Cloning an expressed gene shared by the human sex chromosomes

(Agt11/cell-surface antigen/X and Y chromosomes/transfectants)

S. M. DARLING*, G. S. BANTING*, B. PYM*, J. WOLFE†, and P. N. GOODFELLOW*

*Laboratory of Human Molecular Genetics, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, United Kingdom; and †Department of Genetics and Biometrics, University College London, Wolfson House, 4 Stephenson Way, London NW1 2HE, United Kingdom

Communicated by G. Pontecorvo, August 15, 1985

ABSTRACT The existence of genes shared by mammalian sex chromosomes has been predicted on both evolutionary and functional grounds. However, the only experimental evidence for such genes in humans is the cell-surface antigen encoded by loci on the X and Y chromosomes (MIC2X and MIC2Y, respectively), which is recognized by the monoclonal antibody 12E7. Using the bacteriophage Agt11 expression system in Escherichia coli and immunoscreening techniques, we have isolated a cDNA clone whose primary product is recognized by 12E7. Southern blot analysis using somatic cell hybrids containing only the human X or Y chromosomes shows that the sequences reacting with the cDNA clone are localized to the sex chromosomes. In addition, the clone hybridizes to DNAs isolated from mouse cells that have been transfected with human DNA and selected for 12E7 expression on the fluorescence-activated cell sorter. We conclude that the cDNA clone encodes the 12E7 antigen, which is the primary product of the MIC2 loci. The clone was used to explore sequence homology between MIC2X and MIC2Y; these loci are closely related, if not identical.

The human sex chromosomes have been shown to share a pair of related genes defined by the monoclonal antibody 12E7. This antibody was raised against leukemoid human cell lines that react with a cell-surface antigen expressed on all human cells, and fails to react with the surface of rodent cells (2). Human–rodent somatic cell hybrids, which retain the human X chromosome, express the 12E7 cell-surface antigen. This defines the MIC2X locus (3). Similarly, somatic cell hybrids that retain the human Y chromosome also express the 12E7 antigen (4) and this defines the MIC2Y locus. No autosomal locus is sufficient or necessary, for 12E7 antigen expression in somatic cell hybrids. The MIC2X locus has been localized to Xpter–Yq11, the euchromatic region of the Y chromosome (4). As might be predicted for an expressed X-linked gene with a Y-linked homologue, the MIC2X locus escapes X-inactivation, thus maintaining functional gene dosage for the MIC2 loci in males and females (6). Two other genes, STS and Xg, map to the tip of the X chromosome short arm (7, 8) and escape X-inactivation (9, 10); however, Y-linked homologues of these genes have not been described in humans.

The level of expression of the 12E7 antigen on erythrocytes, but not on nucleated cells, is polymorphic (11). In females, the polymorphism is controlled by the Xg blood group locus, and in males it is controlled by Xg and a Y-linked locus, Yg (11, 12). The precise genetic and biochemical relationship between the MIC2 loci and the Xg and Yg loci has not been resolved.

The 12E7 antibody is a poor reagent for immunoprecipitation experiments; however, in immunoblot analysis it reacts strongly with a 32-kDa cell-surface molecule expressed on the surface of human but not mouse cells. The antibody also recognizes a 29-kDa cytoplasmic molecule found in both mouse and human cells. The cell-surface products of the human X and Y chromosomes cannot be distinguished by either size or charge (13). These results support the contention that the MIC2X and MIC2Y loci are related, but the degree of homology between these genes cannot be gauged without knowing if the 12E7 antigen is the primary polypeptide product of the MIC2 loci. A direct approach for resolving this question is to clone the DNA sequences corresponding to the MIC2 loci. In this report, we describe the isolation of a cDNA clone corresponding to an MIC2 locus. We have used this clone to investigate the relationship between the 12E7 antigen and the MIC2 loci and sequence homology between the MIC2X and MIC2Y loci. We conclude that the 12E7 antigen is the primary product of the MIC2 loci and that MIC2X and MIC2Y are closely related, if not identical.

MATERIALS AND METHODS

Cell Cultures. Cells were grown at 37°C in Dulbecco's modified Eagle's medium, or RPMI 1640 medium, supplemented with 10% fetal bovine serum, penicillin, and streptomycin. Where appropriate (HORL 9X and "12E7-transfectants"), the medium was supplemented with HAT medium (100 µM hypoxanthine/10 µM methotrexate/16 µM thymidine).

