Am. J. Hum. Genet. 41:907--924, 1987

Evolutionary Implications of the Human Aldolase-A, -B, -C, and -Pseudogene Chromosome Locations

DEAN R. TOLAN,* JOSHUA NICLAS,† BARRY D. BRUCE,§ AND ROGER V. LEBO‡·§

*Biological Sciences Center, Boston University, Boston; †Department of Medicine, Hematology Division, and ‡Department of Obstetrics, Gynecology, and Reproductive Sciences, and §Howard Hughes Medical Institute, University of California, San Francisco

SUMMARY

The aldolase genes represent an ancient gene family with tissue-specific isozymic forms expressed only in vertebrates. The chromosomal locations of the aldolase genes provide insight into their tissue-specific and developmentally regulated expression and evolution. DNA probes for the human aldolase-A and -C genes and for an aldolase pseudogene were used to quantify and map the aldolase loci in the haploid human genome. Genomic hybridization of restriction fragments determined that all the aldolase genes exist in single copy in the haploid human genome. Spot-blot analysis of sorted chromosomes mapped human aldolase A to chromosome 16, aldolase C to chromosome 17, the pseudogene to chromosome 10; it previously had mapped the aldolase-B gene to chromosome 9. All loci are unlinked and located on to two pairs of morphologically similar chromosomes, a situation consistent with tetraploidization during isozymic and vertebrate evolution. Sequence comparisons of expressed and flanking regions support this conclusion. These locations on similar chromosome pairs correctly predicted that the aldolase pseudogene arose when sequences from the aldolase-A gene were inserted into the homologous aldolase location on chromosome 10.

INTRODUCTION

Fructose-1,6-diphosphate aldolase (E.C.4.1.2.13) is a ubiquitous glycolytic enzyme that catalyzes the reversible aldol cleavage of fructose-1,6-diphosphate to

Received March 25, 1987.

Address for correspondence and reprints: Dr. Dean R. Tolan, Biological Sciences Center, Boston University, Boston, MA 02215; phone (617) 353-5310.

 0002-9297/87/4105-0011$02.00

glyceraldehyde-3-phosphate and dihydroxyacetone phosphate. The class I aldolases of animals and higher plants all consist of four 40,000-dalton subunits (Penhoet et al. 1967) with highly conserved amino acid and nucleotide sequences, indicating that all arose from a common ancestral gene. Three tissue-specific forms exist in higher vertebrates (Penhoet et al. 1966): aldolase A (muscle and red blood cells), aldolase B (liver, kidney, and intestine), and aldolase C (nervous tissue). A fourth electrophoretically separable isozyme has been reported in advanced fish (Lebherz and Rutter 1969).

The coordinate expression of these isozymes has been reported during ontogeny and carcinogenesis (Rutter et al. 1963). One class of hereditary nonspherocytic hemolytic anemic disorders is caused by aldolase A deficiency. In one patient with an unstable enzyme (Beutler et al. 1973), normal nucleated cells replace aldolase A at a sufficient rate, but aldolase-A activity becomes depleted in enucleated erythrocytes in which transcriptional and translational activity decreases as cells age. Consequently, energy production is impaired and membrane stability decreases with declining ion-transport activity. In addition, a deficiency of the related liver aldolase-B isozyme results in fructose intolerance (Chambers and Pratt 1956). To further understand the structure and expression of these "housekeeping" genes, the complete structures of aldolase mRNA transcripts and genes have been determined from a variety of organisms (Rottmann et al. 1984; Tolan et al. 1984; Burgess and Penhoet 1985; Tsutsumi et al. 1985). Previously we cloned the human aldolase-B gene and mapped it to chromosome 9 (Lebo et al. 1985). In the present paper we describe (1) the isolation, sequence determination, and mapping of both the human aldolase-A cDNA and of an aldolase pseudogene, (2) the mapping of the human aldolase-C gene, and (3) comparison of the aldolase-A, -B, and -C genes and aldolase pseudogene.

The chromosomal location of the human aldolase-A, -B, and -C genes and of aldolase pseudogenes are significant in the context of isozyme evolution, polyploidization, human chromosome morphology, and vertebrate evolution. Theories of vertebrate evolution postulate that vertebrates arose from primitive chordates via an initial tetraploidization and that a second tetraploidization event occurred in ancestral crossopterigian fish before vertebrates left the sea to live on land (Ohno 1970, 1973; Ohno et al. 1986). These theories were based upon the DNA content of chordates, fish, other vertebrates (Ohno 1970, 1973; Ohno et al. 1986), and chromosome morphology (Comings 1972). This is further supported by recent karyotype analyses (Sawyer and Hozier 1986) and the conservation of linkage groups between species (Seuanez 1979; Morizot 1983; O'Brien et al. 1986). The first polyploid organisms would have been selected through increased vigor and stabilization of more-fit intermediate genotypes (Stebbins 1966). Simultaneously, tetraploidization would have generated the additional gene loci from which isozymes evolved, as well as duplicating the chromosome complement to produce morphologically duplicate pairs of homologous chromosome pairs. Subsequent rearrangements or "diploidization" (Ohno 1970) would generate similar but different chromosomes carrying approximately twice as much DNA, while individual genes simultaneously di-

verged into tissue-specific isozymes. Duplicated human chromosomes would continue to look similar provided subsequent chromosome rearrangements were not too extensive. With this kept in mind, the banded human karyotype has been divided into morphologically similar pairs of pairs as follows: 1 and 2, 4 and 5, 7 and 8, 9 and 10, 11 and 12, 14 and 15, 16 and 17, 19 and 20, and 21 and 22. The remaining chromosomes have been paired, with minimal rearrangements (Comings 1972). This theory of vertebrate evolution can be tested by rigorous comparison of aldolase gene sequences and their chromosome map positions, since aldolase isozymes are found in all vertebrates.

MATERIAL AND METHODS

Isolation of Aldolase-A cDNA Clones

Two libraries constructed in pUC9 with poly A^+ RNA from human umbilical cord endothelial cells and human lung fibroblasts by means of the linker-adapter method (Coleclough and Erlitz 1985) were screened (Grunstein and Hogness 1975), first with a 580-bp *Kpn*I-*Pst*I fragment of pRM223, a rabbit aldolase-A cDNA clone (Tolan et al. 1984), and, second, with a 260-bp *Pst*I fragment of the human aldolase-A clone, pHA404, isolated from the endothelial cell library. Positive clones were purified via three rounds of screenings. DNA probes were excised by restriction-endonuclease digestion, gel electrophoresed twice (Maniatis et al. 1982), and nick-translated (Lebo et al. 1985).

Preparation and Characterization of Human Aldolase Clones

Aldolase-pseudogene clones were isolated and characterized from a partial *Hae*III/*Alu*I library constructed in Charon 4A (Lawn et al. 1978; Tolan and Penhoet 1986). Genomic clones were isolated and subcloned into pUC vectors (Vieira and Messing 1982). Purified plasmid subclones and cDNA (Davis et al. 1980) were used for restriction-enzyme analysis.

