bull. **121**: 345-379.

Yeung, E. C., 1980 Embryogeny in Phaseolus: the role of the suspensor. Z. Pflanzenphysiol. **96:** 17-28.

Zhuchenko, A. A., A. B. Korol', I. Yu. Vizir, N. I. Bocharnikova and I. A. Zamorzaeva, 1988 Sex differences in crossover frequency for tomato and thale cress (Arabidopsis thaliana). Soviet Genetics **24:** 1104-1110.

# APPENDIX A

## PRINTOUT OF $F_2$ SEGREGATION DATA FROM THE BUILD AND SUMMARY PROGRAMS

TITLE: W2 X EMB30-1 (10-13-87)

| POT# | PLANT# | AN | DIS1 | ER | EMB30-1 |
|---|---|---|---|---|---|
| 1 | A | + | + | + | M |
| 1 | B | M | M | + | M |
| 1 | C | . | . | . | . |
| 1 | D | + | + | + | M |
| 1 | E | + | + | + | M |
| 1 | F | M | M | + | M |
| 1 | G | + | + | + | M |
| 1 | H | + | + | + | M |
| 1 | I | M | M | + | + |
| 2 | A | M | M | + | + |
| 2 | B | + | + | M | M |
| 2 | C | M | M | + | + |
| 2 | D | M | M | M | + |
| 2 | E | + | + | + | M |
| 2 | F | M | M | + | + |
| 2 | G | + | M | + | + |
| 2 | H | + | + | + | M |
| 2 | I | M | M | + | + |
| 3 | A | + | + | M | M |
| 3 | B | M | M | + | + |
| 3 | C | M | M | + | + |
| 3 | D | + | + | + | M |
| 3 | E | M | + | + | + |
| 3 | F | + | + | + | M |
| 3 | G | + | + | M | M |
| 3 | H | + | + | + | M |
| 3 | I | + | + | M | M |
| 4 | A | M | M | M | + |
| 4 | B | + | + | + | M |
| 4 | C | + | + | M | M |
| 4 | D | M | M | + | + |
| 4 | E | + | + | M | M |
| 4 | F | + | + | + | M |
| 4 | G | + | + | + | M |
| 4 | H | + | + | + | M |
| 4 | I | + | + | + | M |
| 5 | A | . | . | . | . |
| 5 | B | + | + | + | M |
| 5 | C | . | . | . | . |
| 5 | D | + | + | + | M |
| 5 | E | . | . | . | . |
| 5 | F | M | M | M | + |
| 5 | G | M | M | M | + |
| 5 | H | M | + | + | M |
| 5 | I | + | + | + | M |

TITLE: W2 X EMB30-1 (10-13-87)

| POT# | PLANT# | AN | DIS1 | ER | EMB30-1 |
|---|---|---|---|---|---|
| 6 | A | M | M | M | + |
| 6 | B | M | M | M | + |
| 6 | C | + | + | + | M |
| 6 | D | M | M | + | + |
| 6 | E | M | M | + | + |
| 6 | F | + | + | + | M |
| 6 | G | M | + | M | M |
| 6 | H | M | + | + | + |
| 6 | I | + | + | M | M |
| 7 | A | + | + | M | M |
| 7 | B | + | + | + | M |
| 7 | C | + | + | + | M |
| 7 | D | . | . | . | . |
| 7 | E | + | + | + | M |
| 7 | F | + | + | + | M |
| 7 | G | M | M | M | + |
| 7 | H | M | M | + | + |
| 7 | I | M | M | + | + |
| 8 | A | + | + | M | M |
| 8 | B | + | + | + | M |
| 8 | C | . | . | . | . |
| 8 | D | . | . | . | . |
| 8 | E | + | + | M | M |
| 8 | F | + | + | + | M |
| 8 | G | M | M | + | + |
| 8 | H | M | M | M | + |
| 8 | I | + | + | + | M |
| 9 | A | M | M | M | + |
| 9 | B | M | M | + | + |
| 9 | C | + | + | M | M |
| 9 | D | + | + | + | M |
| 9 | E | + | + | + | M |
| 9 | F | M | M | + | + |
| 9 | G | + | M | M | + |
| 9 | H | + | M | + | + |
| 9 | I | M | M | + | + |
| 10 | A | + | + | + | M |
| 10 | B | + | + | + | M |
| 10 | C | M | M | + | + |
| 10 | D | + | + | + | + |
| 10 | E | + | + | + | M |
| 10 | F | + | + | + | M |
| 10 | G | M | + | + | M |
| 10 | H | + | + | + | M |
| 10 | I | + | + | + | M |

TITLE: W2 X EMB30-1 (10-13-87)

| POT# | PLANT# | AN | DIS1 | ER | EMB30-1 |
|---|---|---|---|---|---|
| 11 | A | + | + | + | M |
| 11 | B | + | + | + | M |
| 11 | C | + | M | + | + |
| 11 | D | + | + | + | M |
| 11 | E | M | + | + | M |
| 11 | F | + | M | M | + |
| 11 | G | M | M | + | + |
| 11 | H | + | + | M | M |
| 11 | I | + | + | M | M |
| 12 | A | + | + | + | M |
| 12 | B | + | + | + | M |
| 12 | C | + | + | + | M |
| 12 | D | + | + | + | M |
| 12 | E | . | . | . | . |
| 12 | F | + | + | + | M |
| 12 | G | + | + | + | M |
| 12 | H | M | M | + | + |
| 12 | I | M | + | + | M |
| 13 | A | . | . | . | . |
| 13 | B | + | + | M | M |
| 13 | C | + | + | + | M |
| 13 | D | + | + | + | M |
| 13 | E | + | + | + | M |
| 13 | F | + | + | + | M |
| 13 | G | + | + | + | M |
| 13 | H | + | + | M | M |
| 13 | I | M | M | + | + |
| 14 | A | + | + | + | M |
| 14 | B | . | . | . | . |
| 14 | C | + | + | M | M |
| 14 | D | . | . | . | . |
| 14 | E | + | + | + | M |
| 14 | F | + | + | M | M |
| 14 | G | + | M | M | + |
| 14 | H | + | + | + | M |
| 14 | I | + | + | + | M |
| 15 | A | + | + | + | M |
| 15 | B | + | + | M | M |
| 15 | C | + | + | + | M |
| 15 | D | M | M | + | + |
| 15 | E | + | M | M | + |
| 15 | F | + | + | + | M |
| 15 | G | + | + | + | M |
| 15 | H | + | + | + | M |
| 15 | I | + | M | + | M |

