Immunogenetics (1999) 49:919–930

ORIGINAL PAPER

**Florian W. Velten · Christine Renard
Claire Rogel-Gaillard · Marcel Vaiman
Juergen Schrezenmeir · Patrick Chardon**

Spatial arrangement of pig MHC class I sequences

Received: 1 December 1998 / Revised: 5 May 1999

Abstract Bacterial artificial chromosome (BAC) clones were assigned within the pig major histocompatibility complex (*Mhc*) by polymerase chain reaction-screening and Southern blot hybridization using sequence-tagged site (STS) markers and BAC end-rescued sequences. In all, 35 BAC clones were discovered containing 12 anchor genes of the *SLA* class I region and two genes of the *SLA* class III region. Twenty of these 35 clones comprised two distinct class I gene clusters, each spanning about 100 kilobases. One cluster enclosed three class I related genes (*SLA-6* to *-8*) and two genes (*MIC-1* and *MIC-2*) more distantly related to class I. The other cluster enclosed typical class I genes, of which three (*SLA-1, -2*, and *-3*) were transcribed by fibroblasts homozygous for the *H01* haplotype which we used to construct a pig BAC library. Ordered clones are certainly helpful in isolating agronomically, biologically, and medically important genes. They would also be useful for inducing genetic modifications in pig cell lines.

Key words Pig · Major histocompatibility complex · Class I region · *SLA* · BAC

F.W. Velten · C. Renard · C. Rogel-Gaillard · P. Chardon
Laboratoire Mixte INRA-CEA de Radiobiologie Appliquée, Domaine de Vilvert, 78352 Jouy-en-Josas Cedex, France

F.W. Velten (✉)Institute of Physiology and Biochemistry of Nutrition, Federal Research Center, 24103 Kiel, Germany
e-mail address: velten@bafm.de,
Tel.: +49-431-6092460,
Fax: +49-431-6092472

M. Vaiman
LREG, CEA-INRA, DSV, DRR, 78352 Jouy-en-Josas Cedex, France

The nucleotide sequence data reported in this paper have been submitted to the EMBL, GenBank, and DDBJ nucleotide sequence databases and have been assigned the accession numbers Z97378 to Z97409, AF074418, AF074419, AF074423, AF083661, AF083662, and AF083663

Introduction

Xenotransplantation of single cells and solid organs from pig to humans constitutes a way of overcoming the increasing shortage of human tissues available for transplantation. However, the problems of xenotransplantation are complex, and a combination of approaches is required. Even though the hyperacute process of xenograft tissue rejection seems avoidable, eliminating the obstacle of acute rejection remains a formidable challenge. There is convincing evidence that pig major histocompatibility complex (*Mhc*) *SLA* class I and class II molecules have a prominent although not exclusive role in activating the human immune cells involved in the acute response to xenografts. The *SLA* class II molecules that are present on several cells, including the endothelial cells of vascularized organs, have the capacity to activate the human $CD4^+$ T helper cells both directly and indirectly. Similarly, the classical *SLA* class I molecules, expressed by almost all cell types, are directly recognized by human $CD8^+$ cytolytic T cells (Shishido et al. 1997) and by human natural killer (NK) cell subpopulations (Sullivan et al. 1997). Both T lymphocytes and NK cells initiate a vast field of inflammatory processes leading to graft destruction within several hours.

One way to protect xenografts from acute rejection would be to eliminate, or at least greatly diminish, the *SLA* class I and class II molecules. Efficient manipulation of these molecules requires precise knowledge of the pig *Mhc* complex. Although important advances have already been made towards this goal, many aspects of the pig *Mhc* are still not clear, such as the characterization of all of the functional *SLA* classical and class I-related genes in specific haplotypes. We also need to evaluate the function of these genes, and to identify the locations at which they are transcribed, especially in a xenograft situation.

The pig *SLA* complex is located on Chromosome (SSC) 7 (Geffrotin et al. 1984). Unlike the organization of the adjacent class I, III and II regions of the *Mhc* in

other mammalian species, the pig *Mhc* region is disrupted by the centromere. The class I and class III regions are located on the short arm in SSC 7p1.1, and the class II region is located on the long arm in SSC 7q1.1 (Smith et al. 1995). Taking advantage of the existence of pig yeast artificial chromosome (YAC) and cosmid libraries, we previously constructed the first physical map of the *SLA* class I and class III regions (Peelman et al. 1996; Velten et al. 1998). We found that all 10 identifiable SLA class I genes were located in two clusters of about 200 kilobases (kb) each. One cluster comprised seven classical class I genes, including pseudogenes, and the other, class I-related genes. At present it is not known which genes of the classical class I sequence are involved in the serologically defined *SLA* class I allelic series, neither has the order of the genes within the clusters been determined. We also showed that in contrast to the absence of orthologous relationships between pig and humans prevailing in the class I sequences, the colinearity of the evolutionarily conserved genes was perfectly preserved, with, however, variations in the distance between the genes. The total size of the class I region in pig is half that of its human counterpart.

The number of pig genomic large fragment libraries available worldwide is limited. Moreover, these libraries were usually constructed from pig bearing undefined *SLA* haplotypes, which introduces additional complexity into the analysis of a highly polymorphic region, such as the *Mhc* complex. Despite the great usefulness of YAC clones in mapping complex genomic regions, they display certain inherent shortcomings, principally chimerism and also instability. In addition, because YAC clones are usually large, they are less amenable to direct sequencing procedures. As bacterial artificial chromosome (BAC) clones probably represent the native genome more faithfully than YACs, we recently produced a pig BAC library of more than 107000 clones, comprising five genome equivalents. This library was constructed with the DNA from one of the two pigs, homozygous for *SLA*, used for the previously constructed YAC library, thus making it possible to compare the data obtained with each library. The BAC library was exhaustively analyzed for the genes of the *SLA* class I and class II regions. The results for the class II region are being reported separately. In the present study, the aims were to improve the physical map of the *SLA* class I region, define the order of the genes as accurately as possible, identify those class I genes which are functional, and when feasible, fill in the gaps in our previous map of the class I region.

Materials and methods

Polymerase chain reaction screening of the BAC library

Polymerase chain reaction (PCR) was done within our pig BAC library based on the genome of a homozygous Large White male displaying the *H01* (*W15*, *FJ1*, *B18*) haplotype. As stated above, the BAC library comprised about 107000 clones with an average insert size of 145 kb, and constituted fivefold coverage of the pig haploid genome. Amplifications were performed either on total BAC DNA or *Eco* RI-digested *SLA* class I sequences containing BAC bands, in a 25 μl reaction volume containing 6.25 to 25 pmol of each primer, 200 μM dNTP, 1.5 mM $MgCl_2$, and 1 unit of either Ampli*Taq* DNA polymerase (Perkin Elmer, Roissy CDG Cedex, France) or Goldstar DNA polymerase (Eurogentec, Seraing, Belgium). The amplification reaction included denaturation at 98 °C for 1 min during the initial three cycles, followed by 27–35 cycles at 95 °C for 1 min, a primer-specific annealing temperature for 1–2 min, and 72 °C for 1–2 min. The terminal extension was for 3 min at 72 °C. A positive control was run on human and pig genomic DNA, and a negative control, on pBeloBAC II and *Escherichia coli* DNA, used to construct the pig BAC library.

In addition to using the primer pairs already described (Velten et al. 1998), we designed several new primers that amplified *SLA* class I genes and *SLA*-related anchor sequences (Table 1). Hence, the primer pairs C1, C4, and C6 to C12 amplified various parts of the *SLA-1, -2, -3, -4, -5, -6, -9*, and *-10* genes (Velten et al. 1998). Primers for *ZNF173* and *POU5F1* were derived from human sequences (Chu et al. 1995; Takeda et al. 1992). On the basis of the pig *TNF-alpha* sequence (Drews et al. 1990) we designed the *TNF* primer pairs.

