Theor Appl Genet (1988) 75:259–271

Genetics of allozyme variation in *Gossypium arboreum* L. and *Gossypium herbaceum* L. (Malvaceae)

K. A. Suiter *

University of North Carolina, Department of Biology, Coker Hall 010-A, Chapel Hill, NC 27514, USA

Received March 25, 1986; Accepted July 17, 1987
Communicated by A. L. Kahler

Summary. Seed protein extracts from 90 accessions of *Gossypium arboreum* and 70 accessions of *Gossypium herbaceum* were electrophoretically analyzed for isozyme variation. Eighteen enzyme systems were resolved, ten of which were polymorphic among accessions. No within accession isozyme variation was observed within these highly inbred lines. A minimum of 24 genes encode the isozymes resolved and data is presented for codominant inheritance at 13 loci. Tests for non-random joint segregation in 63 of the 78 possible two-locus combinations from the 13 characterized loci give evidence for four pairs of linked genes (*Lap2/Mel* [$r = 0.160 + /-0.027$], *Lap2/Pgil* [$r = 0.285 + /-0.055$], *Mdh6/Tpil* [$r = 0.197 + /-0.028$], and *6Pgd2/6Pgd3* [$r \approx 0.000$]. Numerous presumptive duplicate isozyme loci were observed and these were usually expressed as patterns of nonsegregating heteromultimers within accessions. Single gene expression was also observed at several loci. The observed results are in agreement with those of previous cytological investigations which have proposed the polyploid origin of the diploid Old World *Gossypiums*.

Key words: *Gossypium* – Allozymes – Linkage – Cotton – Gene duplication

Introduction

Gossypium arboreum L. and *G. herbaceum* L. are the only two members of the genus *Gossypium* in the *A* genome group (Beasley 1940; Fryxell 1979; Valicek 1978a, b). These primitive (*sensu* Hutchinson 1962) Old World cotton species are herbaceous or shrubby, annual or perennial plants which are phytographically restricted to the paleotropical savannahs, semi-arid regions, and deserts of Africa, the Middle East, and Asia (Fryxell 1979; Valicek 1978a, b). They have been cultivated for well over 5,000 years (Hutchinson 1962; Fryxell 1979) and are the only diploid ($n = 13$) members of the genus presently being commercially cultivated for their fiber.

Previous studies of genetically determined traits within the *A* genome have documented the inheritance of a variety of morphological and biochemical characters, and a few linkage groups have been constructed (summarized in Endrizzi et al. 1985). Electrophoretic studies of *G. arboreum* and *G. herbaceum* accessions have shown banding patterns for three enzyme systems [catalase, esterase, and leucine amino peptidase (Cherry et al. 1972)]; however, the inheritance of these systems was not studied.

The analysis of isozyme variation in some polyploid and putative diploid species which have undergone ancient ploidy amplifiction has shown an increase in the number of isozymes present at many loci (Gottlieb 1982). Thus, patterns of isozyme variation can be useful markers when developing hypotheses about past events which may have occurred during the evolution of a species. Numerous researchers have proposed an polyploid origin for the diploid *Gossypiums* based on cytological evidence (Lawrence 1931; Davie 1933; Skovsted 1933; Abraham 1940; Beasley 1942); however this hypothesis has not been widely accepted due to the lack of supporting data. An examination of the isozyme variation in these *Gossypium* species, when viewed in light of these earlier findings, may provide further insight into the evolutionary history of the genus and

* Present address: North Carolina State University, Department of Entomology, Box 7634, Raleigh, NC 27695-7634, USA

Table 1. Composition of starch gels, gel buffers, electrode buffers, electrical parameters and stain references

	Gel A	Gel B	Gel C	Gel D
Starch[a]/Sucrose	22.35 g Electrostarch 22.35 g Connaught 15.00 g Sucrose	8.97 g Electrostarch 35.74 g Connaught 15.00 g Sucrose	22.35 g Electrostarch 22.35 g Connaught 15.00 g Sucrose	32.40 g Electrostarch 32.40 g Connaught 15.00 g Sucrose
Electrode[b] Buffer	0.210 M Tris 0.150 M Borate 0.004 M EDTA pH 8.2	0.065 M Histidine 0.019 M Citrate pH 5.7	0.223 M Tris 0.028 M Citrate pH 8.5	0.135 M Tris 0.043 M Citrate pH 7.0
Gel[b] Buffer	0.021 M Tris 0.020 M Borate 0.001 M EDTA 1 N NaOH to pH 8.2	0.009 M Histidine 0.003 M Citrate pH 5.7	0.022 M Tris 0.003 M Citrate pH 8.5	0.009 M Tris 0.003 M Citrate pH 7.0
Gel Buffer Volume	345 ml	345 ml	345 ml	500 ml
Power (watts)	8.8	9.4	6.9	7.4
Time (h)	4.75	7.50	6.25	6.50
Enzymes[c] Resolved/ Reference	CAT PGI PGM TPI	ACO ACP LAP MDH ME 6PGD	AAT DIA FDH GDH SOD	ADH IDH PER

[a] Electrostarch obtained from Otto Hiller Electrostarch Co., Madison, WI (USA). Connaught Starch obtained from Connaught Laboratories, Ltd., Springhill, PA (USA)

[b] Gel A: Wendel and Parks 1982; Gel B: Cardy et al. 1980; Gel C: Suiter 1984; Gel D: Meizel and Markert 1967

[c] ADH, GDH, IDH, LAP, MDH, ME, 6PGD, PGM, PGI, TPI (Siciliano and Shaw 1976); DIA (Harris and Hopkinson 1976); ACO, ACP, CAT, PER (Shaw and Prasad 1970); AAT (Cardy et al. 1980); FDH (Wendel and Parks 1982); SOD (Baum and Scandalios 1982)

possibly lend stronger support for the polyploid origin of these species.

The purpose of this paper is to present the results of a genetic analysis of the mode of inheritance of electrophoretically detectable enzyme variation within and between *G. arboreum* and *G. herbaceum* accessions. Detailed within are the methodologies utilized for the resolution of 18 enzyme systems, the mode of inheritance of variants for 13 postulated loci, and data regarding the linkage arrangements of these genes. Additionally, evidence is presented which supports the hypothesis of a polyploid origin of the *A* genome species.

Materials and methods

Plant materials

Ninety accessions of *G. arboreum* and 70 accessions of *G. herbaceum* were initially screened for isozyme variation by horizontal starch gel electrophoresis. Ten to twenty seeds of each accession were electrophoresed to ascertain the level of heterogeneity within each accession. In every case, each accession was fixed for all bands scored; variation, however, was observed between accessions. Interspecific and intraspecific crosses were made between accessions to produce heterozygous (F_1) individuals. Hand pollinations were accomplished as detailed in Suiter (1984).

F_1 individuals, heterozygous for a number of enzyme systems, were used to produce reciprocal backcross (BC_1) and F_2 families. Seed from BC_1 and F_2 families were collected and stored at 5°–10°C until they could be electrophoretically analyzed.

Extraction of enzymes from seed tissue

Enzymes were extracted from mature seed tissue. The seed coats of individual seeds were removed, the embryos extracted and homogenized by hand in a pre-chilled motar and pestle with 150 µl of chilled extraction buffer [0.05M Na phosphate buffer, pH 7.0, containing 5% sucrose (w/v) and 0.05% β-metcaptoethanol (v/v)]. Crude enzyme extracts obtained from this procedure were used for electrophoresis. All extractions were performed at 2°–4°C.

Electrophoretic procedures and enzyme detection

Enzymes were separated into discrete bands by horizontal starch gel electrophoresis. Table I shows the starch and sucrose compositions of the gel, gel and electrode buffers, the electrical parameters and run time for each system, the enzyme systems resolved, and the reference for each stain recipe. Gels were pre-cooled to 2°–4°C prioir to electrophoresis. Crude extracts were absorbed onto paper wicks (3 mm × 11 mm, Whatmann, no. 3) and inserted vertically into a slit cut across the width of

the gel. Electrophoresis was carried out at 2°–4°C using the times and wattages given in Table I.

Isozyme phenotypes that were poorly resolved by starch gel electrophoresis were further resolved by acrylamide disc electrophoresis (Shumaker et al. 1982). Samples for acrylamide electrophoresis were prepared by adding 15 µl of bromophenol blue and 20 µl of glycerine to 100 µl of crude protein extract. Fifty µl of this mixture were layered directly onto the surface of the stacking gel. Gels were electrophoresed at 2°–4°C in a water-cooled Gelman vertical disc electrophoresis apparatus for 4 h at 40 milliamps.