TITLE: W2 X EMB30-1 (10-13-87)

| POT# | PLANT# | AN | DIS1 | ER | EMB30-1 |
|---|---|---|---|---|---|
| 16 | A | M | M | + | + |
| 16 | B | + | + | + | M |
| 16 | C | + | + | M | M |
| 16 | D | + | + | M | M |
| 16 | E | + | M | M | + |
| 16 | F | M | + | + | M |
| 16 | G | M | M | + | + |
| 16 | H | M | M | + | + |
| 16 | I | . | . | . | . |
| 17 | A | + | + | + | M |
| 17 | B | M | M | + | + |
| 17 | C | + | + | + | M |
| 17 | D | + | + | + | M |
| 17 | E | + | + | + | M |
| 17 | F | + | + | M | M |
| 17 | G | + | + | + | M |
| 17 | H | + | + | + | M |
| 17 | I | + | + | + | M |
| 18 | A | + | + | M | M |
| 18 | B | M | M | + | + |
| 18 | C | M | M | M | + |
| 18 | D | . | . | . | . |
| 18 | E | + | + | + | M |
| 18 | F | M | M | M | + |
| 18 | G | + | + | M | M |
| 18 | H | + | + | M | M |
| 18 | I | M | + | M | M |
| 19 | A | M | M | M | + |
| 19 | B | M | M | M | + |
| 19 | C | + | + | + | M |
| 19 | D | + | + | + | M |
| 19 | E | + | + | + | M |
| 19 | F | M | M | + | + |
| 19 | G | + | + | + | M |
| 19 | H | + | + | + | M |
| 19 | I | + | + | + | M |
| 20 | A | + | + | + | M |
| 20 | B | + | + | + | M |
| 20 | C | + | M | + | + |
| 20 | D | + | M | + | + |
| 20 | E | + | + | + | M |
| 20 | F | + | + | M | M |
| 20 | G | + | + | + | M |
| 20 | H | M | M | M | + |
| 20 | I | M | + | + | M |

TITLE: W2 X EMB30-1 (10-13-87)

| POT# | PLANT# | AN | DIS1 | ER | EMB30-1 |
|---|---|---|---|---|---|
| 21 | A | M | M | + | + |
| 21 | B | + | + | + | M |
| 21 | C | + | + | + | M |
| 21 | D | + | + | + | M |
| 21 | E | + | + | + | M |
| 21 | F | + | + | M | M |
| 21 | G | M | M | M | + |
| 21 | H | M | M | M | + |
| 21 | I | M | M | + | + |
| 22 | A | + | + | M | M |
| 22 | B | M | M | + | + |
| 22 | C | + | + | M | M |
| 22 | D | + | + | + | M |
| 22 | E | . | . | . | . |
| 22 | F | + | + | + | M |
| 22 | G | + | + | M | M |
| 22 | H | + | + | M | M |
| 22 | I | + | + | + | M |
| 23 | A | + | + | M | M |
| 23 | B | + | M | + | + |
| 23 | C | . | . | . | . |
| 23 | D | + | M | + | + |
| 23 | E | + | + | + | M |
| 23 | F | + | + | + | M |
| 23 | G | M | M | + | + |
| 23 | H | . | . | . | . |
| 23 | I | + | + | M | M |
| 24 | A | + | M | + | + |
| 24 | B | + | + | + | M |
| 24 | C | M | M | + | + |
| 24 | D | . | . | . | . |
| 24 | E | + | + | + | M |
| 24 | F | M | M | + | + |
| 24 | G | + | + | + | M |
| 24 | H | + | M | + | + |
| 24 | I | + | + | + | M |
| 25 | A | + | M | + | + |
| 25 | B | + | + | + | M |
| 25 | C | . | . | . | . |
| 25 | D | M | M | + | + |
| 25 | E | + | + | + | M |
| 25 | F | M | M | + | + |
| 25 | G | M | M | + | M |
| 25 | H | + | + | M | M |
| 25 | I | + | + | M | M |

TITLE: W2 X EMB30-1 (10-13-87)

| POT# | PLANT# | AN | DIS1 | ER | EMB30-1 |
|---|---|---|---|---|---|
| 26 | A | + | + | + | M |
| 26 | B | . | . | . | . |
| 26 | C | M | M | + | + |
| 26 | D | + | + | + | M |
| 26 | E | + | + | + | M |
| 26 | F | + | + | + | M |
| 26 | G | + | + | + | M |
| 26 | H | . | . | . | . |
| 26 | I | + | + | M | M |
| 27 | A | + | + | + | M |
| 27 | B | . | . | . | . |
| 27 | C | M | M | + | + |
| 27 | D | . | . | . | . |
| 27 | E | M | M | M | + |
| 27 | F | . | . | . | . |
| 27 | G | . | . | . | . |
| 27 | H | + | + | + | M |
| 27 | I | + | + | M | M |

SUMS OF:

| | | | | |
|---|---|---|---|---|
| + | 152 | 146 | 159 | 72 |
| M | 67 | 73 | 60 | 147 |
| OTHER | 24 | 24 | 24 | 24 |

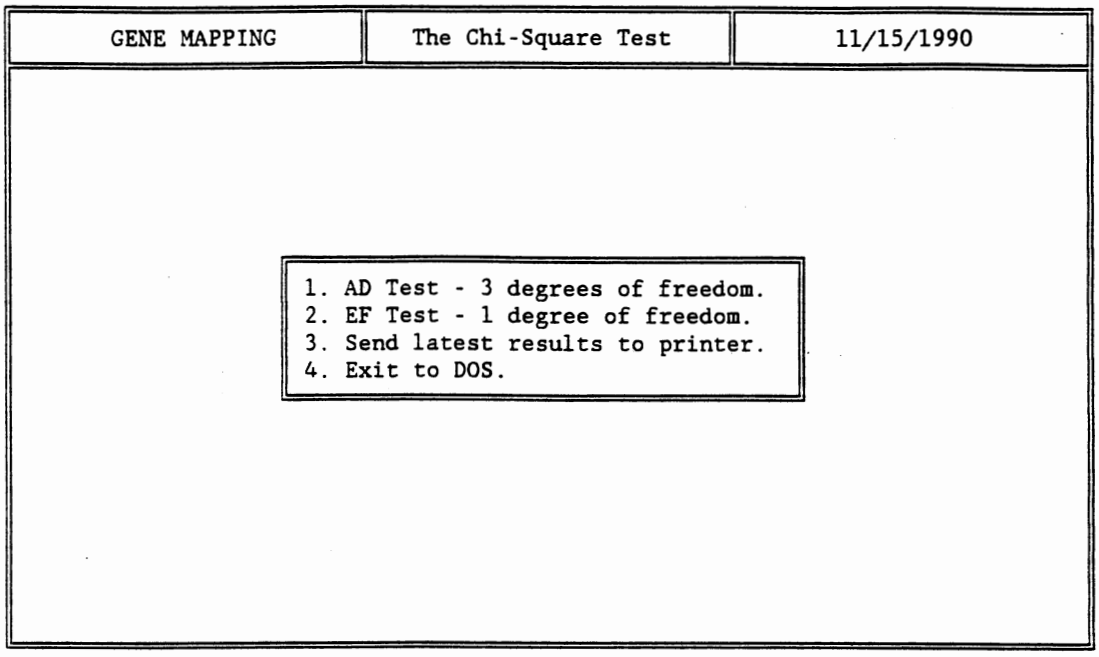
SUMS OF:

