Construction of a Multi-Functional cDNA Library Specific for Mouse Pancreatic Islets and Its Application to Microarray

Motoi Nishimura,† Norihide Yokoi, Takashi Miki, Yukio Horikawa, Hirokazu Yoshioka, Jun Takeda, Osamu Ohara, and Susumu Seino

Department of Cellular and Molecular Medicine, Graduate School of Medicine, Chiba University, 1-8-1 Inohana, Chuo-ku, Chiba 260-8670, Japan, Division of Cellular and Molecular Medicine, Kobe University Graduate School of Medicine, 7-5-1 Kusunoki-cho, Chuo-ku, Kobe 650-0017, Japan, Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Corporation (JST), 4-1-8 Honcho, Kawaguchi 332-0012, Japan, Institute for Molecular and Cellular Regulation, Gunma University, 3-39-15 Showa-machi, Maebashi 371-8512, Japan, Kakengeneqs Corporation, 439-1 Matsuhidai, Matsudo 270-2214, Japan, Department of Endocrinology, Diabetes and Rheumatology, Division of Bioregulatory Medicine, Gifu University School of Medicine, 1-1 Yanagido, Gifu 501-1194, Japan, Department of Human Gene Research, Kazusa DNA Research Institute, 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan, and Laboratory for Immunogenomics, RIKEN Research Center for Allergy and Immunology, RIKEN Yokohama Institute, 1-7-22 Suchiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan

(Received 19 January 2004; revised 24 August 2004)

Abstract

We have constructed a high-quality and multi-applicable cDNA library specific for mouse pancreatic islets. This is the first pancreatic islet cDNA library created using a recombination-based method, which can readily be converted into other applications including yeast two-hybrid and mammalian expression libraries. Based on sequence data of the library, we constructed a sequence database specific for mouse pancreatic islets. Among the 8882 non-redundant clones, 5799 were classified into specific functional categories using a classification system designed by the Gene Ontology Consortium, 10% of which were “molecular function unknown” genes. We also developed cDNA microarray membranes with 8108 non-redundant clones. Analyses of expression profiles of three different cell lines and of MIN6 cells with or without overexpression of transcription factor NeuroD1 established the usefulness and applicability of our microarrays. The mouse pancreatic islet cDNA library, sequence database, set of clones, and microarrays developed in this study should be useful resources for studies of pancreatic islets and related diseases including diabetes mellitus.

Key words: cDNA library; microarray; pancreatic islet; recombination-based method; sequence database

1. Introduction

Pancreatic islets are small inner parts of pancreata, and play the central role in glucose homeostasis. The majority of islet cells are β-cells, which secrete the hypoglycemic hormone insulin; other islet cells (α-cells, δ-cells and PP cells) also secrete hormones affecting glucose homeostasis. Impairment of pancreatic β-cell function readily causes disorders of glucose homeostasis such as diabetes. Studies of knockout mice and a subtype of human diabetes have revealed that impairment of the transcription network in pancreatic β-cells is involved in the pathogenesis of diabetes. In these models, diabetes occurs due to disruptions or mutations in transcription factors that regulate gene expression in pancreatic β-cells. Thus, construction of a gene expression database for pancreatic islets and comprehensive analysis of expression profiles in the pancreatic islets might provide valuable information on the molecular mechanisms of the disease.

To investigate the molecules involved in normal function of pancreatic islets, several groups have constructed cDNA libraries from human pancreatic islets and sequenced clones from the libraries (In the GenBank database (www.ncbi.nlm.nih.gov/), nucleotide sequences of 4559 clones were deposited by...
Table 1. Characteristics of the mouse pancreatic islet cDNA library.

| Fraction by insert size | 1 | 2 | 3 | 4 | 5 | 6 |
|-------------------------|---|---|---|---|---|---|
| Proportion of chimeric clones (%) | 4.23 | 2.76 | 4.60 | 6.45 | 1.84 | 1.10 |
| Proportion of clones with full ORFs (%) | 64 | 56 | 48 | 76 | 72 | 52 |

1) Insert sizes of the fractions are: 1, >8 kbp; 2, 6–8 kbp; 3, 4–6 kbp; 4, 3–4 kbp; 5, 1.5–3 kbp; 6, <1.5 kbp.
2) Proportions of chimeric clones are calculated on the analysis of 3264 clones sequenced from both ends.
3) Proportions of clones with full ORFs are estimated from 25 randomly selected clones from each fraction.

Bell and colleagues, Chicago University, and those of 2055 clones by the I.M.A.G.E. consortium). As mice are widely used as experimental animals in studies of pancreatic islets, a cDNA library specific for mouse pancreatic islets should be useful. Recently, a recombination-based method for cDNA library construction has been developed. With size fractionation of the cDNA inserts, this method provides several improvements. The characteristics of the library are a high degree of complexity, an abundance of full-length sequences, and multiple applications. Libraries constructed by this method can readily be converted into other libraries including yeast two-hybrid and mammalian expression libraries.

