Reprinted with permission from the authors. Copyright © 1976 Nobumichi Hozumi and Susumu Tonegawa.

Proc. Natl. Acad. Sci. USA
Vol. 73, No. 10, pp. 3628–3632, October 1976
Genetics

Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions

(κ-chain mRNA/restriction enzymes/RNA·DNA hybridization)

NOBUMICHI HOZUMI AND SUSUMU TONEGAWA

Basel Institute for Immunology, 487, Grenzacherstrasse, CH-4058 Basel, Switzerland

Communicated by N. K. Jerne, July 2, 1976

ABSTRACT **A high-molecular-weight DNA from Balb/c mouse early embryo or from MOPC 321 plasmacytoma (a κ-chain producer) was digested to completion with *Bacillus amyloliquefaciens* strain H restriction enzyme (*Bam*H I). The resulting DNA fragments were fractionated according to size in preparative agarose gel electrophoresis. DNA fragments carrying gene sequences coding for the variable or constant region of κ chains were detected by hybridization with purified, ^{125}I-labeled, whole MOPC 321 κ mRNA and with its 3'-end half. The pattern of hybridization was completely different in the genomes of embryo cells and of the plasmacytoma. The pattern of embryo DNA showed two components, one of which (molecular weight = 6.0 million) hybridized with C-gene sequences and the other (molecular weight = 3.9 million) with V-gene sequences. The pattern of the tumor DNA showed a single component that hybridized with both V-gene and C-gene sequences and that is smaller (molecular weight = 2.4 million) than either of the components in embryo DNA. The results were interpreted to mean that the V_κ and C_κ genes, which are some distance away from each other in the embryo cells, are joined to form a contiguous polynucleotide stretch during differentiation of lymphocytes. Such joining occurs in both of the homologous chromosomes. Relevance of these findings with respect to models for V–C gene joining, activation of a specific V_κ gene, and allelic exclusion in immunoglobulin gene loci is discussed.**

Both light and heavy chains of immunoglobulin molecules consist of two regions: the variable region (V region) and the constant region (C region) (1, 2). Uniqueness (i.e., one copy per haploid genome) of the genetic material coding for C region ("C gene") has been conjectured from normal Mendelian segregation of allotypic markers (3). Nucleic acid hybridization studies have confirmed this notion (4–10). Hybridization studies have also demonstrated that a group of closely related V regions are somatically generated from a few, conceivably even a single, germline gene(s) (V gene) (4–6). They did not, however, give us any reliable estimate of the *total* number of germ line V genes. However, given the enormous diversity of V regions, the existence of multiple germ line V genes seems likely. If this is so, a problem arises: how is the information in the V and C genes integrated to generate a contiguous polypeptide chain? Since V- and C-gene sequences exist in a single mRNA molecule as a contiguous stretch (11), such integration must take place at either the DNA or the RNA level. Integration at the RNA level could result from a "copy-choice" event during transcription or from joining of two RNA molecules after transcription.

We report here experimental evidence for possible joining of V and C sequences at the DNA level.

MATERIALS AND METHODS

Myeloma Tumors. Balb/c mouse plasmacytoma MOPC 321 was kindly provided by Dr. M. O. Potter. Tumors were maintained in mice as described (4).

Abbreviations: V and C genes, genes coding for variable and constant regions, respectively; M_r, molecular weight.

***Bam*H I Restriction Endonuclease.** *Bacillus amyloliquefaciens* strain H, originally from Dr. F. Young, was obtained from Dr. T. Bickles at the Biozentrum, Basel. Cells were grown in L-broth. The *Bam*H I endonuclease was prepared according to the method of Wilson and Young (12), except that 10% glycerol was present in all buffers during the phosphocellulose step. Five milligrams of high-molecular-weight embryo or MOPC 321 tumor DNA in a buffer consisting of 6 mM Tris-HCl, pH 7.4, 6 mM $MgCl_2$, and 6 mM 2-mercaptoethanol was incubated with 10^4 units (1 unit is defined as the amount of the enzyme sufficient for digesting 1 μg of phage λ DNA in 60 min at 37° under the above conditions) of the purified enzyme at 37° for 4 hr. In order to investigate completeness of digestion we followed the following scheme. An aliquot of the reaction mixture containing mouse DNA was removed before incubation and mixed with a small amount (ratio of λ DNA to mouse DNA, 1:10) of phage λ DNA. This pilot mixture was incubated in parallel with the main reaction mixture under the conditions described above. After incubation the pilot mixture was electrophoresed in 0.9% agarose. DNA was visualized under ultraviolet light after staining with 1 μg/ml of ethidium bromide (13). Digestion of mouse DNA in the main reaction mixture was considered to be complete when the electrophoresis pattern of the λ DNA in the pilot mixture exhibited no indication of incomplete digestion.

Preparative Agarose Gel Electrophoresis. An apparatus conventionally used for separation of serum proteins was adapted for nucleic acids. Agarose (Sigma, electrophoresis grade) was melted in TA buffer (20 mM Tris-CH_3CO_2, pH 8.0, 18 mM NaCl, 2 mM Na_2EDTA, 20 mM $NaCH_3CO_2$) at a final concentration of 0.9% and cast in a Plexiglas tray (40 × 50 × 1 cm) with a central longitudinal partition. Agarose was allowed to solidify at room temperature. With the aid of a scalpel, 5-mm wide slots were made in the agarose at 10 cm from one end. The DNA digested with *Bam*H I was concentrated with 2-butanol by the procedures of Stafford and Bieber (14) and dialyzed against TA buffer. The DNA solution was supplemented with 10 × concentrated TA buffer and bromphenol blue. The mixture was warmed to 47°. Melted agarose was mixed with the warmed DNA solution such that the final mixture contained the original TA buffer, 0.01% dye, and 0.45% agarose. The mixture was transferred to the slot in the agarose sheet and was allowed to solidify at room temperature. In order to avoid overloading, the amount of the DNA was kept at less than 2.1 μg/mm^2 of the cross section of the gel. The gel was covered with plastic wrap and electrophoresis was carried out at 4° for 3 days. For the first 20 hr the electric current was kept at 0.75 mA/cm^2 of gel cross section; it was raised to 3 mA for the remainder of the run.