DNA Sequencing and Computer Analysis

The insert fragments of cDNA and subclones were excised, isolated, and subcloned into M13-derivative vectors (Norrander et al. 1983) for subsequent sequence analysis (Sanger et al. 1977). The analogue, 7-deaza-2′-deoxyguanosine-5′-triphosphate was substituted for dGTP to eliminate G/C compression (Barr et al. 1986). The DNA sequence was analyzed for translation, restriction maps, and homologies on a VAX/UNIX system with programs developed by Hugo Martinez of the University of California, San Francisco. Optimal alignments between two sequences were generated using a program (MALIGN) that aligned segments of ≥ 3 bp that were in common. The program scored the alignment based upon the number of bases in common, with a penalty of 2.0 for each insertion or deletion and of 0.05 for each base in the insertion or deletion. The penalty values were determined empirically by aligning the homologous 3′-untranslated regions of human (Tolan and Penhoet 1986) and rabbit (Amsden 1985) aldolase-B genes, which possess several different-size insertions and deletions. With the same values, the programs further analyzed the optimal

alignments by randomizing the repeat elements and realigning them. Repeated randomizations generated a binomial distribution from which the statistical significance of the optimal alignment was calculated.

Analysis of Genomic and Sorted Chromosomal DNA

DNA purified from lymphoblast cell lines or peripheral blood lymphocytes (Blin and Stafford 1976) was digested with 50 units of *Eco*RI, phenol extracted, ethanol precipitated, electrophoresed, transferred to nitrocellulose paper (Southern 1975), hybridized, and washed (Lebo et al. 1985). Spot-blot panels were sorted, denatured, baked, and prehybridized (Lebo et al. 1985). Previously hybridized panels or blots were denatured by incubation at 65 C for 30 min in 50% (v/v) deionized formamide, 50 mM Na-Hepes (pH 7.0), 3 × SSC (1 × SSC = 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0), 200 μg denatured salmon-sperm DNA/ml, 150 μg yeast RNA/ml, and 1 × Denhardt's solution (20 mg polyvinyl-pyrrolidone-360/ml, 20 mg Ficoll 400/ml, 20 mg bovine serum albumin [Sigma A-4378]/ml). The prehybridization solution was removed, and the filters were then hybridized for 1 day in prehybridization solution plus 10% dextran sulfate and 3 × 10^6 cpm human aldolase probe/ml. The filters were washed three times for 10 min each in 2 × SSC, 0.1% sodium dodecyl sulfate at 20 C; then for 1.5 h in 0.1 × SSC, 0.1% SDS with gentle shaking and one solution change at a higher temperature; then twice briefly in 0.1 × SSC at 25 C, dried, and radioautographed.

RESULTS

Human Aldolase-A and Pseudogene Clones

The largest recombinant plasmid of eight that were isolated from 50,000 transformants, pHA404, was characterized, and the insert was used to probe a human lung-fibroblast cDNA library. One of the 202 isolated clones, pHL-1, was larger than pHA404. The relationship of the two clones to each other was determined by the sequencing strategy illustrated in figure 1, panel *A*. The derived amino acid sequence (fig. 2, panel *A*) was identical to the published partial sequence of the human aldolase-A protein (Freemont et al. 1984). The 881-bp nucleotide sequence starts at the codon for lysine-138 and contains the sequence coding for the last 226 amino acids and the entire 3′-noncoding region. The protein-coding region shares 92.2% nucleotide and 98.7% amino acid homology with the respective region of the rabbit aldolase A (Tolan et al. 1984).

Several overlapping genomic clones encoding an aldolase-related pseudogene were isolated, and two separate subclones were constructed. These were subcloned further, and the DNA sequence was determined. Figure 1*B* depicts the restriction map and sequencing strategy. The sequence is shown in figure 2*B*. This sequence has been aligned with the known amino acid sequence of aldolase A. Amino acid substitutions are shown as blanks in the amino acid sequence. The 2,082-bp sequence shares 80% amino acid homology and 90% nucleotide homology with the above-cited aldolase-A cDNA. This pseudogene lacks the introns known to exist in the human aldolase-B (Tolan and Penhoet

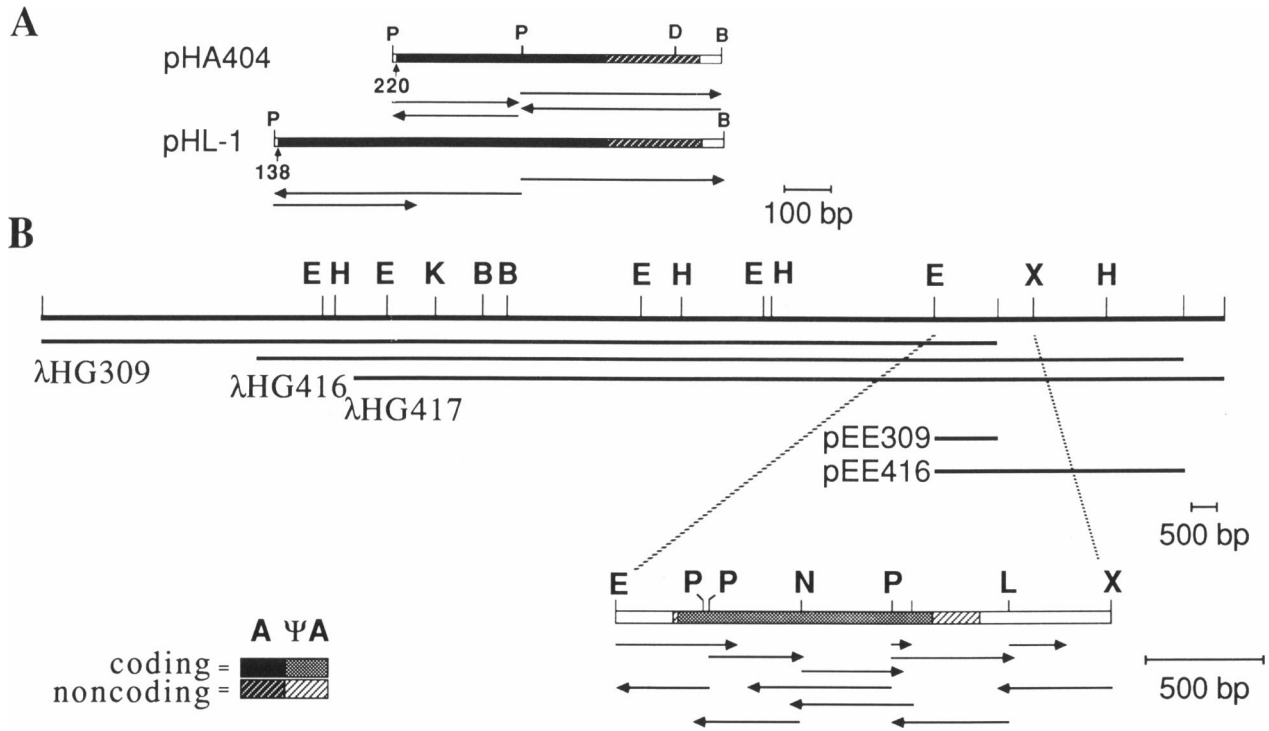


FIG. 1.—Aldolase A–related clones. *Panel A*, Restriction map of the human aldolase-A cDNA clones. The numbers indicate the amino acid positions (Tolan et al. 1984). *Panel B*, Mapped restriction sites of the human aldolase-pseudogene region with the overlapping genomic clones and subclones. Below that is an enlarged restriction map of the gene region for which the DNA sequence was determined. Mapped restriction-enzyme sites are denoted for *Bam*HI (B), *Eco*RI (E), *Hin*dIII (H), *Kpn*I (K), *Nhe*I (N), *Pst*I (P), *Xba*I (X), *Bcl*I (L), and *Dde*I (D). The arrows denote the direction and extent of DNA sequencing. The key describes the coding and untranslated gene regions.