In the present study, we constructed a recombination-based cDNA library specific for mouse pancreatic islets. Based on the sequence data of the library, we built a mouse pancreatic islet sequence database, and also developed cDNA microarray membranes.

2. Materials and Methods

2.1. Construction of a cDNA library specific for mouse pancreatic islets

All animal procedures were approved by the Chiba University Animal Care Committee. Pancreatic islets were isolated from 8- to 10-week-old male C57BL/6 mice by hand-picking under the microscope and collagenase digestion method as described previously. Poly(A)+ RNA was obtained from the islets using RNeasy Mini kit (Qiagen GmbH, Hilden, Germany) and µMACS mRNA isolation kit (Miltenyi Biotech, Bergisch Gladbach, Germany). Two micrograms of poly(A)+ RNA was isolated from approximately 5000 islets, from which cDNA synthesis was carried out as described previously. The synthesized cDNA was size-fractionated by agarose gel electrophoresis into three segments (<1.5, 1.5–3, and >3 kbps, respectively), and subjected to recombination reaction with BP Clonase (Invitrogen, Carlsbad, CA) and plasmid attP pSP73. Recombined plasmids were introduced into Escherichia coli cells (ElectroMAX DH10B cells, Invitrogen) by electroporation.

2.2. Plasmid preparation and DNA sequencing

Plasmids for sequencing the cDNA inserts and spotting on microarray membranes were prepared with MAG-NIA robot (TOYOBO, Osaka, Japan). With ABI3700 sequencers (Applied Biosciences, Mountain View, CA), 21,018 plasmid clones were successfully sequenced with T3 sequencing primer, and 3,264 clones were sequenced with T7 sequencing primer after elimination of redundant clones.

2.3. Construction of a sequence database

The sequence data obtained using T3 or T7 primers were subjected to BLAST search against the GenBank/DDBJ/EMBL nucleotide sequence database (www.ncbi.nlm.nih.gov/genbank/index.html) with a threshold E-value of 1.0e−50. The sequences were also subjected to BLAST search against the Ensembl mouse genome database (www.ensembl.org) to determine their locations on mouse chromosomes. All the sequence data were submitted to the DDBJ/EMBL/GenBank nucleotide sequence database (Accession nos. BP753069-BP777127), and are available from the ftp site at ftp://ftp.kazusa.or.jp/pub/pancreas/. The summary of the homology searches and the genome mapping data of the analyzed cDNAs are also available (Supplemental Information 1, http://www.dna-res.kazusa.or.jp/
11/5/01/supplemental/information1.html).

2.4. cDNA microarray analysis

The method of cDNA microarray analysis was described previously. A total of 8108 cDNA clones were spotted on the microarray membranes. For single microarray analysis, 10 µg of total RNA from each sample was reverse-transcribed in the presence of SuperScriptII Reverse Transcriptase (Invitrogen), [35S]dCTP, oligo(dT)12-18, and oligo(dT)25. The signal intensities in each analysis were normalized against the housekeeping gene Gapd, and the average signal intensities were calculated from more than two independent experiments.

2.5. Cell culture and recombinant adenoviruses

Mouse pancreatic islet \(\beta\)-cell line MIN6 cells were cultured as described previously. Mouse ES cell line R1 cells were cultured on mitomycin C-treated embryonic fibroblasts in a gelatinized dish with complete ES medium: high-glucose Dulbecco’s modified Eagle’s medium (Invitrogen) supplemented with 20% fetal calf serum (FCS) (Invitrogen), 2 mM L-glutamine (Invitrogen), 1× nonessential amino acids (Invitrogen), 1 µM 2-mercaptoethanol (Invitrogen), 100 µg/ml streptomycin sulfate, 60.5 µg/ml penicillin G (Invitrogen), and 1000 U/ml leukemia inhibitory factor (Chemicon, Temecula, CA). Mouse pituitary cell line AtT-20 cells were grown in Ham’s F10 medium (Invitrogen) with 5% FBS (Sigma, St. Louis, MO), and 15% heat-inactivated horse serum (Invitrogen).

Neurogenic differentiation 1 (NeuroD1/BETA2) recombinant adenovirus (Ad.CMVNeuroD1) and LacZ recombinant adenovirus (Ad.CMVLacZ) were constructed to express human NeuroD1 mRNA or E. coli LacZ (β-galactosidase) mRNA under cytomegalovirus (CMV) promoter, using an adenovirus expression vector kit (Takara Bio Inc., Otsu, Japan). MIN6 cells were infected with Ad.CMVNeuroD1 or Ad.CMVLacZ at a range of 5, 10, 30 and 50 multiplicity of infections (MOI), and cultured for another 2 days. Expression of NeuroD1 was confirmed by immuno blot analysis using antibodies specific for NeuroD1/BETA2 (Santa Cruz Biotechnology, Santa Cruz, CA) (Ishizuka et al. unpublished data), and expression of LacZ was confirmed using a galactosidase reporter gene staining detection kit (Sigma).