Cell Lines. The following human cell lines were used: J6, a subclone of the male T-cell line Jurkat (14); MolT-4, a male T-cell line (15); GM1416B, a female B-cell line that contains four X chromosomes (purchased from the Institute for Medical Research, Camden, NJ); OK, a B-cell line containing one X chromosome and four Y chromosomes (see ref. 16; a gift from M. Fellous, Institut Pasteur, Paris).

The following mouse cell lines were used: the L cells IR (17) and LMTK− (18); the adenocarcinoma cell line RAG (19). The human–mouse somatic cell hybrids were used: HORL 9X (mouse parent IR) retains the X chromosome as its only human genetic contribution (20); 3E7 (mouse parent RAG), a hybrid retaining the human Y chromosome only (21). Full details of the 12E7 antigen-positive transfectant lines will be presented elsewhere (unpublished data). Briefly, LMTK− cells were transfected with DNA from MolT-4 and the selectable plasmid pTK1 (22) by using the calcium phosphate precipitation method derived from Wigler et al. (23). The transfectants were sorted for 12E7 antigen expression on the fluorescence-activated cell sorter (FACS) and the primary positive transfectant TKM1EP5 was isolated. This mass positive population was sorted further for increased 12E7 antigen expression. TKM1EP7, TKM1EP9, and TKM1EP13 are mass cultures that express increasing amounts of 12E7 antigen (primary amplified transfectants).

Abbreviation: FACS, fluorescence-activated cell sorter.

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.
DNA from TKM1EP7 was transfected into LMTK cells as described above, and two independent secondary positive transfectant populations, EP5A and EP5C, were isolated by FACS selection.

Genomic DNA was isolated from these cultured cell lines by standard techniques (24).

Construction of cDNA Library. The library, a gift from H. Kataoka and M. Collins (Imperial Cancer Research Fund, London), was constructed as follows: cDNA was synthesized from J6 poly(A)^+ RNA (25), and the RNA-DNA hybrid was converted to double-stranded cDNA by treatment with RNase H, DNA polymerase I, and Escherichia coli DNA ligase (26). The cDNA was methylated with EcoRI methylase and ligated to synthetic EcoRI linkers and oligonemic terminal linkers cleaved with EcoRI. The excess linkers were removed by gel filtration through Bio-Gel A-50 and the cDNA was ligated to $\lambda gt$11 arms (27) previously digested with EcoRI and dephosphorylated with calf intestinal alkaline phosphatase. After in vitro packaging, the library of $2.5 \times 10^6$ plaque-forming units (pfu) was amplified in E. coli Y1088, a lact^+

Antibody Screening of cDNA Library and Immunoblot Analysis. The amplified library was propagated in the E. coli strain Y1090 as described by Young and Davis (29). After induction of the $\beta$-galactosidase-cDNA fusion protein for 3.5 hr at 42°C, the nitrocellulose filters were oriented, and then blocked by incubation in phosphate-buffered saline (PBS) (pH 7.5) supplemented with 0.5% Tween 20 and 50 kallikrein units of aprotinin per ml (PBS/Tween) for 15 min at room temperature with gentle agitation (30). The filters were incubated at room temperature overnight with antibody in heat-sealed plastic bags.

A mixture of two monoclonal antibodies, 12E7 (1) and RFB-1 (31), was used for the screening; both antibodies have been shown to recognize the products of the MIC2 loci by immunoblot analysis (ref. 13; unpublished results). The antibody source was hybridoma tissue culture supernatant that was concentrated 10-fold and then diluted 1:200 with PBS/Tween and incubated for 1 hr with the filters. The filters were washed 2 × 20 min in PBS/Tween and 1 × 10 min in PBS before color detection using 4-chloronaphthol (13, 30). Immunoreactive bacteriophage were picked and purified through two additional rounds of screening. Subsequently, bacteriophage inserts were recloned into the EcoRI sites of either pEX2 (32) or pUC8 (33). E. coli pop 2136 (a gift from K. Stanley, European Molecular Biology Laboratory, Heidelberg) was used for transformation of the pEX2 constructs.

For immunoblot analysis (34), proteins from either cell line lysates or E. coli lysates were separated by electrophoresis through NaDodSO_4/polyacrylamide gels and transferred onto nitrocellulose filters as described (13). The filters were treated as described above for antibody screening. A monoclonal anti-$\beta$-galactosidase antibody was a gift from D. Lane (Imperial College, London). RFB1, a gift from M. Bodger (New Zealand), is a monoclonal antibody that recognizes the same antigens as 12E7.