At the termination of the electrophoretic run, the anodal section of each starch gel was removed from its mold and sliced horizontally into 1 mm–1.2 mm thick slices. Gels were histochemically stained to reveal the isozyme bands according to the methods referenced in Table I. Following acrylamide disc electrophoresis, disc gels were transferred to 10 ml test tubes and stained as above. Gels were incubated at 37 °C until the isozyme bands had developed sufficient intensity to permit scoring.

Allozyme nomenclature

When an enzyme system was controlled by multiple loci, each locus was identified by assigning it an abbreviation of the name given by the International Union of Biochemistry (Webb 1984) [e.g., alcohol dehydrogenase (E.C. 1.1.1.1 = ADH)] and a numerical code. The locus most anodal to the origin was designated as "1", the next "2", etc. At each locus, the fastest migrating (i.e., most anodal) allozyme was given the alphabetical code "*a*". Other allozymes at that locus were assigned codes in alphabetical order relative to their position of migration in the starch gel. Genotypes were designated by superscripting the locus label (e.g., $Adh1^{aa}$ would be the allozyme of *Adh1* migrating the furthest distance from the origin).

Data analysis

Data obtained from BC and F_2 families were analyzed for conformance to expected single-locus Mendelian segregation ratios by chi-square analysis. Contingency chi-square analyses were utilized to test for non-random joint segregation among all pairs of loci. When indications of linked loci were obtained, the recombination fraction and its standard error were estimated using the formulae of Mather (1955) and Allard (1956). These analyses were facilitated by the use of a Pascal computer program (Suiter et al. 1983). Recombination estimates and their standard errors were combined across families using the methods of Kramer and Burnham (1947).

Results and discussion

Eighteen enzyme systems from *Gossypium* protein extracts could be consistently resolved by starch and acrylamide gel electrophoresis. Of the 18 systems resolved, eight were monomorphic in all 160 accessions. The remaining ten systems exhibited allelic variation at one or more enzyme locus. In some cases, this included a null (silent) allele within the zone of enzyme activity.

Monomorphic enzymes

No polymorphisms were detected for eight enzyme systems [(i.e., aconitase (ACO), acid phosphatase (ACP), catalase (CAT), formate dehydrogenase (FDH), glutamate dehydrogenase (GDH), isocitrate dehydrogenase (IDH), peroxidase (PER), and superoxide dismutase (SOD)]. These systems are represented schematically in Fig. 1.

Polymorphic enzymes

The remaining ten enzyme systems displayed polymorphisms and are encoded by a minimum of 24 genes. Of these, segregation data were obtained for 13 loci. The proposed genotypes of the enzyme phenotypes observed at these loci and the relative migration distances of the isozymes within starch gels are illustrated in Fig. 2. A detailed description of these enzyme systems and their associated loci follows.

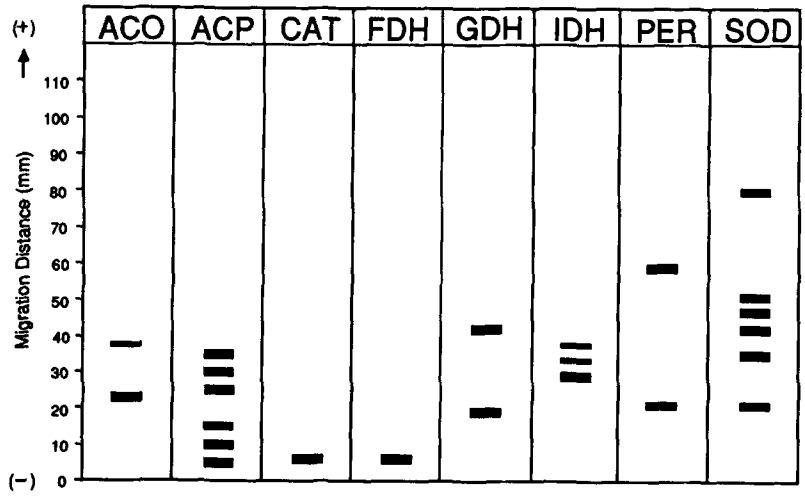

Fig. 1. Schematic representation of banding patterns for eight enzyme systems in which variation was not observed

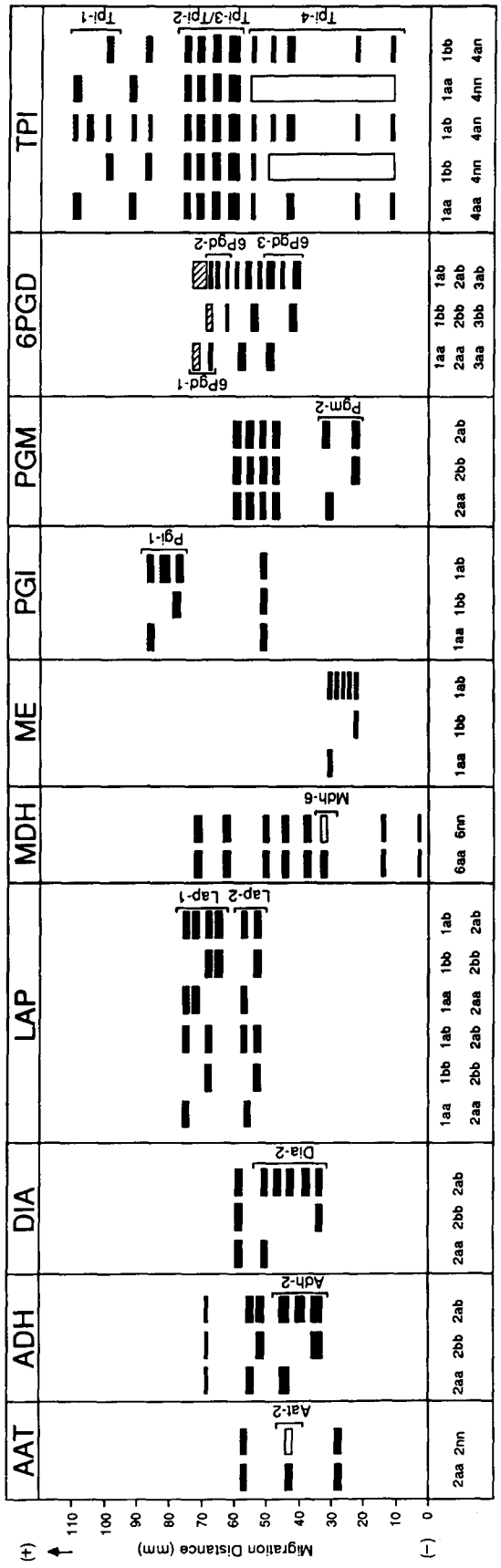

Fig. 2. Representative illustration of allozyme variation in *Gossypium arboreum* and *G. herbaceum* for 14 loci. Phenotypes and genotypes (*below*) are shown. *Clear bands* indicate possible null allelic variants. *Hatched bands* represent areas of poor resolution

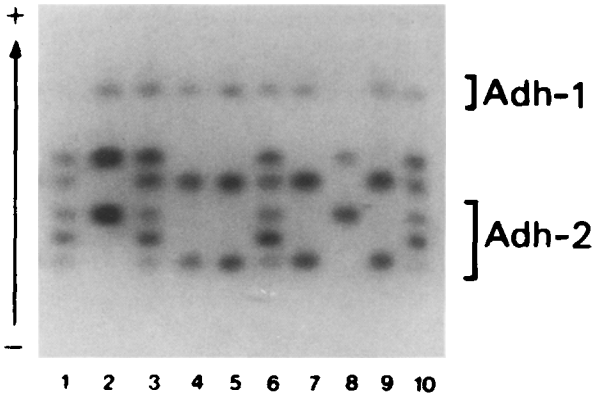

Fig. 3. Photograph of a gel stained for ADH showing segregation at *Adh2* for the F_2 family *G. arboreum* "2485" × *G. herbaceum* "A_2-1". Parental $Adh2^{aa}$ (*tracks 2 and 8*) and $Adh2^{bb}$ (*tracks 4, 5, 7,* and *9*) phenotypes interact with *Adh1* to produce a single heterodimeric band, whereas both intergenic heterodimers and an intragenic heterodimer are observed when *Adh2* is heterozygous (*tracks 1, 3, 6,* and *10*)

Description of polymorphic enzymes

Alcohol dehydrogenase (ADH). Two zones of enzyme activity were detected on gels stained for alcohol dehydrogenase. ADH is dimeric and encoded by two loci, *Adh1* and *Adh2*. Both intergenic and intragenic heterodimers are formed between the products of these two genes (Fig. 3). Allelic variation was detected at the *Adh2* locus only. However, because intergenic dimerization was observed between *Adh1* and *Adh2* in all families analyzed, the dimeric structure of *Adh1* was nonetheless confirmed. The dimeric structure of *Gossypium* ADH is in agreement with that observed in a number of plant species (Schwartz 1969; Torres 1974; Brown 1980; Roose and Gottlieb 1980; Banuette-Bourrillon 1982; Wendel and Parks 1984). Segregation data obtained for *Adh2* conforms to expectations for a single locus governing a functionally dimeric enzyme (Table 2).