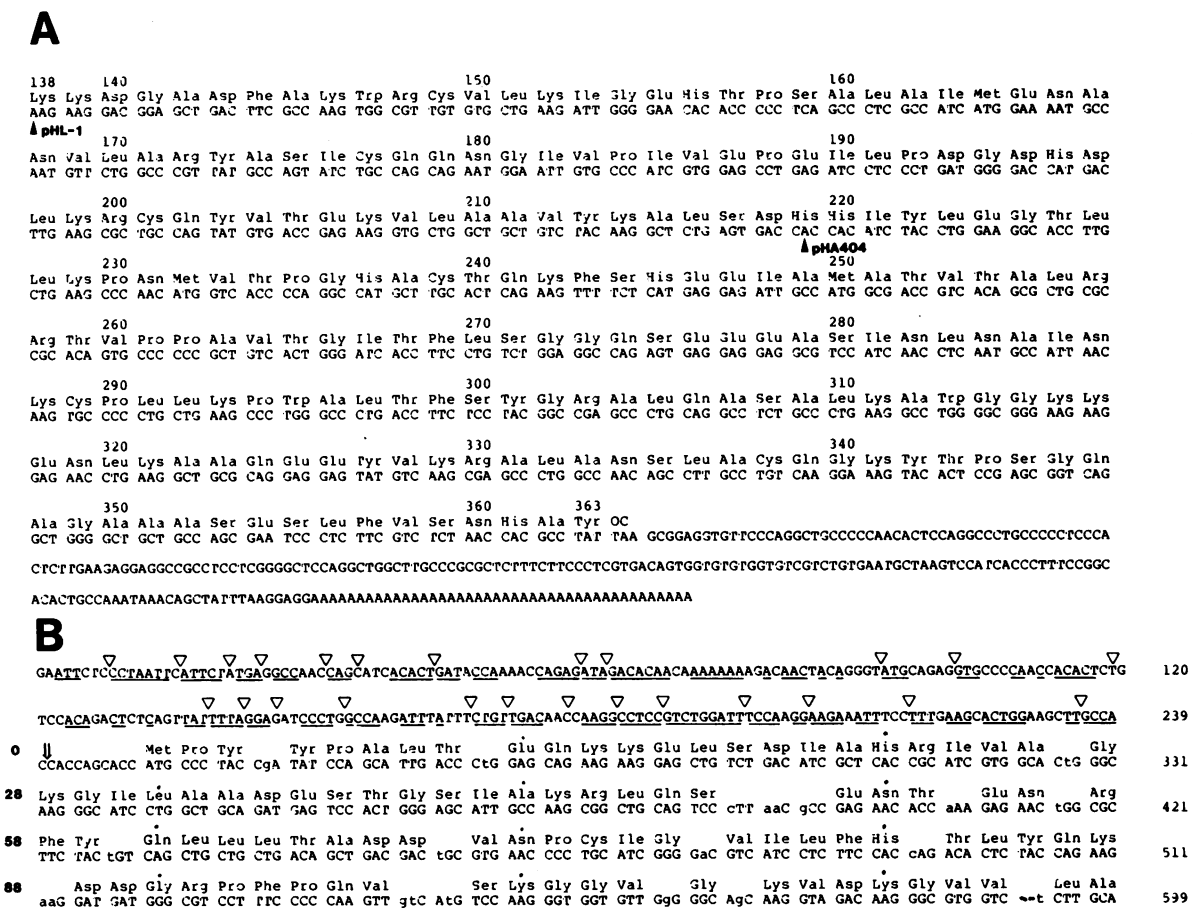

A
pHL-1
pHA404
B

FIG. 2.—Aldolase A nucleotide sequence. *A*, Nucleotide sequence of the overlapping human aldolase-A cDNA clones. The sequence of the entire pHL-1 insert with the deduced amino acid sequence is illustrated above the nucleotide sequence. Arrows indicate the start of the insert region of each clone. There were four differences between pHL-1 and pHA404; all were transitions at the third position in codons 260, 291, and 354 and position 198 in the 3′-untranslated region. *B*, Aldolase pseudogene showing the aldolase-A amino acids above each codon. No amino acid is given where differences exist. Rabbit amino acid sequence was used in cases in which the human sequence was not available (Freemont et al. 1984; Tolan et al. 1984). Lowercase nucleotide bases denote differences between the aldolase-A and -pseudogene sequences. Dashes denote deletions. The vertical arrows in the flanking regions denote the positions where homology to aldolase A stops. At the 5′ end the homology is between rat mRNA sequences; the 3′ end is between human sequences. The wavy line is the boundary where homology to pHL-1 starts. The underlined sequence denotes where there is an alignment with the aldolase-B flanking region (Tolan and Penhoet 1986). The triangles denote where insertions or deletions were made to optimize the alignment.

1986) and -C genes (Rottmann et al. 1987) and in the rat aldolase-A gene (Joh et al. 1986). Indications that this gene does not produce a functional aldolase are as follows: (1) codons 178 and 330 have been mutated to stop codons; (2) the active site lysine-229 has been changed to an asparagine (AAC); and (3) there are three deletions and a single base insertion, some of which cause frameshifts that introduce stop codons in the reading frame. The sequence homology with aldolase-A mRNA decreased before the ends of the mRNA and contained no poly(A). This pseudogene lacks any direct repeats in the 5′- and 3′-flanking regions, repeats that are commonly found in many "processed" pseudogenes (Vanin 1985). Either this gene was not derived from mRNA insertion, or it has been inserted into this location via a mechanism other than retrotransposition.