Southern Analysis. Genomic DNAs (20 $\mu$g) were cleaved with restriction enzymes following the manufacturer's recommended conditions and fractionated through 0.8% agarose gels. The gels were denatured, neutralized, and transferred to nitrocellulose filters as described (24). DNA as probe was labeled with [a-$^32$Pj)dCTP by the method of random priming (35) and hybridized to the filters at 42°C in 40% formamide/6% polyethylene glycol 6000. Filters were washed down to 15 mM NaCl/1.5 mM sodium citrate at 65°C and exposed to XAR-5 film (Kodak) at -70°C with an intensifying screen.

RESULTS

Analysis of the Products Encoded by the MIC2 Loci. The monoclonal antibody 12E7 reacts in a species-specific manner with a protein of 32.5-kDa that is expressed on human cells and on human–rodent somatic cell hybrids when the human X and/or Y chromosome is present. A smaller band (29-kDa) is also recognized by the 12E7 antibody and is present in both human and mouse cell extracts (ref. 13; Fig. 1).

Two polypeptides were detected in immunoblot analysis of primary, primary amplified, and secondary DNA transfectant cell lines (Fig. 1) that had been selected for 12E7 surface-antigen expression. In the amplified primary transfectant, the 32.5-kDa species-specific protein is amplified in expression, whereas the common band of 29-kDa protein remains at approximately the same level.

Screening of the cDNA Library. The poor reaction of the 12E7 antibody in immunoprecipitation (13) precluded classical cDNA cloning approaches. However, its reactivity in immunoblotting suggested that direct recognition of fusion proteins in a prokaryotic expression system might be successful. In bacteriophage $\lambda gt$11, eukaryotic peptides can be expressed as part of a $\beta$-galactosidase fusion protein. Previous experiments had suggested that 12E7 and RFB1 recognize the same antigens (unpublished observations). Since it had been proposed that polyclonal antisera might be more efficient for screening $\lambda gt$11 expression libraries (27), we used a mixture of 12E7 and RFB-1 in an attempt to increase the sensitivity of the detection procedure. We screened 2.5 × 10^6 pfu; 30 bacteriophage plaques produced positive signals. $\lambda$SG1 was isolated after two subsequent rounds of plaque purification. DNA from $\lambda$SG1 was prepared, digested with EcoRI, and found to contain a 1-kilobase insert.

Analysis of Fusion Proteins. Initially, we made lysogens of E. coli strain Y1089 by infection with $\lambda$SG1 and analyzed isopropyl $\beta$-D-thiogalactopyranoside (IPTG)-induced lysates for the presence of the $\beta$-galactosidase-cDNA fusion protein. Since the $\beta$-galactosidase-cDNA fusion protein is a much larger protein than the majority of the other E. coli proteins, detection of the product in Coomassie blue-stained gels is

![Fig. 1. Immunoblot analysis of proteins. Cell lysates were separated on NaDodSO_4/polyacrylamide gel and transferred to nitrocellulose. The filter was incubated with the 12E7 antibody and reacting antigenic determinants were detected as described. Lanes: 1, RAG; 2, 3E7; 3, HLR19X; 4, IR; 5, LMTK-, 6, TKM1EP7; 7, TKM1EP13; 8, EP5C.](image)
possible (27). However, we could not detect a fusion protein in Coomassie blue- or silver-stained gels. After immunoblot analysis of these gels and probing with the 12E7 antibody, a single small band of 27-kDa was seen but no large fusion protein was detected. This polypeptide was not constitutively expressed and its production was IPTG inducible. Therefore, we concluded that there was breakdown of the fusion proteins; similar observations have been reported elsewhere (36).

This problem was circumvented by using another prokaryotic expression system—the pEX vectors (32). These vectors contain a cro-lacZ gene fusion, which is under the control of the P8 promoter of bacteriophage λ. A polylinker sequence has been inserted into the 3' end of the lacZ gene in all three translational reading frames so any open reading frame DNA may be expressed as a fusion protein. Insert from λSG1 was cloned into the EcoRI site of pEX2. The recombinant plasmid was introduced into E. coli pop 2136, amplified at 30°C, and then transient expression was induced at 42°C. A 130-kDa cro-β-galactosidase-cDNA fusion protein was detected after Coomassie staining of the gel. In immunoblot analysis the fusion protein reacted with both 12E7 and a monoclonal anti-β-galactosidase antibody. These results strongly suggest that the 12E7 antigenic determinant is entirely defined by a specific sequence of amino acids and is therefore part of a primary gene product (Fig. 2C).

Chromosomal Localization of Sequences Reacting with the cDNA Probe. The insert from λSG1 was subcloned into the vector pUC8 (33) for easier manipulation and was called pSG1. This probe was hybridized to a panel of DNAs consisting of human–mouse hybrids containing the human X or Y chromosomes and human cell lines with various numbers of sex chromosomes (Fig. 3). pSG1 reacts with DNA from the human X and Y chromosomes. No autosomal sequences react with pSG1 (data not shown).