| | | | |
|---|---|---|---|
| ++ | 16 | 3 | 50 |
| +M | 136 | 143 | 109 |
| M+ | 56 | 69 | 22 |
| MM | 11 | 4 | 38 |

TOTAL NUMBER OF PLANTS: 243
TOTAL NUMBER OF DEAD PLANTS: 24

# APPENDIX B

## EXAMPLE OF HOW TO USE THE CHI PROGRAM TO DETECT LINKAGE



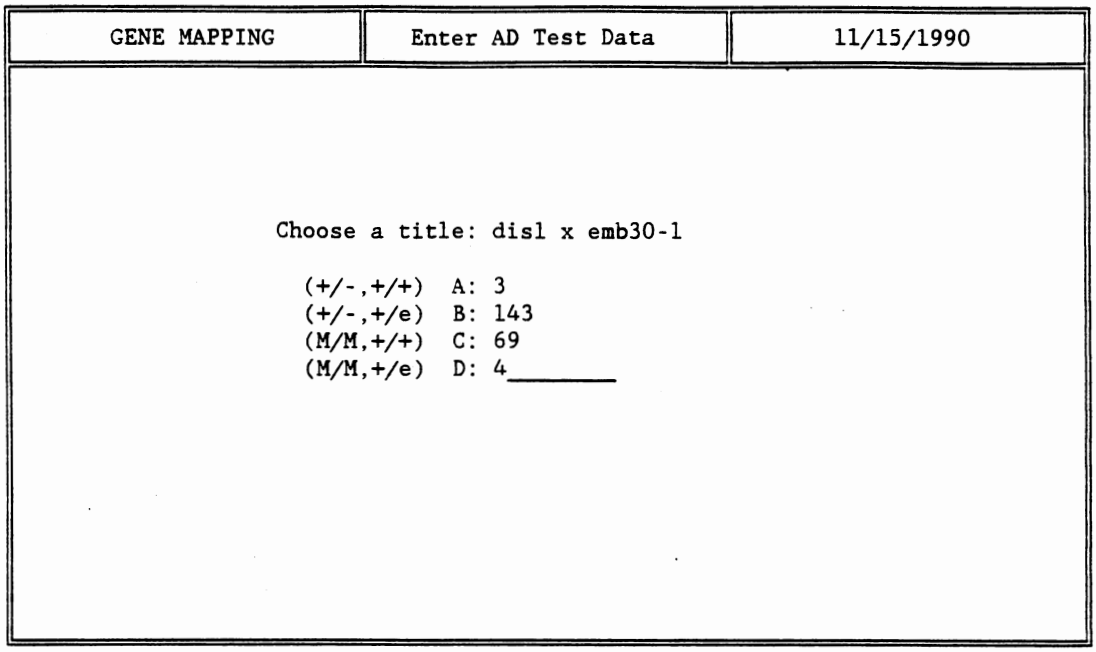

The main menu, shown above, appears on the screen after the CHI program is started. Use the cursor control keys to highlight the type of data and press return, the AD and EF methods are explained in Table 2 (p.66). The AD segregation data from Appendix A for the genes dis1 and emb30 will be used in this example, so we highlight "AD Test" and press return.

| GENE MAPPING | Enter AD Test Data | 11/15/1990 |
|---|---|---|

Choose a title: _____

(+/-,+/+) A:
(+/-,+/e) B:
(M/M,+/+) C:
(M/M,+/e) D:

The menu shown above appears on the screen and prompts the user for a title and the number of plants in each $F_2$ class. Enter the appropriate information. The genotypes of the four classes A, B, C, and D are shown to the left of the appropriate letter.




Once the information is filled in, the screen should appear as shown above. To start the calculation press Control-Q (press control key while holding down the key for the letter "Q" on the keyboard.

| GENE MAPPING | Enter AD Test Data | 11/15/1990 |
|---|---|---|

dis1 x emb30-1

| | Observed | Expected | Difference |
|---|---|---|---|
| A | 3 | 54.75 | -51.75 |
| B | 143 | 109.50 | 33.50 |
| C | 69 | 18.18 | 50.82 |
| D | 4 | 36.57 | -32.57 |
| Total | 219 | | |

Chi-square = 230.28 ***

Press any key to continue.

The screen will then show the chi-square calculation along with the level of statistical significance (* = significantly different at P = 0.05; ** at P = 0.01; and *** at P ≤ 0.005). In this case, the segregation of the dis1 and emb30 genes is significantly different from what is expected for unlinked genes, therefore the genes appear to be linked. To print the results, press any key to return to main menu shown on P.139, then highlight "Send latest results to printer" and press return. The final printout looks exactly like what is inside the box shown above. To quit, highlight "Exit to DOS" and press return.

# APPENDIX C

# USING THE RECF2 COMPUTER PROGRAM

```
******** MAIN MENU ************************
*                                          *
*                                          *
*   choose the phase (coupling/repulsion)  *
*                                          *
*   choose the type of data set            *
*                                          *
*   choose the parameters to estimate      *
*                                          *
*                                          *
*   enter the data (observed frequencies)  *
*                                          *
*   enter the number of iterations         *
*                                          *
*   enter the trial values                 *
*                                          *
*                 --------------------     *
*   RUN           | make your choice |     *
*                 | (use cursor keys)|     *
*   QUIT          | and press ENTER  |     *
*                 --------------------     *
********************************************
```

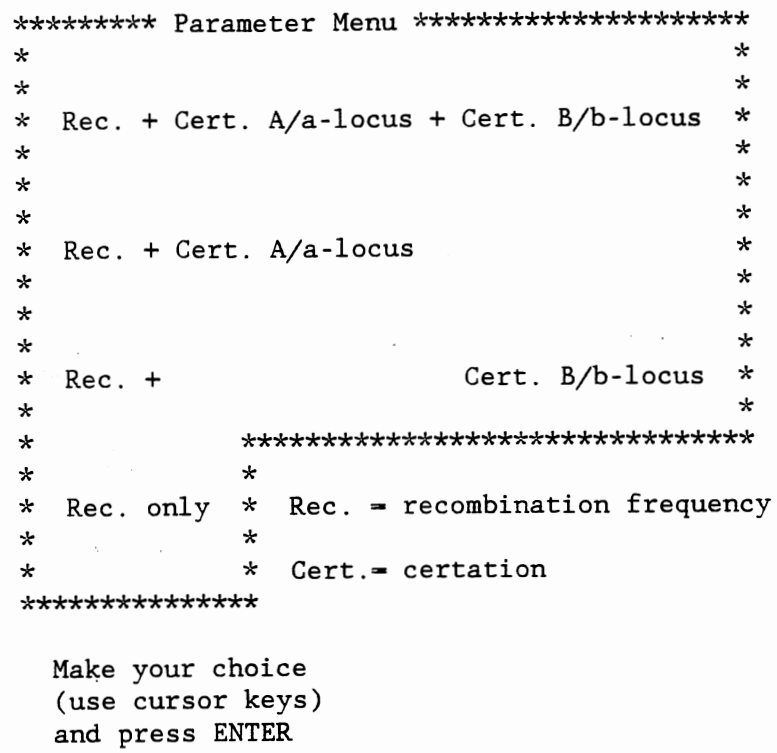
The menu shown above appears on the screen once RECF2 has been started. Each item in this menu must be selected before RECF2 can make recombination estimates. Highlight the first item "choose the phase" with the cursor keys, and press enter.

```
******** PHASE MENU **********
*                            *
*                            *
*    coupling                *
*                            *
*                            *
*    repulsion               *
*                            *
*                            *
******************************

Make your choice
(use cursor keys)
and press ENTER
```