Extraction of DNA from Agarose Gel. After electrophoresis the gel was cut with a scalpel into 5-mm thick slices. Each slice was transferred to a test tube containing one volume of 5 M

$NaClO_4$ (15). The gel slices were melted by heating at 60° for 60 min with occasional shaking by hand. In order to remove dissolved agarose as well as $NaClO_4$, DNA was absorbed to a small column of hydroxyapatite (Bio-Gel, DNA grade) and then eluted with 0.4 M potassium phosphate, pH 6.9.

Hybridization. DNA eluted from hydroxyapatite was sonicated in a MSE sonicator to yield fragments 400–500 base pairs long. Purified, sonicated *Escherichia coli* DNA (Sigma type VIII) was added at a final concentration of 100 μg/ml. The mixture was dialyzed against water overnight, and DNA was precipitated with 2 volumes of ethanol. The precipitate was collected by centrifugation, dried under reduced pressure, and dissolved in 25 μl of a mixture consisting of 0.3 M NaCl, 0.03 M sodium citrate, 0.01 M Tris HCl, pH 7.5, 0.001 M Na_2EDTA, and 700–1500 cpm of ^{125}I-labeled mRNA or its 3′-end fragment. The hybridization mixture was covered with a thin layer of mineral oil and heated at 98–100° for 5 min. Annealing was carried out at 70° for 40–48 hr. The processing of the hybridization mixture has been described (4).

Preparation of High-Molecular-Weight DNA. DNA was purified from either 12- to 13-day-old Balb/c mouse embryos (Bomholtgard, Denmark) or MOPC 321 tumors according to procedures described by Gross-Bellard *et al.* (16).

Other Procedures. Methods for purification of κ-chain mRNA of MOPC 321, for isolation of 3′-end half fragment of the mRNA, and for RNA iodination have all been described (4, 6). The κ mRNA preparation used in the present work is the one described in Fig. 1 of ref. 6.

RESULTS

Detection of V- and C-gene sequences in DNA fragments

A straightforward procedure for quantitation of a specific sequence in a particular DNA preparation is to carry out hybridization either in liquid or preferentially on filter paper under conditions where labeled RNA is well in excess over DNA (RNA excess, RNA-driven hybridization) (17). For this procedure it is necessary that the RNA preparation is completely pure. The κ mRNA preparation used in the present experiments, although as pure as the present technology permits, does not meet this condition. From the "fingerprint" analysis of T1 RNase digest and from gel electrophoresis in 99% formamide, the purity of the κ mRNA preparation was estimated to be about 90% (4, 6). For this reason and also because of the difficulty of obtaining a large amount of purified κ mRNA, the RNA saturation procedure is not suitable for quantitation of immunoglobulin genes.

When the purpose is to detect a particular DNA sequence in relative rather than absolute terms, as in the present case, an alternative method is available which circumvents the above problems. In Fig. 1 a varying amount of MOPC 321 DNA was hybridized with ^{125}I-labeled κ mRNA of MOPC 321 or the half containing its 3′-end. The volume of hybridization mixtures, concentration of RNA, and time of incubation were constant. Under these conditions the level of ^{125}I-labeled RNA in the hybrid is a function of the concentration of the complementary DNA sequences. As shown in Fig. 1, the increment in the level of hybridization decreases as the concentration of complementary DNA sequences increases. Thus, when DNA is subjected to a fractionation procedure such as gel electrophoresis, and each fraction is assayed for a specific DNA sequence by this method, the shape of the resulting hybrid peak is flatter than expected from the actual distribution of complementary DNA sequences. (In principle, it is possible to correct each fraction for the nonlinearity of the assay by using Fig. 1 or a "standard curve". We have not done this, however, in Fig. 2.) This disadvantage is more than counterbalanced by the critical advantage; namely, that hybridization of minor RNA impurity does not overshadow that of the major RNA component. The method is particularly suitable where the impurity is distributed among many different RNA species, for in this case, DNA sequences complementary to the contaminating RNAs would be distributed among many different fractions, and hybridization of any particular RNA impurity would not stand out of the general background level. As previously shown, κ mRNA preparations have impurities of exactly this type (4, 6).

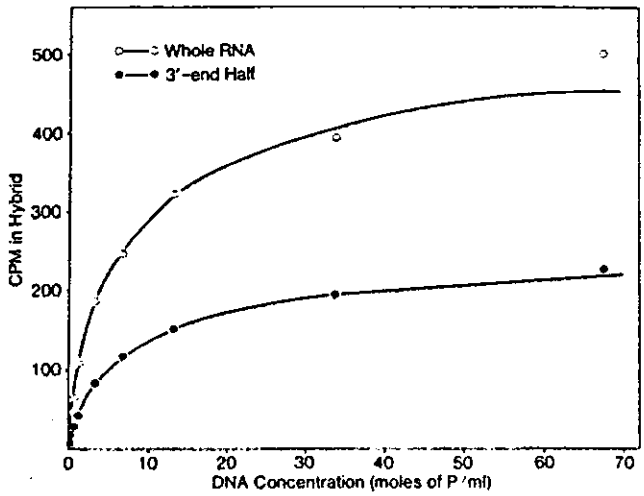


FIG. 1. Dependency of hybridization level on DNA concentration. ^{125}I-Labeled, whole κ mRNA of MOPC 321 (1300 cpm) or its 3′-end half fragment (600 cpm) was annealed with varying indicated amounts of denatured MOPC321 DNA in 0.3 M NaCl, 0.03 M sodium citrate (pH 7) at 70° for 42 hr. The volume of annealing mixture was 0.05 ml. Specific activity of the RNA was 6.5×10^7 cpm/μg. Each point represents an average of duplicate measurement. Intrinsic RNase-resistant counts, 22 cpm for whole RNA and 16 cpm for 3′-end half, are subtracted.