The observations that the pSG1 insert in pEX2 encodes a fusion polypeptide that reacts with the 12E7 antibody and that pSG1-related sequences are present on the X and Y indicates that pSG1 is a cDNA clone homologous to MIC2 sequences. This suggestion was confirmed by hybridization to DNA from the 12E7 expressing genomic DNA transfectants. Primary transfectants and independent secondary transfectants both react with pSG1 (Fig. 4). DNA from the secondary transfectants fails to react with Alu repeat or total human DNA (data not shown), and thus the total amount of human DNA in these cells must be limited. In addition, amplification of 12E7 antigen expression in the transfectants is associated with amplification of the sequences defined by pSG1 (Fig. 4).

Sequence Homology Between MIC2X and MIC2Y. The pSG1 clone has been used to explore the relationship between MIC2X and MIC2Y. The human–rodent hybrids HORL9X and 3E7 containing only the X and Y chromosomes as their human component, respectively, were cleaved with 12 different restriction enzymes and Southern analysis was performed with pSG1. No difference in hybridization pattern

---

**Fig. 2.** Detection of fusion polypeptides. The insert from λSG1 was cloned into the pEX2 vector. The plasmid was amplified at 30°C and transient expression was induced at 42°C. Cell lysates were electrophoresed through NaDodSO4/polyacrylamide gels and the fusion polypeptides were detected by either staining with Coomassie blue (A) or immunoblot analysis and probing either with 12E7 (B) or an anti-β-galactosidase monoclonal antibody (C). Lanes: 1, pEX2; 2, pEX2SG1.

**Fig. 3.** Chromosomal localization of sequences that react with λSG1. Southern analysis of EcoRI-cleaved genomic DNA. Lanes: 1, 3E7; 2, mouse control; 3, HORL9X; 4, GM1416B; 5, OX; 6, J6; M, molecular size markers in kilobase pairs.

**Fig. 4.** Southern analysis of 12E7 transfectant DNAs cleaved with EcoRI. The insert from pSG1 was hybridized to DNAs isolated from mouse cells that had been transfected with human DNA and selected for 12E7 expression on the FACS. Lanes: 1, J6; 2, EP5A (a secondary 12E7 positive transfectant); 3, EPSC (a second independent secondary 12E7 positive transfectant); 4, TKM1EP13 (a primary transfectant showing greater expression of 12E7 antigen than TKM1EP9); 5, TKM1EP9 (a primary transfectant showing increased expression of 12E7 antigen); 6, TKM1EP7 (a primary 12E7 positive transfectant); 7, LMTK-; M, molecular size markers in kilobase pairs.
Evidence from restriction analysis leads us to conclude that the MIC2X and MIC2Y loci are closely related, if not identical. Restriction analysis of the genomic clones for MIC2X and MIC2Y will allow us to investigate their relatedness over a greater distance.

This is the first described probe for a Y-linked expressed gene and the first probe for an X-linked gene that escapes X inactivation. The existence of structural genes on both sex chromosomes makes a relationship between MIC2 and the minor histocompatibility antigen H-Y unlikely.

Although the human sex chromosomes differ extensively in morphology and genetic content (39), they are thought to have evolved from a homologous pair of chromosomes (40). The sharing of genes by the sex chromosomes might reflect this common origin or might reflect a recent exchange of genetic material. It has been proposed that sex chromosomes share sequences to facilitate meiotic chromosomal pairing and that such sequences might be exchanged between the sex chromosomes by normal recombination (41, 42). Such sequences have been termed pseudoautosomal, as they would not be expected to show sex linkage. The existence of pseudoautosomal genes has been supported by the finding that the Srs and abnormal Sxr regions are exchanged between the X and Y chromosomes in the mouse (43–45).

Early studies suggested that the pairing, and consequently the homology, involved the distal third of the short arm of the human X chromosome and almost the entire Y chromosome short arm (46, 47). More recently, experiments have shown that pairing varies extensively between species (48). Homology of single-copy DNA sequences between the X and Y chromosomes was first reported by Page et al. (49) for the marker DXYSI. However, in situ hybridization revealed that DXYS1 maps to the long arm of the X chromosome and to the short arm of the Y (50), outside the postulated pairing region. Other single-copy probes have been mapped to the human X and Y chromosomes, although again none has been shown to reside within the pairing region (51, 52). The MIC2X locus has been mapped to Xp22.32–Xpter (4, 5), and the MIC2Y locus has been mapped to Yq11–Ypter (4), locations that include the postulated pairing regions. Using the cloned cDNA to isolate genomic clones we will be able to investigate the extent of homology surrounding the MIC2 loci and whether these loci are exchanged between the sex chromosomes by normal recombination.

