LETTER

CpG Islands of the Pig

Heather A. McQueen,[1,4] Victoria H. Clark,[1] Adrian P. Bird,[1] Martine Yerle,[2] and Alan L. Archibald[3]

[1]Institute of Cell and Molecular Biology, University of Edinburgh, Edinburgh, EH9 3JR, Scotland; [2]Institut National de la Recherche Agronomique Cytogenetics Laboratory, Toulouse, France; [3]Roslin Institute (Edinburgh), Midlothian, EH25 9PS, Scotland

We describe an analysis of the CpG islands (CGIs) of the pig. We have used both database survey and a porcine genomic library that is enriched for CGIs. Approximately half of 41 pig genomic database sequences had CGIs with an average G + C content of 65.3%, an average CpG observed/expected frequency of 0.85, and an average size of 978 bp. Of 27 CGI library clones, 16 were nonrepetitive, nonribosomal DNA and CGI-like. CGI library clones had similar average values for G + C and CpG frequency to CGIs of database genes, and an average size of 670 bp, as *Mse*I cuts within some islands. Library clones were also shown to be low copy number and unmethylated in genomic DNA. The presence in the library of seven previously known CGI sequences was confirmed as was the absence of one nonisland sequence. The CGI library exhibits an R-band pattern for many chromosomes in FISH analysis. The pig chromosome arms that show the most dense CGI population are homologous to segments of human chromosomes that are known to be gene rich.

[The sequence data described in this paper have been submitted to the EMBL data library under accession nos. Y10585–Y10592, Y10594, Y10596–Y10597.]

CpG islands (CGIs) are short, unmethylated CpG-rich sequences situated at the 5′ end of many vertebrate genes (Fig. 1; Antequera and Bird 1993). CGIs are associated with 60% of human genes, including all housekeeping genes, and 40% of tissue-specific genes. Thus, CGIs can be regarded as useful markers for identifying genes (Larsen et al. 1992; Cross and Bird 1995).

Increasing the gene content of the genome maps is critical to current strategies for isolating economically important trait genes in livestock species where large-scale genomic sequencing is not currently economically feasible. Rather, following genome scans with microsatellite markers for economic trait loci, a positional candidate gene approach (Collins 1995) will be adopted to identify trait genes. However, the majority of markers on published genetic (linkage) maps for the domestic pig are anonymous microsatellites (Archibald et al. 1995; Marklund et al. 1996; Rohrer et al. 1996). For example, of 239 markers on the PiGMaP linkage consortium maps, only 81 represent known genes (Archibald et al. 1995). In the revised versions of these maps currently being prepared for publication, ~200 of the 650 mapped markers are associated with known genes (A.L. Archibald, pers. comm.).

The lack of genes mapped in livestock species can be compensated for in part by exploiting knowledge of the extensive conservation of synteny between, for example, humans and pigs (Rettenberger et al. 1995; Goureau et al. 1996). Nevertheless there is a need to increase the gene content of livestock genome maps, if only to improve the accuracy of alignment with the maps of humans and mice.

As CGIs are effective markers for identifying genes, cloned CGIs would be useful reagents for mapping genes. Furthermore, each CGI is present in genomic DNA in equimolar quantities unlike the variable relative abundance of transcripts represented by cDNA libraries. Thus, the isolation and cloning of CGIs would provide access to a large number of genes independent of developmental stage or tissue of expression. CGI libraries have been established for humans (Cross et al. 1994), mice (Cross et al. 1997), and chickens (McQueen et al. 1996) and have been used to examine the organization of genes and genomes (Cross and Bird 1995; McQueen et al. 1996; Cross et al. 1997). These and other studies (Craig and Bickmore 1994) have demonstrated that CGIs are not evenly distributed throughout the genome. Chickens represent the most extreme example of this nonrandom distribution with proportionately more CGIs found on micro- rather than macrochromosomes (McQueen et al. 1996).

[4]Corresponding author.
E-MAIL heather.mcqueen@ed.ac.uk; FAX 0131-650-5379.

We have developed a pig CGI library as a source of markers associated with genes which can be used to significantly increase the number of mapped pig genes. We describe the production and characterization of the pig CGI library and its use to confirm the predicted genomic organization of pig CGI sequences by fluorescence in situ hybridization (FISH).

RESULTS

Database Analyses

Pig genomic gene sequences were selected from the database and examined for the frequency of the dinucleotide CpG by a moving 100-bp window with a 3-bp shift. Of 41 genes, 20 contained regions where 70 consecutive windows sustained at least six CpGs and were considered likely CGI candidates (Matsuo et al. 1993). Three pig CGI genes are shown in Figure 1. The percent G + C and CpG observed/expected values were calculated for each CGI and are shown in Table 1, along with CGI position relative to coding regions. Eighteen CGIs contained at least part of

exon 1, one was within intron 1, and one was not located. The average G + C content for islands was 65.3%, the average CpG content was 0.85, and the average size was 978 bp.