3. Results and Discussion

3.1. Construction of a mouse pancreatic islet cDNA library based on a recombination-based method

To construct a high quality cDNA library specific for normal mouse pancreatic islets, we applied a recombination-based method for library construction. Using 2 µg of poly(A)+ RNA extracted from approximately 5000 mouse pancreatic islets, we successfully constructed a cDNA library having a complexity of more than \(1.32 \times 10^7\) without amplification (Table 1). The library consists of six size fractions, which minimizes the size-bias effect on the population of cDNAs and results in improved variety. The library is also characterized by an abundance of full-length cDNA clones and fewer chimeric clones. Since each cDNA insert has specific sequences for recombination (compatible with Invitrogen GATEWAY™ system) at both ends, the library can readily be converted into other libraries such as mammalian expression and yeast two-hybrid libraries. This is the first pancreatic islet cDNA library created based on a recombination-based method, and should be a useful tool for pancreatic islet studies.

3.2. Construction of a cDNA sequence database specific for mouse pancreatic islets

We then sequenced a portion of the cDNA library to construct a sequence database specific for mouse pancreatic islets. We collected 4608 plasmid clones from each fraction, resulting in a total of 27,648 clones, and successfully sequenced 21,018 clones from the 3' region of the cDNAs. This raw sequence data contained about 13.7% of insulin 1 or 2 (Ins1/2) sequences and only 0.03% anylase sequences (representative gene of exocrine pancreas), indicating that the library was constructed primarily from pancreatic islets.

After eliminating redundant clones, we obtained a total of 8882 non-redundant clones (Supplemental Information 2, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information2.html). Among them, we also sequenced 3264 clones from the 5’ region of the cDNAs, and detected only 114 chimeric clones by comparison with the Ensembl mouse genome database (www.ensemble.org). By analyzing 150 randomly selected clones, we found that about 60% of the clones contained full open reading frames (ORFs) (Table 1). We collected these results and built a mouse pancreatic islet sequence database.

We compared 8882 non-redundant sequence data with the “all genes” database (www.allgenes.org), and found that 261 had no match in the database. Of 8621 sequences, 5799 were classified into specific functional categories using a classification system designed by the Gene Ontology (GO) Consortium (www.geneontology.org) (Fig. 1, Table 2 and Supplemental Information 3, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information3.html), while 2822 have no GO annotation, suggesting that these were derived from non-coding RNAs or rare transcripts. Of the 5799 sequences having GO annotation, about 42% belong to the largest category “binding,” which includes nucleic acid binding and peptide/protein binding. The second category “catalytic activity” comprises about 25%, and the third, “molecular function un-
3.1. Mouse Pancreatic Islets cDNA Library

3.2. Functional classification of the mouse pancreatic islet cDNA clones. Out of 8882 non-redundant clones, 5799 were classified into specific functional categories using a classification system designed by the Gene Ontology Consortium. Representative genes for each functional category are listed in Table 2. A complete list is available in Supplemental Information 3, http://www.dna-res.kazusa.or.jp/11/5/03/supplemental/information3.html.

3.3. Application to microarray analyses

We then applied our cDNA library to microarray analyses. We selected a total of 8108 cDNA clones from 8882 non-redundant clones and spotted them on microarray membranes with 12 control clones (Supplemental Information 2, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information2.html). To estimate the reproducibility of our microarray analyses, we compared membranes of different lots hybridized with the same probe. Scattered plot analysis showed very similar signal intensities from the two different membranes, indicating that our microarray analyses are reproducible (Supplemental Information 4A, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information4.html).

To evaluate the usefulness and applicability of our microarray membranes, we performed the following two experiments. First, we compared the expression profiles of three cell lines: MIN6 cells (mouse insulin-secreting pancreatic islet β-cell line), mouse embryonic stem (ES) cells, and AtT-20 cells (mouse pituitary endocrine cell line), representing endocrine pancreatic islet cells, undifferentiated cells,
Table 2. A list of functional classifications of the mouse pancreatic islet cDNA clones.