We would like to thank H. Kataoka and M. Collins for the λgt11 library; K. Stanley for the pEX vectors and E. coli pop 2136; D. Lane for the monoclonal anti-β-galactosidase antibody; G. Lam and L. Davis for expert assistance on the FACS; P. Lamb, C. Mondello, C. Pritchard, M. V. Wiles, and B. Williams for helpful suggestions; and C. Middlemiss for typing the manuscript.


**Table 1.** Restriction enzyme analysis of sequence homology between MIC2X and MIC2Y

<table>
<thead>
<tr>
<th>Restriction enzyme</th>
<th>Recognition site</th>
<th>No. of fragments</th>
<th>Minimum bp sites (obs. to hybridize* to λSG1)</th>
<th>No. of bp identical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alu I</td>
<td>AGCT</td>
<td>3</td>
<td>53</td>
<td>90</td>
</tr>
<tr>
<td>Bgl I</td>
<td>GCCNNNNGGCC</td>
<td>3</td>
<td>4</td>
<td>24</td>
</tr>
<tr>
<td>Pvu II</td>
<td>CAGCTG</td>
<td>6</td>
<td>7</td>
<td>32</td>
</tr>
<tr>
<td>HindIII</td>
<td>AAGCTT</td>
<td>3</td>
<td>4</td>
<td>24</td>
</tr>
<tr>
<td>Pst I</td>
<td>CTCGAG</td>
<td>6</td>
<td>7</td>
<td>42</td>
</tr>
<tr>
<td>Sau3A</td>
<td>GATC</td>
<td>3</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>Xba I</td>
<td>TCTAGA</td>
<td>3</td>
<td>4</td>
<td>24</td>
</tr>
<tr>
<td>Bgl II</td>
<td>AGATCT</td>
<td>2</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>EcoRI</td>
<td>GAAATTCC</td>
<td>4</td>
<td>5</td>
<td>30</td>
</tr>
<tr>
<td>HincII</td>
<td>GTTPuAC</td>
<td>4</td>
<td>5</td>
<td>30</td>
</tr>
<tr>
<td>Hinfl</td>
<td>GANTC</td>
<td>3</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>290</td>
</tr>
</tbody>
</table>

*bp, Base pairs.

*Hybridization patterns for HORL9X and 3E7 were identical. Hybridizing fragments were counted after 1-week exposure of the filters at -70°C; longer exposure results in the appearance of more bands.

was found between MIC2X and MIC2Y (Table 1). Two restriction enzymes, *Taq* I and *Msp* I, have been excluded from this analysis because they recognize extensive X- and Y-linked polymorphisms. The 12 restriction enzymes recognize a minimum total of 53 sites and have sampled a minimum total of 290 base pairs. We conclude that the MIC2X and MIC2Y loci are closely related in sequence and may be identical. Because of this close degree of homology and the fact that the RNA used to construct the cDNA library was derived from J6, a male T-cell line, we do not know whether we have isolated a cDNA clone corresponding to MIC2X or MIC2Y.

**DISCUSSION**

We have isolated a cDNA clone that corresponds to a MIC2 locus and conclude that the 12E7 antigen is the primary product of the MIC2 loci. Several experiments lead us to this conclusion. First, a recombinant bacteriophage, λSG1, expressing an antigenic determinant that is recognized by the 12E7 antibody has been isolated. The insert from this bacteriophage, after subcloning into a plasmid expression system, has been shown to express a cro-β-galactosidase–cDNA fusion protein that is recognized by both the 12E7 antibody and a β-galactosidase antibody. Second, hybridization analysis of the insert shows localization to the human X and Y chromosomes. Third, sequences that share homology with an insert from λSG1 are present in DNAs isolated from mouse cells that have been transfected with human DNA and selected for 12E7 antigen expression on the FACS. It has been estimated that L-cell transfectants obtained by using cotransfection with the thymidine kinase gene incorporate ~0.1% of the human genome (37, 38). Thus, the probability of finding by chance a sequence of interest in two independent secondary transfectants isolated by transfection from DNA of a primary transfectant line is 1 in 106. Finally, increased antigen expression in the primary transfectants is associated with an increased amount of the 32.5-kDa polypeptide and increased amounts of the sequences defined by the insert from bacteriophage λSG1.

Previous biochemical analysis had not detected a difference between the MIC2X and MIC2Y gene products (13).
Genetics: Darling et al.