

The human *Thy-1* gene: Structure and chromosomal location

(nucleotide sequence/transmembrane segment/somatic cell hybrid/chromosome 11)

TETSUNORI SEKI*, NIGEL SPURR†, FUMIYA OBATA*, SANNA GOYERT*, PETER GOODFELLOW†,
AND JACK SILVER*

*Cellular and Molecular Biology Unit, Hospital for Joint Diseases, New York, NY 10003; and †Imperial Cancer Research Fund, 44 Lincoln's Inn Fields, London, WC2A3PX, England

Communicated by Alexander G. Bearn, June 6, 1985

ABSTRACT The human *Thy-1* gene has been isolated and sequenced and compared to the rat and mouse *Thy-1* genes. All three genes are organized in the same way: one exon encoding the majority of the signal peptide, another encoding the transmembrane segment, and a third encoding the remainder of the protein. One major structural difference between the human and rodent *Thy-1* glycoproteins is that the former contains two instead of three glycosylation sites. RNA blot analysis of a human T-cell line expressing the T3 complex showed an absence of *Thy-1* mRNA, excluding the possibility that *Thy-1* represents one of the component chains of T3. The structural gene for human *Thy-1* was localized to the long arm of chromosome 11 by nucleic acid hybridization to genomic DNA isolated from somatic cell hybrids.

Thy-1 was originally described as a cell surface differentiation marker expressed predominantly in mouse brain and thymus (1, 2). It is present, however, in substantially lower amounts in other tissues such as bone marrow and epidermal cells (3-7). *Thy-1* analogs have now been described in a number of species including rats (8), dogs (9), chickens (10), frogs (11), and man (12). Although *Thy-1* expression in many tissues is subject to species variation, expression in brain tissue is invariable, implying a crucial role in the functioning of that organ. Indeed, the preferential expression of *Thy-1* on synaptosomes (13-15) and its appearance on neurons concomitant with synaptogenesis and biochemical and morphological maturation of the brain (16-19) suggests a role for *Thy-1* in synapse formation. The structure of *Thy-1* is consistent with such a function. *Thy-1* is a glycoprotein of $M_r \approx 18,000$ with sequence homology to the immunoglobulins (20) and is therefore part of the immunoglobulin supergene family, which includes histocompatibility antigens and the T-cell and polymeric Ig receptors. Because of this homology it has been proposed that *Thy-1*, like these other membrane proteins, plays a role in cellular interactions and that it may function as an adhesion molecule stabilizing the formation of synapses.

One of the more unusual properties of *Thy-1* is that its pattern of tissue expression varies in different species. For example, *Thy-1* is present on peripheral T cells of mice but absent from rat peripheral T cells (21, 22). In man, the question of *Thy-1* expression on T cells takes on special importance in light of the recent suggestion that *Thy-1* represents one of the component chains of the T3 complex (23). However, this question has remained largely unresolved due to contradictory observations regarding the expression of *Thy-1* on T cell lines and peripheral T cells (24-26). In any event, the differential expression of *Thy-1* makes it an intriguing model for the study of gene regulation.

We report here on the structure of the human *Thy-1* gene and compare it to the structures of the rat and mouse *Thy-1* genes previously isolated (27, 28). In addition, we have examined the possibility that *Thy-1* represents one of the component chains of the T3 complex. Finally, we have determined the chromosomal location of the human *Thy-1* gene.

METHODS AND MATERIALS

Isolation and Characterization of the Human *Thy-1* Gene. High molecular weight DNA was isolated from a human B-lymphoblastoid cell line, LG2, and partially digested with *Mbo* I. The DNA was then used to prepare a genomic library in λ Charon 30 (29). The library was probed with a nick-translated *Pst* I fragment corresponding to the rat *Thy-1* coding sequence (30). One positive plaque was obtained and a 6-kilobase (kb) *Eco*RI fragment containing the *Thy-1* gene was subcloned into pBR322 and sequenced by the method of Maxam and Gilbert (31).