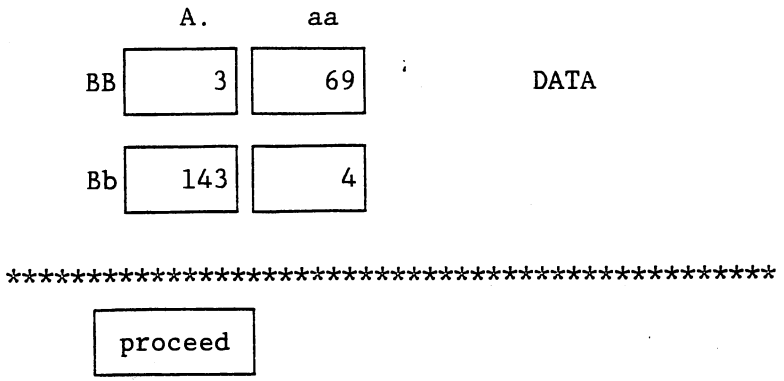
The "phase menu" shown above then appears on the screen. The recessive alleles are either in the coupling or repulsion phase in the $F_1$ plants. When mapping embryonic lethals with visible markers, use repulsion.

```
                                        **********************
     A. aa                              *  TYPE OF DATA      *
                                        *                    *
B. ... ...    Dominance at              *  Make your choice  *
              both loci                 *  (use cursor keys) *
bb ... ...    4 phenotypes              *  and press ENTER   *
                                        **********************

     A. aa                                   AA  Aa  aa

BB ... ...    Dominance at                 BB ... ... ...   No dominance
              one locus
Bb ... ...                                 Bb ... ... ...   9 phenotypes
              6 phenotypes
bb ... ...                                 bb ... ... ...

     A. aa
              Dominance at
BB ... ...    one locus;
              one recessive lethal
Bb ... ...
              4 phenotypes
```

The "type of data" choice is then highlighted in the main menu; the menu shown above then appears on the screen. The choices in this menu refer to the genetic characteristics of the marker and the gene of interest being mapped. In our case, there is dominance at one locus (heterozygotes indistinguishable from wild-type) and one recessive lethal. Highlight the appropriate choice and press enter.




Next, choose the parameters to estimate in the main menu; the "parameter menu" shown above will appear on the screen. The choices here ask whether you want to simultaneously estimate recombination and certation parameters. Usually, "recombination only" is selected in this menu. If certation is to be estimated, the user must manually enter his/her own estimate of certation effects into the "trial values" menu shown later. Certation values range from 0.0 (pollen carrying recessive allele does not contribute to fertilization to 1.0 (pollen carrying recessive allele outcompetes the wild-type allele in all cases). A certation value of 0.5 would represent a "normal" heterozygote where half of the pollen grains that contribute to fertilization carry the wild-type allele and half carry the recessive allele.




The next choice in the main menu is "enter the data"; after this choice is selected, the data menu shown above is displayed. The observed number of plants in each class is entered into the four rectangular boxes. The letter "A" represents the visible marker and "B" is the lethal. The capital letters represent the wild-type allele and the lower case is for recessive. The "A." genotype represents plants with the wild-type phenotype and "aa" is for plants with the recessive phenotype. The "BB" symbol represents plants that are not segregating for the lethal, while "Bb" plants are segregating. The four rectangular boxes represent $F_2$ classes A (upper left), B (lower left), C (upper right), and D (lower right) as defined in table 2 (p.68). The data shown here are from segregation data shown in Appendix A, for the genes dis1 and emb30.

Number of iterations : [ 20 ]

(default=20)

[ proceed ]

The next choice in the main menu is "number of iterations". The number of iterations (or mathematical "tries" to fit your data with a level of recombination) is set at default value of 20 as shown above, which seems to work well as long as your segregation data are fairly close to what is expected for some percent recombination. In most cases, highlight 20 and press return.

TRIAL VALUES

Recombination: [ 0.250 ]

[ proceed ]

The next choice in the main menu is "enter the trial values". The trial values are your best estimate of the level of recombination (and/or certation). The program uses this value as a starting point as it tries to fit your data with what would be expected at different levels of recombination. If you are not sure, just enter 0.25 (25% recombination) as shown above. In nearly all cases this value is sufficient. If not, you may need to try higher or lower values depending upon the data.

```
Estimates after  5 iterations :

Parameter    Estimate        St.Dev.
      1       0.02444        0.00923625


 Covariance matrix:

    0.000085


Observed Expected    Chi-square contribution

       3     3.525           0.0781
      69    69.475           0.0033
     143   142.519           0.0016
       4     3.481           0.0773

Total Chi-square =           0.1603 (d.f.=  2)
```

Finally, choose "RUN" from the main menu and press return, the computer screen will then display the recombination estimate and ask if you want to print (Y/N), press "Y". The covariance matrix calculation is then displayed, press "Y" to print. Finally an internal chi-square value is shown that compares your observed data with what is expected for that level of recombination, again press "Y" to print. The printout shown above is from the data entered into the data menu of this Appendix. The recombination estimate in this case is 0.02444 with a standard deviation of 0.00923625 (2.4 ± 0.92%). The covariance matrix and chi-square calculations are both very low with these data, suggesting that the observed number of $F_2$ plants in each class is very close to what you would expect for two genes separated by 2.4% recombination. When you are finished, highlight "quit" in the main menu and press return to exit the program and return to the disk operating system (DOS).

# APPENDIX D

# HOW TO USE GENMAP

It is important to note that you must have a good idea of the gene order before using GENMAP. The example shown here reconstructs the positions of genes on chromosome 3 which is shown in Figure 23 (p.75). Chromosome 3 was chosen as an example because there are only 11 genes to work with. Once the GENMAP program is started, the user is prompted for an "input file". The input file is a DOS text file that contains all of the available recombination for a given chromosome, note that there must be recombination data in this file for each gene to be mapped to this chromosome. There are 5 of these files, one for each chromosome. The names are chr1.dat, chr2.dat, chr3.dat, chr4.dat, and chr5.dat. The example shown below is for chr3.dat.

| | | | |
|---|---|---|---|
| hy2 | ch6 | 43.0 | 4.0 |
| hy2 | cer7 | 48.6 | 4.1 |
| hy2 | cer7 | 46.0 | 3.6 |
| hy2 | tt6 | 48.1 | 3.5 |
| gl1 | tt5 | 39.4 | 3.0 |
| gl1 | ch6 | 28.3 | 4.3 |
| gl1 | cer7 | 28.4 | 5.5 |
| gl1 | tt6 | 38.0 | 3.8 |
| tt5 | cer7 | 12.8 | 5.6 |
| ch6 | gl1 | 43.0 | 3.0 |
| ch6 | tt5 | 2.9 | 0.8 |
| hy2 | abi3 | 26.1 | 3.0 |
| abi3 | gl1 | 20.1 | 3.1 |
| abi3 | tt5 | 33.0 | 2.4 |
| hy2 | gl1 | 32.0 | 0.8 |
| gl1 | tt5 | 31.6 | 1.3 |
| hy2 | tt5 | 46.5 | 1.6 |
| gl1 | cer7 | 37.0 | 2.6 |
| hy2 | ap3 | 48.0 | 4.6 |
| gl1 | ap3 | 30.3 | 5.4 |
| hy2 | csr | 42.2 | 5.2 |
| tt5 | csr | 9.0 | 2.4 |
| gl1 | csr | 25.5 | 2.5 |
| csr | cer7 | 16.3 | 2.9 |
| gl1 | emb29 | 18.0 | 2.6 |
| hy2 | emb29 | 10.7 | 1.8 |
| cer7 | emb78 | 0.6 | 0.5 |
| gl1 | emb78 | 30.3 | 3.7 |

Note that care must be taken when adding recombination data to these

files: 1) always retrieve and save these files in a DOS text format; 2) make sure there are no extra spaces at the end of a text line or no extra hard returns at the end of the file; and 3) never add any special commands from a word processor such as underline, bold, or superscripts/subscripts. After you have entered the name of the input data file (chr3.dat in this case) the program asks whether you want a printed list of the genes to be mapped. You must print this list (shown below) before going to the next step because each gene will be referred to by the whole number printed next to the gene symbol in the list.

```
 1  hy2
 2  ch6
 3  cer7
 4  tt6
 5  gl1
 6  tt5
 7  abi3
 8  ap3
 9  csr
10  emb29
11  emb78
```

GENMAP then asks: "Do you want a printed list of the data sets that have been read from the input-file? (Y/N)". Normally you will not need to print this list unless you need to verify recently added recombination data. The program then asks: "Should all markers listed be mapped (Y/N)?". Press "Y" to map all genes listed. The next prompt is "Do you want to remove SUSPECT data sets? (Y/N)". Press "N", the recent maps (Koornneef 1990; Patton et al. submitted) include SUSPECT data because they reflect the peculiarities present in the genome of Arabidopsis. GENMAP then prompts you to enter the gene order starting at the top of the chromosome with position 0. See the next page for exactly how the screen looks when gene order is entered and how the program prompts you to enter the correct order.