Aldolase Gene Number and Location

The number of aldolase genes and the specific hybridization conditions for each were determined before gene mapping by spot-blot analysis. Figure 3 depicts the restriction-enzyme fragments detected when aldolase isozyme probes were hybridized to *Eco*RI-digested human genomic DNA. Six fragments hybridized to an aldolase-B cDNA probe (Rottmann et al. 1984) at lower stringency (fig. 3, lane 1). Blots done at higher stringency revealed only the 9.6- and 5.7-kb *Eco*RI fragments (fig. 3, lane 2), a result consistent with the known restriction map of the aldolase-B gene (Tolan and Penhoet 1986). The aldolase-C gene probe, a 1.5-kb *Kpn*I fragment located immediately 3′ to the coding

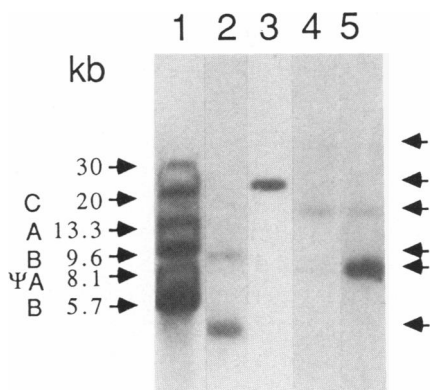
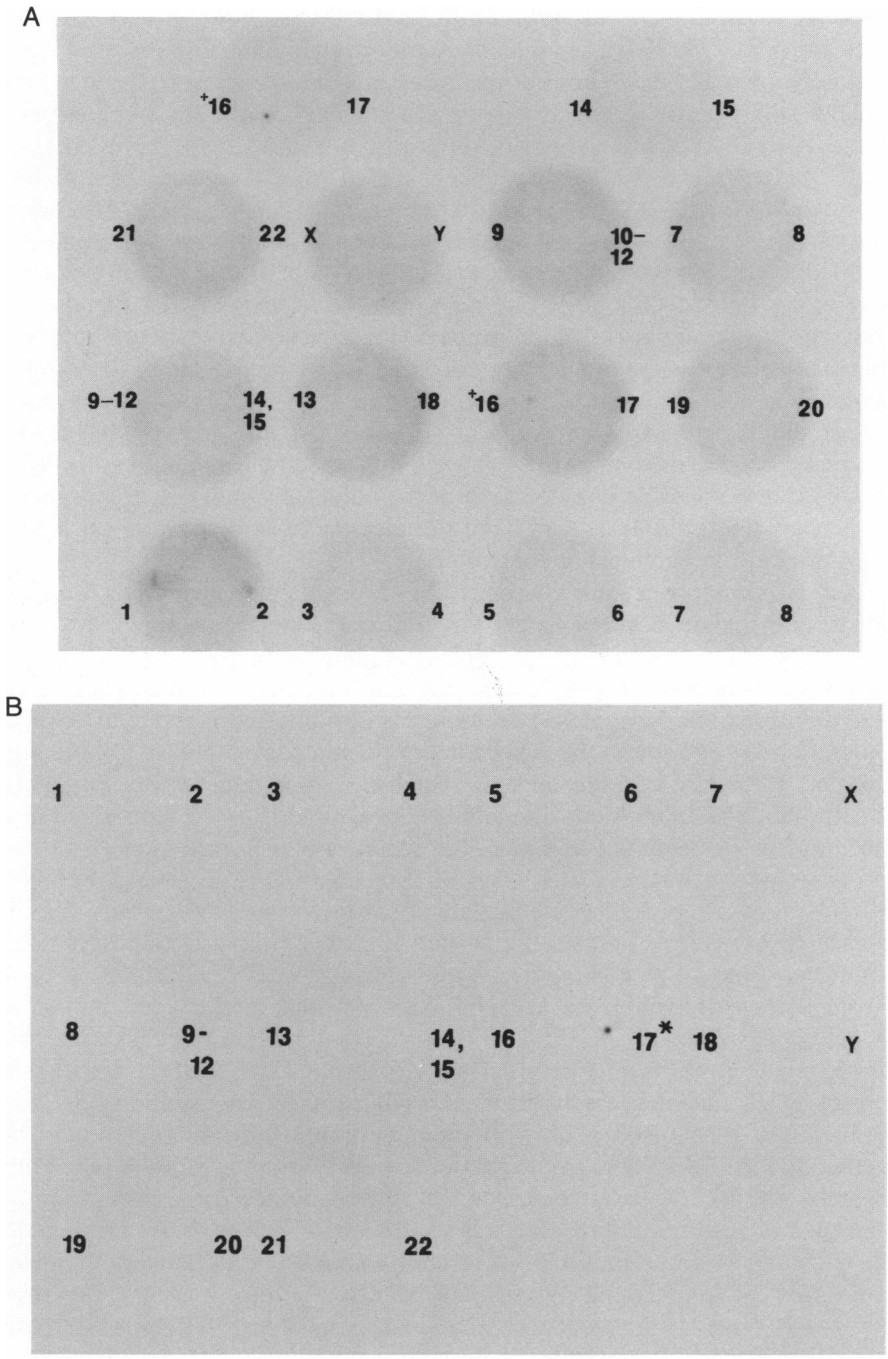


FIG. 3.—Hybridization of specific aldolase gene probes to genomic DNA. Genomic blots used DNA (10 μg) digested with *Eco*RI. Lane 1 represents a different blot than lanes 2–5. The blots were hybridized to a 1.1-kb *Hinf*I fragment of the human aldolase-B cDNA, pHL413 (Rottmann et al. 1984), at 42 C and washed at 45 C in 0.2 × SSC (lane 1) or hybridized at 57 C and washed at 64 C (lane 2). Lane 3 is the same blot as in lane 2 but hybridized with an aldolase-C probe (a 1.5-kb *Kpn*I fragment excised from pKK503; Rottmann et al. 1987) at 42 C and washed at 55 C. Lane 4 is the same blot hybridized with the human aldolase-A cDNA probe (a *Pst*I-*Dde*I fragment excised from pHL-1) at 50 C and washed at 53 C. Lane 5 is the same blot as in lane 4 but hybridized with the human aldolase-A-pseudogene probe (a *Pst*I-*Xba*I fragment excised from pEE416). Radioautographs were exposed for 12–18 h. The arrows denote the sizes of the hybridizing fragments, as calculated using *Hin*dIII-digested lambda-phage DNA as standards.

sequence (Rottmann et al. 1987), hybridized to the 20-kb *Eco*RI restriction fragment (fig. 3, lane 3). The aldolase-A and -pseudogene probes were from the relatively unconserved 3′ ends of the genes, both starting at a *Pst*I site at amino acid 306 (fig. 1). The relatively divergent *Pst*I-*Dde*I restriction fragment in the aldolase-A cDNA 3′ region hybridized primarily to the 13.3-kb *Eco*RI fragment (fig. 3, lane 4). The pseudogene probe, a *Pst*I-*Xba*I fragment from pEE416, hybridized primarily to the 8.1-kb fragment (fig. 3, lane 5). Even at relatively high stringency, these probes cross-hybridized slightly both with each other and with a 30-kb fragment. Since the restriction map of the human aldolase-A gene is not known, this large fragment may reflect two *Eco*RI fragments hybridizing to the aldolase-A gene probe. Alternatively, we cannot exclude the possibility that this fragment may correspond to another aldolase A–related pseudogene, since it hybridized to both the aldolase-A cDNA and pseudogene probes but not to the aldolase-B or -C probes. All of the *Eco*RI genomic fragments can be accounted for using the four specific gene probes. This demonstrated that only one copy of each of the aldolase genes—A, B, C, and one (or perhaps two) aldolase A–related pseudogene(s)—account for all aldolase sequences in the haploid human genome.