Since a RNA probe for V gene sequences is not available (the 3′-end half fragment is a RNA probe for C-gene sequences), V-gene sequences were determined indirectly from the difference in the two hybridization levels obtained by the whole molecule and the half-molecule containing the 3′-end (3′-half). For this purpose two RNA probes with identical specific activity were prepared; hybridization of whole molecules was carried out using twice as much input radioactivity as with 3′-end half-molecules. The final hybridization levels obtained with the whole molecules are just about twice those with the 3′-half in a wide range of DNA concentration (Fig. 1); this justified such an indirect measurement of V-gene sequences.

V- and C-gene sequences in embryo and myeloma DNA fragments generated with *Bam*H I endonuclease

High-molecular-weight embryo and MOPC 321 tumor DNAs were digested with *Bam*H I endonuclease, and the resulting DNA fragments were fractionated according to size by preparative agarose gel electrophoresis. DNA eluted from gel slices was assayed for V- and C-gene sequences by hybridization with ^{125}I-labeled whole or 3′-half κ mRNA of MOPC 321 under the condition described in the last section. The pattern of hybridization is shown in Fig. 2. With embryo DNA, two DNA components of molecular weight (M_r) 6.0 and 3.9 million hybridized with the whole RNA molecules, whereas only the 6.0 million M_r component hybridized with the 3′-half. The fact that only the 6.0 million M_r component exhibits any hybridization

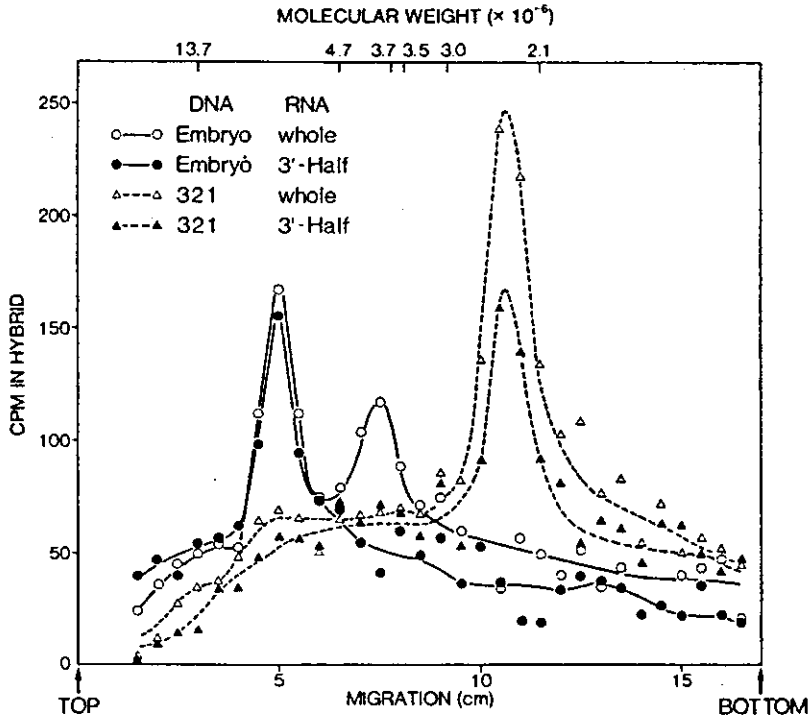


FIG. 2. Gel electrophoresis patterns of mouse DNA fragments, generated by *Bam*H I, carrying V- or C-gene sequences. ^{125}I-Labeled, whole κ mRNA of MOPC 321 (1220 cpm) or its 3'-end half fragment (600 cpm) was annealed with DNA extracted from gel slices. Intrinsic RNase-resistant counts are subtracted. Conditions of electrophoresis and hybridization are as described in *Materials and Methods*. Hybridization patterns with DNA of two different sources, embryo and MOPC 321 tumor, are superimposed. The molecular weight scale was obtained from phage λ DNA, digested by *Eco*RI (*E. coli*) endonuclease, which was electrophoresed in parallel with mouse DNAs.

above the background level with the 3'-end half indicates that not only the C-region sequences, but also the sequences corresponding to the untranslated region at the 3'-end of mRNA molecule, are in this component. Furthermore, since the 6.0 million M_r component hybridizes almost equally with the two RNA probes, it should not contain appreciable V gene sequences. Thus, the 3.9 million M_r component should contain V-gene sequences as well as sequences corresponding to the untranslated region at the 5'-end of mRNA molecule.

Since V and C genes are in separate DNA fragments whose size is much larger than the size of either gene, they are probably some distance away from each other in the embryo genome. However, the possibility that the enzyme cleaved contiguously arranged V and C genes near the boundary is not entirely eliminated. Cleavage sites are possible in the nucleotides coding for the amino acids at positions 93–95 and 99–100 (18). However, the probability that either of these amino acid sequences provides the exact nucleotide sequence required is low.

The pattern of hybridization is completely different in the DNA from the homologous tumor (Fig. 2). Both RNA probes hybridized with a new DNA component of M_r 2.4 million. There are no indications that either of these RNA probes hybridizes with other DNA components above the general background level. These results indicate that both V and C genes, or the entire sequences represented in the mRNA molecule [except for the unlabeled poly(A)sequence], are contained in the 2.4 million M_r component in the tumor genome. The whole RNA hybridizes with this component nearly twice as well as does the 3'-end half, thereby supporting this notion. Hence, the V_κ and C_κ genes, which are most likely some distance away from each other in the embryo genome, are brought together in the plasma cells expressing this particular V_κ gene, presumably to form a contiguous nucleotide stretch. The fact that neither the 6.0 nor the 3.9 million M_r component exists in the plasma cell genome indicates that such rearrangement of immunoglobulin genes takes place in all of the homologous chromosomes.