Subcloning, sequencing, and computer analysis

PCR amplicons were subcloned into the T-A vector pGEM-T Easy (Promega, Charbonnières, France) and then transformed into electrocompetent DH10B cells (Life Technologies SARL, Cergy Pontoise Cedex, France). Sequence reactions were obtained by cycle sequencing with the fluorescent primers used in the dideoxynucleotide chain termination method (Sanger et al. 1977) and analyzed on an ABI 373 A automated sequencer (Perkin Elmer Applied Biosystems, Foster City, Calif.). After two separate PCR amplifications, each sequence was characterized in at least two bacterial clones. These sequences were compared with the EMBL/GenBank database (release 54/106) using the Goad and Kanehisa (1982) algorithm and the FASTA program (Pearson and Lipman 1988).

DNA miniprep

Three ml overnight cultures were centrifuged and resuspended at 4 °C in 200 μl of solution 1 (50 mM Tris-HCl, pH 8.0, 10 mM ethylenediaminetetraacetate (EDTA), 50 mM glucose, and 100 μg/ml RNase A), and lysed for 5 min at room temperature in 200 μl of solution 2 (200 mM NaOH, and 1% SDS). One-hundred and fifty microliters of solution 3 (1.32 M potassium acetate, pH 4.8), was added, and the preparation was incubated for 5 min on ice and centrifuged twice. The DNA in the supernatant was then precipitated by adding 0.6 vol isopropanol, pelleted for 30 min by centrifugation at 20 000 g and 4 °C, washed once in cold 70% ethanol, centrifuged, dried, and resuspended in 50 μl TE.

Pulsed-field gel electrophoresis

For sizing inserts in BAC clones, 1 μg DNA was digested with 10 units of *Not* I for 3 h and submitted to electrophoresis. Pulsed-field gel electrophoresis (PFGE) was performed at 14 °C in 1% agarose gel (Seakem GTG, Biozym Diagnostik, Hess. Oldendorf, Germany), 0.5 × TBE, on a CHEF DRIII apparatus (Bio-Rad, Ivry Sur Seine Cedex, France) using a switch time of 3–15 s, at a voltage of 6 V/cm, an angle of 120°, and a migration time of 18 h. The size standards used included low-range and mid-range PFG markers (New England Biolabs, Beverly, Mass.).

Table 1 Primer pairs used for *SLA* gene amplification. Also indicated are the locus amplified, PCR melting temperatures (Tm), and the amplicon length of DNA (normal type) and mRNA (bold type). h: human, p: pig

Locus	Primer	Forward / reverse primer sequence	Source	Length	Tm	Accession or Ref.
		Class I genes				
SLA-2,-4,-10 (exons 3 to 4)	C1	5′-GATGGGGAAGGACACGTCGC 3′-CCAGGAGGCACCACCAGGGC	(p)	848 bp **235 bp**	65 °C	PD1: M21057 PD7: (Frels et al. 1990) PD14: M21058
SLA-1,-3,-5,-7, and *–9* (exons 1 to 4)	C4	5′-TCTTCCTGCTGCTGTCGGGAACCTTG-GCCCTGAC 3′-CCCATCTCAGGGTGAGGGGCTCCTG-CAGCCCT	(P)	2021 bp	65 °C	PD1: M21057 PD7: (Frels et al. 1990) PD14: M21058
SLA-4 (intron 3/exon 4)	C6	5′-AGACCCCACCCCCATCCCCAGGC 3′-CATGTCCTGGCTCTGGTCCTGGC	(p)	292 bp	67 °C	Z97387
SLA-1 (exon 2)	C7	5′-CTCCCGCTTCATCGCCGTCG 3′-GGCCTCGCTCTGGTTGTAGTAG	(p)	214 bp	67 °C	Z97379
SLA-2 and *-5* (exon 2)	C8	5′-GCCCCGCTTCATCGCCGTCG 3′-GGCCTCGCTCTGGTTGTAGTAG	(p)	214 bp	67 °C	Z97379 Z97389
SLA-3 (exon 2)	C9	5′-AAGCCCCGTTTCATCTCCGTC 3′-GGCCTCGCTCTGGTTGTAGTAG	(p)	216 bp	67 °C	Z97384
SLA-6 (exon 2)	C10	5′-CACTCGCTAAGATACCTCC 3′-CGACTGGCTATGGTTGTGGTTG	(p)	264 bp **264 bp**	60 °C	Z97392
SLA-7 (exon 2)	C11	5′-TCCCCGCTTCTCCGTCGTCG 3′-GGCCTCGCTCTGGTTGTAGTAG	(p)	214 bp	67 °C	Z97394
SLA-9 (exon 2)	C12	5′-GCCCCGTTTCATTGCCATCG 3′-GGCCTCGCTCTGGTTGTAGTAG	(p)	214 bp	67 °C	Z97397
transcribed *SLA* class I genes (exons 2 to 4)	C13	5′-C7 or -C8 or -C9 or -C11 or -C12 3′-CCAGGAGGCACCACCAGGGC	(p)	**708 bp**	67 °C	
		Class I anchor genes				
TNF-alpha	TNF	5′-AATGTCAAAGCCGAGGGACAG 3′-GGCAATGATCCCAAAGTAGAC	(p)	414 bp	65 °C	X54001
BAT1	BAT	5′-AACTTCCCAGCCATTGCCATCCACC 3′-GGAAATGTCTATCTCGTCAGGCAGC	(p)	570 bp	63 °C	Z34846
SC1	SC	5′-GACTTTGCTGCCATTACCATCCAC 3′-CAGATTTCCTCCGATTGCGTTGGGG	(p)	292 bp	65 °C	Z97401
POU5F1	POU	5′-GTTCTGTAACCGGCGCCAGAAGGGC 3′-CTCAGTTTGAATGCATGGGAGAGCC	(h)	254 bp	55 °C	Z11900
S-protein	S	5′-TGTAGCAATGGTGGCCTTCC 3′-CTAGGATATCCGGATGGAGC	(p)	694 bp	72 °C	Z97402
ZNF173	ZNF173	5′-ACCCTGGACCCACAGTCGGCCAGTG 3′-GGGCGATGCCCACTCTCCGGGGCCG	(h)	532 bp	64 °C	U09825
RFB30	RFB	5′-TAATAGACGGGTGCCAAGGGGAAT 3′-TGGGTGAAGGAGGCAGTGAAGGTG	(p)	549 bp	60 °C	Z97403
MOG	MOG	5′-CGGGCTCTGGTCGGGGATGAAG 3′-CATTGCTGCCTCCTCTTGGTAA	(p)	297 bp	55 °C	Z97404
OLF-42	OLF42A	5′-CCAAGCTCCACTCTCCAATG 3′-TGGAGGTAGACAGCAATGACTG	(p)	625 bp	65 °C	Z97405 Z97406
OLF-42	OLF42B	5′-TCTTCATCTTCCTGCTCCTGGGG 3′-GGGCGATGGCACCGTAAGAGATA	(p)	366 bp	67 °C	Z97405
OLF-89	OLF89	5′-TTTCCTCACCAACCTCTCCTT 3′-GCCATAAAATAAGAAAACTAC	(p)	574 bp	55 °C	Z9740/ Z97408

Southern blots and hybridization

After obtaining the DNA miniprep, 10 μg of BAC insert DNA was digested at 37 °C for 3 h with *Eco* RI (2 units/μg DNA). Digested DNA samples underwent electrophoresis for 20 h at 50 V in 1% low-melting agarose gel (Seakem GTG). The DNA in the gel was then depurinated, denaturated, and transferred onto Nytran N^+ membranes (Schleicher & Schuell, Dassel, Germany). The membranes were probed, either with a radiolabeled 0.8 kb probe (Velten et al. 1998) corresponding to the pig exon 3 and 4 regions of the class I gene, or with a 283 base pair (bp) human *MICA* probe corresponding to the exon 4 and 5 regions of this

gene (Bahram et al. 1996). The *MICA* probe was obtained by amplifying human genomic DNA with primers: 5′-GAG CCA CGA CAC CCA GCA GT, and 3′-CCA ATG ACT CTG AAG CAC CA. Both probes were radiolabeled with [a^{32}P] dCTP using a random-primed labeling kit (Boehringer Mannheim GmbH, Mannheim, Germany).