RNA Isolation and RNA Blotting Analysis. Total RNA was isolated from the human neuroblastoma cell line IMR-132 and the human T-leukemic cell line HPB-ALL by using the guanidinium/cesium chloride method (32). Poly(A)⁺ RNA was purified on an oligo(dT)-cellulose column. RNA was electrophoresed on 1% agarose gels containing formaldehyde and blotted on nylon filters (Schleicher & Schuell). Hybridization was carried out at 42°C in 50% formamide using a nick-translated 970-base-pair (bp) *Bam*HI-*Pst* I fragment of the cloned human *Thy-1* gene as a probe followed by washing of the filter at 42°C in 2× standard saline citrate (NaCl/Cit; 1× NaCl/Cit is 0.15 M NaCl/0.015 M sodium citrate, pH 7). To reprobe with T-cell receptor DNA (purchased from Oncor) residual probe was removed by boiling the filter for 20 min in 0.01× SSPE buffer (1× SSPE is 0.18 M NaCl/10 mM NaPO₄, pH 7.7/1 mM EDTA), 1% NaDodSO₄.

Chromosomal Mapping of *Thy-1*. High molecular weight DNA was isolated from a panel of human-mouse somatic cell hybrids and digested with *Pst* I. The digested DNA was blotted onto nitrocellulose (33) and probed with a nick-translated 970-bp *Pst* I-*Bam*HI fragment containing the second coding exon (amino acids -7 to 105). Hybridization was done at 65°C in 6× NaCl/Cit, and filters were washed under stringent conditions (0.1× NaCl/Cit, 65°C).

RESULTS

Identification and Characterization of a *Thy-1* Genomic Clone. A genomic library was prepared from a human B-cell lymphoblastoid cell line in Charon 30 using partially digested *Mbo* I fragments (29). After screening 7×10^5 plaques using a nick-translated fragment of rat *Thy-1* cDNA (30), one positive plaque was obtained. A 6-kb *Eco*RI fragment con-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: bp, base pairs(s); kb, kilobase(s).

taining the *Thy-1* gene was subcloned into pBR322 and subsequently sequenced (Fig. 1). The coding sequence of the protein is divided into three exons separated by introns of 484 and 527 bp. The first exon encodes the first 12 amino acids of the signal peptide, the second encodes the remaining 7 amino acids of the signal peptide plus amino acids 1–105 of the mature protein, and the third exon encodes the remaining 37 amino acids, including a hydrophobic stretch of 20 amino acids at the carboxyl terminus. Polyadenylation signals are located 594 and 1205 bp 3' to the termination codon although presumably only the latter one is recognized (see below). Comparison of the human *Thy-1* gene with the rat and mouse genes reveals that the three are organized in an identical fashion; the only major difference is in the size of the introns (27, 28). However, detailed comparisons of the nucleotide sequences reveal that, although the mouse and rat genes are highly homologous throughout—i.e., in the introns and the 3' untranslated region as well as in the coding regions (data not shown)—the human and rodent *Thy-1* genes display extensive homology only in the coding regions and only a modest degree of homology in the 3' untranslated region (Fig. 2). This

conservation of the 3' untranslated region undoubtedly reflects some important functional role.

Two important points emerge from comparison of the coding sequences of the three genes (Fig. 3). First, the human *Thy-1* gene contains a 20-amino acid hydrophobic segment at the carboxyl terminus analogous to those previously observed for the rat (27) and mouse (28) genes that very likely functions to anchor *Thy-1* to the membrane. This region is highly conserved (>90% homology) in all three species. Second, although the rat and mouse *Thy-1* proteins contain sites of *N*-glycosylation at amino acid positions 23, 75, and 99, the human *Thy-1* protein contains only two sites of *N*-glycosylation; the asparagine residue at position 75 has been replaced by an alanine residue, precluding *N*-glycosylation. Furthermore, the *N*-glycosylation site normally present at position 99 has apparently been moved to amino acid position 101, where the characteristic *N*-glycosylation sequence Asn-X-Ser (X representing any amino acid) is present. It should be noted that there is also a potential *N*-glycosylation site at amino acid position 121 (just prior to the hydrophobic transmembrane segment) in all three *Thy-1*