| GO function group | (subgroup) | annotation | GENBANK accession No (gb) | gene index (gi) | unigene cluster (ug) |
|------------------|------------|------------|--------------------------|----------------|---------------------|
| Binding          | Nucleic acid binding | Mm#S2205607 Mus musculus, H2B histone family, member A, clone MGC:19269 IMAGE:398962, mRNA, complete cds | gb=BC011440 | gi=15030325 | ug=Mm.21579 |
|                  |            | Mm#S937714 Mus musculus H3 histone, family 3A (H3a3a), mRNA | gb=NM_008210 | gi=6680158 | ug=Mm.89136 |
|                  |            | Mm#S937247 Mus musculus poly A binding protein, cytoplasmic 1 (Pabpc1), mRNA | gb=NM_008774 | gi=6679196 | ug=Mm.2642 |
|                  |            | Mm#S2608854 Mus musculus dicer-like protein (Dicer1) mRNA, complete cds | gb=AF430845 | gi=20385912 | ug=Mm.31523 |
|                  |            | Mm#S978931 Mus musculus poly(rC) binding protein 2 (Pcbp2), mRNA | gb=NM_011042 | gi=6997238 | ug=Mm.111 |
|                  | Peptide/ Protein binding | Mm#S1985421 Mus musculus Ran binding protein 5 mRNA, partial cds | gb=AF204327 | gi=12034715 | ug=Mm.151329 |
|                  |            | Mm#S2204674 Mus musculus, TAP binding protein, clone MGC:13789 IMAGE:423292, mRNA, complete cds | gb=BC015074 | gi=15929261 | ug=Mm.14097 |
|                  |            | Mm#S939589 Mus musculus syntaxin binding protein 2 (Stxbp2), mRNA | gb=NM_011503 | gi=7575867 | ug=Mm.7247 |
|                  |            | Mm#S937219 Mus musculus nuclear receptor-binding SET-domain protein 1 (Nsd1), mRNA | gb=NM_008739 | gi=6679137 | ug=Mm.12964 |
| Others binding   |            | Mm#S1660484 Mus musculus piccolo (presynaptic cytomatrix protein) (Pico1), mRNA | gb=NM_011995 | gi=15273339 | ug=Mm.40996 |
| Catalytic activity | hydrolase activity | Mm#S937421 Mus musculus protein tyrosine phosphatase, non-receptor type 2 (Ptpn2), mRNA | gb=NM_008977 | gi=6679552 | ug=Mm.985 |
|                  |            | Mm#S937616 Mus musculus acid beta glucosidase (Gba), mRNA | gb=NM_008594 | gi=6679554 | ug=Mm.5031 |
|                  |            | Mm#S979159 Mus musculus carboxypeptidase E (Cpe), mRNA | gb=NM_013494 | gi=7304972 | ug=Mm.31959 |
|                  |            | Mm#S979344 Mus musculus proprotein convertase subtilisin/kexin type 1 (Pcsk1), mRNA | gb=NM_013628 | gi=703570 | ug=Mm.1333 |
|                  |            | Mm#S979263 Mus musculus proprotein convertase subtilisin/kexin type 2 (Pcsk2), mRNA | gb=NM_008792 | gi=6679228 | ug=Mm.1247 |
| Ligase activity  |            | Mm#S1658986 Mus musculus cytidine 5'-triphosphate synthase 2 (Ctps2), mRNA | gb=NM_018737 | gi=9055197 | ug=Mm.2065 |
|                  |            | Mm#S2003048 Mus musculus propionyl-Coenzyme A carboxylase, beta polypeptide (Pccb), mRNA | gb=NM_025835 | gi=13385309 | ug=Mm.21079 |
|                  |            | Mm#S2003110 Mus musculus methionine adenosyltransferase E (Mae), mRNA | gb=NM_025939 | gi=13383453 | ug=Mm.182931 |
|                  |            | Mm#S2534128 Mus musculus, glutamate-cysteine ligase, catalytic subunit, clone MGC:30487 IMAGE:4195425, mRNA, complete cds | gb=BC019374 | gi=18043914 | ug=Mm.4386 |
|                  |            | Mm#S2551407 Mus musculus, similar to alanyl-tRNA synthetase (H. sapiens), clone MGC:37368 IMAGE:4976684, mRNA, complete cds | gb=CT026613 | gi=20072364 | ug=Mm.24174 |
| Oxidoreductase activity |            | Mm#S1660039 Mus musculus NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 1 (Ndufa1), mRNA | gb=NM_019443 | gi=9506910 | ug=Mm.34869 |
|                  |            | Mm#S1660493 Mus musculus pyruvate-5-carboxylate synthase (pyruvate-ketolase), mRNA | gb=NM_019698 | gi=9790060 | ug=Mm.29751 |
|                  |            | Mm#S2002811 Mus musculus sterol-C4-methyl oxidase-like (Sc4mol), mRNA | gb=NM_025436 | gi=13384833 | ug=Mm.30119 |
|                  |            | Mm#S2534582 Mus musculus cytochrome c oxidase, subunit VIIC (Cox6c), mRNA | gb=NM_053071 | gi=16716342 | ug=Mm.548 |
|                  |            | Mm#S936625 Mus musculus stearyl-Coenzyme A desaturase 2 (Scd2), mRNA | gb=NM_009128 | gi=6677862 | ug=Mm.193096 |
| Transferase activity |            | Mm#S1437353 Mus musculus histone acetyltransferase (Morf-pending), mRNA | gb=NM_017479 | gi=8563735 | ug=Mm.30998 |
|                  |            | Mm#S1660190 Mus musculus mfa methyltransferase (Mfa-pending), mRNA | gb=NM_019721 | gi=9790046 | ug=Mm.36983 |
|                  |            | Mm#S1985402 Mus musculus prenylcytochrome b5 methyltransferase mRNA, partial cds | gb=AF209926 | gi=12082482 | ug=Mm.29356 |
|                  |            | Mm#S2545982 Mus musculus, transglutaminase 2, C polypeptide, clone MGC:6152 IMAGE:1256943, mRNA, complete cds | gb=BC016492 | gi=16741319 | ug=Mm.18843 |
|                  |            | Mm#S2003650 Mus musculus glycogen synthase 1, muscle (Gys1), mRNA | gb=NM_030678 | gi=13507598 | ug=Mm.185247 |
| Molecular function unknown |            | Mm#S1971846 Mus musculus adult male hippocampus CDNA, RIKEN full-length enriched library, clone:290059A22:DlGeorge syndrome chromosome region 6, full insert sequence | gb=AK0119346 | gi=12859501 | ug=Mm.27155 |
|                  |            | Mm#S1973121 Mus musculus 8 days embryo whole body CDNA, RIKEN full-length enriched library, clone:5730433K22:unclassifiable, full insert sequence | gb=AK017611 | gi=12856941 | ug=Mm.158400 |
| GO function group | (subgroup) | annotation | GENBANK accession No (gb) | gene index (gi) | unigene cluster (ug) |
|------------------|-----------|------------|---------------------------|----------------|---------------------|
| transcription regulator activity | | Mm#S937155 Mus musculus myelin transcription factor 1 (Myr1), mRNA | gb=NM_008665 | gi=667899 | ug=Mm.2098 |
| structural molecule activity | | Mm#S121804 Tubulin, alpha 2, mRNA | gb=M13446 | gi=202209 | ug=Mm.1975 |
| transporter activity | | Mm#S979102 Mus musculus ATP-binding cassette, sub-family A (ABC1), member 1 (Abca1), mRNA | gb=NM_013454 | gi=730484 | ug=Mm.369 |
| obsolete molecular function | | Mm#S1660236 Mus musculus SEC23B (S. cerevisiae) (Sec23b), mRNA | gb=NM_019741 | gi=978996 | ug=Mm.3415 |
| signal transducer activity | | Mm#S939854 Mus musculus growth hormone receptor (Ghr), mRNA | gb=NM_019787 | gi=979021 | ug=Mm.2870 |
| enzyme regulator activity | | Mm#S939211 Mus musculus neutral sphingomyelinase (N-SMase) activation associated factor (Nsmaf), mRNA | gb=NM_010945 | gi=675487 | ug=Mm.3059 |
| chaperone activity | | Mm#S1997753 Mus musculus heat shock 70kD protein 5 (Gluc-78kD) (Hsp78), mRNA | gb=NM_022310 | gi=1616248 | ug=Mm.918 |

