Volume 15 Number 13 1987 Nucleic Acids Research

Human cardiac myosin heavy chain genes and their linkage in the genome

Lino J.Saez, Katherine M.Gianola, Elizabeth M.McNally, Rebecca Feghali, Roger Eddy[1], Thomas B.Shows[1] and Leslie A.Leinwand

Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, NY 10461 and [1]Department of Human Genetics, Roswell Park Memorial Institute, Buffalo, NY 14263, USA

Received September 30, 1986; Revised May 12, 1987; Accepted May 29, 1987 Accession no. Y00362

ABSTRACT

Human myosin heavy chains are encoded by a multigene family consisting of at least 10 members. A gene-specific oligonucleotide has been used to isolate the human β myosin heavy chain gene from a group of twelve nonoverlapping genomic clones. We have shown that this gene (which is expressed in both cardiac and skeletal muscle) is located 3.6kb upstream of the α cardiac myosin gene. We find that DNA sequences located upstream of rat and human α cardiac myosin heavy chain genes are very homologous over a 300bp region. Analogous regions of two other myosin genes expressed in different muscles (cardiac and skeletal) show no such homology to each other. While a human skeletal muscle myosin heavy chain gene cluster is located on chromosome 17, we show that the β and α human cardiac myosin heavy chain genes are located on chromosome 14.

INTRODUCTION

Myosin heavy chains (MHC) constitute a family of proteins of 200,000 dalton molecular weight that are widely distributed in nature. Through their interaction with actin, myosins convert chemical energy into mechanical force through the hydrolysis of ATP. This force manifests itself in many processes ranging from muscle contraction to cell division in nonmuscle cells. As such, myosins are present in at least one form in every eukaryotic cell examined (1). Myosin, like most contractile proteins, exists in a number of polymorphic forms in different types of muscle and developmental stages (2,3). The MHC isoenzymes expressed in different muscle fibers differ substantially in sequence, one result of which is quantitative differences in hydrolysis of ATP. This is correlated with the contractile velocity of a particular muscle type (4). These different MHC isoenzymes are encoded by multigene families in vertebrates. Although the exact number of MHC genes has not yet been determined, the estimates vary from 7 to 22 closely related members (5,6). cDNA clone analysis has described numerous forms of MHC expressed in skeletal and cardiac muscle (7,8). For example, adult human

skeletal muscle expresses at least 5 distinct forms (8). One form of MHC that is expressed in skeletal muscle is also expressed in cardiac muscle and appears to be developmentally and hormonally regulated (8,9).

Three MHC isoenzymes can be separated electrophoretically in cardiac muscle (10). These are products of two distinct genes referred to as α and β (11,12). Expression of α- and β-MHC is not restricted to cardiac ventricles, since α-MHC is also expressed predominantly in atrium (13) and β-MHC is expressed in skeletal muscles (9,13). The genes encloding α- and β- cardiac MHC have been well characterized in both rat (7,12) and rabbit (11,14). Separate genes exist for the two MHCs. In addition, these two genes in rat are organized in a tandem arrangement in the genome, separated by 4 kb (12). The myosin V_1, with high ATPase actvity, and myosin V_3, with low ATPase activity are composed of αα and ββ homodimers, respectively. Myosin V_2 is a heterodimer consisting of one α- and one β-MHC (10). The expression of α- and β-MHC follows a defined developmental pattern that varies in different species. Both α and β MHC are usually expressed in normal animals although their proportions differ (9). In all species studied, myosin V_3 is the most abundant myosin in late fetal life (9). In rat and mouse, myosin V_1 predominates (70% of total myosin) in perinatal and adult life (15). In large mammals, myosin V_1 is the most abundant form in perinatal stage and myosin V_3 becomes the most abundant isoenzyme in the adult animal (15).

The physiological status and the hormonal environment of the muscle are known to affect the expression of different MHC isoforms (9,10). Both cardiac and skeletal muscle MHC expression are controlled by thyroid hormone (13). Using gene-specific sequences, it has been demonstrated that thyroid hormone can regulate MHC gene expression depending on the tissue in which it is expressed (13). In rat (9) and rabbits (16) α- and β-MHC genes are regulated in an antithetic fashion by thyroid hormone. That is, hyperthyroidism stimulates the expression of α-MHC, while induction of a hypothyroid state results in a progressive disappearance of α-MHC and the induction of β-MHC (15). Furthermore, the α-β transition of MHC produced by thyroid hormone levels appears to be transcriptionally regulated in rat (9) and rabbits (16).

We previously described a human cDNA clone encoding an MHC mRNA that appears to be coexpressed in skeletal and cardiac muscle (8). By homology of the protein coding as well as untranslated regions of the human clone to a previously described rat β cardiac MHC cDNA, we suggested that this clone is a human β MHC form. Using this as well as other MHC DNAs as probes, we have isolated 12 human genomic clones encoding MHC. An oligonucleotide

complementary to the 3' untranslated region of the human slow skeletal MHC mRNA has been used to isolate the gene encoding this specific MHC. We show that as in the rat, the human slow skeletal gene, hereafter referred to as β, is linked to the α cardiac gene, separated by 3.6 kb of DNA. In addition to the expected protein coding sequence conservation, there is extremely strong homology between the human and rat α cardiac genomic sequences for a 300 bp region upstream of the α MHC promoter. These sequences are good candidates for regions that regulate the α MHC gene. Somatic cell hybrids were used to map the chromosomal location of these two MHC genes. These genes are not linked to the skeletal MHC genes which have been assigned to human chromosome 17 (17). These human MHC genes are located on chromosome 14.

MATERIALS AND METHODS

Screening of Human Genomic Libraries

Two human genomic DNA libraries were screened, each containing human DNA which was partially cleaved with EcoRI and cloned into bacteriophage λ Charon 4A. The libraries were generously provided by Dr. A. Banks (Columbia University, College of Physicians & Surgeons) and Dr. T. Maniatis (Harvard University).

Bacteriophage were grown in five 20 cm^2 plates, each containing approximately 10^5 plaque-forming units of recombinant bacteriophage. Selection of positive clones was performed according to Benton and Davis (18) with the following modifications: phage were allowed to absorb to nitrocellulose filters, denatured in 0.5 N NaOH/ 1M NaCl and neutralized in 1.5 M NaCl/Tris-HCl, pH 7.4 for 4 minutes at each step. Filters were then baked for 2 hours at 80^o under vacuum. Hybridizations were carried out in 5X SSC (1X = 150 mM sodium chloride, 15 mM sodium citrate), 1X Denhardt's, 10 mM $NaPO_4$ and 50 λg/ml of denatured salmon sperm DNA at 65^oC. Filters were always pre-hybridized for at least 3 hours, hybridized for 18 hours and washed in 2X SSC, 0.2% SDS at 65^oC for at least 2 hours. Following plaque purification bacteriophage DNA was isolated according to Maniatis et al. (19).