CGI Library

A CGI library was prepared from pig spleen DNA using a methyl CpG binding column and methods described previously (Cross et al. 1994). Sequence was generated for 27 clones that were selected from the library at random. Of these 27, 20 clones (74%) were CGI-like (Fig. 2), 4 of which were found to constitute ribosomal DNA sequence. The 16 remaining sequences of CGI-like clones had an average G + C content of 63.6%, an average CpG observed/expected value of 0.79, and an average insert size of 607 bp. Accession numbers and insert sizes for eight fully sequenced CGI-like clones are as follows: CGI 21 (Y10591), 246 bp; CGI 35 (Y10585), 314 bp; CGI 37 (Y10586), 478 bp; CGI 39 (Y10596), 132 bp; CGI 40 (Y10587), 743 bp; CGI 41 (Y10588), 333 bp; CGI 43 (Y10589), 440 bp; CGI 45 (Y10590), 180 bp. Larger inserts from eight more CGI-like clones were
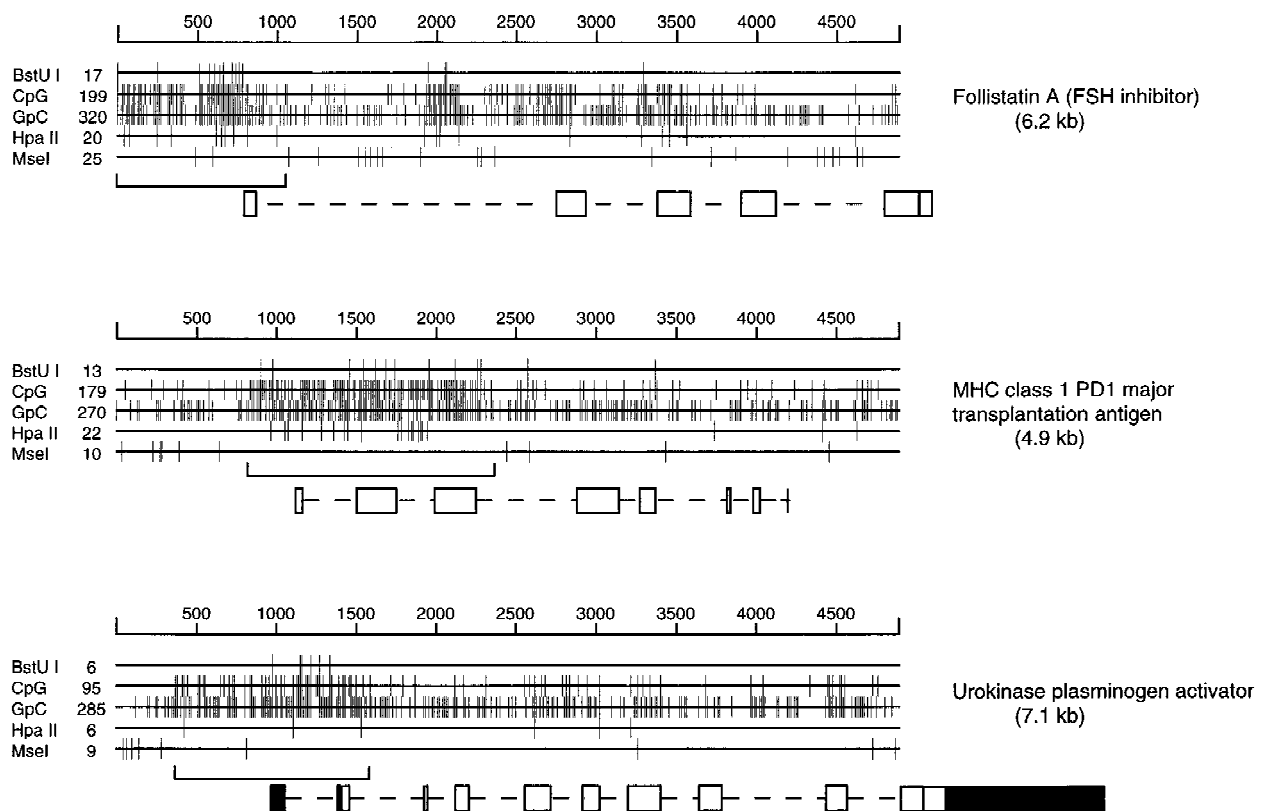

Figure 1 Three pig CpG island genes. The size (in bp) is shown above each genomic sequence, and the CGI is bracketed below. Exons are shown as solid boxes when untranslated. The positions of target sites are shown for two GC-rich cutters (BstUI and HpaII) and for MseI, whose target site is TTAA. The frequencies of CpG and GpC dinucleotides within the sequence are also shown.