Each aldolase probe was hybridized to chromosome-specific DNA on spot-blot filter panels according to the hybridization conditions specific for each gene. In the first experiment the aldolase-A gene probe hybridized only to the chromosome-16 spot on two entire spot-blot filter panels. This result was repeated with another complete panel and a single filter containing just chromosomes 16 and 17 (fig. 4*A*). In the second experiment the aldolase-C gene probe revealed a positive chromosome-17 signal on two complete filter panels (fig. 4*B*). In the third experiment, the aldolase-pseudogene probe hybridized to the chromosome 10/11 spot in an entire filter panel with DIPI-chromomycin stained chromosomes that separated 10/11 from 9/12 (fig. 4*C*) (Lebo and Bruce 1987). This result was repeated twice on individual filters containing these chromosomal DNAs. Subsequently, the pseudogene was mapped to the long arm of chromosome 10 by testing derivative chromosomes (fig. 5). Different-size chromosomes derived from a reciprocal translocation between the normal chromosomes 10 and 7 were sorted from their normal homologues. The probe hybridized to a spot containing derivative chromosome-7 (der7) and chromosome-1 DNA and did not hybridize to chromosome 1, chromosome 7, the der10, or a chromosome 11 with an interstitial short-arm deletion. Thus, specific hybridization was to the part of the der7 derived from the long arm of chromosome 10. The pseudogene was further excluded from chromosome 11 by means of restriction-enzyme analysis of somatic-cell hybrids carrying all of chromosome 11 and a multiple set of chromosome-11 deletions (not shown).

Preliminary somatic-cell hybrid data from two groups assigned the human aldolase-A gene to chromosomes 16 and 22, respectively (Human Gene Mapping 8 1985). This discrepancy could be explained by a chromosome-16 rearrangement that occurred in the series of somatic-cell hybrids used to assign the gene to chromosome 22. Our gene-specific probe and hybridization conditions identified a unique signal on sorted chromosome 16. The reliability of the spot-


A
+16
17
14
15
21
22
X
Y
9
10–12
7
8
9–12
14, 15
13
18
+16
17
19
20
1
2
3
4
5
6
7
8
B
1
2
3
4
5
6
7
X
8
9-12
13
14, 15
16
17*
18
Y
19
20
21
22

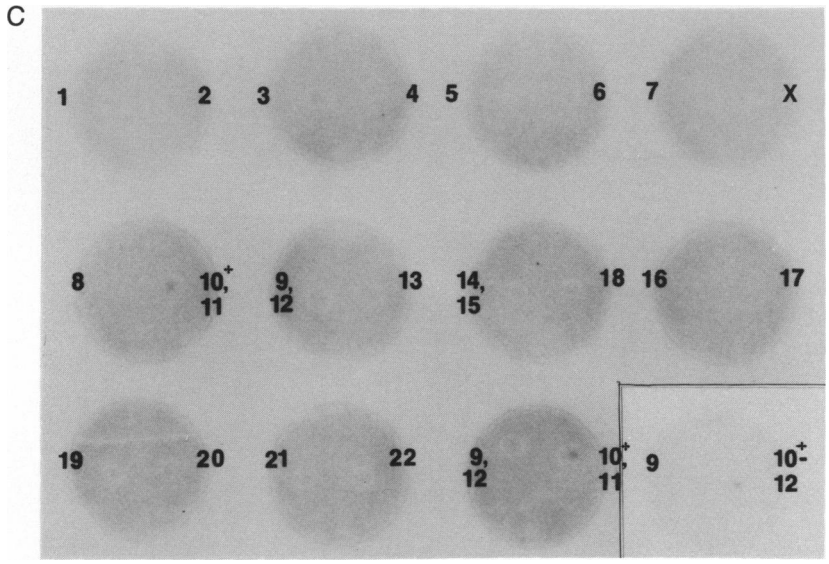


FIG. 4.—Human aldolase spot-blot analysis of whole sorted chromosomes. *A*, Radioautograph of an entire spot-blot filter panel (Lebo et al. 1985) and a partial panel from lymphocyte cell line GM130 hybridized to the human aldolase A cDNA. The positive signal corresponded to a sorted chromosome-16 DNA spot (indicated by a plus [+] sign). The nonspecific signals on the chromosome-1-and-2 filter are not located over the actual chromosome spots. *B*, Radioautograph of a similar spot-blot filter panel hybridized to the aldolase-C-gene probe reveals a positive signal on the chromosome-17 spot. *C*, The radioautograph of a spot-blot filter panel and two separate filters containing DNA from chromosomes 9–12 of the female-derived lymphoctye cell line GM131 hybridized with the aldolase-pseudogene probe. The positive signal corresponds to spots indicated by a plus (+) sign. All filters have ~30,000 chromosomes of each type in each spot. Differences in spot intensity vary inversely with the number of previous hybridizations.

blot method has been demonstrated by confirming the location of 13 previously mapped genes, including the α- and ζ-globin genes, on chromosome 16. To date, 18 new gene assignments made by spot-blot analysis have been confirmed by somatic-cell or in situ hybridization, including our assignment of two homeo-box genes to chromosome 17 (Joyner et al. 1985; Rabin et al. 1985). When a single discrepancy occurred with in situ hybridization, the spot-blot result was confirmed. Thus, we have confidence in the spot-blot assignment of (1) aldolase A to chromosome 16, (2) aldolase C to chromosome 17, (3) a pseudogene to chromosome 10, and (4) the aldolase-B gene to chromosome 9 (Lebo et al. 1985).

Aldolase Gene and Chromosome Evolution

These new localizations and our previous assignment of the aldolase-B gene tc chromosome 9 (Lebo et al. 1985) map the aldolase genes to the following four human chromosomes: 9, 10, 16, and 17. These four genes are found on the following two pairs of morphologically similar human chromosomes: 9 and 10, and 16 and 17 (Comings 1972). These homeologous (i.e., of similar origin [Comings 1972]) chromosome pairs may have arisen from one or two tetra-