DNA and RNA Blot Analysis

Reaction conditions for restriction endonuclease cleavage of plasmid and phage DNA were essentially as suggested by Maniatis et al., (19). Digested DNA samples were electrophoresed on either 0.7% or 1.0% agarose gels in TBE (50 mM Tris-Borate pH 8.0, 1 mM EDTA).

The transfer of DNA from an agarose gel to a filter was performed according to Southern (20), except that Gene Screen was used in place of

nitrocellulose for binding of cloned DNA. Hybridization of filters was carried out as described above. For re-use of filters with cloned DNA samples, signal was removed by incubating the filters twice in 0.5N KOH for 15 min at 42 ^{o}C.

Radioactive plasmid DNA probes were used at a final concentration of 2 x 10^5 cpm/ml. The specific activity of the probes was between 5 x 10^7 and 2 x 10^8 cpm/µg. Oligonucleotide probes were radioactively labelled with γ-^{23}PATP.

RNA was analyzed by the methods described (19) using 0.7% MOPS-formaldehyde agarose gel. Tissues for RNA analysis were a gift from Dr. Frank Veith, Montefiore Hospital.

Oligonucleotide Synthesis and Hybridization

Oligonucleotides were synthesized by the phosphotriester solid-phase method in the Albert Einstein oligonucleotide synthesizing facility. A 0.1 µg sample of the oligomer was radiolabelled in presence of 50 mM Tris-HCl pH 7.6, 10mM $MgCl_2$, 5mM dithiotreitol, 0.1mM spermidine, 350 µCi (3000 Ci/mmole) μ^{32}P-ATP (Amersham Corp.) and 7 units of T4 polynucleotide kinase (PL Biochemicals) for 30 minutes at 37^{o} C.

Oligonucleotide hybridization was carried out according to Geliebter et al., (5), with the following modification. For plasmid DNA, only Gene Screen was utilized. The temperature of hybridization was 3^{0}C below the T_m of the oligonucleotide utilized. The T_m was calculated by the formula Tm = 2n(A+T) + 4n(G+C) (21). The blots were prehybridized for 4 hours and hybridized for 3 hours in 5X SSPE (0.9 M NaCl, 40 mM NaOH, 50mM Na_2HPO4, 5 mM EDTA),2x Denhardt's 0.1% SDS and 50 ug/ml of denatured salmon sperm DNA. After 3 hours of hybridization, the filter was washed in 5X SSPE at room temperature for one hour, for 2 minutes at 50^{0}C and air dried. Autoradiography was performed at -80 ^{o}C with a Dupont Cronex Lightning Plus intensifying screen for times indicated in the Figure Legends.

DNA Sequencing

DNA sequencing was carried out by the dideoxynucleotide method of Sanger et al.,(22). Restriction endonuclease fragments to be sequenced were purified by gel electrophoresis and cloned into M13 mp18 and mp19 (24). For M13 clones containing large inserts (larger than 1.5Kb) a rapid deletion system was used (23). This generated a series of clones containing overlapping inserts by exonuclease digestion of single stranded DNA cloned into M13 vectors. The single strand DNA isolation and sequencing reactions were carried out as described by Messing et al., (24). DNA sequencing was also performed by Maxam and Gilbert method as previously described (25).

Chromosome Mapping

Human-mouse somatic cell hybrids were constructed by fusing four different mouse cell lines to ten different human fibroblast lines and selecting for hybrids in HAT medium. Hybrid cells were analyzed karyotypically and banded by Giemsa-Trypsin staining (26,27). Enzyme markers assigned to each chromosome except the Y were tested on each hybrid, confirming the chromosome analysis. Genomic DNA was prepared from these hybrids, digested with restriction endonucleases, subjected to electrophoresis on a agarose gels, transferred to nitrocellulose and hybridized with human MHC DNA fragments as probes as described previously (17).

Computer Analysis of DNA Sequences

Graphic matrix analysis was carried out with the MBSP dot matrix program written at Albert Einstein College of Medicine. Percent homologies were calculated according to an alignment program called NUCALN described by Wilbur and Lippman (28).

RESULTS AND DISCUSSION

Human MHC Gene Family

It is estimated that there are 7-22 MHC genes in vertebrates (29,30,31). We previously described the isolation and characterization of four nonoverlapping human MHC genomic clones (30). In order to isolate the full complement of human sarcomeric MHC genes, two human genomic libraries were screened with a mixture of three previously described human MHC probes: pSMHCA, pSMHCZ (8) and p10-3 (30). pSMHCA and Z are cDNA clones isolated from an adult human skeletal muscle cDNA library. Their inserts are 2.7Kb and 2.0Kb and they encode the entire light meromyosin regions of fast and slow skeletal fiber myosins, respectively. p10-3 is a 2.0Kb genomic fragment from an adult human skeletal muscle genomic clone (30). Each of these probes hybridizes to multiple fragments in the human genome (data not shown). The entire DNA sequences of the two cDNA clones have been determined and are 73% conserved at the amino acid level (8). Fifteen genomic clones containing MHC gene segments were isolated from this screening. Restriction endonuclease analysis shows that a total of 12 clones are unique and non-overlapping (data not shown). These include the four human genomic clones previously described (30). The inserts in the genomic clones range in size from 12 to 18.0 kb.

In order to identify coding regions in the genomic clones, EcoRI digests of the twelve clones were electrophoresed on agarose gels (Figure 1A) and

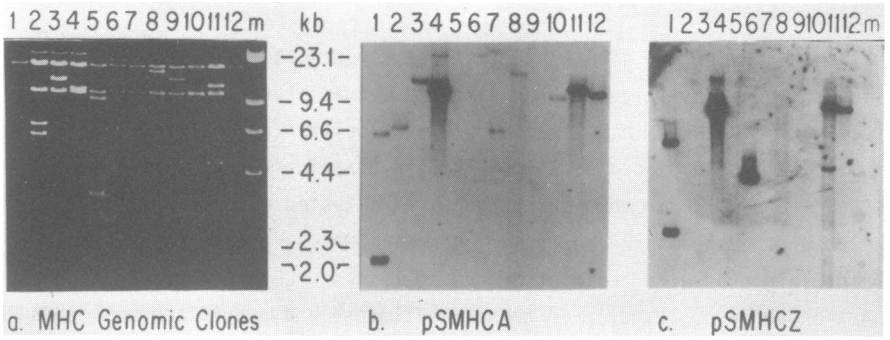

Figure 1. Hybridization analysis of MHC genomic clones with human MHC cDNAs. Twelve human genomic clones were digested with EcoRI, size fractionated by agarose gel electrophoresis (Panel A), transferred to Gene Screen and hybridized to human MHC cDNA pSMHCA (Panel B) and pSMHCZ (Panel C). Numbers above the gel indicate the different genomic clones. Lane M, contains Hind III digested λ DNA molecular markers.

transferred to nitrocellulose. The blot was hybridized to the two human MHC cDNA probes (pSMHCA and pSMHCZ) used in the screening (Figure 1B and C, respectively). Both cDNA clones hybridize to multiple genomic clones because of the conservation of sequences in the MHC gene family. The human fast fiber MHC cDNA clone (pSMHCA) hybridizes to nine of the twelve genomic clones (Figure 1B). Similarly, the human β MHC cDNA clone (pSMHCZ) hybridizes to four of the clones which are recognized by pSMHCA (compare lanes 1, 4, 11, and 12 in Figure 1) and to one genomic clone (lane 6 in Figure 1C) that is not recognized by pSMHCA. Two of the genomic clones (lanes 5 and 9) are not recognized by the cDNA probes but are recognized by the genomic segment used in the screening (data not shown).