Aspartate amino transferase (AAT). Three loci encoding AAT (*Glutamate Oxaloacetate Transaminase, GOT*) were fixed in all accessions of *G. herbaceum* (*Aat1, Aat2,* and *Aat3*), whereas in *G. arboreum,* every accession displayed only the products of *Aat1* and *Aat3*. The loss of activity at *Aat2* in *G. arboreum* could be due to the co-migration of *Aat2* allelic products with those of another locus, or to the presence of a null activity variant ($aat2^{null}$) at this locus. AAT is known a dimeric protein (Schaal 1975; Crawford and Wilson 1979; Skibinski et al. 1984; Arulsekar et al. 1985; Ostergaard et al. 1985). If the products of *Aat2* co-migrate with those of another locus, F_1 or BC *Aat2* heterozy-

gotes should display an intragenic heterodimeric band. However, heterodimers were never observed in starch or acrylamide gels. Thus, it appears that the loss of enzyme activity is due to a null variant at *Aat2*. Interlocus dimerization between the gene product of *Aat1*, *Aat2*, and *Aat3* was not observed. Gene segregation data for *Aat2* conforms to expected Mendelian ratios (Table 2), although these data give no indication to the subunit makeup of in *Gossypium*. To gain additional quantifiable data on this locus, one F_2 family (*G. arboreum* "S2062A" × *G. herbaceum* "2485") segregating in a 1:2:1 expected ratio for $Aat2^{aa}$, $Aat2^{anull}$, and $aat2^{nullnull}$, respectively, was subjected to 7% acrylamide gel electrophoresis. Following electrophoresis, gels were stained and then scanned for AAT activity using a Gilmore scanning spectrophotometer at 500 nm absorbance. The resultant gel scans confirmed that the homozygous $aat2^{nullnull}$ genotype displayed no enzyme activity, whereas the heterozygous $Aat2^{anull}$ genotype exhibited approximately one-half the enzyme activity of the homozygous $Aat2^{aa}$ genotype (Suiter 1984). This data is not unexpected, if we accept the hypothesis that the two copies of the *a* gene carried by the $Aat2^{aa}$ genotype expresses twice the enzyme activity of the heterozygous $Aat2^{anull}$ genotype.

Diaphorase (DIA). Gels stained for diaphorase (= *Menadione Reductase, MNR*) displayed two regions of activity. The fast migrating locus, *Dia1*, was invariant in all material studied. The more cathodal locus, *Dia2*, exhibited two allelic variants, $Dia2^a$ and $Dia2^b$. Intergenic interaction between the two loci was not observed. Starch gel electrophoresis resolved 5-banded phenotypes for heterozygous $Dia2^{ab}$ individuals, suggesting that this enzyme is tetrameric. DIA has been found to be monomeric in Douglas-fir (Neale et al. 1984) and whitebark pine (Furnier et al. 1986). Both monomeric and dimeric loci have been found in maize (Wendel, pers. common.), and DIA is tetrameric in *Camellia* (Wendel and Parks 1982) and in the *D* genome of *Gossypium* (Wendel, pers. common.). Segregation data for *Dia2* is consistent with that of a single locus governing the expression of a codominantly inherited allele (Table 2).

Leucine amino peptidase (LAP). Two regions of staining (i.e., LAP1 and LAP2) were observed in gels stained for LAP. At the *Lap1* locus, homozygous individuals expressed either a one or two-banded phenotype; while at *Lap2* homozygotes were single-banded (Fig. 2). Four-banded heterozygous individuals were synthesized by crossing two-banded $Lap1^a$ and $Lap1^b$ homozygotes (Fig. 4). This result is unusual in that LAP alleles typically produce a single band in most other plant species. The observed two-banded phenotypes could result from post-translational modification, or they could be artifacts of the extraction procedure. Torres et al. (1982) also observed two-banded homozygous LAP phenotypes in *Citrus* spp. Cherry et al. (1972) observed two invariant single-banded zones of LAP activity in *G. arboreum* and *G. herbaceum* following acrylamide electrophoresis. These two regions of staining probably correspond to the *Lap1* and *Lap2* loci described in this study. Both the single-banded and two-banded *Lap1* phenotypes and the single-banded phenotypes at the *Lap2* locus segregate according to Mendelian expectations for loci governing the expression of monomeric enzymes (Table 2).

Malate dehydrogenase (MDH). A total of eight bands was observed for all accessions on gels stained for MDH. The only electrophoretic variation observed was a putative null allele associated with the locus tentatively designated as *Mdh6* (Fig. 2). Due to the possibility that the *Mdh6* isozyme may comigrate with the gene products of another locus, it cannot be ascertained with certainty if this isozyme is a null variant. MDH is known to be dimeric in other plant species and is usually encoded by three or four isozyme loci (Gottlieb 1982). The multiplicity of isozyme bands in *Gossypium* extracts is encoded by an unknown number of genes. Segregation data presented here serves only to substantiate the presence of a putative null allele, and gives no indication of the subunit structure of the enzyme. Segregation data obtained for the *Mdh6* conforms to expected ratios (Table 2).

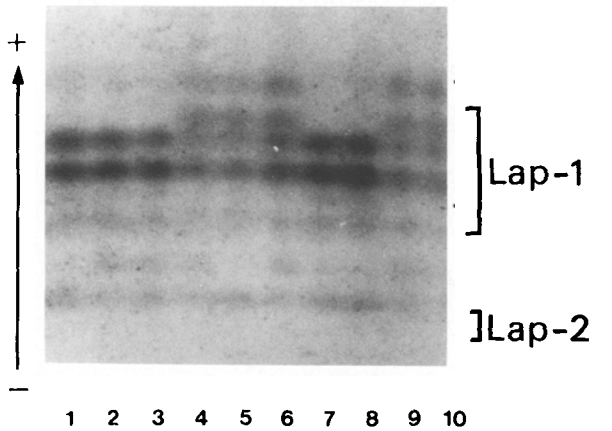

Fig. 4. Portion of a gel stained for LAP exhibiting two- and four-banded isozyme phenotypes in the BC family (*G. herbaceum* "2488" × *G. arboreum* "S2062A") × *G. arboreum* "S2062A". *Tracks 1, 2, 3, 7, and 8* show the homozygous $Lap1^{bb}$ two-banded individuals and *tracks 4, 5, 6, 9,* and *10* four-banded $Lap1^{ab}$ heterozygotes. These two-banded homozygote patterns are unique in that all other families studied produced single-banded phenotypes at *Lap1*

Table 2. Single-locus goodness of-fit tests at 13 polymorphic loci in *Gossypium arboreum* and *G. herbaceum*