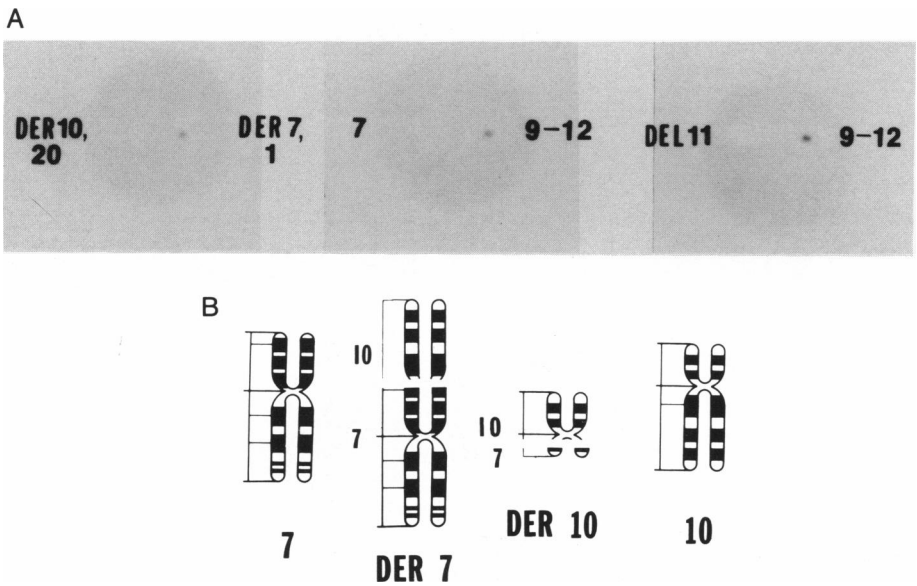


FIG. 5.—Spot-blot analysis of sorted derived chromosomes. *A*, Radioautographs from the two filters on the left are from cell line GM44, which includes a reciprocal translocation, t(7;10) (p21;q11), as illustrated by the ideograms (*c*) (National Foundation–March of Dimes 1972). The radioautograph on the right was a filter sorted from cell line ML2010, which contains a deletion in the short arm of chromosome 11. The filters were hybridized to the human aldolase-pseudogene probe (*Pst*I-*Xba*I) at 50 C and washed at 53 C. The positive signal is shown consistently over spots containing sorted DNA from the long arm of chromosome 10.

ploidization events. If tetraploidization occurred once, then a reciprocal translocation between the ancestral 16/17 chromosome and the ancestral 9/10 chromosome (fig. 6, panel *A*) would have moved the primordial aldolase gene to a second chromosome. Subsequently, a single tetraploidization would have created chromosomes 16 and 17 and chromosomes 9 and 10. Since then, these chromosomes would have diverged morphologically, along with gene structure and function. On the other hand, if tetraploidization occurred twice, the primordial gene would have been duplicated the first time when the ancestral chromosome 9/10/16/17 was duplicated to form the ancestral chromosomes 9/10 and 16/17 (fig. 6, panel *A*), with subsequent diploidization. Alternatively, this topology could have been derived from a nondisjunction leading to duplicate copies of primordial chromosomes 9/10 and 16/17 that became fixed in offspring. Thus nondisjunction would have generated a partial rather than complete genomic duplication. Either scheme predicts that the aldolase-A and -C sequences will share more homology with each other than either of them will with the aldolase-B and -pseudogene sequences.

As predicted by the chromosomal locations (fig. 6, panel *A*), the coding sequences of the expressed aldolase-A and -C genes are more homologous to each other than either of them is to the expressed aldolase-B gene (fig. 6, panel *B*) (Paolella et al. 1986; Rottmann et al. 1987). However, the coding region of

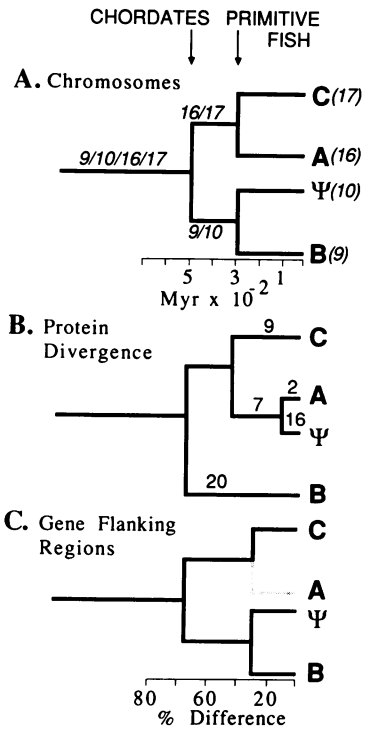


FIG. 6.—Evolutionary trees of aldolase genes. *Panel A*, Evolutionary tree illustrating proposed evolution of the chromosomes containing aldolase loci, as predicted on the basis of chromosomal locations. The arrows point to tetraploidization events as predicted by Ohno (1973). The scale is in millions of years. *Panel B*, Tree based on a difference matrix, including correction for multiple mutations at the same position (Fitch and Margoliash 1967), as derived from the protein sequences. The numbers are the percentage of relative difference along each limb. *Panel C*, Tree based upon alignments that can be made between flanking regions at the aldolase-C, -B, and -pseudogene loci on chromosomes 17, 9, and 10, respectively. The scale is in percentage difference. The grey limb for aldolase A is presumed, since no data exist.

the unexpressed pseudogene shares 80% homology with the aldolase-A gene and only 69% homology with the aldolase-B gene (fig. 6, panel *B*). We have reasoned that if the original pseudogene locus was more homologous to aldolase B, then the pseudogene on chromosome 10 may have been changed to its present condition via recombination, leaving the flanking sequence homologous to aldolase B. Moreover, an ancestral chromosome-10 sequence homologous to aldolase B would have promoted recombination at that locus. The presence of homologous aldolase-B sequences at the chromosome-10 locus would support this hypothesis.

Indeed, comparison of the pseudogene flanking sequences to those of the other aldolase genes supports this hypothesis. The flanking sequences of the pseudogene share some homology with the flanking regions of the human aldolase-B gene but not with those of the aldolase-C gene. Homology between the pseudogene and mammalian aldolase-A mRNA sequences (Tolan et al. 1984;

Joh et al. 1986) declines 10 bp upstream from the start of the protein reading frame and declines again between the pseudogene and the human aldolase-A mRNA sequence 3 bp before the polyadenylation site (see fig. 2*B*). Computer-generated alignment of these gene regions (see Material and Methods) showed significant homology to similar regions upstream and downstream from the aldolase-B gene (fig. 6, panel *C*). Moreover, no such homology was detected between comparable regions of the aldolase-C gene, between aldolase B and aldolase C, or between opposite (e.g., 3′ compared with 5′) aldolase-B-gene flanking regions and the pseudogene. The homologies between the pseudogene and aldolase-B-gene flanking regions were 65% in the 5′-region and 50% in the 3′ region, with a statistical significance of being found by chance of one in 1,000 (0.001) and one in 300 (0.003), respectively. The regions involved in the homology are underlined in fig. 2*B*, with triangles denoting loops or deletions necessary to optimize the alignment. The human aldolase-A-gene flanking sequences are unavailable for comparison. The described homologies are consistent with a recombination event at the pseudogene locus with related aldolase-A sequences. The conversion of this chromosome-10 locus may have been derived via a recombination between a relatively recent processed pseudogene, evidenced by its high degree of homology to the aldolase-A cDNA and inserted elsewhere in the genome via retrotransposition, and the nonfunctional homologous aldolase locus on chromosome 10. Recurrent recombination between these loci on chromosomes 9 and 10 could have preserved and enhanced the similarity at these loci, especially in the noncoding region. The question remains as to what happened to the fourth aldolase in early vertebrates. Clearly, a fourth aldolase isozyme is not expressed in higher vertebrates, although there is electrophoretic evidence for a fourth aldolase in advanced teleost fish, e.g., trout and salmon (Lebherz and Rutter 1969). Taken together, all the data support the evolution of the four homologous aldolase loci via one or two DNA-tetraploidization events that generated four aldolase loci on two pairs of morphologically similar chromosomes (fig. 6, panel *A*).