Since the sizes of vertebrate MHC genes range from 20-30 kb (29,31) and the total insert size of any one of the clones does not exceed 18 kb, none of the clones could contain an intact MHC gene. The genomic clones were isolated with DNA probes encoding the LMM of MHC and therefore contain sequences coding for the rod portion of the MHC molecule. These results suggest that the human MHC gene family is composed of a minimum of 12 closely related genes. Whether all hybridizing genomic fragments are part of functional genes or pseudogenes is not yet known (see below).

Identification of the Human Slow Skeletal MHC Gene

Given the highly conserved nature of the sarcomeric MHC gene family it is difficult to identify the genes encoding specific MHC mRNAs and to distinguish

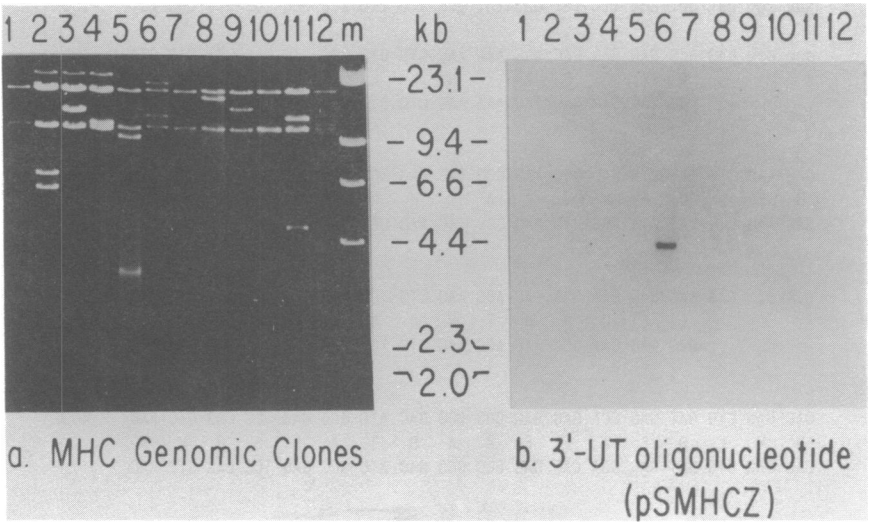

Figure 2. Hybridization analysis of human genomic MHC clones with an MHC gene-specific oligonucleotide.
Panel A. Ethidium bromide stained gel of twelve human genomic clones digested with EcoRI (described in Figure 1) (lanes 1-12).
Panel B. Hybridization of filter corresponding to gel in panel A with an oligonucleotide from the 3' untranslated region of pSMHCZ.

expressed genes and pseudogenes. Although the coding regions of MHC genes are highly homologous, their 3' untranslated regions show no significant homology within an organism (8). Therefore, in order to identify the MHC genes from the 12 genomic clones that encode the specific MHC mRNAs represented by the two cDNA clones, we developed gene-specific oligonucleotide probes. These probes derive from the 3' untranslated region of the two human MHC cDNA clones which were used as probes in the genomic screening (8). In order to determine whether these two oligonucleotides are in fact gene-specific probes, each one was hybridized to a filter containing nine independently isolated MHC cDNA clones. Each oligonucleotide recognized the clone of its origin or cDNAs of different length that encode the same mRNA as determined by restriction enzyme and DNA sequence analysis (data not shown).

To determine if any of the MHC genomic clones isolated correspond to genes encoding either of the two MHC mRNAs represented by the cDNA clones, a Southern blot of the twelve genomic clones was hybridized with the two gene-specific oligonucleotide probes. No positive hybridization was detected with the fast-fiber specific probe from pSMHCA (data not shown). However, the slow skeletal MHC oligonucleotide probe hybridizes exclusively to a 3.8 kb EcoRI

```
gga ggg gat gct acc ttc tat gac tgt gcc atc ttc acc ccc tgc cta ccc tct    105

ggc ccc CAG ACG GAG GAG GAC AGG AAA AAC CTG CTG CGG CTG CAG GCT CTG GTA    159
        ___  T   E   E   D   R   K   N   L   L   R   L   Q   A   L   V
             ACG GAG GAG GAC AGG AAA AAC CTG CTG CGG CTG CAG GCT CTG GTA


GAC AAG CTG CAG CTA AAG GTC AAG GCC TAC AAG CGC CAG GCC GAG GAG GCC GTG..  203
 D   K   L   Q   L   K   V   K   A   Y   K   R   Q   A   E   E   A  ___
GAC AAG CTG CAG CTA AAG GTC AAG GCC TAC AAG CGC CAG GCC GAG GAG GCC


0.1kb.. CAG GAG GAG CAA GCC AAC ACC AAC CTG TCC AAG TTC CGC AAG GTG CAG    351
        ___  E   E   Q   A   N   T   N   L   S   K   F   R   K   V   Q
             GAG GAG CAA GCC AAC ACC AAC CTG TCC AAG TTC CGC AAG GTG CAG


CAC GAG CTG GAT GAG GCA GAG GAG CGG GCG GAC ATC GCC GAG TCC CAG GTC AAC    405
 H   E   L   D   E   A   E   E   R   A   D   I   A   E   S   Q   V   N
CAC GAG CTG GAT GAG GCA GAG GAG CGG GCG GAC ATC GCC GAG TCC CAG GTC AAC


AAG CTG CGG GCC AAG AGC CGT GAC ATT GGC ACG AAG GTG ggt ccc...0.84 kb ..  1695
 K   L   R   A   K   S   R   D   I   G   T   K  ___
AAG CTG CGG GCC AAG AGC CGT GAC ATT GGC ACG AAG


ttt caa AAG GGC TTG AAT GAG GAG TAG CTT TGC CAC ATC TTG ATC TGC TCA GCC   1749
        ___  G   L   N   E   E   *
             GGC TTG AAT GAG GAG TAG CTT TGC CAC ATC TTG ATC TGC TCA GCC


CTG GAG GTG CCA GCA AAG CCC CAT GCT GGA GCC TGT GTA ACA GCT CCT TGG GAG   1803

CTG GAG GTG CCA GCA AAG CCC CAT GCT GGA GCC TGT GAT ACA GCT CCT TGG GAG


GAA GCA GAA TAA AGC AAT TTT CCT TGA AGC CGA gat cct gac tcc...            1848

GAA GCA GAA TAA AGC AAT TTT CCT TGA AGC CGAn
```