Table 1. Pig Genes with and without CGIs

Gene[a]	Accession no.	CGI size (bp)	Percent G + C	CG/GC	Position[b]
MHC class 1 D haplotype	M59750	1605	67	1.08	
HSP	M69100	2365	63	0.91	u, e1
MHC class 1 PD1	M21057	1660	66	1.07	u, i3
MHC class 1 PD14	M21058	1410	67	1.12	u, i3
G protein	M57287	510	72	0.95	u, i1
Urokinase plasm. activ.	L27481	210	68	0.90	i1
C-type natriuretic pept.	M64758	1240	71	0.73	u, i2
L1P	D10616	791	65	0.7	u, i1
Nuclear factor 1	X12764	1025	67	0.94	u, i2
? (overlap compl. st. 21)	M83939	779	67	0.74	u, i3
FOLI (FSH inhib.)	M19529	1010	66	0.88	u, i1
Gprotα	L07317	429	67	1.07	u, e1
cAMP-dep. pk1 reg. sub.	X05943	508	71	0.81	u, i1
α2A adrenergic receptor	J05652	1728	66	0.80	all
c-mos	X78318	1080	65	0.67	u, e1
Cholin-acetyltransferase	L33836	1638	60	0.64	all
Aminopeptidase N	X16088	460	61	0.69	e1
Protamine 1	M80678	532	57	0.64	e1, i2
Protamine 2	M80676	376	70	0.78	u, i1
Follistatin B	M36512	200	51	0.82	e1
Apolipoprotein CIII	L00627	none			
TNFα	X54859	none			
TNFβ	X54859	none			
MHC SLA ster. 21 hydr.	M83939	none			
D-amino acid oxidase	M18444	none			
Relaxin	J02792	none			
IFN-γ	X53085	none			
Uteroferrin	M30283	none			
Slow myosin heavy chain	L10130	none			
Transition protein 1	M80679	none			
Transition protein 2	M80677	none			
Short type-1 IFN-rel. prot.	Z22708	none			
IFN-β-1	M86762	none			
IFN-α	M28623	none			
Insulin-like GF 1	X52388	none			
FSH β-subunit	D00621	none			
Muscarinic acetylch. rec.	X12712	none			
Leuteinizing hormone β-sub.	D00579	none			
Growth hormone	M17704	none			
Osteopontin	X16575	none			
Pituitary glyco. hormone β-sub.	D00766	none			

Of 41 unselected pig genes, 20 had CGIs. The size, percent G + C, frequency of CpG, and 5′ and 3′ boundaries of each island within the gene are shown.

[a](MHC) Major histocompatibility complex; (HSP) heat shock protein 70; (L1P) arachidonate 12-lipoxygenase [? (overlap compl. st. 21)] gene of unknown function found in the complementary strand (and overlapping) the steroid 21 hydroxylase gene; [FOLI (FSH inhib.)] follistatin with follicle-stimulating hormone inhibition; (Gprotα) G protein α-subunit; (cAMP-dep. pK1 reg. sub.) cAMP-dependent protein kinase 1 regulatory subunit; (TNFα, TNFβ) tumor necrosis factor α, β; (mHC SLA ster. 21 hydr.) MHC SLA locus steroid 21 hydroxylase; (IFN) interferon; (short type-1 IFN-rel. prot.) short type-1 interferon-related protein); (insulin-like GF 1) insulin-like growth factor 1; (muscarinic acetylch. rec.) muscarinic acetylcholine receptor; (pituitary glyco. hormone β-sub.) pituitary glycoprotein hormone β-subunit.
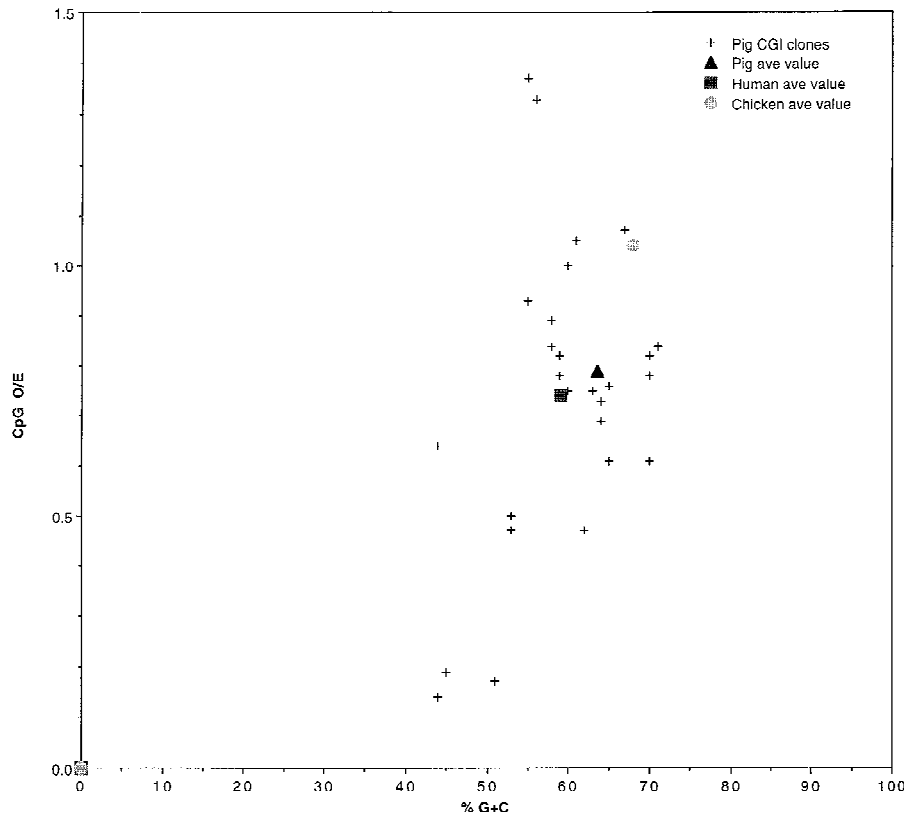
[b](u) Upstream; (e) exon; (i) intron.