The pig SLA and MIC genes containing bands of *Eco* RI-digested BACs were cut out of the agarose gel and purified. SLA class I genes were amplified with the primer pairs used for library screening and sequenced. Bands containing MIC genes were redigested with *Sau* 3A, subcloned in the pUC18 vector, and sequenced.

Reverse transcription – polymerase chain reaction

To identify transcribed class I genes, we produced cDNA from fibroblasts of the *H01* haplotype, for PCR amplification with SLA class I gene-specific primers. Under RNAse protection conditions, the reverse transcription reaction was obtained on 3 μg total fibroblast RNA in a final volume of 50 μl, using standard buffer with 0.5 mM dNTPs, 10 mM DTT, 0.5 μg oligo $(dT)_{12–18}$ and Super-Script II reverse transcriptase (600 units, Life Technologies SARL, Cergy Pontoise Cedex, France). After reverse transcription at 42 °C for 50 min, the transcriptase was inactivated at 70 °C for 15 min. The length of the amplified poly-A mRNA fragments was evaluated in 1% agarose gel, after the amplification of 10 μl of [a^{32}P]dCTP radiolabeled RT-PCR solution.

Pig interspersed repetitive sequence polymerase chain reaction

The alignment of the *SLA* class I and III sequences containing YAC and BAC clones was confirmed by comparing the length of pig inter-SINE elements after pig interspersed repetitive sequence PIRS-PCR amplification. The primers 5′-CCA CTG AGC CAC AAC GGG AAC and 3′-GAC CCC TAG CCT GGG AAC CT were derived from highly conserved pig SINE elements (Yerle et al. 1996) and amplified within clone 4 to 12 inter-SINE fragments whose length ranged from about 0.1 to 1.3 kb.

BAC end sequencing

To detect the BAC end sequences of *SLA* class I DNA fragments containing BAC clones of the contig, BAC-DNA was digested with *Hin* dIII and directly sequenced, using fluorescence-labeled dideoxyterminators and the pBELOBAC II-derived direct and reverse primers 5′-GTA AAA CGA CGG CCA GT and 3′-AGC GGA TAA CAA TTT CAC.

Results

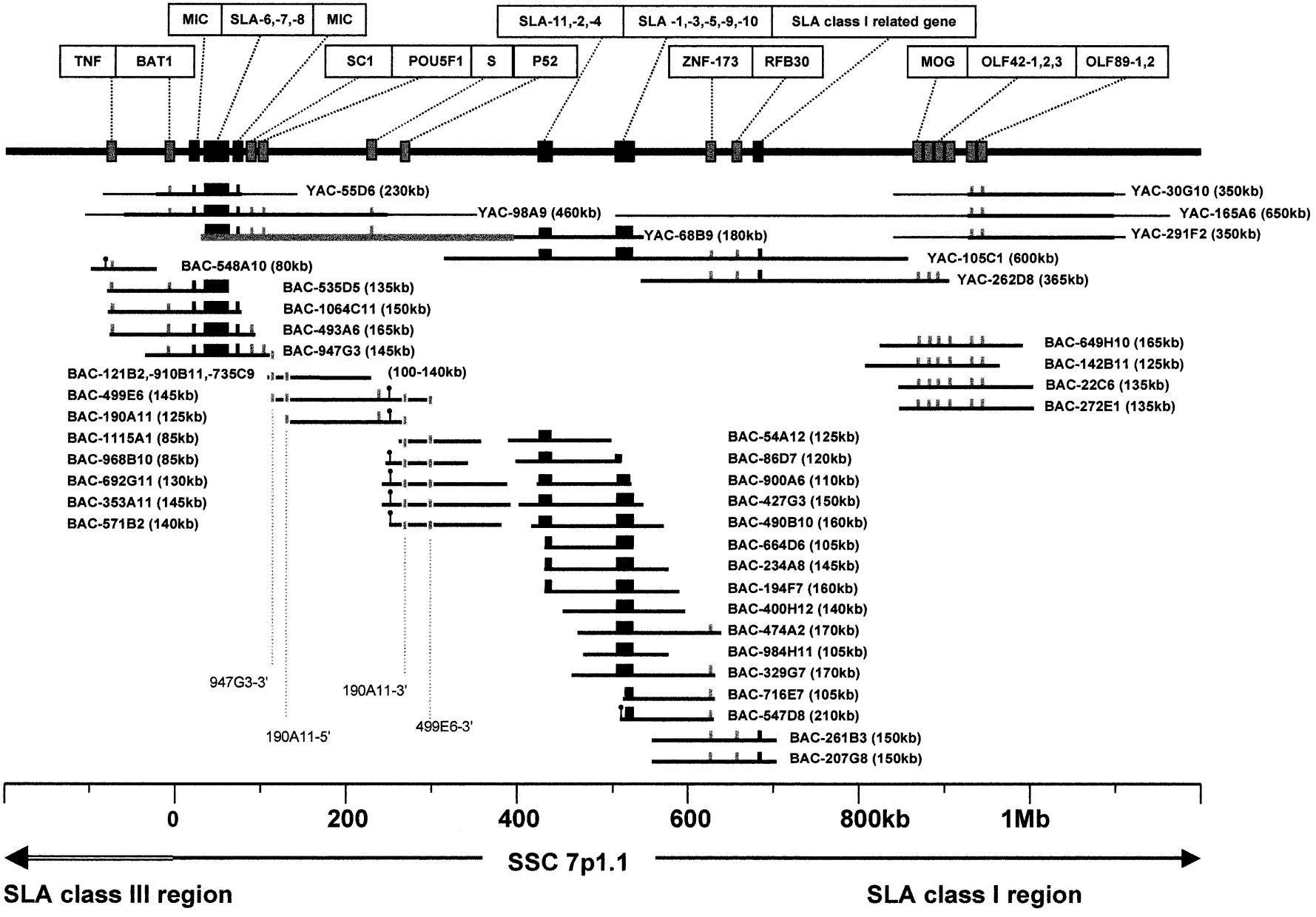
Polymerase chain reaction screening of the BAC library

The pig BAC library comprising over 107000 clones organized in pools was screened by PCR. The primer pairs used were derived (1) from exons 1 to 4, the best conserved sequence among the pig class I genes, (2) from the evolutionarily conserved or anchor genes *TNFA, BAT1, SC1, POU5F1, S, ZNF173, RFB30, MOG, OLF42,* and *OLF89*, and (3) from end-rescued sequences of selected BACs. Eighteen class I-containing BACs whose size ranged from 80 to 210 kb were isolated using the class I primers. Next, screening with the anchor gene primers yielded 17 BAC clones, ten of which also carried one or more class I sequences. Since most of the anchor genes were of human origin, the specificity of the amplified pig products was checked by sequencing. Further screening of the library with primers from BAC end sequences permitted the recovery of eight additional clones. Once aligned according to their sequence content, the clones constituted three contigs of overlapping BACs. As shown in Fig. 1, these contigs were positioned in relation to the previously established YAC physical map. Fifteen BACs made up the most centromeric contig, which spanned more than 400 kb. From the most centromeric end, the order of the genes initially characterized was as follows: the *TNFA* and *BAT1* class III genes, five class I-related genes from the *SLA* class I and *MIC* families, and the *SC1, POU5F1,* and *S* genes. The central contig consisted of 16 class I-containing BACs, the distal ones comprising either the acid finger protein gene *ZNF173* alone, or the *ZNF173* and *RFB30* genes. The size of this contig was about 300 kb. The four BACs of the distal contig spanning about 170 kb all included the *MOG* gene and the *OLF42* and *OLF89* clusters. Systematic comparison of the BAC end-rescued sequences with the GenBank database revealed a similarity of about 90% between one end sequence of BAC number 190A11 and the corresponding sequence of the human *p52* gene. In pig, the *p52* gene was found to be telomeric but was located near the *POU5F1-S* gene cluster. *ZNF173* is a newly mapped gene which lies just in front of the *RFB30* gene. The size of each gap between the BAC contigs was roughly evaluated at 200 kb.