DISCUSSION

We have shown that the pattern of *Bam*H I DNA fragments that carry immunoglobulin V- or C-gene sequences is completely different in the genomes of mouse embryo cells and of a murine plasmacytoma. The pattern of embryo DNA shows two components, one of which (M_r = 6.0 million) hybridizes with C-gene sequences and the other (M_r = 3.9 million) with V-gene sequences. The pattern of tumor DNA shows a single component that hybridizes with both V-gene and C-gene sequences and that is smaller (M_r = 2.4 million) than either of the components in embryo DNA. The straightforward interpretation of these results is that V_κ and C_κ genes, which are some distance away from each other in the embryo cells, are joined to form a contiguous polynucleotide stretch during differentiation of lymphocytes. Such joining occurs in both of the homologous chromosomes.

An alternative explanation of the results, namely, that accumulation of mutations or base modifications leading to either loss or gain of *Bam*H I sites generated the observed pattern difference, is not impossible. On this view, there would have to be a *Bam*H I site close to the V-C junction in embryo DNA. This *Bam*H I site would have to be lost by mutation or by base modification in the MOPC 321 tumor. By itself, such an alteration would cause the appearance of a single 9.9 million M_r component in the tumor. To achieve the M_r of the single component actually observed in the tumor (2.4 million), there would have to be new *Bam*H I sites created by mutation between the V_κ gene and the nearest site on either side. Since there is no reason why there should be any selective pressures in-

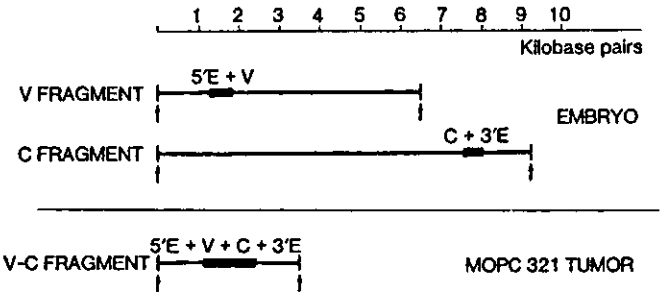


FIG. 3. Mouse DNA fragments carrying V_κ and C_κ genes. DNA fragments were generated by *Bam*H I restriction endonucleases. Arrows indicate *Bam*H I sites. 5'E and 3'E designate base sequences corresponding to untranslated regions of a κ-chain mRNA molecule at the 5'- and 3'-end, respectively. V and C designate base sequences corresponding to variable and constant regions, respectively. The relative position of these sequences within each fragment is deduced from the present results within the framework of either of the latter three models depicted in Fig. 4.

volving *Bam*H I sites, the occurrence of three base alterations would seem to be quite unlikely. Furthermore, such pattern changes are not unique to this particular combination of enzyme and DNA. With *Eco* RI enzyme and embryo and HOPC 2020 DNA (a λ-chain producer) we have obtained results that lead us to a similar conclusion (Hozumi and Tonegawa, unpublished).

There is also an alternative explanation for the absence of the embryonic DNA components in the tumor, namely, that the V–C gene joining took place in only one of the homologous chromosomes and that the other chromosome(s) has been lost during propagation of the tumor. In view of the known chromosome abnormalities of murine plasmacytomas, we cannot eliminate this trivial possibility.

Fig. 3 summarizes our interpretation of the experimental results. What is the mechanism by which the integration of V- and C-gene sequences is brought about? In the past, several models have been proposed, and some of these models are schematically illustrated in Fig. 4. In the "copy-insertion" model a specific V gene is duplicated and the copy is inserted at a site adjacent to a C gene (19). In this model, the embryonic sequence context of a V gene should be retained in the lymphocyte expressing that particular V gene. Since the embryonic DNA fragment carrying the MOPC 321 V gene does not seem to exist in the genome of this tumor, our results are clearly incompatible with this model. In the "excision-insertion" model, a specific V gene is excised into an episome-like structure, and this in turn is integrated adjacent to a C gene (20). In the "deletion" model, DNA in the interval between a particular V gene and the corresponding C gene loops out, is excised, and diluted out upon subsequent cell multiplication (21). In the "inversion" model, V genes and a C gene are arranged in a chromosome in opposite directions and a segment of chromosome between a particular V gene (inclusive) and a C gene (exclusive) is inverted. The latter three models are both consistent with the experimental results presented here.

It is easy to think of variants of these models. In one variant of the copy-insertion model, the C gene is copied and the copy is inserted next to the V gene. In another variant, both V and C genes are copied, and the copies inserted at a third location. These variants are just as incompatible with our results as is the copy-insertion model illustrated in Fig. 4. There are analogous variants of the "excision-insertion" model and they are just as compatible with the experimental results as the excision-insertion model illustrated in Fig. 4.

A "committed" bone marrow-derived (B) lymphocyte and

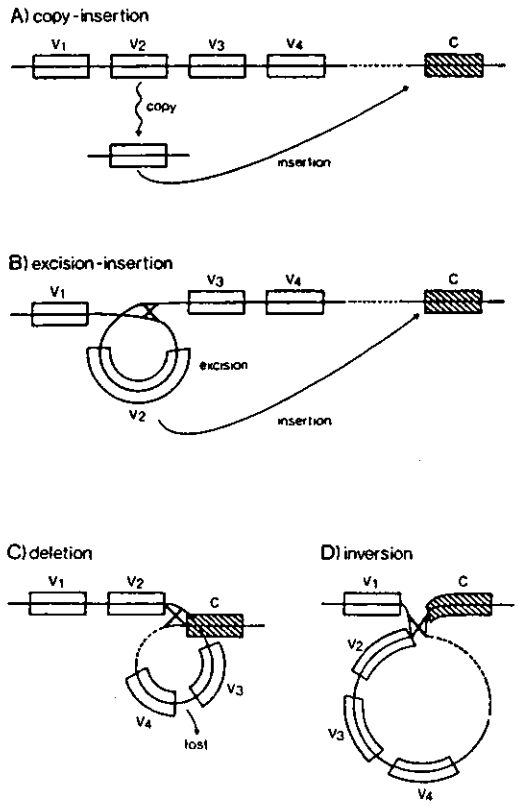


FIG. 4. Models for V–C gene joining at DNA level. See *text* for explanation.