Ordering of the markers.

The following prompts ask for the numerical identifiers of the markers at the successive positions; positioning starts at position 0.

```
position   0:  1
position   1:  10
position   2:  7
position   3:  5
position   4:  9
position   5:  8
position   6:  6
position   7:  2
position   8:  11
position   9:  3
position  10:  4
```

The numbers on the right were entered from the keyboard. The gene order is then printed out on the screen for you to verify before GENMAP calculates the relative positions of genes. The screen will then show the following:

The order of markers currently is as follows:

```
pos.  0 <---> hy2
pos.  1 <---> emb29
pos.  2 <---> abi3
pos.  3 <---> gl1
pos.  4 <---> csr
pos.  5 <---> ap3
pos.  6 <---> tt5
pos.  7 <---> ch6
pos.  8 <---> emb78
pos.  9 <---> cer7
pos. 10 <---> tt6
```

Do you want to change the gene order? (Y/N)

Press "N" if the order is correct. GENMAP will then calculate the positions of the genes for the chromosome and ask whether you want a copy printed out. The printout for chromosome 3, using the above gene order and the recombination data (chr3.dat file) shown previously, is shown on the next page.

Estimated map distances.

| nr | marker_1 | marker_2 | distance | st.dev. | cum.dist. | st.dev. |
|---|---|---|---|---|---|---|
| 1 | hy2 | emb29 | 13.018 | 1.632 | 13.018 | 1.632 |
| 2 | emb29 | abi3 | 12.212 | 2.883 | 25.231 | 2.491 |
| 3 | abi3 | gl1 | 12.024 | 2.393 | 37.254 | 1.213 |
| 4 | gl1 | csr | 26.956 | 2.039 | 64.210 | 2.336 |
| 5 | csr | ap3 | 8.689 | 8.688 | 72.899 | 8.528 |
| 6 | ap3 | tt5 | 0.283 | 8.594 | 73.182 | 1.943 |
| 7 | tt5 | ch6 | 2.889 | 0.794 | 76.072 | 2.070 |
| 8 | ch6 | emb78 | 3.611 | 2.714 | 79.683 | 2.774 |
| 9 | emb78 | cer7 | 0.653 | 0.499 | 80.336 | 2.745 |
| 10 | cer7 | tt6 | 7.138 | 9.180 | 87.474 | 8.913 |

The column labeled "cum.dist." shows the estimated positions of the genes, remember that there will always be one gene at position 0, hy2 in this case. Occasionally GENMAP will not agree with the gene order that you put in, and will place one of the genes in another position. You can manually override this order and reenter your original order if you have proof that this is actually the gene order. The program asks "Do you want to change the order of genes? (Y/N)" If you press "Y" reenter the old order to override the GENMAP estimated order. If you press "N" the program asks "DO you want a list of SUSPECT data sets? (Y/N)". The list of suspect data sets for the above calculation is shown below.

| Used data, original estimates, combined estimates, chi-square values | | | | | | |
|---|---|---|---|---|---|---|
| 10 | ch6 | gl1 | 64.667 | 38.817 | 5.034 | ** SUSPECT ** |
| 13 | abi3 | gl1 | 21.302 | 12.024 | 6.296 | ** SUSPECT ** |

Total Chi-square value= 33.803 (d.f.= 17)

The program then prompts you to hit any key to exit, from there you can reboot GENMAP to calculate another chromosome.

# APPENDIX E

## USING CHROMAP AND PRINTAPLOT COMPUTER PROGRAMS TO PRINT THE LINKAGE MAP

Before starting the CHROMAP program, you must first create a DOS text file that contains information on how the chromosomes will look. The lengths of chromosomes, centromere positions (optional), and gene positions (from GENMAP) are included in the text file. An example of this file is shown below. The second entry (5) is the number of chromosomes while the next five entries are the lengths of the chromosomes. The negative numbers (chosen at random) indicate that centromere positions will not be included in this map. If centromere positions are to be included, the centromere position as a positive number one space after the chromosome length. The remainder of the data file is separated in to three columns, chromosome number, gene position in cM, and gene symbol. This data file was used to generate the map shown in Figure 23 (p.75).

```
*** Arabidopsis Chromosome Map ***
5
123 -60.0
48 -50
88 -57
74 -50
99 -41
1   0.0     an
1   0.4     emb76
1   0.5     rgn
1   0.8     cer1
1   10.6    alb1
1   13.4    chl1
1   16.6    dis1
1   17.8    emb30
1   20.6    Dw1
1   20.9    Gai
1   21.0    ga4
1   32.1    th1
1   32.4    fb
1   43.6    dis2
1   53.8    tt1
1   54.2    chl3
1   58.4    ch1
1   62.2    emb22
1   66.4    le
1   69.8    cer5
```

| | | |
|---|---|---|
| 1 | 88.1 | ft |
| 1 | 91.4 | clv2 |
| 1 | 97.1 | cer6 |
| 1 | 100.0 | ap1 |
| 1 | 101.7 | emb25 |
| 1 | 110.7 | clv1 |
| 1 | 115.8 | emb33 |
| 1 | 119.8 | ga2 |
| 1 | 119.9 | gl2 |
| 1 | 122.4 | fe |
| 2 | 0.0 | Gf |
| 2 | 2.3 | hy3 |
| 2 | 5.0 | chl2 |
| 2 | 7.0 | cp2 |
| 2 | 16.2 | er |
| 2 | 16.9 | hy1 |
| 2 | 22.0 | py |
| 2 | 32.7 | as |
| 2 | 32.8 | emb18 |
| 2 | 34.6 | aux1 |
| 2 | 42.4 | sul |
| 2 | 47.1 | cer8 |
| 2 | 47.4 | emb39 |
| 3 | 0.0 | hy2 |
| 3 | 13.0 | emb29 |
| 3 | 25.2 | abi3 |
| 3 | 37.3 | gl1 |
| 3 | 64.2 | csr |
| 3 | 72.9 | ap3 |
| 3 | 73.2 | tt5 |
| 3 | 76.1 | ch6 |
| 3 | 79.7 | emb78 |
| 3 | 80.3 | cer7 |
| 3 | 87.5 | tt6 |
| 4 | 0.0 | ga1 |
| 4 | 6.3 | hy4 |
| 4 | 9.5 | bp |
| 4 | 24.5 | cp1 |
| 4 | 28.2 | fca |
| 4 | 30.8 | th3 |
| 4 | 37.7 | ag |
| 4 | 39.2 | ch42 |
| 4 | 42.5 | emb28 |
| 4 | 43.5 | im |
| 4 | 47.0 | cer2 |
| 4 | 47.5 | ga5 |
| 4 | 48.7 | emb20 |
| 4 | 49.8 | Abi1 |
| 4 | 56.2 | cer4 |
| 4 | 56.3 | fd |
| 4 | 63.3 | ap2 |
| 4 | 71.6 | cer9 |
| 4 | 73.4 | cp3 |

```
5     0.0     fy
5     4.2     emb24
5     8.4     tt7
5     9.2     ch7
5    11.1     hy5
5    13.7     lu
5    15.1     tt4
5    16.4     fg
5    17.2     alb2
5    23.5     pi
5    25.3     cnx
5    25.6     ms1
5    31.4     ttg
5    35.5     ga3
5    38.1     su
5    45.2     ch5
5    47.0     th2
5    47.2     tt2
5    55.7     gl3
5    57.7     tt3
5    63.9     emb9
5    66.8     pgm
5    77.2     bio1
5    79.0     tz
5    83.6     cer3
5    88.7     yi
5    89.1     emb15
5    89.8     min
5    93.1     emb16
5    98.3     aba
```