Figure 2 A comparison of the 27 pig CGI library clones with average CGI library values for pig, human, and chicken. Clones are plotted with respect to percent G + C on the x-axis and the frequency of the dinucleotide CpG on the y-axis. The seven borderline and non-CGI pig clones (see text) are included. Average CGI library values for human and chicken are calculated from original data supplied by S. Cross and H. McQueen, respectively (pers. comm.).

mous CGIs isolated from a previous version of the pig CGI library (H. McQueen, unpubl.). All seven CGI sequences were found to be present in the library (Fig. 3) indicating that the library contains a high proportion of all pig CGIs. Primers that amplify a portion of interferon α with a single CpG were also tested. This sequence, which is representative of nonisland bulk genomic DNA, was absent from the library as expected.

CGI Distribution in the Porcine Genome

A biotin-labeled probe was made for the pooled library and used in competitive FISH experiments. Varying quantities of porcine competitor DNA were tested. With a 10-fold excess of competitor it was difficult to distinguish specific probe signal from the relatively high background. Increasing the competitor DNA to 1000-fold resulted in total extinction of specific signal on the chromosomes. The clearest results were obtained with 100-fold excess of competitor. At this level of competition, the probe produced an R-band pattern on chromosomes 2, 3p, 5p, 6, 7, 10, 12, 13, 14, and 17 (Fig. 4). R-banding on other chromosomes was not detectable. The Y chromosome was almost completely devoid of signal (Fig. 4). Strongly hybridizing repetitive sequences were seen on all of the acrocentric chromosomes and were not extinguished by 100-fold excess of competitor. The rDNA clusters that contain CGI-like sequence and lie on pig chromosomes 8 and 10 were clearly labeled by the probe as expected.

DISCUSSION

We describe the construction and characterization of a pig CGI library using a methylated DNA-binding column. By comparison to values from a set of pig database CGI genes, the average library values for GC content and CGI size are slightly reduced.
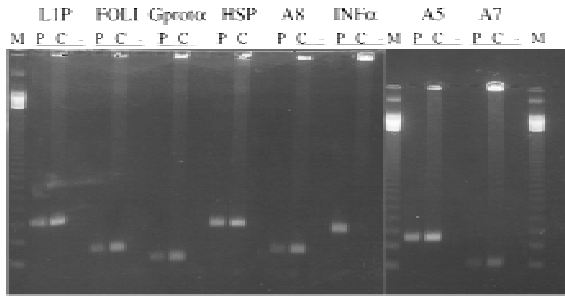
only partially sequenced from both ends. Three clones had GC contents that were borderline for CGIs but contained the pig repetitive element (PRE), whereas four were non-CGI-like (Fig. 2). A plate of 96 library clones was probed with radioactively labeled total genomic DNA, and six repetitive clones (6%) were identified. This indicates the level of repetitive sequence in the library to be low. In addition to their sequence characteristics, CGIs are expected to be unmethylated in the genome. To test this, the inserts of three CGI-like clones were radioactively labeled and used to probe Southern blots of genomic pig DNA digested with *Mse*I, *Mse*I, and *Hpa*II, which is methylation sensitive, and *Mse*I and *Msp*I, which cuts in a methylation-independent manner. All three tested clones were unmethylated in genomic DNA, two were single copy, and one was present in five copies (results not shown).