Table 2. Continued.
Table 2. Continued.

| GO function (subgroup) | annotation | GENBANK accession No (gb) | gene index (gi) | unigene cluster (ug) |
|------------------------|------------|--------------------------|-----------------|----------------------|
| antioxidant activity   | Mm#5937966 Mus musculus peroxiredoxin 3 (Prdx3), mRNA | gb=NM_007452 | gi=6680689 | ug=Mm.29821 |
|                        | Mm#1437442 Mus musculus peroxiredoxin 4 (Prdx4), mRNA | gb=NM_016764 | gi=7948998 | ug=Mm.19127 |
|                        | Mm#993659 Mus musculus thioredoxin 1 (Trx1), mRNA | gb=NM_011660 | gi=6755910 | ug=Mm.1275 |
|                        | Mm#1660289 Mus musculus thioredoxin 2 (Trx2), mRNA | gb=NM_019913 | gi=9903608 | ug=Mm.3533 |
|                        | Mm#2455825 Mus musculus glutaredoxin 1 (thioltransferase) (Glrx1), mRNA | gb=NM_053108 | gi=16716404 | ug=Mm.29728 |

cell adhesion molecule activity

| annotation | GENBANK accession No (gb) | gene index (gi) | unigene cluster (ug) |
|------------|--------------------------|-----------------|----------------------|
| Mm#5979051 Mus musculus integrin alpha 6 (Ig6a), mRNA | gb=NM_008397 | gi=7110658 | ug=Mm.25232 |
| Mm#541562  | Marine mRNA for integrin beta subunit, mRNA | gb=Y00769 | gi=52721 | ug=Mm.4712 |
| Mm#308025  | Procollagen, type I, alpha 1, mRNA | gb=U08020 | gi=475073 | ug=Mm.22621 |
| Mm#208041  | Procollagen, type IV, alpha 1, mRNA | gb=J04694 | gi=556296 | ug=Mm.738 |
| Mm#244281  | Procollagen, type XV, mRNA | gb=AF011450 | gi=2558824 | ug=Mm.4352 |