Once started, the CHROMAP program asks for the name of the input file and the name of the output file. The input file is the DOS text file shown above and the output file is a plotter language file that contains the information needed to print the map. Note that the output file should contain the three extension "plt", for example "map.plt". After the names of the input file and output file have been entered, the program plots the map on the screen, and then moves the overlapping gene symbols up or down until no overlaps are detected. CHROMAP then saves this final plot in a form that can be used to print the map with a plotting program. The plotting program PRINTAPLOT is then used to print the map. Start the PRINTAPLOT program by typing "pp" while in the

PRINTAPLOT (pplot) directory. The menu shown below then appears on the screen.

```
PrintAPlot 1.1                    (C) Copyright 1988 Insight Development Corp.

Plotter   Settings   Files   Go   Configuration   Exit
Select plotter model to emulate

        Use arrow keys to highlight an option, then press Enter to select it
                                - OR -
                    Just enter the first letter of the option

                        Press Esc to revert to previous menu

                                SETTINGS SHEET

PLOTTER EMULATION                                          SETTINGS SHEET FILE
    HP 7475A

  CONFIGURATION          Pen     Size     Color       Work disk      C:
Active printer           1       1        Black       Copy count     1
  HP LaserJet            2       1        Black
Output destination       3       1        Black       Origin  X  0.00 in
  LPT1                   4       1        Black               Y  0.00 in
Paper source             5       1        Black
  Paper tray             6       1        Black       Resolution  150 dpi
                                                      Inverse       No
                                                      Adjust size 100 %
```

Use the arrow keys to highlight the "settings" option then press enter. The option "Get" will then be highlighted automatically, press enter again. The cursor moves to the file "C:\PPLOT\"; type in "standard" to complete the filename C:\PPLOT\STANDARD and press return. This file automatically contains the settings that used to print the map with a Hewlett Packard DeskJet. The screen now shows the proper settings as shown on the next page.

PrintAPlot 1.1 (C) Copyright 1988 Insight Development Corp.

Plotter Settings Files Go Configuration Exit
Get/Save/Change values in the SETTINGS SHEET

Use arrow keys to highlight an option, then press Enter to select it
- OR -
Just enter the first letter of the option

Press Esc to revert to previous menu

SETTINGS SHEET

| PLOTTER EMULATION HP 7475A | | | | SETTINGS SHEET FILE standard |
|---|---|---|---|---|
| CONFIGURATION | Pen | Size | Color | Work disk C: |
| Active printer | 1 | 2 | Black | |
| HP DeskJet | 2 | 3 | Black | |
| Output destination | 3 | 4 | Black | Origin X 0.00 in |
| LPT1 | 4 | 1 | Black | Y 0.00 in |
| | 5 | 1 | Black | |
| | 6 | 1 | Black | Resolution 300 dpi |
| | | | | Inverse No |
| | | | | Adjust size 99 % |

Use the arrow keys to highlight "files" and press enter. The cursor moves to the filename "C:\PPLOT\"; use the delete (←) key to delete this filename and replace with the name of the output file from CHROMAP (eg. map.plt). Be sure to use the proper Drive:directory address for the output file, for example A:\map.plt or C:\CHROMAP\map.plt. The name of the file to be printed will appear at the top of the screen, press return to confirm file name or press the space bar to cancel. Use the arrow keys to highlight "Go" and press return. The data file will then be read and plotted by the printer. When the printer is finished, highlight "Exit" to quit and return to DOS.

# APPENDIX F

## CONVERSION OF RECOMBINATION FREQUENCY (P) INTO CENTIMORGANS USING THE KOSAMBI (1944) MAPPING FUNCTION

| (P) | (cM) |
|---|---|
| 0.01 | 1.00 |
| 0.02 | 2.00 |
| 0.03 | 3.00 |
| 0.04 | 4.01 |
| 0.05 | 5.02 |
| 0.06 | 6.03 |
| 0.07 | 7.05 |
| 0.08 | 8.07 |
| 0.09 | 9.10 |
| 0.10 | 10.13 |
| 0.11 | 11.18 |
| 0.12 | 12.24 |
| 0.13 | 13.31 |
| 0.14 | 14.38 |
| 0.15 | 15.48 |
| 0.16 | 16.58 |
| 0.17 | 17.71 |
| 0.18 | 18.84 |
| 0.19 | 20.00 |
| 0.20 | 21.18 |
| 0.21 | 22.38 |
| 0.22 | 23.61 |
| 0.23 | 24.87 |
| 0.24 | 26.15 |
| 0.25 | 27.47 |
| 0.26 | 28.82 |
| 0.27 | 30.21 |
| 0.28 | 31.64 |
| 0.29 | 33.12 |
| 0.30 | 34.66 |
| 0.31 | 36.25 |
| 0.32 | 37.91 |
| 0.33 | 39.64 |
| 0.34 | 41.46 |
| 0.35 | 43.37 |
| 0.36 | 45.38 |
| 0.37 | 47.52 |
| 0.38 | 49.81 |
| 0.39 | 52.27 |
| 0.40 | 54.93 |
| 0.41 | 57.84 |
| 0.42 | 61.05 |
| 0.43 | 64.67 |
| 0.44 | 68.79 |
| 0.45 | 73.61 |
| 0.46 | 79.45 |
| 0.47 | 86.90 |
| 0.48 | 97.30 |
| 0.49 | 114.88 |

# APPENDIX G

## DERIVATION OF FORMULAE USED TO CALCULATE MINIMUM NUMBER OF $F_2$ PLANTS REQUIRED TO DETECT LINKAGE

AD equation (3 degress of freedom):

where: $\chi^2$ is a Chi-squared critical value for P = 0.05, 0.01 or 0.001

N is minimum number of plants required for linkage detection

p is recombination frequency (0.1, 0.2, 0.3 or 0.4)

$p_o$ is recombination frequency for unlinked genes (0.5)

$$\chi^2 = \frac{\{N[1/3(p)(2-p)] - N[1/3(p_o)(2-p_o)]\}^2}{N[1/3(p_o)(2-p_o)]} + \frac{\{N[2/3(1-p+p^2)] - N[2/3(1-p_o+p_o^2)]\}^2}{N[2/3(1-p_o+p_o^2)]}$$

$$+ \frac{\{N[1/3(1-p)^2] - N[1/3(1-p_o)^2]\}^2}{N[1/3(1-p_o)^2]} + \frac{\{N[2/3(p)(1-p)] - N[2/3(p_o)(1-p_o)]\}^2}{N[2/3(p_o)(1-p_o)]}$$

when solved for N, this simplifies to:

$$N = \frac{\chi^2}{K_{AD}} \quad \text{where} \quad K_{AD} = \frac{48p^4-128p^3+136p^2-64p+11}{9}$$