DISCUSSION

The analysis of the clones for both human aldolase-A cDNA and a pseudogene have completed the description of the relationship and genomic organization of this set of isozymes. The single-copy genes A, B, C, and a pseudogene map to chromosomes 16, 9, 17, and 10, respectively. These loci reflect the evolution of this ancient enzyme and perhaps the process by which the vertebrate chromosomes carrying them have evolved.

This set of genes for the functionally divergent aldolase isozymes maps to two pairs of morphologically similar, homeologous chromosomes. This relationship between homologous genes adds direct functional evidence for chromosomal evolution via tetraploidization, which has been proposed previously on the basis of chromosome morphology and total cellular DNA content. In fact, additional evidence exists for a homeologous relationship between chromosomes 9, 10, 16, and 17. The mitochondrial and cytosolic forms of thymidine kinase map to chromosomes 16 and 17, respectively (Boone et al.

1977; Willecke et al. 1977). Moreover, lysosomal acid lipase B and glutamate oxaloacetate transaminase-2 map to chromosome 16, and the homologous lysosomal acid lipase A and glutamate oxaloacetate transaminase-1 are on chromosome 10 (Human Gene Mapping 8 1985). The precise nature of the homology of these genes is not as well defined as that of the aldolase isozymes, but clearly these data (plus those for aldolase) strongly indicate a common origin for these chromosomes. Homeologous chromosomes have been found to carry other genes with isoforms. For instance, chromosomes 11 and 12 carry the lactate dehydrogenase A and B genes (McKusick 1985) and the H-ras and K-ras genes, respectively (Huerre et al. 1983; Popescu et al. 1985). In addition, the homologous UDP-glucose pyrophosphorylase-1 and -2 genes have been mapped to homeologous chromosomes 1 and 2 (McKusick 1985). These additional locations support the occurrence of at least a single tetraploidization in chromosome evolution. The four aldolase loci require us to consider the occurrence of two tetraploidization events consistent with total DNA content.

Persistence of ancient chromosomal locations presupposes that chromosomes or parts thereof have been frozen throughout evolution. The plausibility of this hypothesis could have been obscured by the evidence that chromosomes have diverged considerably since vertebrates arose (Searle 1976). During this divergence numerous linkage groups on pairs of morphologically similar chromosomes would have been separated via translocations, deletions, and insertions. However, numerous subchromosomal groups of genes have been conserved between primate (man) and rodent (mouse) chromosomes and even in such distantly related genera as fish (Seuanez 1979; Morizot 1983). Moreover, the conservation of five linked genes on both human chromosome 11 and the feline chromosome D1 indicates that this chromosomal linkage group has been conserved in two additional orders of mammals, Primates and Felidae (O'Brien et al. 1986). The same group of genes has been dispersed to three different mouse chromosomes (O'Brien et al. 1986). If the same trend continues between human and cat chromosomes, the morphology of both human and cat chromosomes will have changed less than that of mouse chromosomes, consistent with the molecular clock running faster in rodents than in primates (Li and Tanimura 1987).

Evolution of vertebrates that involves tetraploidization suggests this mechanism as the origin of all enzyme isozymes. The divergence and maintenance of tissue-specific isozyme genes reflects the usefulness of isozymes for organismal adaptation and development. The mechanism of gene divergence of duplicate aldolase genes to tissue-specific aldolase isozyme loci remains to be established. Most adaptive evolution at the organismal level has been proposed as resulting from mutations affecting gene expression (Wilson et al. 1977) rather than from mutations affecting protein structure. For example, tissue-specific activation of a gene could result from insertion of previously existing regulatory-DNA sequences. This would partially explain the increased fitness of polyploid organisms. Tetraploidization followed by different regulatory-sequence insertion into multiple duplicate-gene sites would have generated more adaptive offspring more rapidly. The highly homologous nature of the

aldolase isozymes, in contrast to their drastically different tissue-specific expression, supports this evolutionary scheme.

ACKNOWLEDGMENTS

We appreciate the support of Dr. Edward E. Penhoet, in whose laboratory at the University of California, Berkeley, the aldolase probes were isolated and characterized. We are grateful to Rob Hallewell for his gifts of the cDNA libraries. We also acknowledge William Rottmann for his early contributions to this work. We thank Alan C. Wilson and Barbara Chapman for helpful discussions and Tom Camerato, Kelly DeSelms, and Lee Anderson for technical assistance. R.V.L. is a consultant to the Howard Hughes Medical Institute. This work was supported in part by grant GM32344 from the National Institutes of Health.

REFERENCES

Amsden, A. B. 1985. FDP aldolase genes of the rabbit. Ph.D. thesis, University of California, Berkeley.

Barr, P. J., R. M. Thayer, P. Laybourn, R. C. Najarian, F. Seela, and D. R. Tolan. 1986. 7-Deaza-2′-deoxyguanosine-5′-triphosphate: enhanced resolution in M13 dideoxy sequencing. BioTechniques **4**:428–432.

Beutler, E., S. Scott, A. Bishop, W. Margolis, F. Matsumoto, and W. Kuhl. 1973. Red cell aldolase deficiency and hemolytic anemia: a new syndrome. Trans. Assoc. Am. Physicians **86**:154–166.

Blin, N., and D. W. Stafford. 1976. A general method for isolation of high-molecular-weight DNA from eukaryotes. Nucleic Acids Res. **3**:2303–2308.

Boone, C., T.-R. Chen, and F. H. Ruddle. 1977. Assignment of three human genes to chromosomes (LDH-A to 11, TK to 17, and IDH to 20) and evidence for translocation between human and mouse chromosomes in somatic cell hybrids. Proc. Natl. Acad. Sci. USA **69**:510–514.

Burgess, D. G., and E. E. Penhoet. 1985. Characterization of the chicken aldolase B gene. J. Biol. Chem. **260**:4604–4614.

Chambers, R. A., and R. T. C. Pratt. 1956. Idiosyncrasy to fructose. Lancet **2**:340.

Coleclough, C., and F. L. Erlitz. 1985. Use of primer-restriction-end adapters in a novel cDNA cloning strategy. Gene **34**:305–314.

Comings, D. E. 1972. Evidence of ancient tetraploidy and conservation of linkage groups in mammalian chromosomes. Nature **238**:455–457.

Davis, R. W., D. Botstein, and J. R. Roth. 1980. Isolation of bacterial and plasmid DNA. Pp. 124–125 *in* Advanced bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Fitch, W. M., and E. Margoliash. 1967. Construction of phylogenetic trees. Science **155**:279–284.