Enzyme locus	Family[a]	Family type	Parental genotypes	Offspring genotypes	Expected ratio	df	χ^2	P
Aat2	(2488 × S2062A) × S2062A	BC	an × nn	an : nn 33 : 26	1 : 1	1	0.831	0.362
	(2485 × A_2-1) SELF[b]	F_2	an × an	a- : nn 49 : 14	3 : 1	1	0.259	0.610
	(2485 × 2689) SELF[b]	F_2	an × an	a- : nn 17 : 10	3 : 1	1	2.086	0.147
	(2485 × 2700) SELF[b]	F_2	an × an	a- : nn 23 : 7	3 : 1	1	0.044	0.833
	(2485 × S2062A) SELF[b]	F_2	an × an	a- : nn 25 : 8	3 : 1	1	0.010	0.920
	(2488 × S2062A) SELF[b]	F_2	an × an	a- : nn 66 : 30	3 : 1	1	2.000	0.157
Adh2	(S4004 × RA33/6A) × RA33/6A	BC	ab × aa	aa : ab 11 : 9	1 : 1	1	0.200	0.655
	(GOA2 × SMA4) × GOA2	BC	ab × aa	aa : ab 11 : 13	1 : 1	1	0.167	0.683
	(2700 × b) × b	BC	ab × aa	aa : ab 10 : 10	1 : 1	1	0.000	1.000
	(2485 × A_2-1) SELF	F_2	ab × ab	aa : ab : bb 18 : 47 : 18	1 : 2 : 1	2	1.458	0.482
Dia2	(WAGAD × S4004) × WAGAD	BC	ab × aa	aa : ab 14 : 17	1 : 1	1	0.290	0.590
	(WAGAD × 2700) × WAGAD	BC	ab × aa	aa : ab 16 : 8	1 : 1	1	2.667	0.102
	(2488 × S2062A) × S2062A	BC	ab × bb	ab : bb 12 : 10	1 : 1	1	0.182	0.670
	(WAGAD × S4004) SELF	F_2	ab × ab	aa : ab : bb 17 : 29 : 15	1 : 2 : 1	2	0.279	0.870
	(2485 × 2689) SELF	F_2	ab × ab	aa : ab : bb 6 : 15 : 11	1 : 2 : 1	2	1.687	0.430
	(2485 × S5001) SELF	F_2	ab × ab	aa : ab : bb 11 : 16 : 5	1 : 2 : 1	2	2.250	0.325
	(2485 × 2700) SELF	F_2	ab × ab	aa : ab : bb 5 : 17 : 8	1 : 2 : 1	2	1.133	0.567
	(2485 × S2062A) SELF	F_2	ab × ab	aa : ab : bb 9 : 14 : 10	1 : 2 : 1	2	0.818	0.664
	(2488 × S2062A) SELF	F_2	ab × ab	aa : ab : bb 22 : 42 : 32	1 : 2 : 1	2	3.583	0.167
Lap1	(2488 × S2062A) × S2062A[c]	BC	ab × bb	ab : bb 21 : 36	1 : 1	1	3.947	0.047
	(2700 × b) × b	BC	ab × bb	ab : bb 8 : 12	1 : 1	1	0.800	0.371
	(2485 × 2689) SELF	F_2	ab × ab	aa : ab : bb 8 : 12 : 12	1 : 2 : 1	2	3.000	0.223
	(2485 × 2700) SELF	F_2	ab × ab	aa : ab : bb 5 : 16 : 9	1 : 2 : 1	2	1.200	0.549
	(2485 × S2062A) SELF	F_2	ab × ab	aa : ab : bb 8 : 16 : 9	1 : 2 : 1	2	0.091	0.956
	(2488 × S2062A) SELF[c]	F_2	ab × ab	aa : ab : bb 8 : 12 : 10	1 : 2 : 1	2	1.467	0.480
Lap 2	(GOA2 × SMA4) × GOA2	BC	ab × aa	aa : ab 10 : 14	1 : 1	1	0.667	0.414
	(WAGAD × S4004) × S4004	BC	ab × aa	aa : ab 20 : 21	1 : 1	1	0.024	0.876
	(WAGAD × 2700) × WAGAD	BC	ab × aa	aa : ab 15 : 9	1 : 1	1	1.500	0.221
	(2488 × S2062A) × S2062A	BC	ab × bb	ab : bb 18 : 17	1 : 1	1	0.028	0.866
	2485 × 2700) SELF	F_2	ab × ab	aa : ab : bb 9 : 15 : 6	1 : 2 : 1	2	0.600	0.741

Table 2 (continued)

Enzyme locus	Family[a]	Family type	Parental genotypes	Offspring genotypes	Expected ratio	df	χ^2	P
	(2485 × S2062A) SELF	F_2	ab × ab	aa : ab : bb 7 : 17 : 9	1 : 2 : 1	2	0.273	0.873
	(2488 × S2062A) SELF	F_2	ab × ab	aa : ab : bb 16 : 26 : 12	1 : 2 : 1	2	0.667	0.717
Mdh6	(S4004 × RA33/6A) × RA33/6A	BC	an × nn	an : nn 11 : 9	1 : 1	1	0.200	0.655
	(2700 × b) × b	BC	an × nn	an : nn 8 : 12	1 : 1	1	0.800	0.371
	(2488 × S2062A) × S2062A	BC	an × nn	an : nn 44 : 37	1 : 1	1	0.605	0.437
	(2485 × A_2-1) SELF[d]	F_2	an × an	a- : nn 61 : 21	3 : 1	1	0.061	0.899
	(2485 × S2062A) SELF[d]	F_2	an × an	a- : nn 25 : 8	3 : 1	1	0.010	0.920
	(2488 × S2062A) SELF[d]	F_2	an × an	a- : nn 72 : 24	3 : 1	1	0.000	1.000
Me1	(WAGAD × 2700) × WAGAD	BC	ab × aa	aa : ab 16 : 8	1 : 1	1	1.500	0.221
	(2488 × S2062A) × S2062A	BC	ab × bb	ab : bb 12 : 10	1 : 1	1	0.309	0.579
	(2485 × A_2-1) SELF	F_2	ab × ab	aa : ab : bb 17 : 33 : 13	1 : 2 : 1	2	0.651	0.722
	(2485 × S5001) SELF	F_2	ab × ab	aa : ab : bb 11 : 18 : 3	1 : 2 : 1	2	4.500	0.105
	(2488 × S2062A) SELF	F_2	ab × ab	aa : ab : bb 21 : 27 : 15	1 : 2 : 1	2	2.429	0.297
6pgd2	(GOA2 × SMA4) × GOA2	BC	ab × aa	aa : ab 9 : 15	1 : 1	1	1.500	0.221
	(WAGAD × S4004) × S4004	BC	ab × aa	aa : ab 5 : 6	1 : 1	1	0.091	0.763
6Pgd3	(GOA2 × SMA4) × GOA2	BC	ab × aa	aa : ab 9 : 15	1 : 1	1	1.500	0.221
	(WAGAD × S4004) × S4004	BC	ab × aa	aa : ab 5 : 6	1 : 1	1	0.091	0.763
Pgi1	(WAGAD × S4004) × WAGAD	BC	ab × aa	aa : ab 22 : 19	1 : 1	1	0.220	0.639
	(WAGAD × 2700) × WAGAD	BC	ab × aa	aa : ab 14 : 10	1 : 1	1	0.667	0.414
	(WAGAD × S4004) SELF	F_2	ab × ab	aa : ab : bb 16 : 31 : 14	1 : 2 : 1	2	0.148	0.929
Pgm2	(WAGAD × 2700) × WAGAD	BC	ab × aa	aa : ab 13 : 11	1 : 1	1	0.167	0.683
	(WAGAD × S4004) SELF	F_2	ab × ab	aa : ab : bb 6 : 18 : 7	1 : 2 : 1	2	0.871	0.647
Tpi1	(WAGAD × S4004) × WAGAD	BC	ab × aa	aa : ab 13 : 28	1 : 1	1	5.488	0.019
	(WAGAD × 2700) × WAGAD	BC	ab × aa	aa : ab 11 : 13	1 : 1	1	0.167	0.683
	(2488 × S2062A) × S2062A	BC	ab × bb	ab : bb 40 : 41	1 : 1	1	0.012	0.912
	(WAGAD × S4004) SELF	F_2	ab × ab	aa : ab : bb 23 : 24 : 14	1 : 2 : 1	2	5.426	0.066
	(2485 × 2689) SELF	F_2	ab × ab	aa : ab : bb 8 : 12 : 12	1 : 2 : 1	2	1.188	0.552
	(2485 × A_2-1) SELF	F_2	ab × ab	aa : ab : bb 25 : 39 : 18	1 : 2 : 1	2	1.390	0.499
	(2485 × S5001) SELF	F_2	ab × ab	aa : ab : bb 8 : 20 : 4	1 : 2 : 1	2	3.000	0.223
	(2488 × S2062A) SELF	F_2	ab × ab	aa : ab : bb 16 : 38 : 9	1 : 2 : 1	2	4.328	0.120

continued overleaf

Table 2 (continued)

Enzyme locus	Family[a]	Family type	Parental genotypes	Offspring genotypes	Expected ratio	df	χ^2	P
Tpi4	(2485×2689) SELF[e]	F_2	an×an	a-:nn 26:6	3:1	1	0.667	0.414
	(2485×A_2-1) SELF[e]	F_2	an×an	a-:nn 64:19	3:1	1	0.197	0.657
	(2485×S5001) SELF[e]	F_2	an×an	a-:nn 22:10	3:1	1	0.667	0.414
	(2485×2700) SELF[e]	F_2	an×an	a-:nn 22:8	3:1	1	0.044	0.833
	(2488×S2062A) SELF[e]	F_2	an×an	a-:nn 41:22	3:1	1	3.307	0.069

[a] *Gossypium arboreum* accessions: A_2-1, S2062A, S4004, S5001, SMA4 from the S.G. Stephens collection maintained by C. R. Parks; GOA2 RA33/6A, b obtained from India Council of Agricultural Research, India; 2689, and 2700 obtained from the United States Department of Agriculture collection, College Station, TX. *Gossypium herbaceum* accessions: WAGAD, 2485, and 2488 obtained from the United States Department of Agriculture collection, College Station, TX.
[b] $Aat2^{aa}$ and $Aat2^{anull}$ offspring phenotypes were pooled into one common $Aat2^{a-}$class
[c] Two-banded phenotypes were observed in this family
[d] $Mdh6^{aa}$ and $Mdh6^{anull}$ offspring phenotypes were pooled into one common $Mdh6^{a-}$class
[e] $Tpi4^{aa}$ and $Tpi4^{anull}$ offspring phenotypes were pooled into one common $Tpi4^{a-}$class

Malic enzyme (ME). One zone of activity was observed for malic enzyme (Fig. 2). Because starch gel electrophoresis did not clearly resolve the tightly spaced ME bands, heterozygous individuals were electrophoresed on a 7% acrylamide gel. Heterozygotes displayed five-banded phenotypes, suggesting that this enzyme is tetrameric. Goodman and Stuber (1983) also report tetrameric ME in maize coleoptile tissue. Segregation data are consistent with a codominantly inherited enzyme (Table 2).