Fig. 1 Alignment of the SLA class I sequences containing BACs and YACs (—— nonchimerical clones, —▬— chimerical clones, —— deletion; ❣: *Not* I restriction site; *TNFa:* tumor necrosis factor alpha gene; *BAT1*: HLA-B associated transcript 1 gene; *SLA*: swine leukocyte antigen; *MIC-1,2*: *Mhc* class I molecule-related genes 1 and 2; *SC1*: cell growth regulated gene 1; *POU5F1*: octamer transcription factor 3 gene*; P52:* fifth subunit of the core of the transcription/DNA repair factor TFIIH*; S*: skin-associated protein gene; *ZNF173*: zinc finger protein 173 gene; *RFB30*: ring finger protein 30 gene; *MOG*: myelin oligodendrocyte glycoprotein gene; *OLF42/89*: olfactory receptor-like genes. BAC clones recovered using the class 1 primers: 535D5, 1064C11, 493A6, 947G3, 54A12, 86D7, 900A6, 427G3, 490B10, 664D6, 234A8, 194F7, 400H12, 474A2, 984H11, 329G7, 716E7, and 547D8. BAC clones containing exclusively anchor genes: 548A10, 499E6, 190A11, 649H10, 142B11, 22C6, 272E1. The BACs 261B3 and 207G8 were subsequently shown to contain a putative related class I gene by Southern blot. BAC clones recovered using end-rescued sequences : 121B2, 910B11, 735C9, 1115A1, 968B10, 692G11, 353A11, and 571B2. YACs clones were aligned previously (Velten et al 1998)

BAC alignment

Southern blot analysis: Purified DNA of all 20 centromeric and central contig BACs containing a class I sequence as well as of three BACs without a class I se-


MIC
SLA-6,-7,-8
MIC
SLA-11,-2,-4
SLA -1,-3,-5,-9,-10
SLA class I related gene
TNF
BAT1
SC1
POU5F1
S
P52
ZNF-173
RFB30
MOG
OLF42-1,2,3
OLF89-1,2
YAC-55D6 (230kb)
YAC-98A9 (460kb)
YAC-68B9 (180kb)
YAC-105C1 (600kb)
YAC-262D8 (365kb)
YAC-30G10 (350kb)
YAC-165A6 (650kb)
YAC-291F2 (350kb)
BAC-548A10 (80kb)
BAC-535D5 (135kb)
BAC-1064C11 (150kb)
BAC-493A6 (165kb)
BAC-947G3 (145kb)
BAC-121B2,-910B11,-735C9 (100-140kb)
BAC-499E6 (145kb)
BAC-190A11 (125kb)
BAC-1115A1 (85kb)
BAC-968B10 (85kb)
BAC-692G11 (130kb)
BAC-353A11 (145kb)
BAC-571B2 (140kb)
BAC-649H10 (165kb)
BAC-142B11 (125kb)
BAC-22C6 (135kb)
BAC-272E1 (135kb)
BAC-54A12 (125kb)
BAC-86D7 (120kb)
BAC-900A6 (110kb)
BAC-427G3 (150kb)
BAC-490B10 (160kb)
BAC-664D6 (105kb)
BAC-234A8 (145kb)
BAC-194F7 (160kb)
BAC-400H12 (140kb)
BAC-474A2 (170kb)
BAC-984H11 (105kb)
BAC-329G7 (170kb)
BAC-716E7 (105kb)
BAC-547D8 (210kb)
BAC-261B3 (150kb)
BAC-207G8 (150kb)
947G3-3'
190A11-5'
190A11-3'
499E6-3'
0
200
400
600
800kb
1Mb
SSC 7p1.1
SLA class III region
SLA class I region

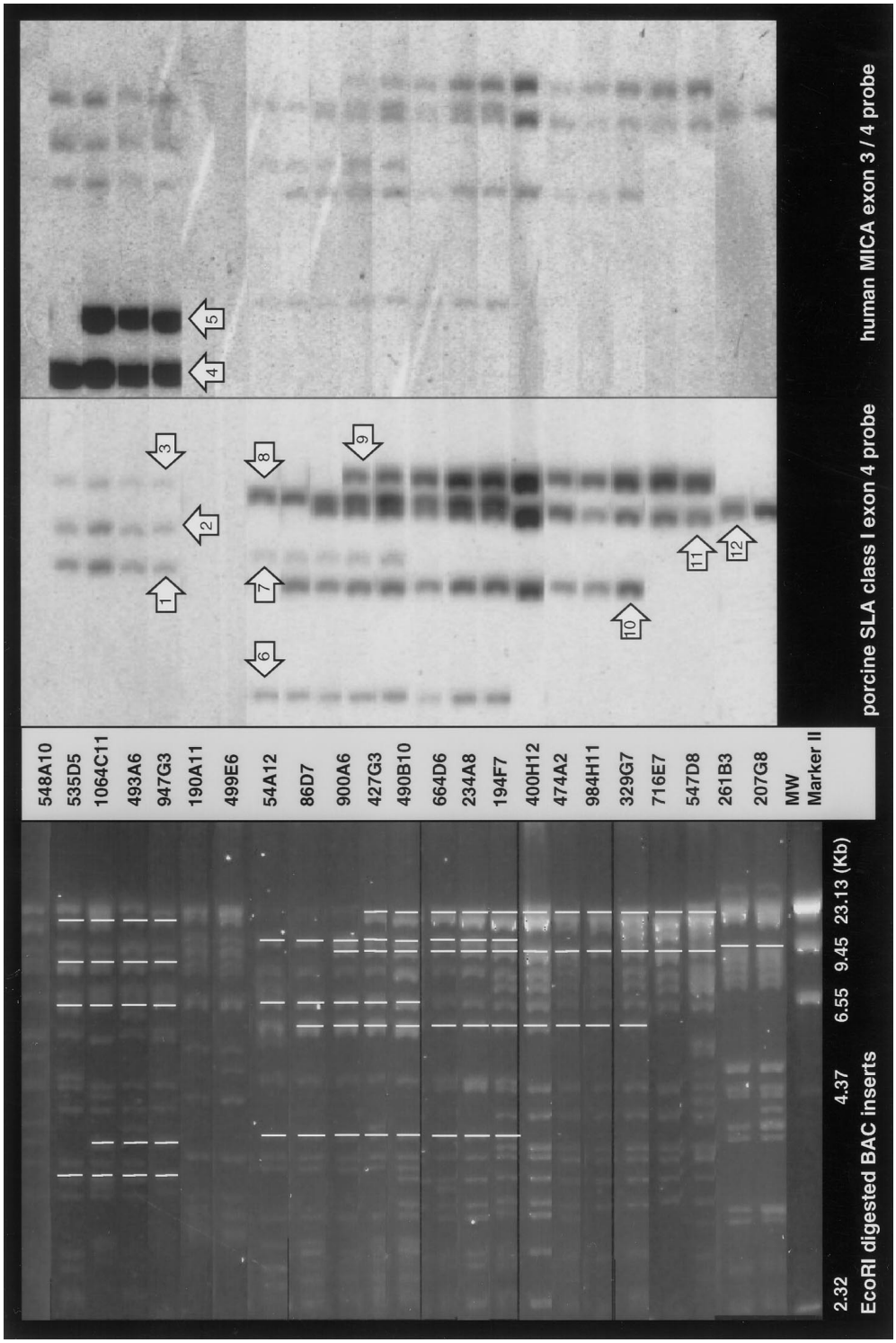


Fig. 2 Southern blot of *SLA* class I gene fragments containing BACs digested with *Eco* RI. Hybridization was carried out with an *SLA* class I gene exon 4 probe (*middle panel*) and an HLA *MICA* gene exon 4 probe (*top panel*). The location of the *SLA* class I-containing bands are *underlined* on the ethidium bromide-colored gel (*bottom panel*). *Arrows* number: *1*: *SLA-7, 2*: *SLA-6, 3*: *SLA-8, 4*: swine *MIC-1, 5*: swine *MIC-2, 6*: *SLA-4, 7*: *SLA-11, 8*: *SLA-2, 9*: *SLA-1* and *SLA-10, 10*: *SLA-3, 11*: *SLA-9, 12*: uncharacterized class I-related gene

quence underwent total digestion by the restriction enzyme *Eco* RI. The electrophoretically separated fragments were blotted onto nylon membrane and hybridized with a radiolabeled 0.8 kb probe corresponding to the conserved exon 3 and 4 regions of *SLA* class I genes. In all, 10 bands were revealed, ranging in size from 4.6 to about 17 kb (Fig. 2, middle panel, nrs. 1–3 and 6 to 12). Each clone displayed a characteristic reproducible pattern of one to six bands of variable intensity. Two-by-two band pattern comparisons permitted the ordering of all the BACs in the central contig, except for the two most distal BACs 261B3 and 207G8 (Fig. 2, middle panel). These two BAC clones displayed the same single class I band not shared by any other clone of the contig, and the *ZNF173* and *RFB30* anchor genes. As shown in Fig. 1, these two clones overlapped