a plasma cell produce antibody of only one specificity (22–24). In particular, it expresses only one light chain V gene. If there are multiple V_κ genes, there must exist a mechanism for the activation of one particular V_κ gene. In the light of the present findings, one intriguing possibility is that activation of a V gene is intimately coupled with its joining to a C gene. For instance, in the "excision-insertion" model, a promotor site may be created by the insertion of the excised V-DNA fragment. This would activate that particular V gene for transcription. In fact, on this view the current terminology of "V genes" and "C genes" is inappropriate. Rather, there are two segments of DNA, one specifying the V region and the other specifying the C region. The gene is *created* by joining.

For immunoglobulin loci, only one allele is expressed in any given lymphocyte (22). This is not the case with most autosomal genes. Our results suggest an interesting explanation for allelic exclusion—that is, that the two homologous chromosomes are, in any given plasma cell, homozygous. Homozygosis could result from the loss of one homologue followed by reduplication of the other or from somatic recombination between the centromere and the immunoglobulin locus (25, 26). The alternative view, that joining takes place in both chromosomes, presents a problem if there is more than one V segment. There is no intrinsic reason why the same V segment should be joined on both homologues.

The present results, when combined with those of our previous reports, demonstrate that both content and context of immunoglobulin genes are altered during differentiation of lymphocytes. Whether such a genetic event is common in other eukaryotic gene systems remains to be seen.

We thank Mr. G. R. Dastoornikoo, Ms. R. Schuller, and Ms. S. Stutz for their expert technical assistance. The "inversion model" was suggested by Dr. C. Steinberg.

1. Hilschmann, N. & Craig, L. C. (1965) *Proc. Natl. Acad. Sci. USA* **53**, 1403–1409.
2. Titani, K., Whitley, E., Avogardo, L. & Putnam, F. W. (1965) *Science* **149**, 1090–1092.
3. Milstein, C. & Munro, A. J. (1970) *Annu. Rev. Microbiol.* **24**, 335–358.
4. Tonegawa, S., Steinberg, C., Dube, S. & Bernardini, A. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 4027–4031.
5. Tonegawa, S., Bernardini, A., Weimann, B. J. & Steinberg, C. (1974) *FEBS Lett.* **40**, 92–96.
6. Tonegawa, S. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 203–207.
7. Faust, C. H., Diggelmann, H. & Mach, B. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 2491–2495.
8. Rabbitts, T. H. (1974) *FEBS Lett.* **42**, 323–326.
9. Stavnezer, J., Huang, R. C. C., Stavnezer, E. & Bishop, J. M. (1974) *J. Mol. Biol.* **88**, 43–63.
10. Honjo, T., Packman, S., Swan, D., Nau, M. & Leder, P. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 3659–3663.
11. Milstein, C., Brownlee, G. G., Cartwright, E. M., Javis, J. M. & Proudfoot, N. J. (1974) *Nature* **252**, 354–359.
12. Wilson, C. A. & Young, F. E. (1975) *J. Mol. Biol.* **97**, 123–125.
13. Sharp, P. A., Sugden, B. & Sambrook, J. (1973) *Biochemistry* **12**, 3055–3063.
14. Stafford, D. W. & Bieber, D. (1975) *Biochim. Biophys. Acta* **378**, 18–21.
15. Fukes, M. & Thomas, C. A., Jr. (1970) *J. Mol. Biol.* **52**, 395–397.
16. Gross-Bellard, M., Dudet, P. & Chambon, P. (1973) *Eur. J. Biochem.* **36**, 32–38.
17. Birnstiel, M. L., Sells, B. H. & Purdon, I. F. (1972) *J. Mol. Biol.* **63**, 21–39.
18. McKean, D., Potter, M. & Hood, L. (1973) *Biochemistry* **12**, 760–771.
19. Dreyer, W. J., Gray, W. R. & Hood, L. (1967) *Cold Spring Harbor Symp. Quant. Biol.* **32**, 353–367.
20. Gally, J. A. & Edelman, G. M. (1970) *Nature* **227**, 341–348.
21. Kabat, D. (1972) *Science* **175**, 134–140.
22. Pernis, B., Chiappino, G., Kelus, A. S. & Gell, P. G. H. (1965) *J. Exp. Med.* **122**, 853–876.
23. Mäkelä, O. (1967) *Cold Spring Harbor Symp. Quant. Biol.* **32**, 423–430.
24. Raff, M. C., Feldman, M. & De Petris, S. (1973) *J. Exp. Med.* **137**, 1024–1030.
25. Ohno, S. (1974) in *Chromosomes and Cancer*, ed. German, J. (John Wiley & Sons, New York) pp. 77–94.
26. Gally, J. A. (1967) *Cold Spring Harbor Symp. Quant. Biol.* **32**, 166–168.