6-phosphogluconate dehydrogenase (6PGD). Gels stained for 6PGD exhibited three zones of enzyme activity. The most anodal region of staining, 6PGD1, was poorly resolved and was excluded from further analysis. The two remaining loci, *6Pgd2* and *6Pgd3,* appeared to code for subunits which interact to produce both intergenic and intragenic heterodimers (Fig. 2). In higher plants, 6PGD interlocus dimerization has been reported in *Zea mays* (Goodman et al. 1980), *Lupinus alba* (Green et al. 1980), and *Lycopersicon esculentum* (Tanksley and Kuehn 1985). It is worth noting that most other diploid plant species examined to date contain only two zones of 6PGD activity, one chloroplastic isozyme and one cytosolic isozyme, between which there is no intergenic dimerization (Ireland and Dennis 1980; Gottlieb 1982). In this study, intergenic interaction between *6Pgd1* and the other two loci was not observed. These observations, in addition to evidence of linkage between *6Pgd2* and *6Pgd3* (see Linkage Analyses), suggest that *6Pgd2* and *6Pgd3* are a duplicated gene pair. Segregation data for *6Pgd2* and *6Pgd3* conform to expectations for dimeric alleles (Table 2).

Phosphoglucose isomerase (PGI). PGI exhibited two zones of enzymatic activity. The most anodal locus, *Pgi1,* was polymorphic for two alleles, whereas the slower migrating *Pgi2* was invariant (Fig. 2). Heterozygous *Pgi1* individuals displayed three-banded phenotypes, indicating that this locus codes for a functionally dimeric protein, as has been found in all other plant species examined (Gottlieb 1982; Wendel 1983; Cheliak and Pitel 1984; Kahler and Lay 1985; Ostergaard et al. 1985; Quiros and McHale 1985). Segregation data obtained for *Pgi1* are concordant with this hypothesis (Table 2).

Phosphoglucomutase (PGM). Two zones of activity were detected in all material studied. The fast migrating region, PGM1, was invariant for four strongly staining bands in all accessions. The more cathodal locus, *Pgm2,* exhibited either a single slow or a single fast isozyme phenotype (Fig. 5). Heterozygous F_1 individuals displayed two bands of activity for *Pgm2.* In diploid plant species two loci are usually present, however three or more loci have been observed. The products of these loci have invariably been shown to be monomeric (Wendel and Parks 1982; Arus and Orton 1984; Cheliak et al. 1984; Kahler and Lay 1985; Furnier et al. 1986), thus the invariant, multiple bands at PGM1 suggest the activity of more than a single locus. Backcross and F_2 segregation ratios for *Pgm2* support the hypothesis that a single locus codes for this monomeric allozyme (Table 2).

Triose phosphate isomerase (TPI). It appears that at least four loci govern the expression of TPI in *G.*

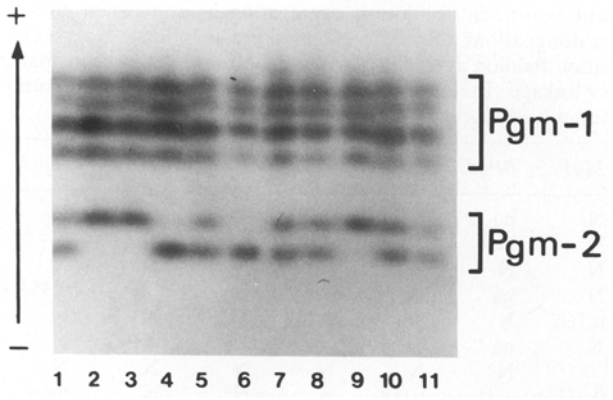

Fig. 5. Photograph of segregation for *Pgm2* in the F_2 family *G. herbaceum* "WAGAD" × *G. arboreum* "S4004". Homozygous $Pgm2^{aa}$ and $Pgm2^{bb}$ individuals produce single-banded phenotypes (*tracks 2, 3, 4, 6,* and *9*). Two-banded heterozygotes are shown in *tracks 1, 5, 7, 8, 10,* and *11*. The more anodal *Pgm1* is seen as a series of four isozyme bands, invariant in all accessions analyzed

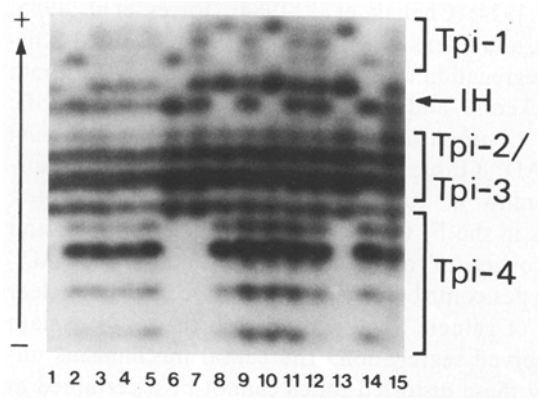

Fig. 6. Photograph of the F_2 family *G. herbaceum* "2488" × *G. arboreum* "S2062A" segregating for both *Tpi1* and *Tpi4*. *Tpi1* is a dimeric protein, producing three isozyme bands in heterozygous $Tpi1^{ab}$ individuals (*tracks 1, 3, 4, 5, 7, 9, 11, 12,* and *15*). Parental $Tpi1^{aa}$ and $Tpi1^{bb}$ homozygous phenotypes are present in *tracks 2, 6, 8, 10, 13,* and *14*. Overlap at the *Tpi2* and *Tpi3* loci in the middle region of the gel prevented these loci from being scored reliably. Note the intermediate location of the intragenic heterodimeric (IH) bands produced through the random interaction of *Tpi1* and *Tpi2* alleles. Variation at *Tpi4* is observed as either a four-banded phenotype ($Tpi4^{aa}$ and $Tpi4^{anull}$ phenotypes are indistinguishable, *tracks 1, 2, 3, 4, 8, 9, 10, 11, 14,* and *15*) or as a zone of null enzyme activity ($tpi4^{nullnull}$) in this same region (*tracks 6, 7,* and *13*)

arboreum and *G. herbaceum* seed tissue. At *Tpi1*, there are two alleles, which in heterozygotes interact to form three-banded phenotypes (Figs. 2 and 6). Segregation of allelic products at *Tpi1* conforms to Mendelian expectations for a dimeric allele (Table 2).

The bands produced by the *Tpi2* locus and those heterodimers formed from the interaction of *Tpi1* and *Tpi2* overlap within the same region of the starch gel as those of *Tpi3*, rendering the scoring of *Tpi1* and *Tpi2* difficult. However, the allelic products of *Tpi2* interact with those of *Tpi1* to produce visible interlocus heterodimers (Fig. 6), indicating a dimeric structure for this enzyme. Triose phosphate isomerase has been found to be dimeric in all other plant species studied (Gottlieb 1982; Wendel and Parks 1982; Pichersky and Gottlieb 1983). Analysis of TPI in other diploid plant species has usually detected two isozymes which correspond to the plastid and cytosolic forms, the products of which do not produce hybrid gene products (Bukowiecki and Anderson 1974). The relative position of *Tpi1* and *Tpi2* in the starch gel (Figs. 2 and 6) and their intergenic dimerization suggests a duplicate gene pair. This situation is analogous to that of *Clarkia,* which has duplicate sets of nuclear genes coding for multi-banded plastid and cytosolic isozymes (Pichersky and Gottlieb 1983), the products of which form intergenic and intragenic heterodimers within each compartment, but not between compartments. Because *Tpi2* cannot be scored accurately and the heterodimers formed with *Tpi1* overlap with *Tpi3*, any possible linkage cannot be ascertained.

The *Tpi3* zone of activity was seen as a darkly staining, multi-banded, invariant region on the starch gel which overlapped with the products of *Tpi2*.