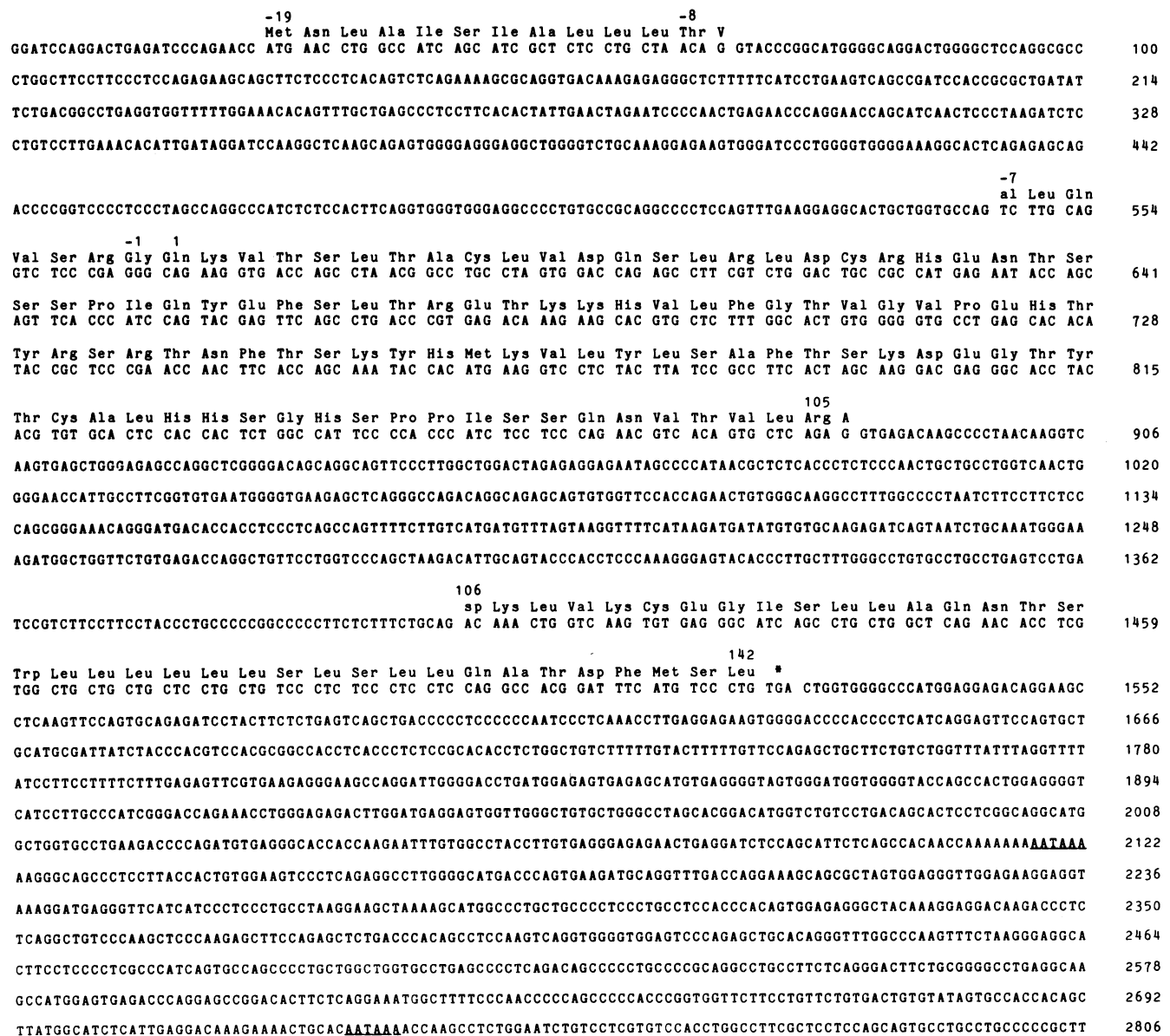


FIG. 1. Nucleotide and predicted amino acid sequences of the human *Thy-1* gene. The two polyadenylation signals are underlined and the termination codon is denoted by an asterisk.