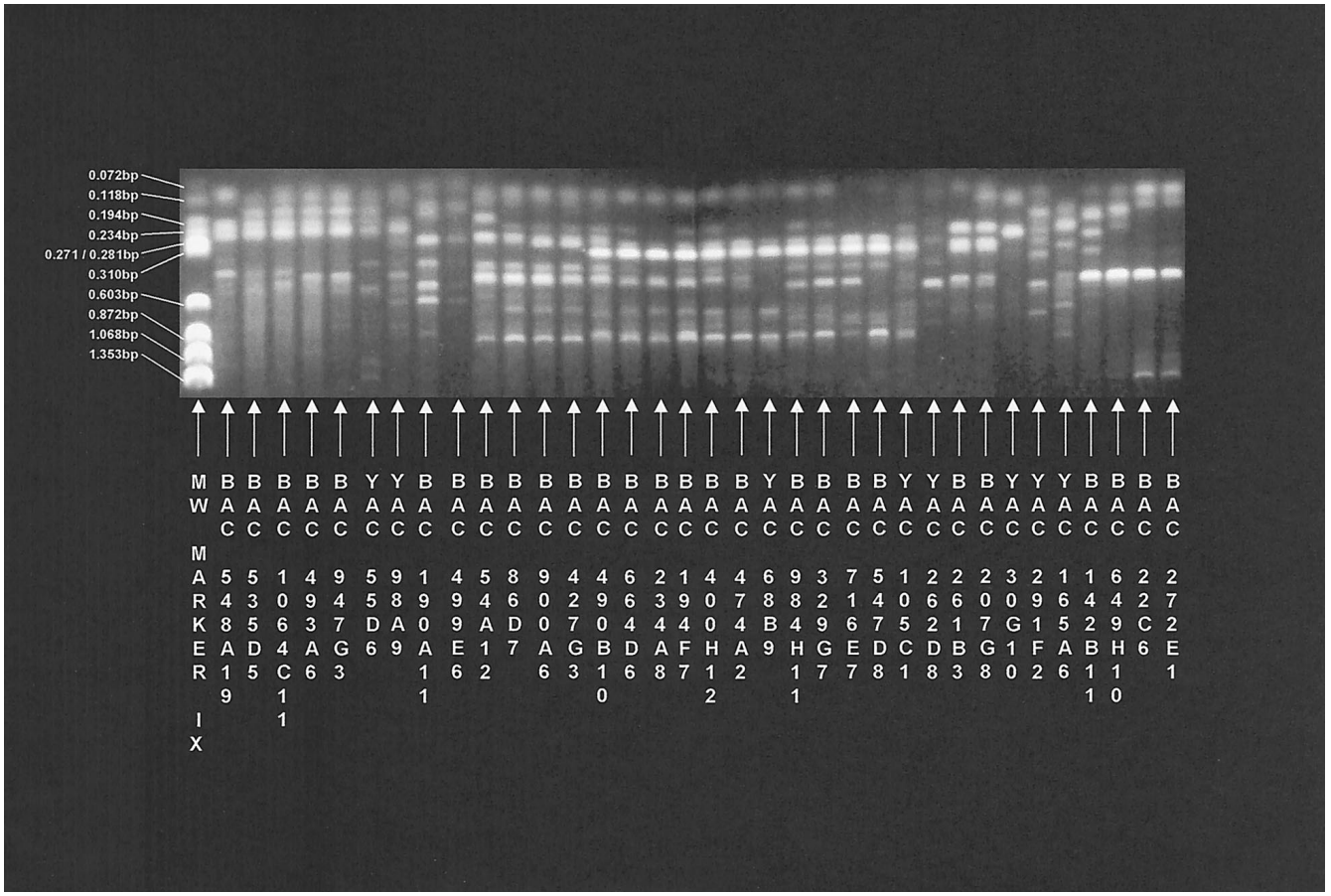


Fig. 3 Alignment of *SLA* class I gene inserts containing BAC and YAC clones by PIRS-PCR (pig interspersed repetitive sequence PCR)

with central contig BACs in the *ZNF173* region. Neither of the two non-class-I-containing BACs revealed any bands.

PIRS-PCR: The BAC order obtained with the class I Southern blot band pattern was confirmed independently, using the pig inter SINE fingerprinting pattern (Fig. 3). Note that the SINE band patterns of the above-mentioned BAC clones 261B3 and 207G8 were very similar, but only exhibited one or two bands common to other clones of the central contig.

Allocation of class I sequences within the BACs

After the class I bands had been identified on Southern blots, the corresponding gel bands visualized by ethidium bromide were recovered (Fig. 2, bottom panel). Each band was cut out of the gel from the most centromeric and telomeric clones of the BAC subset of interest, and then amplified separately. For example, the band labeled 6 was recovered from clones 54A12 and 194F7 (Fig. 2, middle panel). The purified DNA was then subjected to PCR amplification using all available class I primer pairs, i.e., C1, C4, and C6 to C12. Sequence analysis was performed on four independently subcloned PCR products for each band amplified. Amplification products and their sequences were readily obtained for bands 1, 2, 6, 8, 9, and 11, but not for the DNA material from bands 3, 7, and 12. (Fig. 2, middle panel). Comparison of these sequences to known *SLA* class I sequences permitted the assignment of bands 1, 2, 6, 8, 9, and 11 to specific class I genes including both classical and class I-related genes (Fig. 4).

SLA classical class I genes

Bands 8, 10, and 1 each generated single amplification products. Thus, band 8 was shown to correspond to the *SLA-2* gene, band 10, to the *SLA-3* gene, band 11, to the *SLA-9*, and band 6, to the *SLA-4* pseudogene. In contrast, the DNA material recovered from band 9 generated at least two amplification products, respectively, corresponding to the *SLA-1* and *SLA-5* genes. As mentioned above, no amplification products were obtained for bands 3, 7, or 12.

SLA class I-related genes

The amplification products obtained for bands 1 and 2, respectively, corresponded to class I-related genes *SLA-7* and *SLA-6.*