The locus tentatively designated as *Tpi4* displayed either a four-banded, five-banded, or null zone of enzymatic activity (Figs. 2 and 6). In segregating progenies, the four-banded or five-banded phenotypes segregated as a single unit. Possible explanations for this phenomenon include: (1) they are a series of tightly linked genes, (2) that one of the slow migrating bands interacts with other unresolved bands in the *Tpi3* region to produce the multiple bands, or (3) that a primary *Tpi4* isozyme is post-translationally modified, either in vivo or in vitro to produce the extra bands. To my knowledge, the segregation of a series of isozyme bands like those associated with *Tpi4* has not been observed in other diploid plant species. Additional information concerning the origin of these bands and their genetic substructure must await further experimentation. Presently, the segregation data obtained for *Tpi4* is consistent with that of a single locus (Table 2).

Linkage analyses

Data obtained from segregating families were available to test 63 of the 78 possible two-locus combinations for evidence of non-random joint segregation. The results of contingency chi-square analyses of these data are presented in Table 3. Evidence for linkage was detected for four pairs of genes: *Lap2/Me1, Lap2/Pgi1, Mdh6/Tpi1* and *6Pgd2/6Pgd3*. Table 4 presents maximum

Table 3. Summary of contingency Chi-square analysis of independent assortment for jointly segregating locus pairs in crosses among various *Gossypium arboreum* and *G. herbaceum* accessions. The designations above the diagonal are as follows: N = no evidence of linkage was detected; numerical value = weighted recombination fraction averaged across all families for that locus pair; na = no families segregating for these two loci were available to test for linkage. The number below the diagonal denotes the number of families (and the total number of individuals) tested for non-random joint segregation

	Aat2	*Adh2*	*Dia2*	*Lap1*	*Lap2*	*Mdh6*	*Me1*	*6Pgd2*	*6Pgd3*	*Pgi1*	*Pgm2*	*Tpi1*	*Tpi4*
Aat2		N	N	0.336	N	N	N	na	na	na	na	N	N
Adh2	2 (63)		N	N	N	N	N	N	N	na	na	N	N
Dia2	5 (221)	1 (30)		N	N	N	N	N	N	N	N	N	N
Lap1	6 (185)	1 (30)	7 (208)		0.385	N	N	na	na	na	na	N	0.353
Lap2	4 (162)	1 (24)	5 (181)	4 (127)		N	0.160	N	N	0.285	N	N	N
Mdh6	5 (283)	2 (103)	4 (181)	4 (150)	3 (131)		N	na	na	na	na	0.197	N
Me1	4 (217)	1 (63)	7 (214)	5 (181)	4 (134)	5 (208)		N	N	N	N	N	N
6Pgd2	–	1 (24)	2 (35)	–	3 (59)	–	2 (35)		0.000	N	N	N	na
6Pgd3	–	1 (24)	2 (35)	–	3 (59)	–	2 (35)	2 (35)		N	N	N	na
Pgi1	–	–	4 (119)	–	2 (65)	–	3 (45)	2 (34)	1 (24)		N	N	na
Pgm2	–	–	4 (119)	–	2 (55)	–	3 (45)	2 (34)	1 (24)	3 (85)		N	na
Tpi1	5 (247)	1 (83)	10 (308)	5 (181)	6 (150)	3 (226)	2 (228)	2 (34)	1 (24)	3 (85)	3 (85)		N
Tpi4	4 (188)	1 (83)	5 (187)	5 (154)	2 (93)	2 (146)	4 (157)	–	–	–	–	5 (207)	

likelihood estimates of the recombination fractions between all locus pairs which exhibited evidence of linkage. The locus pair *Lap2/Me1* showed significant deviation from expected joint segregation ratios in three of the four families analyzed. The combined recombination fraction across all families is 0.160 + /– 0.027, $\chi^2 = 70.712$, $P = 0.0001$ (Table 4). For the gene pair *Lap2* and *Pgi1*, significant chi-square values were obtained in two of the four families. Estimation of the combined recombination fraction was 0.285 + /–0.055, $\chi^2 = 12.308$, $P = 0.005$ (Table 4). Linkage of *Mdh6* and *Tpi1* was detected in all families jointly segregating for this gene pair; the combined recombination estimate is 0.197 + /–0.028, $\chi^2 = 67.201$, $P = 0.0001$ (Table 4). Recombination was not detected in the two families which were jointly segregating for *6Pgd2* and *6Pgd3*. The absence of recombination suggests that these two loci may have arisen by a tandem duplication, perhaps by unequal crossing over between chromosomal segments.

Inconsistent non-random joint segregation was obtained for three additional pairs of loci (Table 4). Weighted estimates of recombination fractions averaged across segregating families (Table 4) suggest weak linkage between *Aat2/Lap1* ($r = 0.336 + /-0.038$), *Lap1/Lap2* ($r = 0.385 + /-0.040$), and *Lap1/Tpi4* ($r = 0. + /-0.044$). Individual linkage estimates for these loci were heterogeneous, however, suggesting that these deviations should probably be ascribed to chance occurrences.

Distorted segregation ratios

Disturbed segregation ratios have been reported for a number of enzyme systems in other plant species (Torres 1974; Cheliak et al. 1984; Torres et al. 1985; Wendel and Parks 1984; Kahler and Lay 1985). Distorted segregation ratios were observed for *Tpi1* in both the backcross and F_2 progenies of the interspecific cross, *G. arboreum* "S4004" and *G. herbaceum* "WAGAD" (Table 2). In the backcross family the distorted ratio is due to an excess of heterozygotes, whereas in the F_2 family the lack of heterozygotes and the recovery of an excess of parental "WAGAD" phenotypes contributes to the distortion. Thus, no clear pattern of gamete loss is evident which would explain the observed segregation. The causal mechanisms underlying these distorted ratios cannot be ascertained at this time. However, segregation for six other loci in these two families shows ratios consistent with expectations (Table 2). Gerstel (1953) has shown that synthesis of *G. arboreum* and *G. herbaceum* interspecific hybrids can lead to the formation of translocation heterozygote complexes. Formation of translocation heterozygotes of chromosome segments involving enzyme loci may produce gametic deficiencies in segregating progenies. This may be the case here, although data addressing this hypothesis is not available at this time.

Distorted segregation was also observed at the *Tpi4* locus in the F_2 family *G. arboreum* "S2062A" × *G. herbaceum* "2488" (Table 2). This single deviation from expected ratios can probably be attributed to chance.

Evidence for the polyploid origin of Gossypium arboreum and G. herbaceum

Electrophoretic banding patterns of *Gossypium arboreum* and *G. herbaceum* seed protein extracts exhibit

Table 4. Maximum likelihood estimates of recombination fraction between segregating loci