Out of 8882 non-redundant clones, 5799 were classified into specific functional categories using a classification system designed by the Gene Ontology Consortium. Representative genes for each functional category are listed in this table. A complete list is available in Supplemental Information 3, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information3.html.

Figure 2. Comparison of expression profiles in MIN6, ES and AtT-20 cells. Genes with signal intensity of one-tenth or more of the housekeeping gene Gapd were compared in the three cell lines. Numbers of genes are shown in this diagram. A complete list of the expressed genes is available in Supplemental Information 5, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information5.html.

Out of 8882 non-redundant clones, 5799 were classified into specific functional categories using a classification system designed by the Gene Ontology Consortium. Representative genes for each functional category are listed in this table. A complete list is available in Supplemental Information 3, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information3.html.

We used a threshold signal intensity of one-tenth that of the housekeeping gene Gapd, and compared the expressed genes in the three cell lines (Fig. 2). A list of the expressed genes is available (Supplemental Information 5, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information5.html). Among a total of 879 genes, 738 (84.0%) were expressed in MIN6 cells, while only 319 (36.3%) and 359 (40.8%) were expressed in AtT-20 and ES cells, respectively.

Comparison of expression profiles of three cell lines revealed both differentially and similarly expressed genes. A subset of 135 genes was expressed in both MIN6 and AtT-20 cells but not in ES cells. As both MIN6 and AtT-20 cells were derived from endocrine cells and preserve regulated exocytotic ability, the genes involved should be expressed in both endocrine cell types. Indeed, well-known endocrine-specific genes such as chromogranins,23,24 prohormone convertase 1/3,25 and carboxypeptidase E26,27 were in this subset (Fig. 2 and Supplemental Information 5, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information5.html). In addition, 28 genes with unknown functions were contained in this subset, suggesting they may have other roles in regulated exocytosis.

By contrast, a subset of 356 genes was found to be expressed only in MIN6 cells. This subset included known pancreatic β-cell-specific genes, such as Ins1/2, Iapp, and Pdx1 as well as many uncharacterized genes (Supplemental Information 5, http://www.dna-res.kazusa.or.jp/11/5/01/supplemental/information5.html) that may be involved in phenomena specific to pancreatic β-cells such as insulin synthesis. Thus, comparison of the mRNA expression patterns of various cell types should be useful in investigating the biological function of the clones newly identified in our cDNA library.

As the second experiment, we compared expression profiles between MIN6 cells and those infected with aden-
Table 3. A list of genes upregulated by overexpression of NeuroD1 in MIN6 cells.

| Signal intensity No. | +LacZ 1) | +NeuroD1 1) | Ratio | Gene name 3) / GenBank accession number |
|----------------------|----------|-------------|-------|-----------------------------------------|
| 1                    | 4.70     | 84.31       | 17.9324 | gi: 26101046 / AK083076                  |
| 2                    | 14.59    | 251.26      | 17.2207 | Neurod1 (NeuroD1) 0) / NM_010894         |
| 3                    | 3.60     | 24.37       | 6.7660  | gi: 38328296 / BC062185                  |
| 4                    | 1.61     | 7.60        | 4.7122  | Scnn1b / NM_011325                       |
| 5                    | 4.26     | 18.02       | 4.2319  | Ipfl (Pdx1) / NM_008814                  |
| 6                    | 16.42    | 62.79       | 3.8232  | Cdkn1a (p21) / NM_007669                 |
| 7                    | 2.07     | 7.51        | 3.6218  | Jund1 / NM_010592                       |
| 8                    | 7.48     | 24.16       | 3.2285  | Ddx5 / BC009142                         |
| 9                    | 5.25     | 15.33       | 2.9198  | Rangap1 / NM_011241                      |
| 10                   | 4.98     | 14.42       | 2.8952  | gi: 12833685 / AK003180                  |
| 11                   | 4.47     | 12.80       | 2.8639  | Atp2a2 / NM_009722                      |
| 12                   | 2.65     | 7.53        | 2.8440  | Scly / NM_016717                        |
| 13                   | 7.50     | 19.71       | 2.6274  | gi: 17391148 / BC018486                 |
| 14                   | 6.66     | 17.30       | 2.5976  | Actn4 / NM_021895                       |

Signal intensities are normalized by two independent experiments.