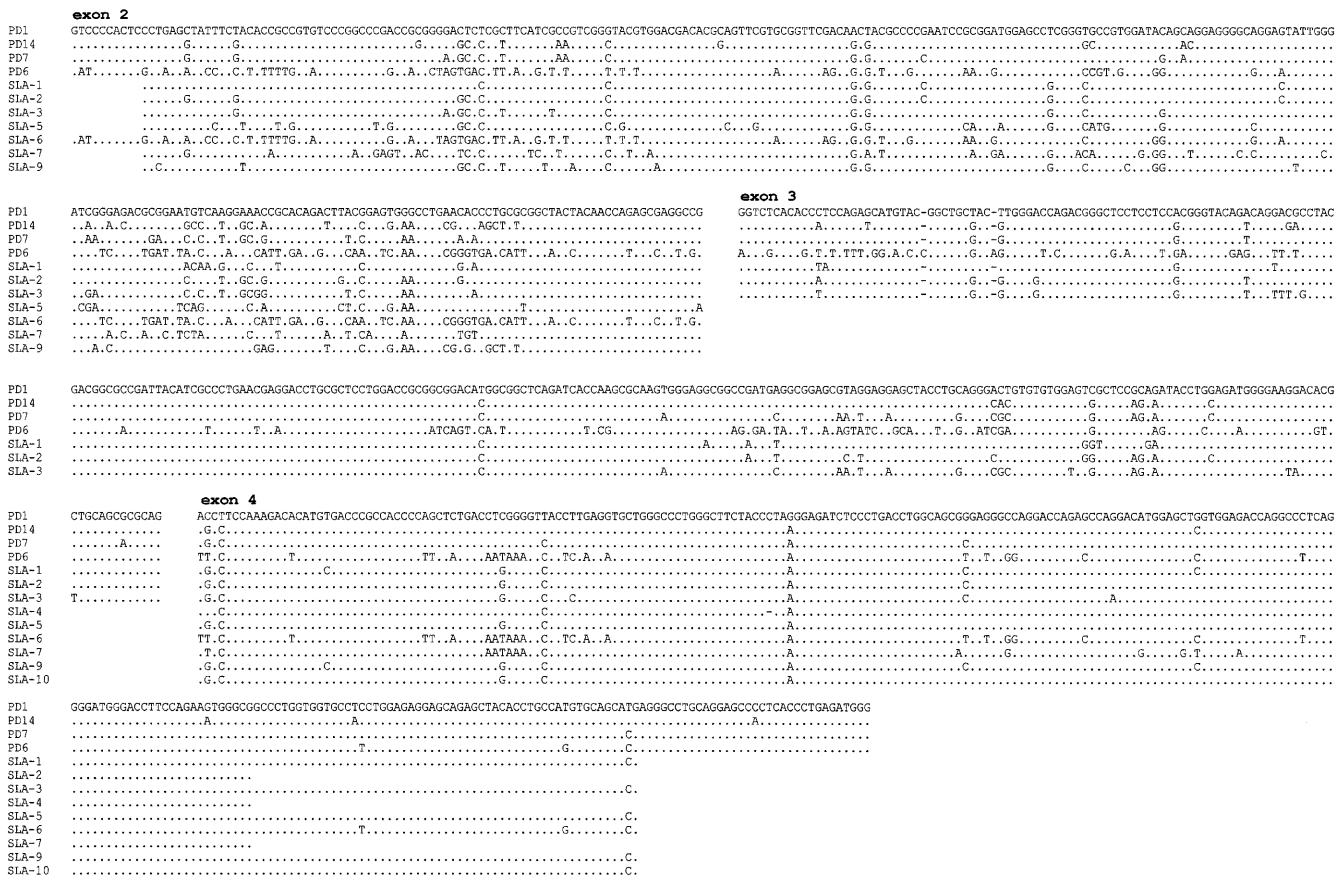


Fig. 4 Comparison of the *SLA* class 1 sequences obtained with the NIH minipig class I-related genes (identities among the sequences are indicated by *points* and gaps are shown by *dashes*)

The pig distantly homologous class I genes MIC-1 and *MIC-2*

The existence of *MIC* sequences in pig was previously suggested in zooblot experiments (Bahram et al. 1994). To find out whether such sequences are present in the pig *Mhc* we designed several primer pairs from the human *MIC-A* and *MIC-B* genes, but these primers did not permit any amplification products to be obtained. On the other hand, a human 283 bp human *MICA* probe, hybridized at low stringency with the *Eco* RI-digested *SLA* class I-containing BACs, revealed two intense bands at about 4.4 and 4.0 kb in the BACs harboring class III and class I-related genes (Fig. 2, top panel). This probe also cross-hybridized with all class I and class I-related sequences. The *MIC* 4.4 kb band (Fig. 2, band 5) was present within three BACs, namely BACs 947G3, 493A6, and 1064C11, while the 4.0 kb band (band 4) also appeared within a fourth overlapping BAC, number 535D5. Sequence analysis was performed after cutting out the *MIC-1*- and *2*-containing bands from the gel, partial digestion of these bands with the *Sau* 3A enzyme, and their cloning. The sequence alignments of 12 subclones for each band revealed a 484 bp segment within band 4 designated *MIC-1* (GenBank accession number AF083661), and a 375 bp segment within band 5, designated *MIC-2* (GenBank accession number AF083662), both of which displayed about 70% similarity to exon 2 of the human *MIC-A* and *MIC-B* genes. The similarity between the part of the pig *MIC-1* and *MIC-2* sequences corresponding to the exon 2 region of the human *MIC* genes was about 80%. The *MIC-1* gene was seen to be telomeric to the class I-related gene cluster, but the location of the *MIC-2* gene was not established.

Transcription study

After the physical *SLA* class I gene alignment, we attempted to identify the genes actually transcribed. For this purpose, we cultured fibroblasts from the SLA H01 homozygous male used to construct the BAC library and part of the YAC library. The cDNA was obtained after reverse transcription of the mRNA recovered from the fibroblasts. The presence of the *SLA-2, -4,* and *-10* genes was explored using the C1 primer couples which amplified a 235 bp fragment of the *SLA* class I exon 3 and 4 regions. The C10 primers were designed to amplify a 264 bp fragment of the *SLA-6* exon 2 region. The design of the primers of the C13 combinations (Table 1) was based on the exon 2 and exon 4 regions of genomic class I sequences of the *H01* haplo-

type. These primers allowed the amplification of a 708 bp cDNA fragment of transcribed class I genes.

Analysis of the sequence obtained revealed that the genes transcribed in fibroblasts are the*SLA-1, -2,* and -3 genes, a result with the fact that the *SLA H01* haplotype is serologically characterized by three class 1 series. Among the *SLA* class I-related genes, transcripts for the *SLA-7* but not for the *SLA-6* gene were readily identified. Transcription of the *SLA-7* gene was not observed in lymphocytes carrying an *H01* haplotype of an unrelated pig.

Discussion

The advent of large insert libraries of complex genomes, especially genomes of mammals, has permitted the construction of high-resolution physical maps spanning large chromosomal segments. Thus, using a pig YAC library, we recently reported the first physical map of the pig *SLA* class-I region ever established. Five overlapping YACs containing all the detectable class-I gene sequences and numerous non-class-I evolutionarily conserved genes constituted a contig of about 1.0 to 1.2-Mb. The most distal genes identified in this contig were the olfactory receptor-like genes *OLF42-1, -2,* and *-3*. Three *OLF89-1,-2* sequences-containing YACs were also mapped by fluorescent *in situ* hybridization (FISH) close to the *SLA* region but did not overlap with the contig. One important achievement of this recent study has been to show that most *SLA* class I genes are tightly clustered in two distinct subregions of about 100 kb each. It is noteworthy that one of these clusters comprised typical class I genes, while the other only contained class I-related genes. However, because of the small number of overlapping YACs of the *SLA* class I region, and of the risks of undetected rearrangements, deletions, or chimerism within these YACs, it was of paramount importance to analyze this region in greater detail by complementary means. In this respect, the screening of our recently produced pig BAC library proved rewarding. In all, its screening for several genes, including class III genes known to be contiguous to the class I region, for genes located within the class I region and for BAC end-rescued sequences led to the recovery of 35 BAC clones. In general, the results reported here confirmed the overall *SLA* class I region organization established previously. In particular, they not only confirmed the existence of two main clusters of class I genes, but also provided more precise insight into the order of the class I sequences within the clusters. We also uncovered a putative new class I divergent gene that had escaped detection in our previous study. Although it is located telomeric to the pig *ZNF173* gene, its precise mapping is not known. In mice, an *H2*-related gene, *H2-M5*, was characterized between the *ZNF173* and *MOG* genes. In humans too, the *HLA-A* locus is located between these two evolutionarily conserved genes. The pig putative new class I gene might be closer to *H2-M5* than to a classical class I locus, as it could not be amplified by any of the class I primer pairs tested. This study also demonstrated the existence within the *SLA* class I region of several new genes, namely *MIC-1*, *MIC-2, p52*, and *ZNF173*. As shown in Fig. 1, the 35 overlapping BAC clones constituted three distinct clusters. The size of the two most centromeric clusters was about 350 kb, and that of the third distal cluster, about 170 kb. The sizes of the two corresponding gaps were not evaluated precisely, but on the basis of the available *SLA* YAC contig, they might both span 200 kb. Studies are underway to fill these gaps by further screening of the library, with both new expressed sequence tags and BAC end-rescued sequences.