Locus pair	Family[a]	Family type	N	Genotype[b]									χ^2	df	P	$r \pm SE$
				A_1B_1	A_1B_2	A_1B_3	A_2B_1	A_2B_2	A_2B_3	A_3B_1	A_3B_2	A_3B_3				
Aat2/Lap1	(2488×S2062A)×S2062A	BC	35	10	08		02	15					7.441	1	0.006	0.286±0.076
Aat2/Lap1	(2485×2689) SELF	F_2	27	04	07	06	03	04	03				0.157	2	0.925	0.475±0.118
Aat2/Lap1	(2485×2700) SELF	F_2	30	05	10	08	00	06	01				4.068	2	0.131	0.500±0.112
Aat2/Lap1	(2485×A_2-1) SELF	F_2	30	03	15	05	00	06	01				1.384	2	0.501	0.474±0.112
Aat2/Lap1	(2485×S2062A) SELF	F_2	33	07	14	04	01	02	05				6.607	2	0.037	0.260±0.087
Aat2/Lap1	(2488×S2062A) SELF	F_2	30	08	09	04	00	03	06				7.857	2	0.020	0.223±0.085
Totals			185	37	63	27	06	36	16				27.514	11	0.004	0.336±0.038
Lap1/Lap2	(2488×S2062A)×S2062A	BC	35	06	12		06	11					0.015	1	0.903	0.482±0.084
Lap1/Lap2	(2485×2700) SELF	F_2	30	01	07	01	01	06	08	03	03	00	12.180	4	0.016	0.418±0.088
Lap1/Lap2	(2485×S2062A) SELF	F_2	30	05	10	08	00	06	01	02	05	02	1.845	4	0.764	0.474±0.087
Lap1/Lap2	(2488×S2062A) SELF	F_2	30	03	15	05	00	07	01	01	03	06	14.328	4	0.006	0.250±0.067
Totals			125	15	44	14	07	30	10	06	11	08	28.368	13	0.008	0.385±0.040
Lap1/Tpi4	(2485×2689) SELF	F_2	32	06	02		11	01		09	03		1.368	2	0.505	0.500±0.108
Lap1/Tpi4	(2485×2700) SELF	F_2	30	05	00		12	04		05	04		3.285	2	0.192	0.305±0.098
Lap1/Tpi4	(2485×A_2-1) SELF	F_2	30	01	02		19	02		05	01		6.171	2	0.046	0.213±0.083
Lap1/Tpi4	(2485×S5001) SELF	F_2	32	08	06		09	01		05	03		3.125	2	0.210	0.431±0.106
Lap1/Tpi4	(2488×S2062A) SELF	F_2	30	02	06		11	01		05	05		9.514	2	0.009	0.419±0.109
Totals			154	22	16		62	09		29	16		24.463	10	0.006	0.353±0.044
Lap2/Me1	(2488×S2062A)×S2062A	BC	35	15	03		00	17					24.792	1	0.000	0.086±0.047
Lap2/Me1	(WAGAD×2700)×WAGAD	BC	24	13	02		02	07					9.967	1	0.002	0.167±0.076
Lap2/Me1	(WAGAD×S4004)×WAGAD	BC	21	03	07		03	08					0.019	1	0.890	0.476±0.109
Lap2/Me1	(2488×S2062A) SELF	F_2	54	14	02	00	05	16	05	01	03	08	35.934	4	0.000	0.169±0.040
Totals			134	45	14	00	10	48	05	01	03	08	70.712	7	0.000	0.160±0.027
Lap2/Pgi1	(WAGAD×2700)×WAGAD	BC	24	10	05		04	05					1.143	1	0.285	0.375±0.099
Lap2/Pgi1	(WAGAD×S4004)×WAGAD	BC	41	16	04		06	15					10.986	1	0.001	0.244±0.067
Totals			65	26	09		10	20					12.129	2	0.002	0.285±0.055
Mdh6/Tpi1	(2488×S2062A)×S2062A	BC	80	34	10		06	30					29.091	1	0.000	0.200±0.045
Mdh6/Tpi1	(2485×A_2-1) SELF	F_2	82	25	28	08	00	12	09				15.677	2	0.000	0.263±0.055
Mdh6/Tpi1	(2488×S2062A) SELF	F_2	63	16	32	02	00	06	07				22.649	2	0.000	0.145±0.047
Totals			225	75	70	10	06	48	16				67.417	5	0.000	0.197±0.028
6Pgd2/6Pgd3	(GOA2×SMA4)×GOA2	BC	24	09	00		00	15					24.000	1	0.000	0.000± –
6Pgd2/6Pgd3	(WAGAD×S4004)×WAGAD	BC	11	05	00		00	06					11.000	1	0.000	0.000± –
Totals			35	14	00		00	21					35.000	2	0.000	0.000± –

[a] Accessions are the same as those listed in Table 2

[b] *Aat2/Lap1:* $A_1 = Aat2^{anull}$ (except in the F_2 families, $A_1 = Aat2^{aa} + Aat2^{anull}$); $A_2 = aat2^{nullnull}$; $B_1 = Lap1^{aa}$; $B_2 = Lap1^{ab}$; $B_3 = Lap1^{bb}$
Lap1/Lap2: $A_1 = Lap1^{aa}$; $A_2 = Lap1^{ab}$; $A_3 = Lap1^{bb}$; $B_1 = Lap2^{aa}$; $B_2 = Lap2^{ab}$; $B_3 = Lap2^{bb}$
Lap1/Tpi4: $A_1 = Lap1^{aa}$; $A_2 = Lap1^{ab}$; $A_3 = Lap1^{bb}$; $B_1 = Tpi4^{aa} + Tpi4^{anull}$; $B_2 = tpi4^{nullnull}$
Lap2/Me1: $A_1 = Lap2^{aa}$; $A_2 = Lap2^{ab}$; $A_3 = Lap2^{bb}$; $B_1 = Me1^{aa}$; $B_2 = Me1^{ab}$; $B_3 = Me1^{bb}$
Lap2/Pgi1: $A_1 = Lap2^{aa}$; $A_2 = Lap2^{ab}$; $B_1 = Pgi1^{aa}$; $B_2 = Pgi1^{ab}$
Mdh6/Tpi1: $A_1 = Mdh6^{anull}$ (except in the F_2 families, $A_1 = Mdh6^{aa} + Mdh6^{anull}$); $A_2 = mdh6^{nullnull}$; $B_1 = Tpi1^{aa}$; $B_2 = Tpi1^{ab}$; $B_3 = Tpi1^{bb}$
6Pgd2/6Pgd3: $A_1 = 6Pgd2^{aa}$; $A_2 = 6Pgd2^{ab}$; $B_1 = 6Pgd3^{aa}$; $B_2 = 6Pgd3^{ab}$

more isozymes for some enzymes than would be expected for a diploid species (Gottlieb 1982). The presence of duplicate genes for *Tpi1/Tpi2* and the multiple isozyme bands observed for MDH, PGM, and TPI are probably the result of an ancient polyploidy event. The patterns of variation for these enzymes are suggestive of the kinds of enzyme variation found in polyploid species (Gottlieb 1982); but, many other enzyme systems studied displayed typical diploid isozyme patterns. Additonally, null zones of enzyme activity were present at *Aat2, Mdh6,* and *Tpi4.* The apparent diploid patterns at many isozyme loci, but the polyploid patterns at others, may be explained by paleopolyploidation followed by the loss of duplicate gene expression at *Aat2, Mdh6,* and *Tpi4.*

There is additional evidence which supports a polyploid origin of these species. Cytological data support the hypothesis that *G. arboreum* and *G. herbaceum* are polyploid in origin and arose from an ancestral allopolyploid species (Lawrence 1931; Davie 1933; Skovsted 1933; Abraham 1940; Beasley 1942). Davie (1933), Skovsted (1933), and Beasley (1942) have proposed that the base chromosome number of the genus, $n = 13$, is secondarily polyploid, derived from an ancestral base number of six or seven.

These results are consistent with the hypotheses of Endrizzi et al. (1985), in which at least two ploidy amplification events are proposed to have occurred during the evolution of the genus. The first event, involving the common ancestors of the present diploid species of *Gossypium*, is hypothesized to have occurred approximately 70 M.Y.B.P. (million years before present).

The genus *Gossypium* is considered to be ancient; however, because fossil data is lacking (Fryxell 1979), it is difficult to estimate the approximate time-frame for the first polyploidation event. The earliest fossil records of the Malvales occur in the Upper Eocene, approximately 40 M.Y.B.P. (Muller 1981). However, fossil fruits of possible Malvalean origin have been collected from late Cretaceous strata (> 65 M.Y.B.P., B. Tiffney, pers. commun.). Additionally, evidence obtained from an analysis of the present-day geographical distribution patterns of *Gossypium* species (Fryxell 1979; Endrizzi et al. 1985), suggests that the genus may be much older than 40 million years. The ancient development of the genus has been previously proposed by a number of authors (Skovsted 1934; Hutchinson et al. 1947; Saunders 1961; Phillips 1963; Fryxell 1979) but see Davie (1935), Hutchinson et al. (1947), Hutchinson (1959) and Johnson (1975).

Diploidization, or the loss of duplicate gene expression, has been observed in the allotetraploid *AD* genome of *Gossypium* for many morphological markers (see Endrizzi et al. 1985, pp 347–349) and within some enzyme systems [(e.g., PGM and TPI (Suiter, unpubl. data)]. A more detailed survey of the diploid and allotetraploid cotton species may expose additional patterns of diploidization and gene loss like those observed in *Gossypium arboreum* and *G. herbaceum.*

Acknowledgements. I would like to thank Dr. C. R. Parks for support provided throughout the course of this project. I am especially indebted to Drs. J. F. Wendel, M. M. Goodman, and three anonymous referees for critically reviewing the manuscript and for helpful suggestions. Drs. P. Bretting, M. Edwards, F. Gould, and Ms. C. Tempte provided valuable editorial comments. Dr. Parks, the Indian Council of Agricultural Research, and the United States Department of Agriculture provided seed of the accessions used in this study.