EF equation (1 degree of freedom):

$$\chi^2 = \frac{\{\{N[1/3(p)(2-p)]+N[2/3(1-p+p^2)]\} - \{N[1/3(p_o)(2-p_o)]+N[2/3(1-p_o+p_o^2)]\}\}^2}{N[1/3(p_o)(2-p_o)]+N[2/3(1-p_o+p_o^2)]}$$

$$+ \frac{\{\{N[1/3(1-p)^2]+N[2/3(p)(1-p)]\} - \{N[1/3(1-p_o)^2]+N[2/3(1-p_o+p_o^2)]\}\}^2}{N[1/3(1-p_o)^2]+N[2/3(1-p_o+p_o^2)]}$$

when solved for N, this simplifies to:

$$N = \frac{\chi^2}{K_{EF}} \quad \text{where} \quad K_{EF} = \frac{16p^4-8p^2+1}{27}$$

# APPENDIX H

## CHI-SQUARE CALCULATION FOR THE DISTRIBUTION OF $F_2$ PROGENY IN CLASSES A-D FOR THE CROSS *an* x *emb22*

| | $F_2$ class | | | |
|---|---|---|---|---|
| | A | B | C | D |
| Observed | 108 | 252 | 54 | 53 |
| Expected | 116.75 | 233.50 | 38.76 | 77.99 |

$$\chi^2 = \frac{(108-116.75)^2}{116.75} + \frac{(252-233.5)^2}{233.5} + \frac{(54-38.76)^2}{38.76} + \frac{(53-77.99)^2}{77.99}$$

$= 16.1$ Significantly different at $P < 0.005$.

6:3:2:1

6:3:2:1

# APPENDIX I

## CHI-SQUARE CALCULATION FOR THE DISTRIBUTION OF PROGENY IN CLASSES E and F FOR THE CROSS *ch1* x *emb22*

| | $F_2$ class | |
|---|---|---|
| | E | F |
| Observed | 328 | 339.75 |
| Expected | 125 | 113.25 |

$$\chi^2 = \frac{(328-339.75)2}{339.75} + \frac{(125-113.25)2}{113.25}$$

= 1.6 Not significantly different.

# APPENDIX J

## CHI-SQUARE CALCULATION FOR THE DISTRIBUTION OF $F_2$ PROGENY IN CLASSES E AND F FOR THE CROSS *ttg* x *biol*

| | $F_2$ class | |
|---|---|---|
| | E | F |
| Observed | 932 | 432 |
| Expected | 1023 | 341 |

$$\chi^2 = \frac{(932-1023)2}{1023} + \frac{(432-341)2}{341}$$

$= 32.4$ Significantly different at $P < 0.005$.

# APPENDIX K

## CHI-SQUARE CALCULATION FOR THE DISTRIBUTION OF $F_2$ PROGENY IN CLASSES I AND II FOR THE BACKCROSS W2 X emb30

| | Class | |
|---|---|---|
| | I | II |
| Observed | 59 | 101 |
| Expected | 80 | 80 |

$$\chi^2 = \frac{(59-80)2}{80} + \frac{(101-80)2}{80}$$

= 11.0 Significantly different at P <0.005.

# APPENDIX L

## MOLECULAR BIOLOGY PROTOCOLS USED IN RFLP MAPPING

Plant DNA Isolation - 5 plants/pool

Extraction Buffer: 0.1 M Tris (8.0)
0.1 M EDTA (8.0)
0.25 M NaCl
100 μg/ml Proteinase K

1) Grind 5 plants to a powder in liquid $N_2$ and transfer frozen powder to weigh boat on dry ice. Estimate fresh weight of tissue if not available.

2) Use spatula to transfer frozen powder to 50 ml Oak Ridge Tube or 250 ml centrifuge bottle then add 5-10 mls extraction buffer per gram fresh weight, mix to suspend cells.

3) Add enough 10% Sarkosyl to give final concentration of 1% sarkosyl, mix gently via inversion to lyse cells and incubate 1-2 hours at 55°C.

4) Spin 7K rpm (JA-20 rotor) 10 min; 4°C.

5) Save supernatant, add 0.6 vol cold isopropanol, gently invert to mix, incubate 30 min -20°C.

6) Spin 9K (JA-20) 15 min save nucleic acid pellet and resuspend pellet in 4.5 mls TE80.

7) Add 4.85 g CsCl; mix well to dissolve; incubate on ice 30 min; transfer to 15 ml Corex tube.

8) Spin 9K for 10 min; use sterile pasteur pipet to transfer supernatant (minus pellicle) to new 50 ml Oak Ridge Tube.

9) Add 250 μl of 10 mg/ml EtBr (WEAR GLOVES); incubate 30 min on ice.

10) Pellet RNA with 9K, 10 min spin; transfer supernatant to 5 ml quick seal ultracentrifuge tube.

11) Spin 60K, 10-12 hours, 0 deceleration, 20°C, VTi 65.2 rotor.

12) Use UV to illuminate tube, puncture top of tube with 21g needle, then insert 18.5 g needle (on a syringe) just below fluorescent DNA band; draw off DNA slowly, remove needle and squirt DNA into Corning 15 ml tube.

13) Extract DNA to 2X past colorless with isopropanol over CsCl water in 15 ml Corning screw cap tube.

14) Add 2vol $H_2O$ and 6vol cold EtOH, mix, incubate 1hr at -20°C. (DNA should appear as stringy, white ppt.)

15) Transfer to 15 ml corex tube, spin 10 min 9K; resuspend DNA pellet in 400 μl TE80 transfer to microfuge tube.

in 400 μl TE80 transfer to microfuge tube.

16) Add 40 μl (0.1vol) 3M Na Acetate and 800 μl (2vol) cold EtOH mix and incubate 1hr -20°C to reprecipitate DNA.

17) Spin 10 min in microfuge, wash DNA pellet with 1 ml 80% EtOH, briefly air dry pellet, resuspend in 50 μl TE80. Let dissolve overnight at 4° C. Estimate concentration and purity using 260/280 readings from spectrophotometer.

# HIGH-EFFICIENCY VACUUM BLOTTING

## REAGENTS

0.25 N HCl (20 mls concentrated HCl and $H_2O$ up to 1 liter)

Denaturation solution (0.5 M NaOH; 1.5 M NaCl; 2 mM EDTA)

20.00 g NaOH
87.66 g NaCl
4.00 mls 0.5 M EDTA
$dH_2O$ up to 1 liter

Neutralization solution
(0.5 M Tris-HCl {pH=7}; 3 M NaCl; 2 mM EDTA)

500 mls 1 M Tris-HCl {pH = 7}
175.3 g NaCl
4.0 mls 0.5 M EDTA
$dH_2O$ up to 1 liter

20X SSC (3 M NaCl; 0.3 M sodium citrate)

175.3 g NaCl
88.2 g sodium citrate (citric acid trisodium salt)
$dH_2O$ up to 800 mls
adjust pH to 7.0 with NaOH
adjust final volume to 1 liter with $dH_2O$
FILTER THROUGH 0.45 μm MESH FILTER BEFORE USE

Note: All of the following steps require gentle shaking of the gel in 3 gel volumes of each solution.

1) 15-20 min in 3 volumes 0.25 N HCl (15 for 0.75 mm gels 20 for thicker) DO NOT OVER-FRAGMENT THE DNA!