The CGI library was tested by PCR for the presence of four known gene CGIs and three anony-

Figure 3 The presence within the library of seven known CGIs was shown by PCR amplification from pooled library stocks (C). Product was also amplified from pig DNA (P) as a positive control, and a no-DNA negative control (−) was always included. Primers used were arachidonate 12-lipoxygenase, follistatin, G protein α subunit, heat shock protein 70, and three anonymous pig CGIs. Primers for a CpG poor region of interferon α were also used to show the absence of this DNA from the library. (M) Marker.

The library is made from *Mse*I-digested DNA, which occasionally cuts within the CGI or includes non-CGI flank in the fragment (Fig. 1). In this way the

CGI library gives a slight underestimation of the GC content and size of CGIs. Regardless of this, our two independent estimates both suggest pig CGIs to have similar average GC content to human CGIs (Matsuo et al. 1993; Cross et al. 1994) and lower GC content than chicken CGIs (Fig. 2; McQueen et al. 1996). The wide scatter of values for the pig (Fig. 2) may mean that pig CGIs are heterogeneous in character or may result from small sample bias. We found 20 of 41 pig genes (49%) to have CGIs. Although these numbers are small, this value is intermediate between previous estimates of 56% for human and 46% for mouse (Antequera and Bird 1993). It is interesting to note that at least four of the mouse homologs (urokinase plasminogen activator, c-mos, choline acetyltransferase, and protamine 2) and one human homolog (protamine 1) of pig CGI genes have no island.

FISH analysis with the CGI library probe revealed a strong signal on acrocentric centromeres (Fig. 4). This may be accounted for by a moderately GC-rich centromeric repeat sequence (X70941; J.R. Miller and P.D. Thomsen, unpubl.) that is specific to
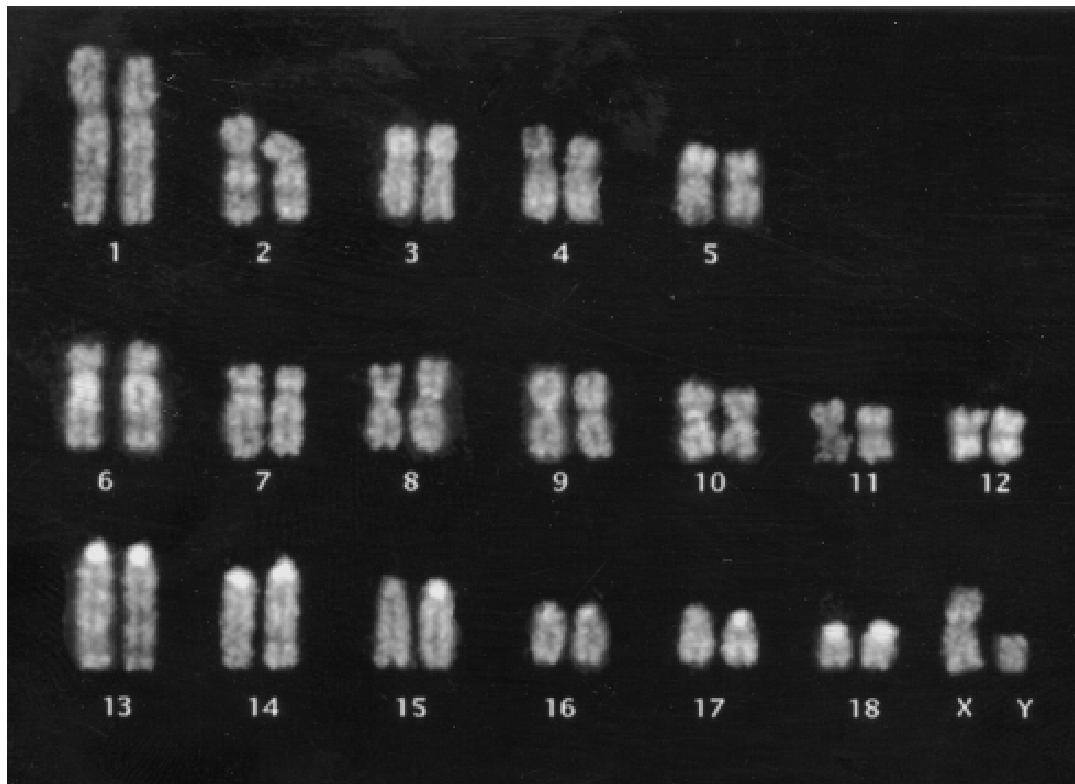