References

Abraham P (1940) Cytological studies in Gossypium. I. Chromosome behavior in interspecific hybrid *G. arboreum* × *G. stocksii.* Indian J Agric Sci 10:285–298

Allard RW (1956) Formulas and tables to facilitate the calculation of recombination values in heredity. Hilgardia 24:235–278

Arulsekar S, Parfitt DE, McGranahan GH (1985) Isozyme markers in *Juglans* species. J Hered 76:103–106

Arus P, Orton TJ (1984) Inheritance patterns and linkage relationships of eight genes in celery. J Hered 75:11–14

Banuette-Bourillon F (1982) Linkage of alcohol dehydrogenase structural genes in pearl millet (*Pennisetum typhoides*). Biochem Genet 20:359–367

Baum JA, Scandelios, JG (1982) Developmental expression and intracellular localization of superoxide dismutases in maize. Differentiation 13:133–140

Beasley J (1940) The origin of the American tetraploid species of *Gossypium* species. Am Nat 74:285–286

Beasley J (1942) Meiotic chromosome behavior in species, species species hybrids, haploids, and induced polyploids of *Gossypium.* Genetics 27:25–54

Brown ADH (1980) Genetic basis of alcohol dehydrogenase polymorphism in *Hordeum spontaneum.* J Hered 71:127–128

Bukowiecki AC, Anderson LE (1974) Multiple forms of aldolase and triose phosphate isomerase in diverse plant species. Plant Sci Lett 3:381–386

Cardy BJ, Stuber CW, Goodman MM (1980) Techniques for starch gel electrophoresis of enzymes from maize (*Zea mays* L.). Institute of Statistics, Mimeograph Ser 1317, North Carolina State University

Cheliak WM, Morgan K, Dancik BP, Strobeck C, Yeh FCH (1984) Segregation of allozymes in megagametophytes of viable seed from a natural population of jack pine, *Pinus banksiana* Lamb. Theor Appl Genet 69:145–151

Cheliak WW, Pitel P (1984) Genetic control of allozyme variants in mature tissues of White Spruce trees. J Hered 75:34–40

Cherry JP, Katterman FRH, Endrizzi JE (1972) Seed esterases, leucine aminopeptidase and catalases of species of the genus *Gossypium.* Theor Appl Genet 42:218–226

Crawford DJ, Wilson H (1979) Allozyme variation in several closely related diploid species of *Chenopodium* in the western United States. Am J Bot 66:237–244

Davie JH (1983) Cytological studies in the Malvaceae and certain related families. J Genet 28:33067

Davie JH (1935) Chromosome studies in the Malvaceae and certain related families. II. Genetica 17:487–498

Endrizzi JE, Turcotte EL, Kohel RJ (1985) Genetics, cytology, and evolution of *Gossypium*. Adv Genet 23:271–375

Fryxell PA (1979) The natural history of the cotton tribe. Texas A & M University Press, College Station, Texas, p 245

Furnier, GR, Knowles P, Aleksiuk MA, Dancik BD (1986) Inheritance and linkage of allozymes in seed tissue of whitebark pine. Can J Genet Cytol 28:601–604

Gerstel DU (1953) Chromosomal translocations in interspecific hybrids of the genus *Gossypium*. Evolution 7: 234–244

Goodman MM, Stuber CW (1983) Maize. In: Tanksley, SD, Orton, TL (eds) Isozymes in plant genetics and breeding, part B, Elsevier, Amsterdam, pp 1–33

Goodman MM, Stuber CW, Newton KJ, Weissinger HH (1980) Linkage relationships of 19 enzyme loci im maize. Genetics 96:697–710

Gottlieb LD (1982) Conservation and duplication of isozymes in plants. Science 216:373–380

Green AD, Brown ADH, Oram RN (1980) Determination of outcrossing rate in a breeding population of *Lupinus alba* L. (White Lupin). Z Pflanzenzücht 84:181–191

Harris H, Hopkinson DA (1976) Handbook of enzyme electrophoresis in human genetics. North Holland, Amsterdam, p 500

Hutchinson JB (1959) The application of genetics to cotton improvement. Cambridge University Press, London

Hutchinson JB (1962) The history and relationships of the world's cottons. Endeavour 21:5–15

Hutchinson JB, Silow RA, Stephens SG (1947) The evolution of *Gossypium*. Oxford University Press, London, p 161

Ireland RJ, Dennis DT (1980) Isozymes of the glycolytic and pentose-phosphate pathways in storage tissues of different oilseeds. Planta 149:476–479

Johnson BL (1975) *Gossypium palmeri* and a polyphyletic origin of the New World cottons. Bull Torrey Bot Club 102:340–349

Kahler AL, Lay CL (1985) Genetics of electrophoretic variants in the annual sunflower. J Hered 76:335–340

Kramer HH, Burnham CR (1947) Methods for combining linkage intensity values from backcross F_2 and F_3 genetic data. Genetics 32:379–390

Lawrence WJC (1931) The secondary association of chromosomes. Cytologia 2:352–384

Mather K (1955) The measurement of linkage in heredity. J Wiley and Sons, New York, p 149

Meizel S, Markert CL (1967) Malate dehydrogenase isozymes of the marine snail *Ilyanassa obsoleta*. Arch Biochem Biophys 122:753–765

Muller J (1981) Fossil pollen records of extant angiosperms. Bot Rev 47:1–142

Neale DB, Weber JC, Adams WT (1984). Inheritance of needle tissue in Douglas-fir. Can J Genet Cytol 26: 459–468

Ostergaard H, Neilsen G, Johansen H (1985) Genetic variation in cultivars of diploid ryegrass, *Lolium perenne* and *L. multiflorum*, at five enzyme systems. Theor Appl Genet 69: 409–421

Phillips LL (1963) The cytogenetics of *Gossypium* and the origin of the New World cottons. Evolution 17:460–469

Pichersky E, Gottlieb LD (1983) Evidence for duplication of the structural genes coding plastid and cytosolic triose phosphate isomerase in diploid species of *Clarkia*. Genetics 105:421–436

Quiros CF, McHale N (1985) Genetic analysis of isozyme variants in diploid and tetraploid potatoes. Genetics 111: 131–145

Roose ML, Gottlieb LD (1980) Biochemical properties and levels of expression of alcohol dehydrogenases in the allotetraploid plant *Tragopogon miscellus* and its diploid progenitors. Biochem Genet 18:1065–1085

Saunders JH (1961) The wild species of *Gossypium* and their evolutionary history. Oxford University Press, London, p 62

Schaal B (1975) Population structure and local differentiation in *Liatris cylindrica*. Am Nat 109:511–528

Schwartz D (1969) Alcohol dehydrogenase in maize: genetic basis for multiple isozymes. Science 164:585–586

Shaw CR, Prasad R (1970) Starch gel electrophoresis of enzymes – a compilation of recipes. Biochem Genet 4: 297–320

Shumaker KM, Allard RW, Kahler AL (1982) Cryptic variability at enzyme loci in three plant species, *Avena barbata, Hordeum vulgare*, and *Zea mays*. J Hered 73:73–90

Siciliano MJ, Shaw CR (1976) Separation and visualization of enzymes on gels. In: Smith, I (ed) Chromatographic and electrophoretic techniques, vol 2. Heinemann Medical Books, London, pp 185–209

Skibinski DOF, Rasool D, Erskine W (1984) Aspartate aminotransferase allozyme variation in a germplasm collection of the domesticated lentil (*Lens culinaris*). Theor Appl Genet 68:441–448

Skovsted A (1933) Cytological studies in cotton. I. The mitosis and meiosis in diploid and triploid Asiatic cotton. Ann Bot 47:227–251

Skovsted A (1934) Cytological studies in cotton. II. Two interspecific hybrids between Asiatic and New World cottons. J Genet 28:407–424

Suiter KA (1984) Genetics of Alloenzyme Variation in *Gossypium arboreum* L. and *Gossypium herbaceum* L. (Malvaceae). Ms Thesis, University of North Carolina, Chapel Hill

Suiter KA, Wendel JF, Case JS (1983) LINKAGE-1: a PASCAL computer program for the detection and analysis of genetic linkage. J Hered 74:203–204

Tanksley SD, Kuehn GD (1985) Genetics, subcellular localization, and molecular characterization of 6-phosphogluconate dehydrogenase isozymes in tomato. Biochem Genet 23: 441–454

Torres AM (1974) An intergenic alcohol dehydrogenase isozyme in sunflowers. Biochem Genet 11:301–308

Torres AM, Soost RK, Mau-Lastovicka T (1982) Citrus isozymes: genetics and distinguishing nucellar from zygotic seedlings. J Hered 73:355–359

Torres AM, Mau-Lastovicka T, Williams TE, Soost RK (1985) Segregation distortion and linkage of *Citrus* and *Poncirus* isozyme genes. J Hered 76:289–294

Valicek P (1978a) Wild and cultivated cottons. Coton Fibres Trop 33:363–387

Valicek P (1978b) Wild and cultivated cottons. Coton Fibres Trop 33:431–448

Webb EC (1984) Enzyme nomenclature. International Union of Biochemistry. Nomenclature Committee. Academic Press, Orlando, Fla, p 646

Wendel JF (1983) Electrophoretic analysis of genetic variation in wild and cultivated *Camellia japonica* L. PhD Thesis, University of North Carolina, Chapel Hill

Wendel JF, Parks CR (1982) Genetic control of isozyme variation in *Camellia japonica* L. (Theaceae). J Hered 73: 197–204

Wendel JF, Parks CR (1984) Distorted segregation and linkage of alcohol dehydrogenase genes in *Camellia japonica* L. (Theaceae). Biochem Genet 22:739–748