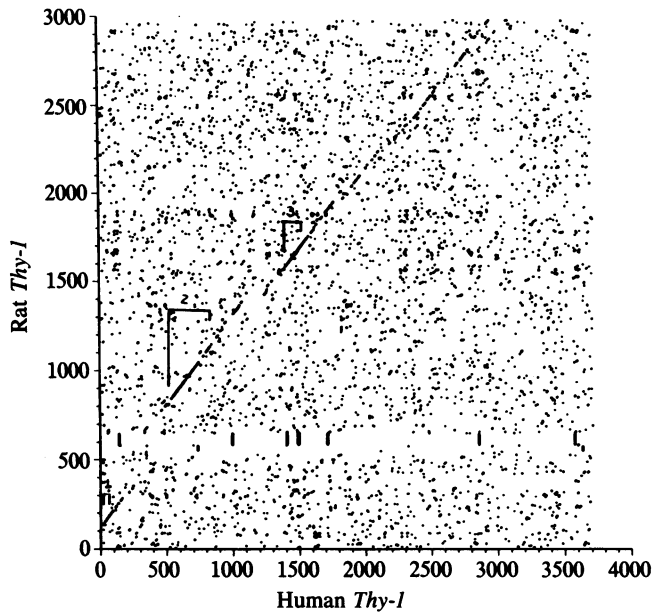


FIG. 2. Dot-matrix plot of the rat versus the human *Thy-1* gene. Dots indicate regions where the two sequences are identical in at least 8 of 10 consecutive nucleotides. The three exons containing coding sequences are bracketed. Analysis was performed using the Steele program of the Albert Einstein College of Medicine molecular biology software package.

proteins but whether this site is actually glycosylated is unknown.

RNA Blotting Analysis. The expression of *Thy-1* mRNA in the T-cell line HPB-ALL, which expresses the T3 complex and from which the gene for the δ subunit of T3 has recently been cloned (34), was examined by RNA blotting analysis. As shown in Fig. 4, no *Thy-1* mRNA was detected in poly(A)⁺ RNA isolated from HPB-ALL cells (lane 2) although *Thy-1* mRNA could be observed in the human neuroblastoma cell line IMR-132 (lane 1). To rule out the possibility that the

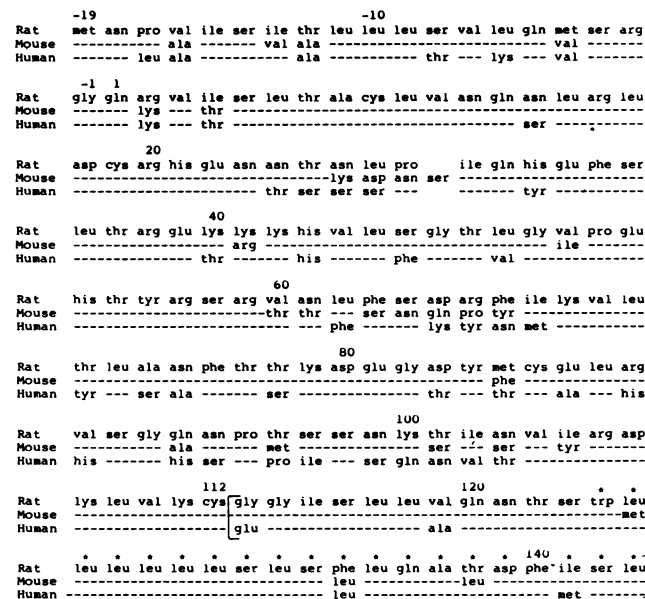


FIG. 3. Protein sequence comparisons of rat, mouse, and human *Thy-1*. Note the insertion of a gap at position 29 to align all three sequences. The additional 31 amino acids predicted from the DNA sequence are bracketed while the 20-amino acid hydrophobic segment is indicated by asterisks.