Figure 4 FISH of pig CGI library to pig metaphase chromosomes. The CGI probe lights up R-bands of most chromosomes as well as satellite sequences of the acrocentrics. Hybridization to the rDNA on chromosomes 8 and 10 was also seen. The probe was preannealed with 100-fold excess of pig Cot-1 DNA.

the acrocentrics. In a previous version of the pig CGI library we found a clone that showed 93% sequence identity to a portion of this repeat sequence. Because densely methylated DNA was removed before we cloned our GC-rich fraction, our results may suggest that this repeat can be unmethylated. Alternatively, as the centromeric sequence is only moderately GC-rich, its presence in the library may result from contamination.

FISH with the CGI probe also revealed a pattern that resembled R-banding of pig chromosomes (Fig. 4; Committee for the Standardized Karyotype of the Domestic Pig 1988). It was known previously that human chromosomes have regions of high island density within early replicating R-bands and that both human and chicken chromosomes are nonuniform with respect to CGI density (Craig and Bickmore 1994; McQueen et al. 1996). In our pig CGI paint we observed particularly bright signals corresponding to complete R-banding patterns on the long arms of chromosomes 2 and 6 (Fig. 4). Human chromosome 19, which is the most CGI-rich human chromosome (Craig and Bickmore 1994; Cross and Bird 1995), has been shown, by bidirectional heterologous painting or zoo FISH, to be homologous to two blocks of pig chromosomes on pig 2q11–q21 and 6q11–21 (Goureau et al. 1996). The strongly staining pig chromosome arms 2q and 6q also have extensive regions of homology to human 5q14-qter and 1pter-p31, respectively. Other pig chromosomes that hybridize strongly with the CGI probe include pig chromosomes 3p and 12, which are homologous to the CGI-rich human chromosome 16p and 17, respectively. Our findings therefore support the predictions made by combining human CGI data with human-to-pig heterologous painting data. Using this information, regions of CGI (and probably gene)-dense pig chromosome can be identified and selectively targeted for map refinement by the pig genome mapping community.

The distribution of CGIs within the mouse genome recently has been analyzed similarly (Cross et al. 1997). In the case of the mouse genome, a relatively homogeneous distribution of CGIs among chromosomes was interpreted as evidence of genome scrambling. In contrast, the notable R-band pattern detected with pig CGIs on 10 of the 18 chromosome pairs supports the idea of relatively few and large blocks of conserved synteny between pig and human (Rettenberger et al. 1995; Fronicke et al. 1996).

Finally, because nearly all database CGIs examined contained part of exon 1, the sequences of CGI library clones should allow identification of the associated genes. CGI clones also contained noncoding upstream and/or intron sequence that could be exploited to design genetic markers either via internal polymorphic repeats or by single-strand conformation polymorphisms. Thus, in combination with full-length cDNAs, the cloned CGIs provide access to many genes in small, readily characterized fragments suitable for linkage analysis and radiation hybrid mapping. The CGI library described therefore represents a uniquely favorable resource for the integration of transcribed sequences into genetic maps of the pig genome.

METHODS

Methyl-binding domain Column and Cloning

The methyl-CpG binding column was prepared and tested according to a method described previously (Cross et al. 1994). Briefly, 30 mg of methyl-binding domain (MBD) protein, purified from crude bacterial cell extract, was coupled with 1 ml of nickel NTA agarose (QIAGEN) and packed into an HR 5/5 column (Pharmacia). DNA was loaded onto the column, washed, and eluted in 20 mM HEPES (pH 7.9), 10% glycerol, 0.1% Triton X-100, 0.5 mM PMSF, and NaCl concen-

Table 2. Accession Numbers and Primer Sequences for Pig CGIs

Gene[a]	Accession no.	Primers
HSP70	M69100	5′-CCGCTTTGTACTTCTCAGCC-3′ 5′-ACCAAGCAGACGCAGATCTT-3′
FOLI	M19529	5′-TTCCCACTCCTCTCCATCC-3′ 5′-CCAGTTCATGGAGGACCG-3′
Gprotα	L07317	5′-CCAGCTCCTCACCAAGTAGC-3′ 5′-TTGAGCGCCGAAGACAAG-3′
INF-α	M28623	5′-TAGAACTGGTGCAGGAGGCT-3′ 5′-TACTCAGCTGCAATGCCATC-3′
L1P	D10616	5′-TTTGGGAACCTGCGTAGAAC-3′ 5′-GTGCCAGGTGTCAATTCCTT-3′
CGI A5	Y10594	5′-GAAGAGGCGGAGGAGGAG-3′ 5′-CTCACCGCTCTAGGTGAAGC-3′
CGI A7	Y10597	5′-GGCCTTGTGGACTACGATGT-3′ 5′-GAATTGAGACGGTGCAGTGA-3′
CGI A8	Y10592	5′-TTTTCGATGGTTCTTTTGTCG-3′ 5′-CGTCTTAGGTGTCGTGAAAGC-3′