Figure 3. Nucleotide sequence of the 3' end of the human β cardiac MHC gene. The sequence of the last 3 exons of the β-MHC gene (top line) and the coding strand of cDNA clone pSMHCZ (bottom line) are compared. The deduced amino acid sequences within exons are shown in the single letter code under each codon. The translation termination codon is represented by an asterisk. The sequence in the box represents the poly(A) addition signal. The underlined sequences represent the putative splice junctions.

fragment of the genomic clone λHMHC8 and not to DNA from the other genomic clones (Figure 2). In order to establish unequivocally that this genomic sequence indeed encodes the mRNA represented by the cDNA clone designated pSMHCZ, the 3.8 kb EcoRI fragment was subjected to partial nucleotide sequence

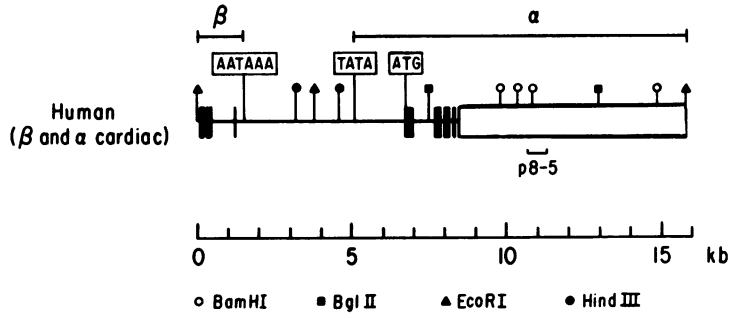

Figure 4. Structural map of human cardiac MHC genes.
The α and β genes are represented and restriction sites indicated. The initiation codon, ATG, and "TATA" boxes for the α gene are shown. Black boxes and solid lines represent exon and introns, respectively. Open boxes represent sequences not yet determined.

analysis by the method of Maxam and Gilbert (25). As shown in Figure 3, the 3' end of the cDNA sequence of pSMHCZ is represented in the cloned genomic sequences. There are 3 exons of 32, 45 and 5 codons each. The fact that there is 100% homology between the 3' untranslated region and coding sequences of pSMHCZ and the corresponding exon sequences in λHMHC8 demonstrates that this cDNA clone is transcribed from the gene represented by λHMHC8. A genomic clone isolated and characterized by another group overlapping λHMHC8 which encompasses more than half this same gene confirms this conclusion (32). These investigators also show through S1 nuclease analysis that this gene is coexpressed in human cardiac and skeletal muscle. Results presented in Figure 6 also show that this gene is expressed in both cardiac and skeletal muscle. The 3' untranslated region oligonucleotide derived from the slow skeletal cDNA clone (psMHCZ) was also used as a hybridization probe against a genomic blot to demonstrate that there is a single homologous sequence in the genome, corresponding to the cloned fragment we have obtained (data not shown).

Tandem Organization of 2 Human MHC Genes

The genes specifying α and β cardiac MHC have been well characterized in the rat and rabbit (11,12). In rat, these two genes are organized in the genome in a head to tail fashion. The α gene is located 4 kb from the 3' end of the β gene (12). To determine if the same organization is observed in human cardiac genes, a DNA fragment of a chicken fast skeletal MHC clone (AFW1) which contains sequences encoding the first three exons was used as a probe against DNA from λHMHC8. The chicken probe was kindly provided by Dr. J. Robbins and is described in (29). If the human β and α genes are organized

```
CGGTGTGAGAAGGTCCAGTCTTCCCAGCTATCTGCTCATCAGCCCTTTGAAGGGGAGGAATGTGCCCAAGGACTAAAAAAAGGCCGTGGA   90
GCCAGAGAGGCTGGGGCAGCAGACCTTTCAAGGGCAAATCAGGGGCCCTGCTGTCCTCCTGTCACCTCCAGAGCCAAAGGATCAAAGGAG  180
GAGGAGCCAGGAGGGAGAGAGGTGGGAGGGAGGGTCCTCCGGAAGGACTCCAAATTTAGACAGAGGGTGGGGGAAACGGGATATAAAGGA  270
ACTGGAGCTTTGAGGAGAGATAGAGAGACTCTGCGGCCCAGGTAAGAGGAGGTTTGGGGTGGGATGCCCTGCAGCCCGTCCACAGAGCCC  360
CCACCGTGAGGGACCTCCTTCACCAGGAGTGGGGTGCAGGTCAGTTGGAGGCCTAAGGGCTCTATTAAAACTGCCTATCTCCAGGCCCAG  450
GGAAGTTCCCCCTGACACAGGAGGTTCCACAGGAAACCCAGAAACCTCTTTTCTCCTTCTCTGACTCTCCATTTCTTTCTCTGCATCATT  540
CTGAGTCTCCTGCATGTTGTCTCCATCTTTCCATCTTCACTTCCTCCTTTGGATGGCTTCCTTCCCTTGATCCTGGCTTTTATCTTGCCT  630
CTTGGTCTTCATCGACACTTGTCACAATCATGCTTCTTTGTCTCTCTCCCTTGTCCTTCCTTCTTGGCACTTGTTCTCACCTCCCTGCCT  720
CTCTGCTTCTAACCCTGTTTCCACACCCCGTCCCACCTGGGGCTGCCTCCATCCCCGGGTGGCCTGCCTNTGGTGTTCTTCACTCTCCTC  810
ATTTGTTCTTCTCTCTGCCCGGCTCTACCTNTGGTGTTCCTTGCTCCACCCACGGTCCAGATTCTTCAGGATTCTCCGTGAAGGGATAAC  900
CAGGTGAGAACTGCCCCCATTTTCTCTGCAGAGACTGGGGCATGCTTCTCCTGGGAGCCGGATTGCTGGACCAGGGGTCTGCTGTCCCAA  990
GCACTCAGCGCCAACCCTTAGCATACTCCAGCCAATGCCACCCCAGGGAAACCCCTTACAGAGATTGTCCTTCAGCATCACCTCAGAGGG 1080
CAGGAGAAGCAGAGCCCTGAGTAGGGGAGGTGCAACAGCAGTGCCTCTCCCAGGGTGGAGGAGAGGAGCGGGGGTAGGGAGGGGGTCTGC 1170
AGAGGACAAAGCCACTCGCTGGAGCCTGGGCTCCCTCAGGAGTAACATAGCCCTCCTGTCTCTGACCCAGGGAAGCACCAAGATGACCGA 1260
                                                                                  M  T  D
TAGCCAGATGGCTGACTTTGGCACGGCCCAGTACCTCCGCAAGTCAGAGAAGGAGCGTCTAGAGGCCCAGACCCGGCCCTTTGACATTCG 1350
  S  Q  M  A  D  F  G  T  A  Q  Y  L  R  K  S  E  K  E  R  L  E  A  Q  T  R  P  F  D  I  R
CACTGAGTGCTTCGTGCCCGATGACAAGGAAGAGTTTGTCAAAGCCAAGATTTTGTCCCGGGAGGGAGGCAAGGTCATTGCTGAAACCGA 1440
  T  E  C  F  V  P  D  D  K  E  E  F  V  K  A  K  I  L  S  R  E  G  G  K  V  I  A  E  T  E
GAATGGGAAGGTGAGTAGGGCATGGCGCCGGGGCAGAAGGGAAGGAGGTCTGGGAAAGAAGATGCAGGAGGAGGTGCCACTTGCAGGGGG 1530
  N  G  K
AGCTGAGAGGGCTGGAGAAAAGCCAAGGCCAGTGGGGATGCCAGGACATGCTCCTTTGAGGAGCCCAGAATCTGATCCCTCTCAAATTAG 1620
CCTGAGCTGGTGCAACAGGTGCCACCCAGGGCCATGTTCCCCCTGCCAGAGAGGATGCTGAGGAAGAAGAACCTCAGTGTTCGCCTAAGA 1710
GGGGTCTTGTAGATAAAGAGGGCACAGACACAGCATTAAATGATGCCCCCTTCTTGCACTTGTATCCCTCCACCCTGTGCCTCAGTTCCT 1800
CCATGAGTCCACCCTCTCAAATTCCGTTCACCCAAATCAAGAGCAATTCTTAGACCCAGATGAACACAAAGATCAGAAACTTTTGAGCTG 1890
ATGCACTCTCCTTGACTGGCGACTCAGAAGCTCTGGTCCCTGGTTTGCTCACACCAGCCAATAGAATCACCCCTGGTTACCAGCTGCGGC 1980
TCAGGCTGTGTGCCTCATGAACTCGTTGACTGAATGTTACAACCCATTGAAGTGTAGAATAACAGGCCACAATCCCCTGGGGCTTTTGAC 2070
TCTGATCCCAGCTCAGCCACCCGCTAGTCACTGTGCAGGCAAATCATTTAGTCATTTAGAGCCACGGATTTCTCCACTATAAAAAACACT 2160
GGAATACCTACTGGCAGGATCTAATGACATCAGGGCATGGCAAACTGACTGCTGCCAATCAAACCACACCAACAGTGATGGATGGGGAGT 2250
GTGGAGTAGATGGGTGAACTACTTTTCCAGCAGGGGTGAAGGTTTGCCCTGAGCAACAGATACCCTAAAGGCCTGCCCGCGGGAGACAGC 2340
CTCGGGGTCAGCATAAGGTGTGCACA                                                              2366
```