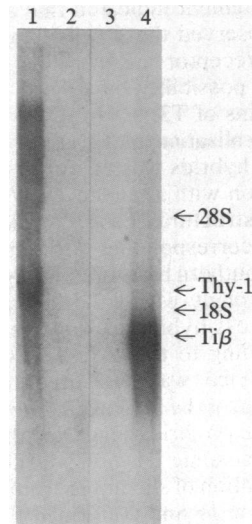


FIG. 4. RNA blot analysis of *Thy-1*. Ten micrograms of total RNA from the human neuroblastoma cell line IMR-132 (lane 1) and 1 μ g of poly(A)⁺ RNA from the human T-leukemic cell line HPB-ALL (lane 2) were subjected to RNA blot analysis using the nick-translated *Bam*HI-*Pst* I fragment of the cloned human *Thy-1* gene as a probe (specific activity, 10⁸ cpm/ μ g). The position of the *Thy-1* mRNA present in the neuroblastoma cell line but absent from the T-cell line is indicated by an arrow. After the residual probe from the filter was removed by boiling, the blot was reprobed with a nick-translated fragment corresponding to the constant region of the β chain of the T-cell receptor. The position of the mRNA for the receptor present in the T-cell line HPB-ALL (lane 4) but absent from the neuroblastoma cell line (lane 3) (*Ti* β) as well as those of ribosomal RNA (28S and 18S) are indicated by arrows.

HPB-ALL cells being analyzed for *Thy-1* had lost expression of T3 they were tested by immunofluorescence using a T3 monoclonal antibody and found to be positive (data not shown). In addition, reprobing of the HPB-ALL mRNA with

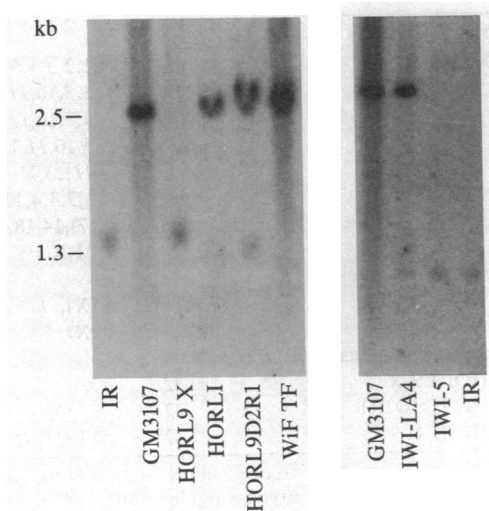


FIG. 5. Chromosomal mapping of the human *Thy-1* gene. High molecular weight DNA from two sets of human-mouse chimeras, HORL9 and IW1, was prepared and approximately 20 μ g of DNA from each subclone was digested with *Pst* I. DNA fragments were separated on a 0.7% agarose gel, transferred to nitrocellulose, and probed with a nick-translated human *Thy-1* gene fragment. The mouse and human fusion partners were IR and WIF TF, respectively. An additional control using the human cell line GM3107 was also included. The human chromosomal composition for each hybrid is described in Table 1.

a T-cell receptor fragment indicated that although no Thy-1 mRNA could be observed these cells expressed substantial amounts of T-cell receptor mRNA (Fig. 4, lane 4). These results rule out the possibility that Thy-1 represents one of the component chains of T3.