FULL PAPER *Biochemistry*

Analysis of a Full-Length cDNA Library Constructed from Swine Olfactory Bulb for Elucidation of Expressed Genes and Their Transcription Initiation Sites

Seiichiro FUJISAKI[1], Akio SUGIYAMA[2], Tomoko EGUCHI[1], Yasuko WATANABE[1]*, Hideki HIRAIWA[1], Daisuke HONMA[1], Toshiyuki SAITO[3] and Hiroshi YASUE[1]**

[1]*Genome Research Department, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305–8602,* [2]*Tsuruga Institute of Biotechnology, Toyobo Co., Ltd., Tsuruga, Fukui 914–0047 and* [3]*Physiology and Genetic Regulation Department, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305–8602, Japan*

(Received 28 April 2003/Accepted 20 August 2003)

ABSTRACT. The olfactory system is indispensable to the survival of animals in finding foods and for the reproductive process. Odorant signals are conveyed through olfactory sensory neurons to the olfactory bulb, which modifies the signals and relays them to the neocortex. In the present study, a "full-length" cDNA library was constructed from the main and accessory olfactory bulbs of 5-week-old male pigs, in order to elucidate the expressed genes. The average insert size of the library was estimated to be 1.7 kb based on 54 randomly-selected clones. One thousand randomly selected clones were subjected to sequencing, and the resulting 883 sequences were then clustered into 753 sequences based on similarity. Since 723 of the 753 sequences had sufficient sequence information for homology analysis, the 723 sequences were subjected to BLAST analysis against GenBank/EMBL/DDBJ; 655 out of the 723 sequences showed similarities with known genes, and the remaining 68 were indicated to be novel sequences. The full-length rate of the library was estimated to be ca. 80%, using 70 sequences corresponding to human full-length cDNAs. The full-length cDNA sequences of a single gene appearing more than 6 times in the analysis were aligned to determine major transcription initiation sites for *SLC25A, CKB, TUBB4, TUBB, YWHAH, TUBB2,* and *CNP* genes.

KEY WORDS: full-length cDNA library, olfactory bulb, swine, transcription initiation site.

J. Vet. Med. Sci. 66(1): 15–23, 2004

The olfactory system receives and conveys a wide variety of odor molecule information surrounding animals to the neocortex. This is an indispensable process for food seeking, reproduction and other behaviors by which animals survive. The system consists of two different pathways for information transfer; one is from sensory neurons in the olfactory epithelium through the main olfactory bulb to the olfactory cortex, and the other is from sensory neurons in the vomeronasal organ through the accessory olfactory bulb to the amygdala, distinct from the zone that receives information from the main olfactory bulb [4, 6, 9, 21]. Olfactory epithelium contains more than 2 million olfactory sensory neurons in mice, each of which expresses one type of odorant receptor out of ca. 1000 receptor genes [1, 5, 10]. It has been shown that each neuron is connected to a single glomerulus in the main olfactory bulb through its axon, and that the axons present zone-to-zone projection from the olfactory epithelium to the main olfactory bulb [11, 14, 15, 17, 18]. Then, the main olfactory bulb projects the processed information to the olfactory cortex [29].

The accessory olfactory bulb receives pheromone signals from the vomeronasal organ and projects the information mainly to the amygdala. The signals provide information about gender, dominance, or reproductive status and elicit innate social and sexual behaviors [7, 9, 20]. For example, in male rodents, removal of the vomeronasal organ in virgin animals severely impairs sexual responses [7, 13]. The functional difference between the main olfactory bulb and the accessory olfactory bulb were also indicated by the differential expression patterns of metabotropic glutamate receptors in the respective bulbs [19].

Alternative or multi-transcriptional initiation sites have been studied [8, 25, 26] and alternative initiation sites for *cytochrome c* mRNA was demonstrated to be tissue specific [8]. These transcriptional controls were often directed by DNA binding elements specific for cell functions, such as Ets-1 in the functions of lymphoid cells [2]. Therefore, the elucidation of the possible transcriptional patterns in the olfactory bulbs would provide a clue to understand the functions of cells in the olfactory bulbs. In order to obtain such alternative or multi-transcriptional initiation sites for genes, a "full-length" cDNA library was constructed in the present study from the porcine main and accessory olfactory bulbs and clones of the library were subjected to sequence analysis.

MATERIALS AND METHODS

Preparation of swine olfactory bulb: For the preparation of olfactory bulb, two 5-week-old Landrace male pigs weighing about 10 kg were used. All animals received humane care as described in the guide for the Care and Use of Experimental Animals (National Institute of Agrobiological Sciences Care Committee). The procedure for the prep-

* PRESENT ADDRESS: WATANABE, Y., Department of Welfare Engineering, Faculty of Engineering Iwate University, Morioka 020–8551, Japan.

* CORRESPONDENCE TO: Yasue, Y., Genome Research Department, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305–0901, Japan.

aration of the olfactory bulb followed the guideline of animal ethics at the National Institute of Agrobiological Sciences, Japan. The pigs were fasted overnight before preparation. On the day of preparation, the pigs were sedated by intramuscular injection of ketamine (Sankyo Pharmaceutical, Tokyo, Japan; 10 mg/kg body weight) and midazolam (Yamanouchi Pharmaceutical, Tokyo, Japan; 1 mg/kg body weight). After tracheal intubation, the pigs were anesthetized with 2–3% halothane and 0.5 *l*/min nitrous oxide. Under deep general anesthesia, a cannula was inserted into a carotid vein; and the skull was cut and removed. Then, the pigs were euthanized by intravenous injection of 10 m*l* of sodium pentobarbital. Immediately after respiration and heartbeat were stopped, the brain was removed and rinsed in an ice-cold saline solution. The main olfactory bulb together with accessory olfactory bulb was dissected on ice, and the blocks were frozen in liquid nitrogen and stored at –80°C until use.