Figure 5. Nucleotide sequence of the 5' end region of human α-MHC gene. The DNA and deduced amino sequence of the 5' flanking region, first exon and intervening sequence are presented. Putative "CAAT" and "TATA" boxes are underlined. The sequence corresponding to the oligonucleotide used in RNA analysis is underlined.

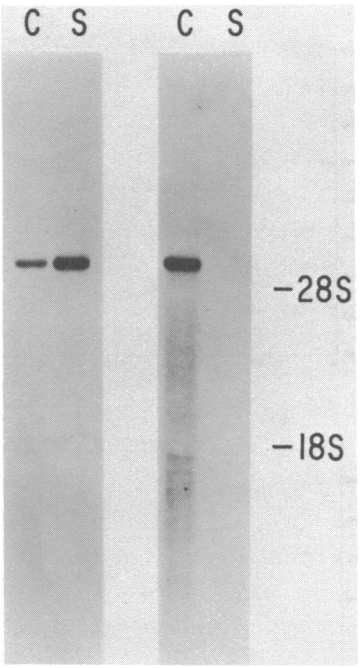

Figure 6. Oligonucleotide hybridization of the human α cardiac MHC sequences to cardiac and skeletal RNA.
10μg of poly(A)$^+$ RNA from cardiac (C) and skeletal (S) muscle were electrophoresed and hybridized to the β oligonucleotide derived from cDNA sequence and to an oligonucleotide derived from the first exon of the α cardiac genomic sequence. 28S and 18S refer to the positions of migration of ribosomal RNA.

in tandem, another MHC gene should be located downstream from the sequences encoding the 3' end of the slow skeletal MHC gene. The chicken probe does identify a sequence in the λHMHC8 genomic clone (see below). Restriction mapping and DNA sequence analysis localize a second MHC gene to a 3.5 kb HindIII/BglII restriction fragment as shown in Figure 4. Portions of this fragment of the genomic clone (coordinates 3.8-13.0 in Figure 4) were subjected to DNA sequence analysis. Putative transcriptional regulatory elements and sequences of the first protein coding exon and intervening sequence were located. The 5' flanking sequence and first coding exon and intervening sequence are presented in Figure 5. Comparison of the first exon of the human α cardiac gene to the analogous sequences for rat cardiac MHC