Freemont, P. S., B. Dunbar, and L. A. Fothergill. 1984. Human skeletal muscle aldolase: N-terminal sequence analysis of CNBr- and o-iodosobenzoic acid cleavage fragments. Arch. Biochem. Biophys. **228**:342–352.

Grunstein, M., and D. S. Hogness. 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. USA **72**:3961–3965.

Huerre, C., S. Despoisse, S. Gilgenkrantz, G. M. Lenoir, and C. Junien. 1983. c-Ha-ras-1 is not deleted in aniridia-Wilms' tumour associations. Nature **305**:638–641.

Human Gene Mapping 8: Eighth International Workshop on Human Gene Mapping. 1985. Cytogenet. Cell Genet. **40**:1–832.

Joh, K., Y. Arai, T. Mukai, and K. Hori. 1986. Expression of three mRNA species from a single rat aldolase A gene, differing in their 5′ non-coding regions. J. Mol. Biol. **190**:401–410.

Joyner, A. L., R. V. Lebo, Y. W. Kan, R. Tjian, D. R. Cox, and G. R. Martin. 1985. Comparative chromosomal mapping of a conserved homeo box region in mouse and human. Nature **314**:173–175.

Lawn, R. M., E. F. Fritsch, R. C. Parker, G. Blake, and T. Maniatis. 1978. The isolation and characterization of linked delta- and beta-globin genes from a cloned library of human DNA. Cell **15**:1157–1174.

Lebherz, H. G., and W. J. Rutter. 1969. Distribution of fructose diphosphate aldolase variants in biological systems. Biochemistry **8**:109–121.

Lebo, R. V., and B. D. Bruce. 1987. Gene mapping with sorted chromosomes. Methods Enzymol. **151**:292–313.

Lebo, R. V., D. R. Tolan, B. D. Bruce, M.-C. Cheung, and Y. W. Kan. 1985. Spot-blot analysis of sorted chromosomes assigns a fructose intolerance disease to chromosome 9. Cytometry **6**:478–483.

Li, W-H., and M. Tanimura. 1987. The molecular clock runs more slowly in man than in apes and monkeys. Nature **326**:93–96.

McKusick, V. A. 1985. The human gene map 1 December 1984. Clin. Genet. **27**:207–239.

Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Agarose gel electrophoresis. Pp. 150–164 *in* Molecular cloning. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Morizot, D. C. 1983. Tracing linkage groups from fish to mammals. J. Hered. **74**:413–416.

National Foundation–March of Dimes, Paris conference (1971). 1972. Standardization in human cytogenetics. Birth Defects **8**:18–19.

Norrander, J., T. Kempe, and J. Messing. 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene **26**:101–106.

O'Brien, S. J., M. E. Haskins, C. A. Winkler, W. G. Nash, and D. F. Patterson. 1986. Chromosome mapping of beta-globin and albino loci in the domestic cat. J. Hered. **77**:374–378.

Ohno, S. 1970. Evolution by gene duplication. Springer, Heidelberg.

———. 1973. Ancient linkage groups and frozen accidents. Nature **244**:259–262.

Ohno, S., U. Wolf, and N. B. Atkin. 1986. Evolution from fish to mammals by gene duplication. Hereditas **59**:169–187.

Paolella, G., P. Buono, F. P. Mancini, P. Izzo, and F. Salvatore. 1986. Structure and expression of mouse aldolase genes: brain-specific aldolase amino acid sequence is closely related to aldolase A. Eur. J. Biochem. **156**:229–235.

Penhoet, E. E., M. Kochman, R. Valentine, and W. J. Rutter. 1967. The subunit structure of mammalian fructose diphosphate aldolase. Biochemistry **6**:2940–2949.

Penhoet, E. E., T. Rajkumar, and W. J. Rutter. 1966. Multiple forms of fructose diphosphate aldolase in mammalian tissues. Proc. Natl. Acad. Sci. USA **56**:1275–1282.

Popescu, N. C., S. C. Amsbaugh, J. A. DiPaolo, S. R. Tronick, S. W. Aaronson, and D. C. Swan. 1985. Chromosomal location of three human *ras* genes in in situ molecular hybridization. Somatic Cell Genet. **11**:149–155.

Rabin, M., C. P. Hart, A. Ferguson-Smith, W. McGinnis, M. Levine, and F. H. Ruddle. 1985. Two homeo box loci mapped in evolutionarily related mouse and human chromosomes. Nature **314**:175–177.

Rottmann, W. H., K. R. DeSelms, J. Niclas, T. Camerato, P. S. Holman, C. J. Green, and D. R. Tolan. 1987. The complete amino acid sequence of the human aldolase C isozyme derived from genomic clones. Biochimie **69**:137–145.

Rottmann, W. H., D. R. Tolan, and E. E. Penhoet. 1984. Complete amino acid sequence for human aldolase B derived from cDNA and genomic clones. Proc. Natl. Acad. Sci. USA **81**:2738–2742.

Rutter, W. J., R. F. Blostein, B. M. Woodfin, and C. S. Weber. 1963. Enzyme variants and metabolic diversification. Adv. Enzyme Regul. **1**:39–56.

Sanger, F., S. Nicklen, and A. R. Coulsen. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5487.

Sawyer, J. R., and J. C. Hozier. 1986. High resolution of mouse chromosomes: banding conservation between man and mouse. Science **232**:1632–1635.

Searle, A. G. 1976. Baltimore conference: Third International Workshop on Human Gene Mapping Birth Defects. Cytogenet. Cell Genet. **16**:430–435.

Seuanez, H. N. 1979. The phylogeny of human chromosomes. Springer, Berlin.

Southern, E. 1975. Detection of specific sequences among fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.

Stebbins, G. L. 1966. Chromosomal variation and evolution. Science **152**:1463–1469.

Tolan, D. R., A. B. Amsden, S. D. Putney, M. S. Urdea, and E. E. Penhoet. 1984. The complete nucleotide sequence for the rabbit muscle aldolase A messenger RNA. J. Biol. Chem. **259**:1127–1131.

Tolan, D. R., and E. E. Penhoet. 1986. Characterization of the human aldolase B gene. Mol. Biol. Med. **3**:245–264.

Tsutsumi, K., T. Mutai, R. Tsutsumi, S. Hidaka, Y. Arai, K. Hori, and K. Ishikawa. 1985. Structure and genomic organization of the rat aldolase B gene. J. Mol. Biol. **181**:153–160.

Vanin, E. F. 1985. Processed pseudogenes: characterization and evolution. Annu. Rev. Genet. **19**:253–257.

Vieira, J., and J. Messing. 1982. The pUC plasmids, and M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259–268.

Willecke, K., T. Reuber, R. S. Kucherlapati, and F. H. Ruddle. 1977. Human mitochondrial thymidine kinase is coded for by a gene on chromosome 16 of the nucleus. Somatic Cell Genet. **3**:237–245.

Wilson, A. C., S. S. Carlson, and T. J. White. 1977. Biochemical evolution. Annu. Rev. Biochem. **46**:573–639.