Construction of oligo-capped cDNA library ("full-length" cDNA library): An oligo-capped cDNA library was constructed as described by Suzuki and Sugano [24] with some modifications. Briefly, 20 μg of poly(A)$^+$RNA isolated from the main and accessory olfactory bulbs was treated with 2.5 units of bacterial alkaline phosphatase (BAP; Takara Co., Ohtsu, Shiga, Japan) in 100 μ*l* of 100 mM Tris-HCl (pH 7.0), 10 mM 2-mercaptoethanol, and 100 units of RNase inhibitor at 37°C for 60 min. The RNA was extracted with phenol:chloroform twice and precipitated with ethanol. The BAP-treated RNA was further treated with 20 units of tobacco acid pyrophosphatase (TAP)[12] in 100 μ*l* of 50 mM sodium acetate (pH 5.5), 1 mM EDTA, 10 mM 2-mercaptoethanol, and 100 units of RNase inhibitor at 37°C for 60 min. After phenol:chloroform extraction and ethanol precipitation, the BAP-TAP-treated RNA was then ligated with 0.4 μg of oligoribonucleotide (5'-AGC AUC GAG UCG GCC UUG UUG GCC UAC UGG-3') using 250 units of RNA ligase (Takara Co., Ohtsu, Shiga, Japan) in 100 μ*l* of 50 mM Tris-HCl (pH 7.0), 5 mM $MgCl_2$, 5 mM 2-mercaptoethanol, 0.5 mM ATP, 25% PEG8000, 100 units RNase inhibitor at 20°C for 3 hr. After removing free oligoribonucleotide, cDNA was synthesized using RNaseH free reverse-transcriptase (Superscript II; Invitrogen Co., Carlsbad, CA, U.S.A) with 10 pmol of oligo-dT adapter-primer (5'-GCG GCT GAA GAC GGC CTA TGT GGC CTT TTT TTT TTT TTT TTT-3') in a condition as recommended by the supplier. After removal of the template RNA by alkaline degradation, the PCR reaction was performed in a volume of 100 μ*l* using an XL PCR kit (Applied Biosystems Co., Foster City, CA, U.S.A) with 16 pmol of 5' (5'-AGC ATC GAG TCG GCC TTG TTG-3') and 3' (5'-GCG GCT GAA GAC GGC CTA TGT-3') PCR primers. PCR amplification cycles were 12 cycles of 94°C for 1 min, 58°C for 1 min, and 72°C for 10 min. PCR products were then extracted with phenol:chloroform, ethanol-precipitated and digested with restriction enzyme *Sfi* I. *Sfi* I-digested PCR products were separated by agarose gel electrophoresis, and cDNA fragments longer than 1,000 bp were isolated and inserted into *Dra*III-digested pME18S-FL3 (Accession No. AB009864). The resulting pME18S-FL3 was introduced into DH5α (Takara Co., Ohtsu, Shiga, Japan) following the procedure described by the manufacturer, and the transformed DH5α was plated on LB-plates containing ampicillin at final concentration of 50 μg/m*l* to develop colonies on the plates.

Approximately 50 colonies were randomly selected from the plates and used to measure sizes of cDNA inserts in the plasmid with InsertCheck (Toyobo Co., Osaka, Japan), in order to estimate the average insert size.

Sequencing and characterization of cDNAs: Colonies were selected randomly and cultured in the LB medium containing ampicilin at the same concentration for the plate. The plasmid DNA was prepared from each culture using MultiScreen (Millipore Co., Bedford, MA, U.S.A) together with Biomek 2000 (Beckman Co., Palo Alto, CA, U.S.A.), and cDNA inserts were sequenced once from the 5' site of cDNA using BigDye™ Terminator Cycle Sequencing Ready Reaction Kit V2.0 and ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems Co., Foster City, CA , U.S.A). The DNA sequences thus obtained were clustered using the software "G-PROF" (Genetyx, Tokyo, Japan), and the consensus sequence of each cluster was subjected to BLAST analysis against DNA sequences in the DNA databases (GenBank/EMBL/DDBJ/) (http://www.ncbi.nlm.nih.gov/Genbank/index.html) and then to ontological analysis using TIGR database (http://tigr-blast.tigr.org/tgi/).

In order to characterize the library, the sequences obtained were examined to determine whether they contained a putative translation start site in comparison with human orthologs using databases (http://elmo.ims.u-tokyo.ac.jp/dbtss/; and http://www.ncbi.nlm.nih.gov/) [27]. In the present examination, when a sequence contained the translation start site, the sequence was defined as "full-length" cDNA. For the calculation of the full-length rate of the library, an appropriate number of the sequences corresponding to the human genes with Expect (e) values smaller than e^{-100} (e-value: http://www.ncbi.nlm.nih.gov/BLAST/tutorial/Altschul-1.html) were selected under the conditions that the cDNAs of the corresponding human genes had sequence information upstream of translation start site. The swine cDNA sequences thus selected were compared to the respective human cDNA sequences to estimate the full-length rate of the library.

The full-length cDNA sequences of a single gene appearing multiple times in the present analysis were aligned to determine major transcription initiation sites of the gene. Then, the initiation sites were compared with those of human orthologs, if available, using databases (http://elmo.ims.u-tokyo.ac.jp/dbtss/; and http://www.ncbi.nlm.nih.gov/) [23, 27].

RESULTS AND DISCUSSION

Sequencing and characterization of cDNAs: The com-

plexity of the cDNA library constructed according to the procedure described above was examined, indicating that the library consisted of 2.16×10^5 clones. Then, 54 clones randomly selected from the cDNA library were characterized to reveal that the average cDNA insert size of the library was 1.7 kb (data not shown), not smaller than for other full-length cDNA libraries [22]. Concerning 3' region, it was considered that cDNA clones of the library contained polyA sequences judged from the procedure for the library construction. However, to make sure, this was examined by sequencing 200 randomly-selected clones and the examination revealed that all the clones contained polyA sequences (data not shown), strongly indicating that all the clones of the library contain polyA at their 3' termini. Therefore, the present library was subjected to further analysis. One thousand clones selected randomly from the library were sequenced, 883 of which were found to contain sequences of cDNA insert. These sequences were clustered into 753 by using the software G-PROF and were reported to the DNA databases with the successive accession numbers AV295960 to AV296712. Seven hundred twenty-three out of 753 sequences, having sufficient sequence information for homology analysis, were then subjected to BLAST analysis against sequences in the DNA databases (GenBank/EMBL/DDBJ at the level of December, 10, 2001).