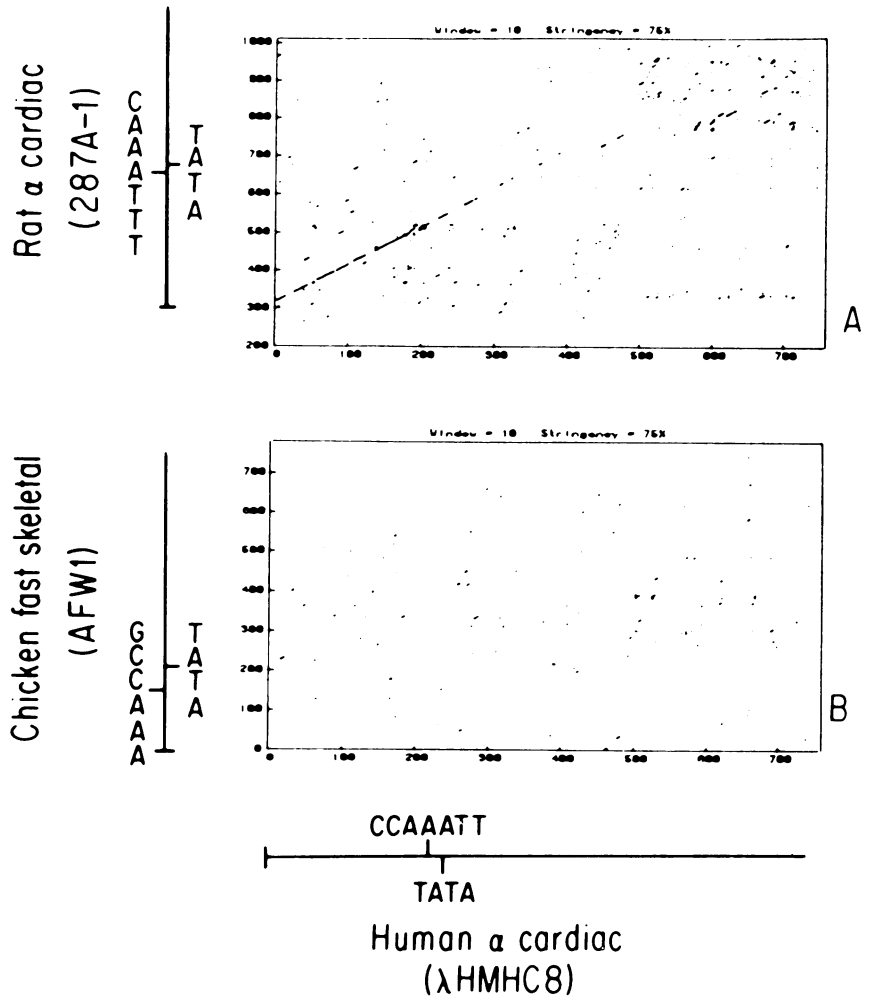

Figure 7. Graphic matrix comparison of 5' flanking regions of MHC genes. Computer assisted DNA comparisons were performed at 75% stringency level with a window of 10. Sequences are oriented 5' to 3', left to right and bottom to top. The comparisons include the 5' flanking region of human α-MHC gene (this report) versus Panel A: rat α-MHC (20) and versus chicken fast fiber skeletal MHC gene (11) (Panel B). The "CAAT" and "TATA" elements are indicated in the side of each panel. Sequences included in this comparison are residues 1-750 from Figure 5 for human and residues 200-1000 from (20) for rat α cardiac gene.

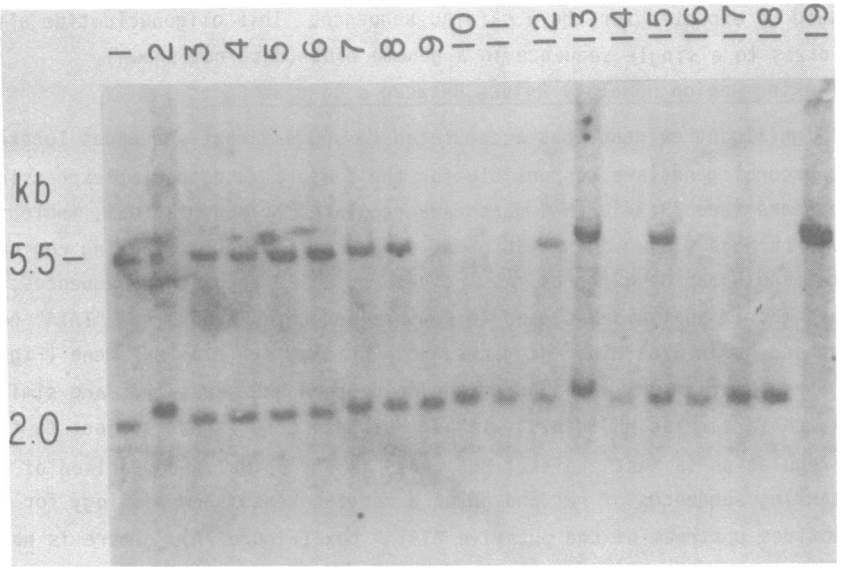

Figure 8. Hybridization of a human α cardiac genomic fragment to human x mouse hybrid cell DNAs.
Hybridization of a human α MHC subcloned DNA fragment (p8-5) to human, mouse and somatic cell hybrid DNA. Cellular DNA was digested with BglII restriction endonuclease, size fractionated on an agarose gel, transferred to nitrocellulose and hybridized to probes as described in Materials and Methods. Each lane contains 10 μg of DNA. For the karyotypic analysis of each human x mouse hybrid cell line see Table I. (lane 1), REW-5; (lane 2), XER-11; (lane 3) JSR 175; (lane 4), ATR-13; (lane 5), REX 11BSAgB; (lane 6), DUM-13; (lane 7), SIR-8; (lane 8), REW-8D; (lane 9), REX-26; (lane 13), XER-8; (lane 14), JSR-14; (lane 15), JWR-260; (lane 16), XTR-2; (lane 17), GAR-1; (lane 18) mouse and (lane 19), human.

(12) and chicken fast white MHC (29) reveals strong homology. The human and rat α cardiac genes differ by nine residues in the first coding exon which consists of 66 amino acids in human and 67 amino acids in the rat. The size of the first intron is conserved as well as some sequence homology between the rat α cardiac MHC gene and the human α cardiac sequences (data not shown; 12).

In order to prove that the gene downstream from the slow skeletal MHC gene is an α cardiac sequence, an oligonucleotide was synthesized from the first exon indicated by underlining in Figure 5 and used as a probe against cardiac and skeletal RNA. Figure 6 shows that the putative α-specific oligonucleotide hybridizes exclusively to cardiac RNA and not to skeletal RNA,

as would be expected for the α cardiac sequence. This oligonucleotide also hybridizes to a single sequence in a genome blot (data not shown).

5' Flanking Region Homology Exists Between α Cardiac Genes

Significant evidence has accumulated to suggest that sequences located 5' to structural genes are responsible for the temporal and spatial expression of these genes (see 33;34). MHC genes are regulated by many factors, including thyroid hormone (15). Conservation of sequences in the 5' flanking regions of the rat and human α-MHC genes may indicate putative regulatory sequences. Graphic matrix analysis was used to compare sequences around the "TATA" boxes of rat and human α cardiac MHC genes and a chicken skeletal MHC gene (Figure 7). These comparisons were undertaken to compare MHC genes that are similarly regulated (α cardiac MHC genes) with an MHC gene that is not subject to the same regulation (a fast skeletal MHC gene) in the animal. Comparison of the 5' flanking sequences of rat and human α cardiac genes show homology for 250 nucleotides upstream of the putative "TATA" box (Figure 7A). There is no homology for the next 700 nucleotides toward the ATG. Neither the human nor the rat cardiac flanking regions show homology with the analogous fast fiber MHC chicken sequence (Figure 7B). This homology between 5' flanking regions of the α cardiac genes suggests that sequences upstream of the "TATA" box may be important in the expression of these genes. Comparison of the 5' flanking regions of other MHC genes as well as genes that are under similar control will give more information about the importance of these regions. Ultimately, transfection experiments using the 5' flanking region of α-MHC genes will determine the biological significance of this homology. We are now in the process of sequencing the entire α/β intergenic region to establish if there are other segments that show sequence conservation.