Chromosomal Localization of *Thy-1*. A series of human-mouse somatic cell hybrids was analyzed by Southern blotting and hybridization with a human *Thy-1* probe to localize the human *Thy-1* structural gene. A nick-translated *Pst* I-BamHI fragment corresponding to the human *Thy-1* gene was used to probe Southern blots of *Pst* I-digested DNA. This probe hybridizes strongly with a 2.5-kb fragment in human genomic DNA but cross-hybridizes only weakly with a 1.3-kb fragment corresponding to the mouse *Thy-1* gene (Fig. 5). This species difference was exploited to determine the chromosomal location of the human *Thy-1* gene. Results from a panel of eight primary hybrids suggested that the *Thy-1* gene is located on chromosome 11 or 15 (Table 1). Southern blotting and hybridization of subclones from IWI and HORN9 excluded chromosome 15 and confirmed the localization to chromosome 11 (Fig. 5). One clone, HORN9I, which contains an undefined part of the long arm of chromosome 11, was also positive, suggesting a localization to this arm (Table 1).

DISCUSSION

We have isolated and sequenced the human *Thy-1* gene and determined its intron-exon organization by comparing it to the rat and mouse *Thy-1* genes. The coding sequence is separated into three exons with the majority of the signal peptide and the transmembrane segment separated from the main coding sequence by two introns. This is analogous to the organization of other genes that belong to the immunoglobulin supergene family—i.e., histocompatibility antigens and T-cell and polymeric Ig receptors.

As previously described for the rat and mouse *Thy-1* gene products, there is a hydrophobic stretch of 20 amino acids at

the carboxyl end of human Thy-1 that probably represents the transmembrane segment; this region is highly conserved in all three species. It is especially intriguing that the aspartic acid residue at position 139 in the transmembrane segment is also conserved in all three species. Although ionic amino acids have previously been observed in transmembrane segments [e.g., rhodopsin (44) and glycoporphin A (45)], their presence in such is highly unusual. The conservation of the aspartic acid residue may therefore be indicative of some important function. In addition, the sequence of the 3' untranslated region has been significantly conserved despite the lack of conservation of the introns; this may also signify some functional role.

The structural gene for human Thy-1 was localized to the long arm of chromosome 11 by probing Southern blots of human-mouse somatic cell hybrids. Despite structural homology none of the other members of the immunoglobulin supergene family are located on this chromosome; this is also true in mice, where *Thy-1* is located on chromosome 9 (46) whereas the major histocompatibility complex and immunoglobulin heavy and light chains are located on other chromosomes.

Although the function of Thy-1 is still unknown, previous studies have suggested that Thy-1 may be part of the T3 molecular complex, which is associated with the T-cell receptor. However, the absence of *Thy-1* expression in the human T-cell line HPB-ALL, from which a cDNA clone encoding the δ chain of T3 has recently been isolated, eliminates this possibility. In addition, the absence of Thy-1 on this T-cell line is consistent with previous studies indicating the lack of expression of Thy-1 on human T-cells. Thus, although the structure of the *Thy-1* gene is highly conserved in rodents and man, its expression on T cells differs dramatically in these species. This may reflect fundamental differences in the way the *Thy-1* gene is regulated. Studies of the regulatory elements involved in *Thy-1* gene

Table 1. Chromosomal localization of the human *Thy-1* gene

	Ref.*	Human chromosomes present†	Human <i>Thy-1</i> gene‡
Primary hybrids			
SIR19A	35	1,2,3,4,5,7,8,9,10,11,12,13,14,15,17,18,19,20,21,22,X	Present
DUR4.3	36	3,5,10,11,12,13,14,(15),17,18,20,21,22,(X)	Present
HORP27R C14	37	4,7,10,11,12,14,15,21	Present
3W4 C15	38	7,10,11,12,14,15,17,21,X	Present
HORN9D2	37	11,15,X	Present
SIR74ii	35	1,2,3,4,12,14,18,21,22,X	Absent
FIR5	39	(7),14,18,(X)	Absent
ThyB13	40	21,X	Absent
Secondary clones IWI§			
IWI-LA4	41, 42	(X),11	Present
IWI-5	41, 42	(X)	Absent
Secondary clones HORN9¶			
HORN9 X	42	X	Absent
HORN9D2R1	42	11 (X)	Present
HORLI	42	15,(11) (X)	Present

*References given are to original production of the human-mouse hybrids. Many of the hybrids have been subcloned and reanalyzed since the original publication.