The most centromeric BAC contig contained the *TNFA* and *BAT1* class III genes, the cluster of the three class I-related *SLA-6, 7* and *-8* sequences, the two newly discovered pig *MIC*-like genes, the previously described *SC1, POU5F1*, and *S* genes, and the new pig counterpart of the human *p52* gene. Although both pig *MIC* genes are located in the same BAC clone, their spatial relationship, both to each other and to the other genes in the BACs, has not yet been established. What is known is that *MIC-1* lies telomeric to the *SLA-6, -7,* and *-8* sequences. The parts of the pig *MIC* gene sequences so far identified exhibited about 70% similarity to the human *MIC* genes, and less than 20% to the *SLA* classical class I or class I-related genes. When the human *MIC-A* exon 4 probe was hybridized to Southern blots of pig BAC clones, it exhibited strong signals with the pig *MIC* genes. However, at low hybridization stringency it also gave weak but specific signals with all other pig class I sequences. In humans, five *MIC*-related sequences have been detected in the HLA complex (Bahram et al. 1994). In pig, there seem to be only two *MIC* genes, and further study will be necessary to determine whether one or both of these genes are functional. Similar genes have been detected in all primate species tested, in ruminants such the goat and cow, and in dog and hamster, but seem to have been lost in mice (Bahram et al. 1994). In humans, the *MIC-A* and *MIC-B* genes give rise to protein chains which fold like a normal class I chain. They do not bind to β_2-microglobulin, and it has not yet been established whether they present peptide or another moiety. The human MIC-A and MIC-B chains are to some extent expressed by fibroblasts and keratinocytes, but mainly by intestinal epithelial cells, where they predominantly interact with T cells expressing the $\gamma\delta$ receptors (Groh et al. 1998). Although not upregulated by γ-interferon, the expression of human MIC chains was greatly increased by heat shock. Accordingly, it has been postulated that both MIC and $\gamma\delta$ T cells have a role in self-antigen recognition, and may fulfil an immune surveillance function for the maintenance of epithelial homeostasis. Considering the numerous and often life-threatening intestinal syndromes in pig, especially piglets, a thorough analysis of pig *MIC* genes and their products is highly desirable. Moreover, further analysis of the ex-

pression pattern of pig *MIC* genes and their interaction with human T and NK cells seems to be indispensable before any xenotransplantation.

In the first *SLA* class I physical map, the results of Southern blots of YACs 55D6 and 98A9 digested with *Eco* RI, and hybridized with an *SLA* class I exon 3–4 probe, revealed, as well as the bands corresponding to the *SLA-6* and *SLA-7* class I-related gene at 7.2 and 8.2 kb, a third band of about 9.8 kb, designated *SLA-8*. Application of a similar procedure to the BAC clones of the most centromeric contig described here confirmed the existence of three bands at 7.2, 8.2, and 17 kb, respectively. The only discrepancy observed concerned the putative *SLA-8* sequence, which was associated with a 17 kb band in BACs but with a 9.8 kb band in YACs. Checking of the YAC data revealed that the 9.8 kb band in YACs was due to typing error.

As already mentioned, we mapped the *p52* gene near and telomeric to the *POU5F1-S* gene cluster. This gene was identified by systematic sequence comparison to the database of a BAC rescued-end product of 400 bp. It actually corresponded to the third exon, with 90% similarity to the recently described human *p52* gene (Marinoni et al. 1997), which was characterized as the fifth subunit of the core of the transcription and DNA repair factor TFIIH. In humans, the *p52* gene has been mapped at 6p21.3, i.e., in the same band as the HLA complex. On the basis of the results for pig, we predict that in humans too, the *p52* gene belongs to the class I region and lies between the *POU5F1* and *S* genes. The presence within the pig Mhc of a *p52* gene involved in DNA repair might be relevant to the predisposition of some herds of pig to develop the malignant melanoma syndrome; The more so as the SLA complex has been shown to be involved in the outcome of this syndrome (Tissot et al. 1993).

The central contig consisted of as many as 16 overlapping BACs, thus enabling us to clarify the order of all but the typical *SLA-4* and *SLA-2* genes, as they were always found together. Either these genes are very close, or, more likely, inserts carrying only one of them have not been cloned. Centromeric to these two genes lies the gene *SLA-11*. The remaining four class I genes *SLA-3, -9, -5,* and *-1* as well as a class I-related sequence lie, in this order, in a telomeric position in relation to the *SLA-2,* and *-4* genes. Thus, in contrast to what our previous results suggested, the *SLA-2, -4,* and *-11* genes are more centromeric than the other SLA classical class I genes. The order of *SLA-2* and *-4* in the

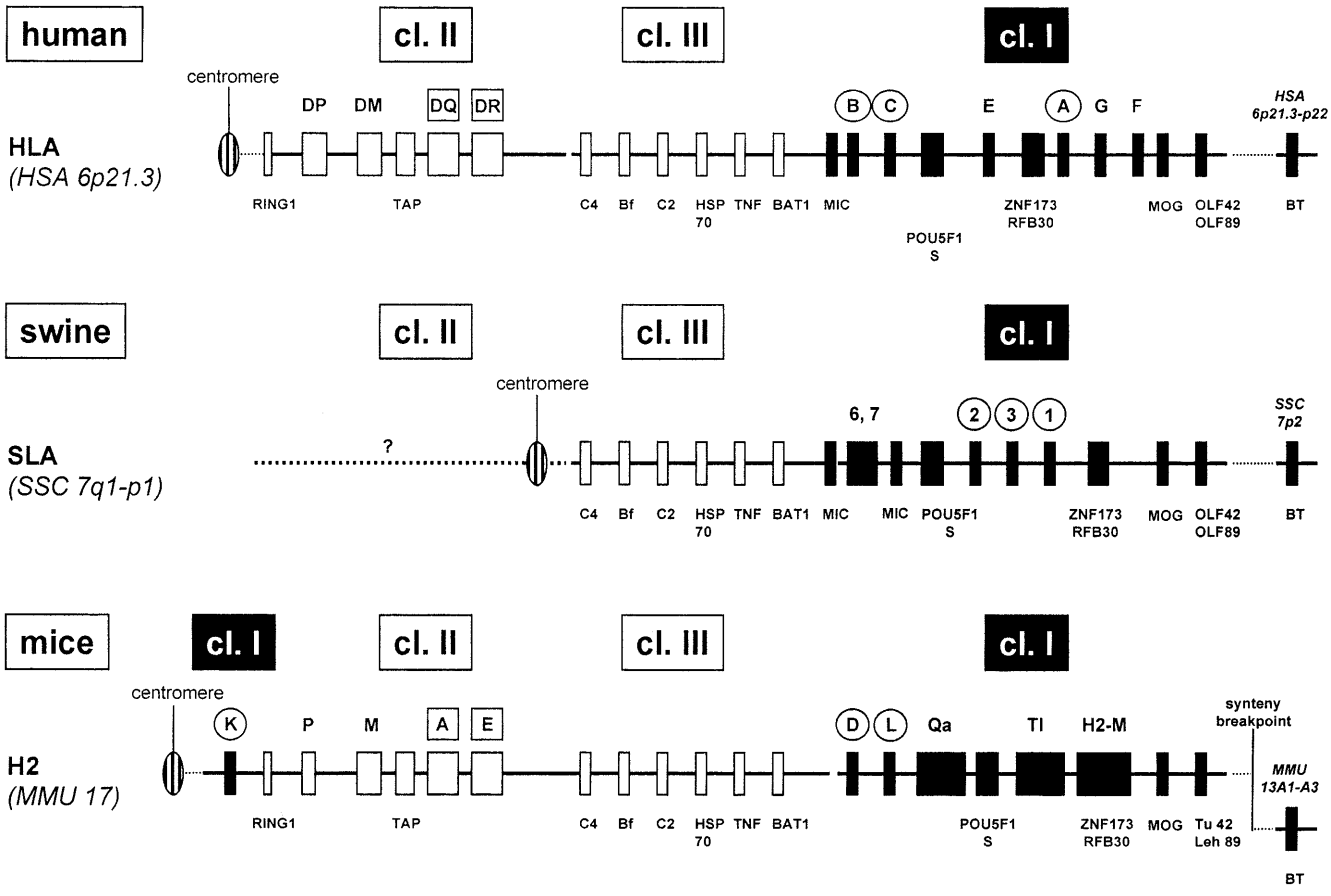


Fig. 5 Comparison of the pig, human, and mouse *Mhc* complexes. ▭ class II genes; □ class III genes; ■ class I genes; *HSA*: *homo sapiens* chromosome; *SSC*: *sus scrofa* chromosome, *MMU*: *mus musculus* chromosome; classical class I genes and antigen-presenting class II genes are *encircled*

contig will be obtained from the sequencing of an appropriate BAC.