[a]See Table 1 for abbreviations.

trations of up to 1 M. DNA (250 μg) isolated from the spleen of a large white/Meishan F_1 male 9591 from the Edinburgh PiGMaP reference pedigrees (Archibald et al. 1995), was digested to completion with *Mse*I and passed over the MBD column three times at 0.55 M NaCl to remove methylated fragments. Pig DNA thus stripped of heavily methylated sequences was methylated de novo with *Sss*I methylase (NEB) and loaded twice onto the MBD column at 0.55 M NaCl. Weakly bound fragments were eliminated by eluting at 0.7 M while strongly bound fragments were collected by eluting at 1.0 M NaCl. The strongly bound fraction was diluted back to 0.55 M and reloaded onto the column for a final elution gradient of 0.55 to 1.0 M NaCl. Fifteen fractions each of 2 ml were collected from the gradient and putative CpG island-enriched DNA was precipitated. The DNAs were spot dialyzed on VSWP filters (Millipore) and ligated to the vector pGEM 5zf at a 2:1 molar ratio. Vector DNA had been linearized previously at the *Nde*I site and treated with bacterial alkaline phosphatase. Ligations were desalted before electroporation into SURE *Escherichia coli* cells (Stratagene). After titration, 10^6 colonies were plated on six bioassay 20-cm^2 plates and scraped into Luria broth with 15% glycerol. The pooled library was stored at −80°C and is freely available from the Human Genome Mapping Project Resource Centre at Hinxton Hall, Cambridgeshire, UK.

Database, PCR, and Sequence Analysis

Candidate CGI sequences were selected from the EMBL database according to the criteria of a G + C content of >50% and an observed/overexpected CpG frequency of at least 0.6 using the GCG8 COMPOSITION program. CGI boundaries were defined, using the GCG8 WINDOWS program, as the center of any 100-bp window where the number of CpG residues fell below 6 and remained below 6 for adjacent windows shifted by 3 bp. Custom-made pig CGI primers (Imperial Cancer Research Fund DNA Synthesis Service) and accession numbers of the sequences from which they were derived are shown in Table 2. Conditions for all primer pairs were as follows: 100 ng of DNA template or 2 μl of library glycerol stock and 10 pmoles of each primer in the presence of 2.5 mM magnesium, 200 μM dNTP, 10% DMSO and 1 unit Pro-ZYME polymerase (Bioline, London, UK), with supplied buffer in a total volume of 50 μl. Cycling conditions were 95°C for 2 min, 30 cycles of 94°C for 45 sec, 56°C for 30 sec, and 72°C for 45 sec; and a final 10-min extension at 72°C. pGEM inserts were amplified using 50 pmoles of previously described pGEM primers, which flank the *Nde*I site (Cross et al. 1994), in 50-μl reactions as above. Inserts were amplified directly from picked colonies, 1 μl of glycerol stocks, or 50–250 ng of purifed clone DNA. Cycling conditions were as follows: 95°C for 5 min, 30 cycles of 95°C for 30 sec, 55°C for 1 min, 72°C for 3 min; and a final 10-min extension at 72°C. Plasmid DNA was prepared using Wizard columns (Promega), and sequenced with an Applied Biosystems 373A DNA sequencer. Sequences were analyzed using Gene Jockey (Biosoft, Cambridge, UK), MacVector (International Biotechnologies), Wisconsin, and BLAST (Altschul et al. 1990) programs for GC content and nucleotide and amino acid homology. Restriction digests, Southern blots, and hybridizations were carried out by standard procedures.