2) Decant acid; rinse gel ≈ 1 min in $dH_2O$; replace with 3 volumes of denaturation solution; agitate for 30 min; can be longer

3) Decant; rinse in $dH_2O$; replace with 3 volumes neutralization solution; shake for 30 min (can be longer)

4) Transfer gel to prepared vacuum blotter; seal edges of gel with 1% molten agar or agarose. Turn on vacuum and set to 40 cm. Add filtered 20X SSC to immerse the gel to approximately twice its thickness. Vacuum blot for 2 hrs or longer. Photograph restained gel. Crosslink DNA to Nytran on 3MM filter paper (wetted with 2X SSC) at a setting of 1,200 (x100) μJ with Stratalinker. Briefly rinse membrane with 2X SSC before drying or use.

# GENIUS PROTOCOL - HYBRIDIZATION

Solutions:

**20X SSC** = 3M NaCl; 0.3M sodium citrate (citric acid sodium salt) pH = 7.0

**Hybridization solution** = 5X SSC; 1.5% Blocking reagent, 0.1% sarkosyl; 0.02% SDS; 50% formamide; 10 % dextran sulfate

| | |
|---|---|
| 13 mls | $H_2O$ |
| 50 mls | formamide |
| 25 mls | 20X SSC |
| 1.5g | blocking reagent (vial #11) |
| 200 µl | 10% SDS |
| 1 ml | 10% sarkosyl |
| 10 g | Dextran sulfate |

put on stir plate 1-2 hrs at 60° C to dissolve blocking reagent

**2X SSC; 0.1 % SDS**

**0.5X SSC; 0.1 % SDS**

For 130 cm2 membranes use the following volumes:

1) Run DNA out on gel; fragment the DNA with HCl; denature with NaOH; and neutralize. Blot 2hrs with 40 cm H20 vacuum with 20X SSC overlay. Crosslink DNA to Nytran with 1200 (x 100 µJ) using Stratalinker. Wash membrane briefly with 2X SSC. **SEE VACUUM BLOTTING PROTOCOL.**

2) Pre-hybridize membrane at 42° C for at least 1 hr with 26 mls hybridization solution in a sealed bag.

3) Replace pre-hybridization solution with 3.25 mls hybridization solution containing freshly denatured probe (10 min at 95° C and quick chill at -70° C for 10 min.). Hybridize at 42° C at least 6 hrs; can go up to 24 hrs.

** Note that you should begin making Buffer #2 prior to step 4 if you are going to go directly to the immunological detection procedure.

4) Wash membrane 2 x 5 min. at room T° with 65 mls 2X SSC; 0.1% SDS.

5) Wash membrane 2 x 15 min. at 65° C with 65 mls 0.5X SSC; 0.1% SDS.

6) Membrane can then be used directly for immunological detection or stored air dried for later use.

# GENIUS PROTOCOL - DNA LABELING

1) Isolate and purify linearized template DNA (may use phenol/chloroform). Bring final pellet up in 37.5 μl TE-80.

2) Denature template DNA by heating in a 95° C water bath for 10 min. then quickly chill tube on dry ice. Complete denaturation is essential.

3) Add the following to a microfuge tube on ice:

   1 μg template DNA (or desired amount) in 37.5 μl TE-80
   5 μl hexanucleotide mixture (random primers) vial #5
   5 μl dNTP labeling mixture vial #6
   2.5 μl Klenow enzyme (2U/μl) vial #7

4) Incubate at 37° C for at least 1hr., may go as long as 20 hrs. Stop the reaction by adding 5 μl of 0.2M EDTA (pH = 8.0).

5) Add: 2.5 μl of 20 mg/ml glycogen (acts as a carrier), 22.5 μl $H_2O$ and 7.5 μl of 4M LiCl, mix.

6) Add 250 μl of cold absolute EtOH; mix and transfer to siliconized tube (dig-deoxy-labeled DNA pellets will be difficult to get into sol'n otherwise). Incubate tube on dry ice or at -70° C for 30 min.

7) Spin in microfuge at room T° for 10 min. drain tube and briefly dry pellet.

8) Resuspend pellet with 50 μl of TE-80 + 0.1% SDS, incubate at 37° C for 10 min. with frequent vortexing. Add contents of tube to 10 mls of hybridization solution and store at -20° C. Use 3.25 mls for 130 $cm^2$ blots. Boil and quick-freeze hybridization solution before use.

# GENIUS PROTOCOL - IMMUNOLOGICAL DETECTION

Solutions: (note that vloumes are for 130 $cm^2$ membranes)

Buffer #1 = 100 mM Tris-HCl; 150 mM NaCl; pH = 7.5

Buffer #2 = 1.25% Blocking reagent and 50 $\mu$g/ml ssDNA in Buffer #1
(2.5 g blocking reagent and 1ml of 10 mg/ml ssDNA per 200mls buffer)

Buffer #3 = 100 mM Tris-HCl; 100 mM NaCl; 50 mM $MgCl_2$; pH = 9.5
Make fresh by adding 10mls each of 1M Tris-HCl (pH = 9.5) and 1M NaCl to 75 mls $H_2O$; then, while stirring, slowly add 5mls 1M $MgCl_2$.

Buffer #4 = 10 mM Tris-HCl; 1mM EDTA; pH = 8.0

(Prepare the following just prior to use)

Antibody conjugate sol'n = 5.2$\mu$l (vial #8) in 26 mls buffer #2

Color solution = 58.5 $\mu$l NBT (vial #9)
45.5 $\mu$l x-phos (vial #10)
13 mls buffer #3

All of the following incubations are carried out at room T° with gentle agitation (except for the color development). Volumes are for 130 $cm^2$ membranes.

1) Wash filter briefly (≈1 min.) in 130 mls buffer #1.

2) Incubate membrane for 30 min. with 130 mls buffer #2.

3) Make Antibody conjugate sol'n.

4) Incubate membrane for 30 min. with 26 mls Antibody conjugate (in sealed bag).

5) Wash 2 x 15 min. with 130 mls Buffer #1.

6) Equilibrate membrane for 2 min. with 26 mls buffer #3.

7) Incubate filter in the dark with 13 mls color solution (in sealed bag) may go as long as 1 day.

8) When (and if) desired bands are detected stop reaction by rinsing membrane with 65 mls buffer #4.

9) Photocopy and/or photograph blot as soon as possible to document results.

# VITA

David A. Patton

Candidate for the Degree of

Doctor of Philosophy

Thesis: MAPPING GENES ESSENTIAL FOR EMBRYO DEVELOPMENT IN ARABIDOPSIS THALIANA

Major Field: Botany

Biographical:

Personal Data: Born in New Orleans, Louisiana, September 2, 1962, the son of Dean and Vera Jo Patton.

Education: Graduated from Cairo American College Senior High School, Cairo, Egypt, in May 1980; attended Northern Oklahoma Junior College, Tonkawa, Oklahoma, from August 1980 to May 1982; received Bachelor of Science Degree in Arts and Sciences from Oklahoma State University, Stillwater, Oklahoma in May 1984; received Master of Science Degree in Botany from Oklahoma State University in December 1986; completed requirements for the degree of Doctor of Philosophy at Oklahoma State University in May 1991.

Professional Experience: Teaching Assistant, Department of Botany, Oklahoma State University, Fall 1986, Fall 1987, and Fall 1988. Member of the Oklahoma Society for Electron Microscopy, American Association for the Advancement of Science, and the International Society for Plant Molecular Biology.