†Human chromosomal contributions of the hybrids were deduced from a combination of karyotypic and marker analysis (reviewed in ref. 43). Subchromosomal fragments are given in parentheses.

‡The presence of the human *Thy-1* gene was based on detection of a 2.5-kb hybridizing fragment (see Fig. 5).

§IWI-LA4 and IWI-5, both derived from IWI (38), contain the long arm of the human X chromosome. IWI-LA4 contains in addition a normal chromosome 11; no other human genetic material has been detected in either hybrid.

¶HORN9X, HORN9D2R1, and HORLI are all derived from HORN9 (37). HORN9X contains only the human X chromosome. HORN9D2R1 has a fragment derived from the human X chromosome and a complete human chromosome 11. No markers from chromosome 15 are present in this hybrid. HORLI contains a normal human chromosome 15 and fragments from the X chromosome as well as a karyotypically undefined fragment derived from chromosome 11. No chromosome 11 short arm markers are present in this hybrid (ref. 42 and unpublished results). However, several markers for the long arm of human chromosome 11 are present (42).

expression should provide insight into the molecular mechanisms that determine this differential expression.

We wish to express thanks to Jeff Mordkowitz for his computer and graphic assistance and to Mrs. Roslyn Berger for her excellent secretarial assistance. This work was supported by a grant from the National Institutes of Health.