Our results suggest that the human cardiac MHC genes have a similar genomic organization to that seen in rat and that the genes may have arisen by a duplication event. Given this information it was important to determine whether the cardiac MHC genes are chromosomally linked to the skeletal MHC genes.

Chromosomal Localization of Cardiac MHC Genes

Earlier, we showed that skeletal MHC genes are clustered on a single chromosome in mouse and human (17). A recent study has determined that mouse cardiac MHC genes are unlinked to skeletal MHC genes (6). To determine the chromosomal location of the tandem array of human MHC genes, we analyzed genomic DNA of human x mouse somatic cell hybrids. The probe used is a

Table 1

SEGREGATION OF A HUMAN DNA PROBE SPECIFIC FOR CARDIAC MYOSIN HEAVY CHAIN GENE WITH HUMAN CHROMOSOMES IN HUMAN-MOUSE CELL HYBRIDS

HYBRID	Cardiac MHC Gene	Human chromosomes 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	Translocations
ATR-13	+	+	+	+	+	+	+	+	+	-	+	-	+	+	+	+	+	+	+	+	-	-	-	t	5/X
DUA-3BSAGA	+	-	+	-	-	-	-	+	+	-	-	-	-	+	+	-	-	+	-	-	-	-	-	-	
DUA-5BSAGA	+	-	-	+	-	+	-	-	-	-	+	-	-	-	+	-	-	+	+	-	-	+	-	-	
DUM-13	+	+	+	+	-	+	+	+	-	-	+	+	+	-	+	t	+	+	+	+	+	+	+	t	X/15 15/X
ICL-15	-	-	-	-	-	-	-	-	+	-	-	-	+	-	-	-	-	+	-	-	+	+	-	-	
JSR-14	-	-	+	+	+	+	-	-	-	-	-	+	+	-	-	-	-	+	-	-	+	+	-	+	
JSR-17S	+	+	+	+	-	+	-	t	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	-	7/9
JWR-26C	+	t	+	+	+	+	+	+	-	+	+	+	+	-	+	+	+	+	+	-	+	+	-	+	1/2
NSL-L6	+	-	-	+	+	+	-	+	+	t	+	-	+	-	+	+	+	+	+	-	+	+	-	-	17/9
REW-5	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	-	+	+	+	-	+	+	+	
REW-8D	+	-	-	-	+	-	-	-	+	-	-	-	-	-	+	-	-	+	-	-	+	+	+	+	
REX-11BSAgA	+	-	-	+	-	-	-	-	-	-	+	-	-	-	+	+	-	-	+	-	-	-	-	-	
REX-11BSHF	+	-	-	+	-	-	-	-	-	-	+	-	-	-	+	-	-	-	+	-	-	-	t	t	22/X
REX-26	+	+	+	+	+	-	-	+	+	+	+	+	-	+	+	+	+	+	+	+	-	+	t	t	22/X
SIR-11	-	-	-	-	-	-	-	+	-	-	-	-	-	+	-	-	-	-	-	-	-	+	+	+	
SIR-8	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+	+	+	
TSL-2	-	-	+	t	-	+	+	-	-	+	-	+	-	-	-	-	-	t	+	-	+	+	-	+	17/3 3/17
VTL-7	+	-	-	-	-	-	-	t	-	-	+	-	-	+	-	-	+	-	+	+	+	+	+	-	7q-
VTL-8	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	-	+	-	-	+	+	+	+	-	
WIL-1	+	-	-	-	-	-	-	-	+	-	-	+	-	+	-	-	+	+	-	-	+	-	+		
WIL-13	-	-	-	+	-	+	-	-	-	-	-	-	-	-	-	-	+	+	-	-	+	+	-		
WIL-14	+	-	-	+	-	-	-	+	+	-	+	-	-	-	+	+	-	+	-	-	-	-	-	+	
WIL-15	+	-	+	+	+	-	+	+	-	-	+	+	+	+	+	+	-	+	+	-	+	+	-	+	
WIL-2	-	-	-	-	-	-	-	+	-	-	-	+	-	-	+	-	+	-	+	-	-	+	-	+	
WIL-5	-	-	-	-	+	-	-	-	+	-	+	-	-	-	-	-	-	+	+	-	-	+	-	+	
WIL-6	+	-	+	-	+	+	+	+	+	-	+	+	-	-	+	-	-	+	-	+	+	+	-	+	
WIL-7	+	-	+	+	-	+	+	-	+	-	+	+	-	+	+	-	-	+	+	-	-	+	-	+	
WIL-8X	+	-	-	+	+	+	-	+	+	-	+	+	+	-	+	-	-	+	+	+	+	+	-	+	
XER-11	-	+	-	+	+	-	+	+	+	-	+	t	+	+	-	+	+	+	+	+	+	+	+	t	11/X X/11
XER-8	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	-	+	+	+	+	+	+	+	t	11/X X/11

		Chromosome	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X
Concordant # of Hybrids	(+/+)		7	12	16	11	12	8	13	14	5	15	13	12	8	21	10	8	19	16	8	10	16	7	10
	(-/-)		8	7	5	6	6	6	7	5	9	6	8	4	5	9	6	8	1	5	8	4	0	5	3
Discordant # of Hybrids	(+/-)		13	9	5	10	9	13	6	7	15	6	8	9	13	0	10	13	2	5	13	11	5	12	6
	(-/+)		1	2	3	3	3	3	2	4	0	3	0	5	4	0	3	1	7	4	1	5	9	4	5
% Discordancy	(-/+)		48	37	28	43	40	53	29	37	52	30	28	47	57	0	45	47	31	30	47	53	47	57	46

The human specific DNA probe for cardiac MHC gene was hybridized to Souther blots containing BglII digested DNA from the human-mouse hybrids listed in the Table. The human cardiac MHC gene was determined by scoring the presence (+) or absence (-) of human bands in the hybrids on the blots. Concordant hybrids have either retained or lost the human cardiac MHC gene together with a specific chromosome or the reverse. Percent discordancy indicates the degree of discordant segregation for a marker and a chromosome. A 0% discordancy is the basis for chromosome assignment. Notes: "t" in the table indicates a translocated piece of a chromosome is present in the hybrid, but no intact chromosome (See under Translocations).

subclone, p8-5, which contains an intervening sequence of the human α-cardiac gene (see Figure 5, coordinates 10.5-11.3). Twenty-four hybrid cell lines were examined for the presence of the human α cardiac gene and its correlation with a specific human chromosome and chromosome-specific isoenzyme markers. The pattern of hybridization of p8-5 to parental and hybrid genomic DNA is shown in Figure 8. The mouse parental DNA has a 2.0 kb band that is

homologous to p8-5 (lanes 17 and 18). The human fragment recognized by p8-5 is 5.0 kb (lane 19). The human cardiac probe segregated without exception with human chromosome 14 and with specific enzyme markers (see Table 1). There were no examples of discordance in these analyses with the human cardiac MHC sequence and chromosome 14. For example, lane 5 in Figure 8 contains DNA from hybrid cell line REX-11BSAgB, that contains human chromosomes 3, 10, 14, 15, and 18 (see Table 1). An earlier report had suggested the presence of human MHC sequences on chromosome 7 (6). No hybridization was detected in somatic cell hybrids containing human chromosome 7, as shown in Figure 8 lane 11, indicating the absence of the α cardiac MHC gene on that chromosome.