FISH

Pig chromosomes were prepared according to standard procedures. To identify the chromosomes, metaphase spreads were G-banded before hybridization and compared to the standard pig karyotype (Committee for the Standardized Karyotype of the Domestic Pig 1988). The CGI library probe was made by amplifying inserts from 30 ng of pGEM-ligated insert using pGEM primers in 50 μl as described above. Biotin 16–dUTP (Boehringer) was incorporated by a second round of amplification from 1 μl of product in a final volume of 25 μl with 1 unit of TAQ polymerase in supplied buffer (GIBCO), 0.2 mM dATP, dCTP, dGTP, 0.02 mM dTTP, and 2 μM of each pGEM primer. The reaction was denatured for 4 min at 94°C, followed by 15 cycles of 94°C for 1 min, 55°C for 1 min, 72°C for 1 min. Three independent reactions were pooled and purified through a G50 Sephadex column. The resultant probe was ethanol precipitated in the presence of salmon sperm DNA as a carrier and porcine Cot-1 competitor DNA at 0, 10-fold, 100-fold, or 1000-fold excess. The pellet was resuspended in 25 μl of hybridization mix and was denatured and preannealed for 3 hr at 37°C to block repetitive sequences. FISH and probe detection were performed as described previously (Yerle et al. 1993). Briefly, hybridization to pig metaphase chromosomes was carried out at 37°C for 48 hr. Biotinylated probes were detected with avidin–FITC followed by biotinylated anti-avidin and an additional layer of avidin–FITC. Slides were counterstained with propidium iodide and examined with a fluorescence microscope.

REFERENCES

Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215:** 403–410.

Antequera, F. and A. Bird. 1993. Number of CpG islands and genes in human and mouse. *Proc. Natl. Acad. Sci.* **90:** 11995–11999.

Archibald, A.L., C.S. Haley, J.F Brown, S. Couperwhite, H.A. McQueen, D. Nicholson, W. Coppieters, A. Van de Wegher, A. Stratil, A.K. Wintero et al. 1995. The PiGMaP consortium linkage map of the pig (sus scrofa). *Mamm. Genome* **6:** 157–175.

Collins, F.S. 1995. Positional cloning moves from perditional to traditional. *Nature Genet.* **9:** 347–350.

Committee for the Standardized Karyotype of the Domestic Pig. 1988. Standard karyotype of the domestic pig. *Hereditas* **109:** 151–157.

Craig, J.M. and W.A. Bickmore. 1994. The distribution of CpG islands in mammalian chromosomes. *Nature Genet.* **7:** 376–382.

Cross, S.H. and A.P. Bird. 1995. CpG islands and genes. *Curr. Opin. Genet. Dev.* **5:** 309–314.

Cross, S.H., J.A. Charlton, X. Nan, and A.P. Bird. 1994. Purification of CpG islands using a methylated DNA binding column. *Nature Genet.* **6:** 236–244.

Cross, S.H., M. Lee, V.H. Clark, J.M. Craig, A.P. Bird, and W.A. Bickmore. 1997. The chromosomal distribution of CpG islands in the mouse: Evidence for genome scrambling in the rodent lineage. *Genomics* **40:** 454–461.

Fronicke, L., B.P. Chowdhary, H. Scherthan, and I. Gustavsson. 1996. A comparative map of the porcine and human genomes demonstrates ZOO-FISH and gene mapping-based chromosomal homologies. *Mamm. Genome* **7:** 285–290.

Goureau, A., M. Yerle, A. Schmitz, J. Riquet, D. Milan, P. Pinton, G. Frelat, and J. Gellin. 1996. Human and porcine correspondence of chromosome segments using bidirectional chromosome painting. *Genomics* **36:** 252–262.

Larsen, F., G. Gundersen, R. Lopez, and H. Prydz. 1992. CpG islands as gene markers in the human genome. *Genomics* **13:** 1095–1107.

McQueen, H.A., J. Fantes, S.H. Cross, V.H. Clark, A.L. Archibald, and A.P. Bird. 1996. CpG islands of chicken are concentrated on microchromosomes. *Nature Genet.* **12:** 321–324.

Marklund, L., M. Johansson Moller, B. Hoyheim, W. Davies, M. Fredholm, R.K. Juneja, P. Mariani, W. Coppieters, H. Ellegren, and L. Andersson. 1996. A comprehensive linkage map of the pig based on a wild pig-Large White intercross. *Anim. Genet.* **27:** 255–269.

Matsuo, K., O. Clay, T. Takahashi, J. Silke, and W. Schaffner. 1993. Evidence for erosion of mouse CpG islands during mammalian evolution. *Somat. Cell Mol. Genet.* **19:** 543–555.

Rettenberger, G., C. Klett, U. Zechner, J. Kunz, W. Vogel, and H. Hameister. 1995. Visualization of the conservation of synteny between humans and pigs by heterologous chromosome painting. *Genomics* **26:** 372–378.

Rohrer, G.A., L.J. Alexander, Z. Hu, T.P. Smith, J.W. Keele, and C.W. Beattie. 1996. A comprehensive map of the porcine genome. *Genome Res.* **6:** 371–391.

Yerle, M., A. Schmitz, D. Milan, B. Chaput, L. Monteagudo, M. Vaiman, G. Frelat, and J. Gellin. 1993. Accurate characterization of porcine bivariate flow karyotype by PCR and fluorescence in situ hybridization. *Genomics* **16:** 97–103.