1. Reif, A. E. & Allen, J. M. (1964) *J. Exp. Med.* **120**, 413–433.
2. Reif, A. E. & Allen, J. M. (1966) *Nature (London)* **209**, 521–523.
3. Basch, R. S. & Berman, J. W. (1982) *Eur. J. Immunol.* **12**, 359–364.
4. Hunt, S. V., Mason, D. W. & Williams, A. F. (1977) *Eur. J. Immunol.* **7**, 817–823.
5. Chambers, D. A., Cohen, R. L. & Heiss, M. A. (1984) *Exp. Cell Biol.* **52**, 125–132.
6. Bergstresser, P. R., Tigelaar, R. E., Dees, J. H. & Streilein, J. W. (1983) *J. Invest. Dermatol.* **81**, 286–288.
7. Tschachler, E., Schuler, G., Hutterer, J., Leibl, H. & Stingl, G. J. (1983) *J. Invest. Dermatol.* **81**, 282–285.
8. Douglas, T. C. (1972) *J. Exp. Med.* **126**, 1054–1062.
9. McKenzie, J. L. & Fabre, J. W. (1981) *Transplantation* **31**, 275–282.
10. Rostas, J. A. P., Shevenan, T. A., Sinclair, C. M. & Jeffrey, P. L. (1983) *Biochem. J.* **213**, 143–152.
11. Mansour, M. H. & Cooper, E. L. (1984) *J. Immunol.* **132**, 2515–2523.
12. Cotmore, S. F., Crowhurst, S. A. & Waterfield, M. D. (1981) *Eur. J. Immunol.* **11**, 597–603.
13. Barclay, A. N. & Hyden, H. J. (1978) *Neurochemistry* **32**, 1583–1586.
14. Stohl, W. & Gonatas, N. K. (1977) *J. Immunol.* **119**, 422–427.
15. Acton, R. T., Addis, J., Carl, G. F., McClain, L. D. & Bridgers, W. F. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3283–3287.
16. Zwerner, R. K., Acton, R. T. & Seeds, N. W. (1977) *Dev. Biol.* **60**, 331–335.
17. Honneger, P. & Richelson, E. (1976) *Brain Res.* **109**, 335–354.
18. Seeds, N. W. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 1858–1861.
19. Seeds, N. W. (1975) *J. Biol. Chem.* **250**, 5455–5458.
20. Campbell, D. G., Gagnon, J., Reid, K. B. M. & Williams, A. F. (1981) *Biochem. J.* **195**, 15–30.
21. Raff, M. C. (1971) *Transplant. Rev.* **6**, 52–80.
22. Acton, R. I., Morris, R. J. & Williams, A. F. (1974) *Eur. J. Immunol.* **4**, 598–602.
23. Gunter, K. C., Malek, T. R. & Shevach, E. M. (1984) *J. Exp. Med.* **159**, 7216–7230.
24. Ades, E. W., Zwerner, R. K., Acton, R. T. & Balch, C. M. (1980) *J. Exp. Med.* **151**, 400–406.
25. Saji, F. & Tanigaki, N. (1982) *Immunogenetics* **15**, 551–563.
26. McKenzie, J. I. & Fabre, J. W. (1981) *J. Immunol.* **126**, 843–850.
27. Seki, T., Moriuchi, T., Chang, H. C., Denome, R. & Silver, J. (1985) *Nature (London)* **313**, 485–487.
28. Seki, T., Chang, H. C., Moriuchi, T., Denome, R., Pleogh, H. & Silver, J. (1985) *Science* **227**, 649–651.
29. Maniatis, T., Fritsch, E. F. & Sambrook, J., eds. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 270–291.
30. Moriuchi, T., Chang, H. C., Denome, R. & Silver, J. (1983) *Nature (London)* **301**, 80–82.
31. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
32. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
33. Southern, E. (1975) *J. Mol. Biol.* **98**, 503–517.
34. vanden Elsen, P., Shepley, B., Borst, J., Coligan, J. E., Markham, A. F., Orkim, S. & Terhorst, C. (1984) *Nature (London)* **312**, 413–418.
35. Whitehead, A. S., Solomon, E., Chambers, S., Bodmer, W. F., Povey, S. & Fey, G. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 5021–5025.
36. Solomon, E., Bobrow, M., Goodfellow, P. N., Bodmer, W. F., Swallow, D. M., Povey, S. & Noel, R. (1976) *Somat. Cell Genet.* **2**, 125–140.
37. van Heyningen, V., Bobrow, M., Bodmer, W. F., Cardner, S. E., Povey, S. & Hopkinson, D. A. (1975) *Ann. Hum. Genet.* **38**, 295–302.
38. Nabholz, M., Miggiano, V. & Bodmer, W. F. (1969) *Nature (London)* **223**, 358–363.
39. Hobart, M. J., Rabbits, T. H., Goodfellow, P. N., Solomon, E., Chambers, S., Spurr, N. & Povey, S. (1981) *Ann. Hum. Genet.* **45**, 331–335.
40. Goodfellow, P. N., Banting, G., Levy, R., Povey, S. & McMichael, A. (1980) *Somat. Cell Genet.* **6**, 777–787.
41. Goodfellow, P. N., Banting, G., Wiles, M. V., Tunnacliffe, A., Parkar, M., Solomon, E., Dalchau, R. & Fabre, J. W. (1982) *Eur. J. Immunol.* **12**, 659–663.
42. Tunnacliffe, A., Goodfellow, P., Banting, G., Solomon, E., Knowles, B. B. & Andrews, P. (1983) *Somat. Cell Genet.* **9**, 629–642.
43. Tunnacliffe, A., Benham, F. & Goodfellow, P. N. (1984) *Trends Biochem. Sci.* **9**, 5–7.
44. Ross, A. H., Radhakrishnan, R., Robson, R. J. & Khorana, H. G. (1982) *J. Biol. Chem.* **257**, 4152–4161.
45. Engelman, D. M., Henderson, R., McLachlan, A. D. & Wallace, B. A. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 2023–2027.
46. Itakura, K., Hutton, J. J., Boyse, E. A. & Old, L. J. (1971) *Nature (London) New Biol.* **230**, 126.