1) Signal intensities from MIN6 cells infected with Ad.CMVLacZ.
2) Signal intensities from MIN6 cells infected with Ad.CMVNeuroD1.
3) gi (gene index number in GenBank) or Gene name registered in NCBI Entrez Gene database.
4) Signal intensity of NeuroD1 from MIN6 cells infected with Ad.CMVNeuroD1 was influenced by the exogenous introduction of NeuroD1.

oviral vector to express the transcription factor NeuroD1 exogenously. Because NeuroD1 is critical in pancreatic islet β-cell development and β-cell functions, and a mutation in the human NeuroD1 gene is responsible for maturity-onset diabetes of the young (MODY6), we overexpressed NeuroD1 in MIN6 cells to evaluate the usefulness of our microarray analyses in the search for novel target genes. So far, Pdx1, p21, and Sur1 have been reported to be transactivated by NeuroD1.

Infection of the adenoviral vector with a range of 5 or 10 MOI did not alter the expression profiles, but infection with 30 or 50 MOI caused significant alterations (data not shown). The results of infection with 50 MOI are shown in Table 3 and Supplemental Information 4D, http://www.dna-res. kazusa.or.jp/11/5/01/supplemental/information4.html. We confirmed that the genes known to be induced by NeuroD1, including Pdx1 and p21 were upregulated, and also found that expression levels of another 11 genes increased significantly. None of these genes has previously been associated with NeuroD1, and all are candidates for transactivation by NeuroD1. This second experiment shows that our microarray analyses will be useful in the search for novel molecular targets of various β-cell transcription factors and in investigations of pancreatic islets and related diseases including diabetes.

Acknowledgements: We thank Y. Murahashi for assistance with computational analysis of nucleotide sequences. We also thank T. Watanabe and T. Hirano for technical assistance in constructing the mouse pancreatic islet library and in the microarray analysis. This work was supported by Grants-in-Aid for Specially Promoted Research and for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan; by a Grant from the Ministry of Economy, Trade and Industry, Japan, and the Futaba Electronics Memorial Foundation.

References

1. Kahn, C. R., Gordon, C., Weir, G. C. et al. 1994, *Joslin’s Diabetes Mellitus*, 13th Ed., Lippincott Williams & Wilkins, Philadelphia, PA.
2. Jonsson, J., Carlsson, L., Edlund, T., Edlund, H. 1994, Insulin-promoter-factor 1 is required for pancreas development in mice, *Nature*, 13, 606–609.
3. Ahlgren, U., Jonsson, J., Jonsson, L., Simu, K., and Edlund, H. 1998, Beta-cell-specific inactivation of the mouse Lpfl/Pdx1 gene results in loss of the beta-cell phenotype and maturity onset diabetes, *Genes Dev.*, 12, 1763–1768.
4. Lee, Y. H., Sauers, B., and Gonzalez, F. J. 1998, Laron dwarfism and non-insulin-dependent diabetes mellitus in the Hnf-1alpha knockout mouse, *Mol Cell Biol.*, 18, 3059–3068.
5. Shih, D. Q., Screenan, S., Munoz, K. N. et al. 2001, Loss of HNF-1alpha function in mice leads to abnormal expression of genes involved in pancreatic islet development and metabolism, *Diabetes*, **50**, 2472–2480.

6. Naya, F. J., Huang, H. P., Qu, Y. et al. 1997, Diabetes, defective pancreatic morphogenesis, and abnormal endoenteroendocrine differentiation in BETAG2/NeuroD-deficient mice, *Genes Dev.*, **11**, 2323–2334.

7. Stoffers, D. A., Ferrer, J., Clarke, W. L., and Habener J. F. 1997, Early-onset type-II diabetes mellitus (MODY4) linked to IPF1, *Nat Genet.*, **17**, 138–139.

8. Yamagata, K., Oda, N., Kaisaki, P. J. et al. 1996, Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY4) linked to IPF1, *Nat Genet.*, **17**, 138–139.

9. Yamagata, K., Furuta, H., Oda, N. et al. 1996, Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1), *Nature*, **384**, 455–458.

10. Horikawa, Y., Iwasaki, N., Haral, M. et al. 1997, Mutation in hepatocyte nuclear factor-1 beta gene (TCF2) associated with MODY, *Nat Genet.*, **17**, 384–385.

11. Malecki, M. T., Jhala, U. S., Antonellis, A. et al. 1999, Mutations in NEUROD1 are associated with the development of type 2 diabetes mellitus, *Nat Genet.*, **23**, 323–328.