Finally, at the distal end of the class I contig, we confirmed the presence of the *RFB30* gene and identified the *ZNF173* gene, as well as the putative class I-divergent gene referred to above. The respective locations of the *RFB30* and *ZNF173* genes in the contig fully agree with what is found in humans and mice (Fig. 5). Altogether, 12 typical and divergent class I sequences have been defined in the pig *H01* haplotype. It remains to be seen whether all other *SLA* haplotypes are similarly organized.

In our previous pig class I YAC contig, the most telomeric YAC clone, 262D8, comprised the gene *RFB30* and two olfactory receptor-like genes, *OLF42-1, -2*. Three other YACs were found to contain two genes of the *OLF89* family, which, however, were not linked physically to the class I YAC contig. The pig *OLF89* sequences exhibited a great similarity to the *Mhc*-linked *OLF89* sequences in humans and mice. Moreover, FISH experiments with all three YACs showed strong signals in SSC 7p1.1, as well as additional signals on some other chromosomes. In the present study, we identified four overlapping BACs found to contain all the *OLF42* and *89* sequences, as well as the most centromeric *MOG* anchor gene. Although these four BACs constitute a separate contig, they actually filled the gap between the *OLF42* and *OLF89* genes of the YAC contig, and therefore confirmed that the *OLF89* genes also belong to the *SLA* class I region.

Analysis of the class I transcription products of the cultured *SLA* H01 boar fibroblast used to build the YAC and BAC libraries demonstrated that three classical class I genes, *SLA-1, -2,* and *-3,* were transcribed. These results were expected, since all serological and biochemical studies in pig homozygous for the *H01* haplotype have always led to the characterization of three distinct *SLA* class I molecules. Less expected, however, was the presence, in these fibroblasts, of apparently abundant transcripts specific to the class I-related *SLA-7* gene. Similar SLA-7 transcripts were not recovered in lymphocytes from pig of another origin, indicating that the *SLA-7* gene probably has a physiological role which differs from that of the second well-defined class I-related *SLA-6* gene, the equivalent of the NIH minipig *PD6* gene. The latter has been shown to be highly transcribed in peripheral blood lymphocytes, especially by cells of the T lineage. Interestingly, the *SLA-6* gene and the human nonclassical class I *HLA-E* gene have common features. Thus, both are mainly transcribed in quiescent T cells and exhibit, within their fifth intron, a 300 bp insert that is absent from the other classical and nonclassical class I genes. The molecular structure of the *SLA-7* gene remains to be established. There is also the need to ascertain whether, like *SLA-6, SLA-7* is essentially monomorphic or if it displays some polymorphism. The putative nonclassical class I *SLA-8* gene also has to be fully characterized. As the pig is indeed a potential donor of tissues and organs to humans, it seems very important to define the expression pattern of all class I genes, including divergent genes. From the results of our present study, one cannot in fact exclude the possibility that class I genes other than the *SLA-1, -2,* and *-3* were transcribed, as low transcription levels may well have escaped detection.

In conclusion, the refined class I physical map presented here is a further step towards a comprehensive view of the pig *Mhc*. Moreover, the presence of all *SLA* class I genes of interest in individual BAC clones has permitted the start of complete sequencing of the two regions harboring the typical class I and class I-divergent genes.

Acknowledgments The excellent technical assistance of N. Borgeaux, J.C. Save, and P. Laffont was much appreciated. This work was partly funded by the PIGMaP II project from EC Biotech and supported by a Postdoctoral Fellowship (Ve 191/1-1) awarded by the Deutsche Forschungsgemeinschaft to F. Velten.

References

Bahram S, Bresnahan M, Geraghty DE, Spies T (1994) A second lineage of mammalian major histocompatibility complex class I genes. Proc Natl Acad Sci USA 91:6259–6263

Bahram S, Mizuki N, Inoko H, Spies T (1996) Nucleotide sequence of the human MHC class I *MICA* gene. Immunogenetics 44:80–81

Chu TW, Capossela A, Coleman R, Goei VL, Nallur G, Gruen JR (1995) Cloning of a new 'finger' protein gene (ZNF173) within the class I region of the human MHC. Genomics 29:229–239

Drews RT, Coffee BW, Prestwood AK, McGraw RA (1990) Gene sequence of porcine tumor necrosis factor alpha. Nucleic Acids Res 18:5564

Frels WI, Bordallo C, Golding H, Rosenberg A, Rudikoff S, Singer DS (1990) Expression of a class I MHC transgene: regulation by a tissue-specific negative regulatory DNA sequence element. New Biol 2:1024–1033

Geffrotin C, Popescu CP, Cribiu EP, Boscher J, Renard C, Chardon P, Vaiman M (1984): Assignment of MHC in swine to Chromosome 7 by in situ hybridization and serological typing. Ann Genet 27:213–219

Goad WB, Kanehisa MI (1982) Pattern recognition in nucleic acid sequences. I. A general method for finding local homologies and symmetries. Nucleic Acids Res 10:247–263

Groh V, Steinle A, Bauer S, Spies T (1998) Recognition of stress-induced MHC molecules by intestinal epithelial gammadelta T cells. Science 279:1737–1740

Marinoni JC, Roy R, Vermeulen W, Miniou P, Lutz Y, Weeda G, Seroz T, Gomez DM, Hoeijmakers JH, Egly JM (1997) Cloning and characterization of p52, the fifth subunit of the core of the transcription/DNA repair factor TFIIH. EMBO J 16:1093–1102

Pearson WR, Lipman DJ: Improved tools for biological sequence comparison (1988) Proc Natl Acad Sci USA 85:2444–2448

Peelman LJ, Chardon P, Vaiman M, Mattheeuws M, Van Zeveren A, Van de Weghe A, Bouquet Y, Campbell RD (1996) A detailed physical map of the porcine major histocompatibility complex (MHC) class III region: comparison with human and mouse MHC class III regions. Mamm Genome 7:363–367

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Shishido S, Naziruddin B, Howard T, Mohanakumar T (1997) Recognition of porcine major histocompatibility complex class I antigens by human $CD8^+$ cytolytic T cell clones. Transplantation 64:340–346

Smith TP, Rohrer GA, Alexander LJ, Troyer DL, Kirby-Dobbels KR, Janzen MA, Cornwell DL, Louis CF, Schook LB, Beattie CW (1995) Directed integration of the physical and genetic linkage maps of swine chromosome 7 reveals that the SLA spans the centromere. Genome Res 5:259–271

Sullivan JA, Oettinger HF, Sachs DH, Edge AS (1997) Analysis of polymorphism in porcine MHC class I genes: alterations in signals recognized by human cytotoxic lymphocytes. J Immunol 159:2318–2326

Takeda J, Seino S, Bell GI (1992) Human Oct3 gene family: cDNA sequences, alternative splicing, gene organization, chromosomal location, and expression at low levels in adult tissues. Nucleic Acids Res 20:4613–4620

Tissot RG, Beattie CW, Amoss MS Jr, Williams JD, Schumacher J (1993) Common swine leucocyte antigen (SLA) haplotypes in NIH and Sinclair miniature swine have similar effects on the expression of an inherited melanoma. Anim Gen 24:191–193

Velten F, Rogel-Gaillard C, Renard C, Pontarotti P, Tazi-Ahnini R, Vaiman M, Chardon P (1998) A first map of the porcine major histocompatibility complex class I region. Tissue Antigens 51:183–194

Yerle M, Echard G, Robic A, Mairal A, Dubut-Fontana C, Riquet J, Pinton P, Milan D, Lahbib-Mansais Y, Gellin J (1996) A somatic cell hybrid panel for pig regional gene mapping characterized by molecular cytogenetics. Cytogenet Cell Genet 73:194–202