We have shown that the human skeletal and cardiac MHC genes are localized on different chromosomes. In human, skeletal and cardiac MHC genes map to chromosomes 17 and 14 respectively. This is paralleled by the actin gene organization in which cardiac α actin gene maps to human chromosome 1 and skeletal α actin genes map to chromosome 15. As with the β MHC genes, the cardiac actin gene is co-expressed in adult skeletal muscle (35). These results suggest that chromosomal linkage of MHC genes is not required for co-expression of members of these two multigene families. Sequential expression of cardiac genes or of skeletal genes during development may require linkage, perhaps because they may be under the control of the same cis-acting regulatory mechanisms.

ACKNOWLEDGEMENTS

We would like to thank R. Schwartz for help with the restriction map. We would like to acknowledge the excellent secretarial assistance of D. Jackson and L. Vitagliano. This work was supported by grants from: AHA to L.L., NIH (GM29090 to L.L.; GM20454, HD05196 to T.S.), ACS CD62 to T.S. L.L. is an American Heart Established Investigator. E.M. is supported by M.S.T.P. Grant, 5T32 GM7288-12.

REFERENCES

1. **Harrington, W.F. and Rodgers, M.E.** (1984). Ann. Rev. Biochem. 53, 35-73.
2. **Whalen, R.G., Sell, S.M., Butler-Browne, G.S., Schwartz, K., Bouveret, D. and Pinset-Harstrom, I.** (1981). Nature 292, 805-809.
3. **Whalen, R.G., Sell, S.M., Eriksson, A. and Thornell, L.** (1982). Dev. Biol. 91, 478-484.
4. **Barany, M.** (1967). J. Gen. Physiol. 50, 197-218.
5. **Geliebter, J., Zeff, R.A., Schulze, D.H., Pease, L.R., Weiss, E.H., Hallor, A.L., Flavell, R.A. and Nathenson, S.G.** (1986). Mol. Cell Biol. 6, 645-652.

6. Weydert, A., Daubas, P., Lazaridis, I., Barton, P., Garner, I. Leader, D.P., Bonhomme, F., Catalan, J., Simon, D., Guenet, J.L., Gros, F. and Buckingham, M.E. (1985). Proc. Natl. Acad. Sci. USA 82, 7183-7187.
7. Mahdavi, V., Periasamy, M. and Nadal-Ginard, B. (1982). Nature 297, 659-665.
8. Saez, L. and Leinwand, L.A. (1986). Nucl. Acid Res. 14, 2951-2969.
9. Lompre, A.M. (1984). J. Biol. Chem. 259, 6437- .
10. Hoh, J.F.Y., Mcrath, P.A. and Hale, H.T. (1978). J. Mol. Cell Card. 11, 1053-1076.
11. Friedman, D.J., Umeda, P.K., Shine, A.M., Hsu, H.J., Jakovcic, S. and Rabinowitz, M. (1984). Proc. Nat'l. Acad. Sci. U.S.A. 81, 3044-3048.
12. Mahdavi, V., Chambers, A.P. and Nadal-Ginard, B. (1984). Proc. Nat'l. Sci. U.S.A. 81, 2626-2630.
13. Izumo, S., Nadal-Ginard, B. and Mahdavi, V. (1986). Science 231, 597-600.
14. Sinha, A.M., Umeda, P.K., Kavinsky, C.J., Rajamanickam, C., Hsu, H-J., Jakovic, S. and Rabinowitz, M. (1982). Proc. Natl. Acad. Sci. USA 79, 5847-5851.
15. Dillmann, W.H. (1984). Mol. Cell. Endocrinol. 34, 169-187.
16. Everett, A.W., Shine, A.M., Omada, P.K., Jakovic, S., Rabinowitz, N. and Zak, R. (1984). Biochemistry 23, 1596-1599.
17. Leinwand, L.A., Fournier, R.E.K., Nadal-Ginard, B. and Shows, T.B. (1983). Science 221, 766-768.
18. Benton, W.D. and Davis, R. W. (1977). Science 196, 180-182.
19. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982). Molecular Cloning. Cold Spring Harbor Lab, Cold Spring Harbor, N.Y.
20. Southern, E. (1975). J. Mol. Biol. 98, 503-517.
21. Suggs, S.V., Hirose, T., Miyake, T., Kawashima, E.H., Johnson, M.J., Itakura, K. and Wallace, R.B. (1981). ICN-UCLA Symp. Mol. Cell Biol. 23, 683-693.
22. Sanger, F., Nicheen, S. and Coulson, A.R. (1977). Proc. Natl. Acad. Sci. USA 74, 5463-5467.
23. Dale, R.M.K., McClura, B.A. and Houchins, J.P. (1985). Plasmid 13, 31-40.
24. Messing, J., Crea, R. and Seeburg, D.H. (1981). Nucl. Acids Res. 9, 309-322.
25. Maxam, A.M. and Gilbert, W. (1980). Methods Enzymol. 65, 499-560.
26. Shows, T.B., Sakaguchi, A.Y. and Naylor, S.L. (1982). Adv. Human Genet. 12, 341-352.
27. Shows, T.B., Sakaguchi, A.Y., Naylor, S.L., Goeddel, D.V. and Lawn, DM. (1982). Science 218, 373-378.
28. Wilbur, W.J. and Lippman, D.J. (1983). Proc. Natl. Acad. Sci. USA 80, 726-730.
29. Gulick, J., Kropp, K. and Robbins, J. (1985). J. Biol. Chem. 260, 14513-14520.
30. Leinwand, L.A., Saez, L., McNally, E. and Nadal-Ginard, B. (1982). Proc. Nat'l. Acad. Sci. U.S.A. 30, 3716-3720.
31. Wydro, R.M., Nguyen, H.T., Gubits, R.M. and Nadal-Ginard, B. (1983). J. Biol. Chem. 258, 670-678.
32. Lichter, P., P. Umeda, J. Levin and Vosberg, H-P. (1986). Eur. J. Biochem. 160, 419-426.
33. Dean, D.C., Knoll, B.J., Riser, M.E. and O'Malley, B.M. (1983). Nature 305, 551-554.
34. Robins, D.M., Peak, I., Seeburg, P.H. and Axel, R. (1982). Cell 29, 623-631.
35. Gunning, P., Ponte, P., Kedes, L.,. Eddy, R. and Shows, T. (1984). Proc. Nat'l. Acad. Sci. U.S.A. 81, 1813-1817.