12. Bell G. I. and Polonsky, K. S. 2001, Diabetes mellitus and genetically programmed defects in beta-cell function, *Nature*, **414**, 788–791.

13. Ohara, O. and Temple, G. 2001, Directional cDNA library construction assisted by the in vitro recombination reaction, *Nucleic Acids Res.*, **29**, E22.

14. Ohara, O., Nagase, T., Mitsui, G. et al. 2002, Characterization of size-fractionated cDNA libraries generated by the in vitro recombination-assisted method, *DNA Res.*, **9**, 47–57.

15. Nakayama, M., Kikuno, R., and Ohara, O. 2002, Protein-protein interactions between large proteins: two-hybrid screening using a functionally classified library composed of long cDNAs, *Genome Res.*, **12**, 1773–1784.

16. Miki, T., Nagashima, K., Tashiro, F. et al. 1998, Defective insulin secretion and enhanced insulin action in KATP channel-deficient mice, *Proc. Natl. Acad. Sci. USA*, **95**, 10402–10406.

17. Mitsui, G., Mitsui, K., Hirano, T., Ohara, O., Kato M., and Niwano, Y. 2003, Kinetic profiles of sequential gene expressions for chemokines in mice with contact hypersensitivity, *Immunol. Lett.*, **86**, 191–197.

18. Kawaki, J., Nagashima, K., Tanaka, J. et al. 1999, Unresponsiveness to glibenclamide during chronic treatment induced by reduction of ATP-sensitive K⁺ channel activity, *Diabetes*, **48**, 2001–2006.

19. Suzuki, M., Mizutani-Koseki, Y., Fujimura, Y. et al. 2002, Involvement of the Polycomb-group gene Ring1B in the specification of the anterior-posterior axis in mice, *Development*, **129**, 4171–4183.

20. Buonassisi, V., Sato, G., and Cohen, A. 1962, Hormone-producing cultures of adrenal and pituitary tumor origin, *Proc. Natl. Acad. Sci. USA*, **48**, 1184–1190.

21. Scearce, L. M., Brestelli, J. E., McWeeney, S. K. et al. 2002, Functional genomics of the endocrine pancreas: the pancreas clone set and PancChip, new resources for diabetes research, *Diabetes*, **51**, 1997–2004.

22. Kaestner, K. H., Lee, C. S., Scearce, L. M. et al. 2003, Transcriptional program of the endocrine pancreas in mice and humans, *Diabetes*, **52**, 1604–1610.

23. Ehrhart, M., Grube, D., Bader, M. F., Aunis, D., and Gratzi, M. 1986, Chromogranin A in the pancreatic islet: cellular and subcellular distribution, *J. Histochem. Cytochem.*, **34**, 1673–1682.

24. Iacangelo, A., Affolter, H. U., Eiden, L. E., Herbert, E., and Grimes, M. 1986, Bovine chromogranin A sequence and distribution of its messenger RNA in endocrine tissues, *Nature*, **323**, 82–86.

25. Zhu, X., Zhou, A., Dey, A. et al. 2002, Disruption of PC1/3 expression in mice causes dwarfism and multiple neuroendocrine peptide processing defects, *Proc. Natl. Acad. Sci. USA*, **99**, 10293–10298.

26. Fricker, L. D., Adelman, J. P., Douglass, J., Thompson, R. C., von Strandmann, R.P., and Hutton, J. 1989, Isolation and sequence analysis of cDNA for rat carboxypeptidase E [EC 3.4.17.10], a neuropeptide processing enzyme, *Mol. Endocrinol.*, **3**, 666–673.

27. Nagert, J. K., Fricker, L. D., Vuralamon, O. et al. 1995, Hyperinsulinaemia in obese fat/fat mice associated with a carboxypeptidase E mutation which reduces enzyme activity, *Nat. Genet.*, **10**, 135–142.

28. Huang, H. P., Chu, K., Nemza-Gaillard, E., Elberg, D., and Tsai, M. J. 2002, Neogenesis of beta-cells in adult BETAG2/NeuroD-deficient mice, *Mol. Endocrinol.*, **16**, 541–551.

29. Kojima, H., Fujimura, M., Matsumura, K. et al. 2003, NeuroD-betacellulin gene therapy induces islet neogenesis in the liver and reverses diabetes in mice, *Nat. Med.*, **9**, 596–603.

30. Mutoh, H., Naya, F. J., Tsai, M. J., and Leiter, A. B. 1998, The basic helix-loop-helix protein BETAG2 interacts with p300 to coordinate differentiation of secretin-expressing enteroendocrine cells, *Genes Dev.*, **12**, 820–830.

31. Kim J. W., Seghers V., Cho J. H., et al. 2002, Transactivation of the mouse sulfonylurea receptor 1 gene by BETAG2/NeuroD, *Mol. Endocrinol.*, **16**, 1097–1107.