The sequences of a single gene appearing more than twice in the 883 sequences were shown in Table 1. The cluster with the largest number of sequences was revealed to correspond to *ADP/ATP carrier*; that with the second largest, to *creatine kinase B*; and that with the third largest, to *tubulin beta 4*. The BLAST analysis has revealed that 655 clusters consisting of 812 sequences showed similarities to known genes with e-values smaller than e^{-100} or similarities exclusively to cDNAs of a specific gene of various species though their e-values were not smaller than e^{-100}. The remaining 68 clusters consisting of 71 sequences showed little similarity with the sequences of various species including human registered in the DNA databases (data not shown), indicating that orthologous sequences have not yet been identified in other species including human, or that the sequences of 68 clusters might be specific to genomes of swine and its phylogenetically-related species. In the latter, it is surmised that the sequences, which had existed in a common ancestor of mammalian species including swine and human, were deleted at least in the ancestry species of human.

Ontological analysis was performed on the 723 independent sequences against the TIGR database at the level as of December, 10, 2001 (http://tigrblast.tigr.org/tgi/). As shown in Fig. 1A, 35% of the sequences were categorized in terms of their function (http://www.godatabase.org/cgi-bin/go.cgi); 45% were not categorized in the database, though their functions were identified; and 20% were unknown or showed no similarities with sequences in the database. When the sequences grouped in the functionally categorized cluster were further classified using the database, 25.6% of

Table 1. Sequences of a single gene appearing more than twice in the 883 sequences

Swine cDNA accession No.	Number of clones	NCBI BLAST Accession	Gene name	e-value
AU295961	31	NM_001152.1	Homo sapiens solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 5 (SLC25A5), nuclear gene encoding mitochondrial protein, mRNA	$1e^{-60}$
		M24103.1	Bovine ADP/ATP translocase T2 mRNA, complete cds	0.0
AU295962	10	NM_001823.1	Homo sapiens creatine kinase, brain (CKB), mRNA	0.0
AU295963	9	NM_006086.1	Homo sapiens tubulin, beta, 4 (TUBB4), mRNA	0.0
AU295964	8	NM_001069.1	Homo sapiens tubulin, beta polypeptide (TUBB), mRNA	0.0
AU295965	8	NM_003405.1	Homo sapiens tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta polypeptide (YWHAH), mRNA	0.0
AU295966	7	NM_006088.1	Homo sapiens tubulin, beta, 2 (TUBB2), mRNA	0.0
AU295967	7	NM_033133.1	Homo sapiens 2',3'-cyclic nucleotide 3' phosphodiesterase (CNP), mRNA	0.0
AU295968	6	NM_005165.1	Homo sapiens aldolase C, fructose-bisphosphate (ALDOC), mRNA	0.0
AU295969	6	NM_002866.1	Homo sapiens RAB3A, member RAS oncogene family (RAB3A), mRNA	0.0
AU295975	6	NM_002116.2	Homo sapiens major histocompatibility complex, class I, A (HLA-A), mRNA	$1e^{-141}$
AU295970	5	NM_015710.1	Homo sapiens glioma tumor suppressor candidate region gene 2(GLTSCR2), mRNA	0.0
AU295971	5	NM_030926.1	Homo sapiens integral membrane protein 3 (ITM3), mRNA	0.0
AU295973	4	NM_014225.1	Homo sapiens protein phosphatase 2 (formerly 2A), regulatory subunit A (PR 65), alpha isoform (PPP2R1A), mRNA	0.0
AU295974	4	NM_005514.2	Homo sapiens major histocompatibility complex, class I, B (HLA-B), mRNA	$1e^{-143}$
AU295976	3	NM_001975.2	Homo sapiens enolase 2, (gamma, neuronal) (ENO2), mRNA	0.0
AU295977	3	NM_006211.1	Homo sapiens proenkephalin (PENK), mRNA	$6e^{-95}$
AU295978	3	NM_057161.2	Homo sapiens testis intracellular mediator protein (PEAS), mRNA	0.0
AU295979	3	M10098	Human 18S rRNA gene, complete	0.0
AU295980	3	NM_006374.1	Homo sapiens serine/threonine kinase 25 (STE20 homolog, yeast)(STK25), mRNA	$1e^{-134}$
AU295982	3	NM_012109.1	Homo sapiens brain-specific membrane-anchored protein (BSMAP), mRNA	$3e^{-52}$
AU295983	3	NM_003768.1	Homo sapiens phosphoprotein enriched in astrocytes 15 (PEA15), mRNA	0.0
AU295984	3	NM_020465.1	Homo sapiens NDRG family, member 4 (NDRG4), mRNA	$1e^{-167}$

them were found to be for ligand binding or carrier; 25.2%, for intracellular proteins; 16.8%, for proteins of cell fraction (defined as parts of cells prepared by disruptive biochemical techniques), and 13.9%, for proteins involved in cell growth and maintenance (Fig. 1B). These largest 4 fractions totaled up to more than 75% of the sequences. The fractions shown in Fig.1B was quite different from those reported in swine liver [16] and from those in human tissues other than olfactory bulb [26], indicating the expressed sequences were closely related to the function of olfactory bulbs.

The transcript of *ADP/ATP carrier* gene (*SLC25A5*) with the largest amount in the 883 sequences (Table 1) was reported to decrease when the cells were induced to differentiate in human [3]. In addition, tubulin genes, the products of which are involved in cell structure and basic cell functions, were expressed in the 3rd, 4th, and 6th largest amounts (Table 1). These findings, taken together with the fact that 13.9% of expressed genes were related to proteins of cell growth and maintenance as described above, led us to infer that the olfactory bulbs in the 5-week-old pigs used in the present study were at the stage of their volume increase, that is, of their development. This inference is in good agreement with the report that the brains of 5-week-old pigs were in the stage of volume increase [28].

Full-length rate of the library and determination of transcription initiation sites using full-length cDNAs: In order to estimate the full-length rate of the library, 70 sequences

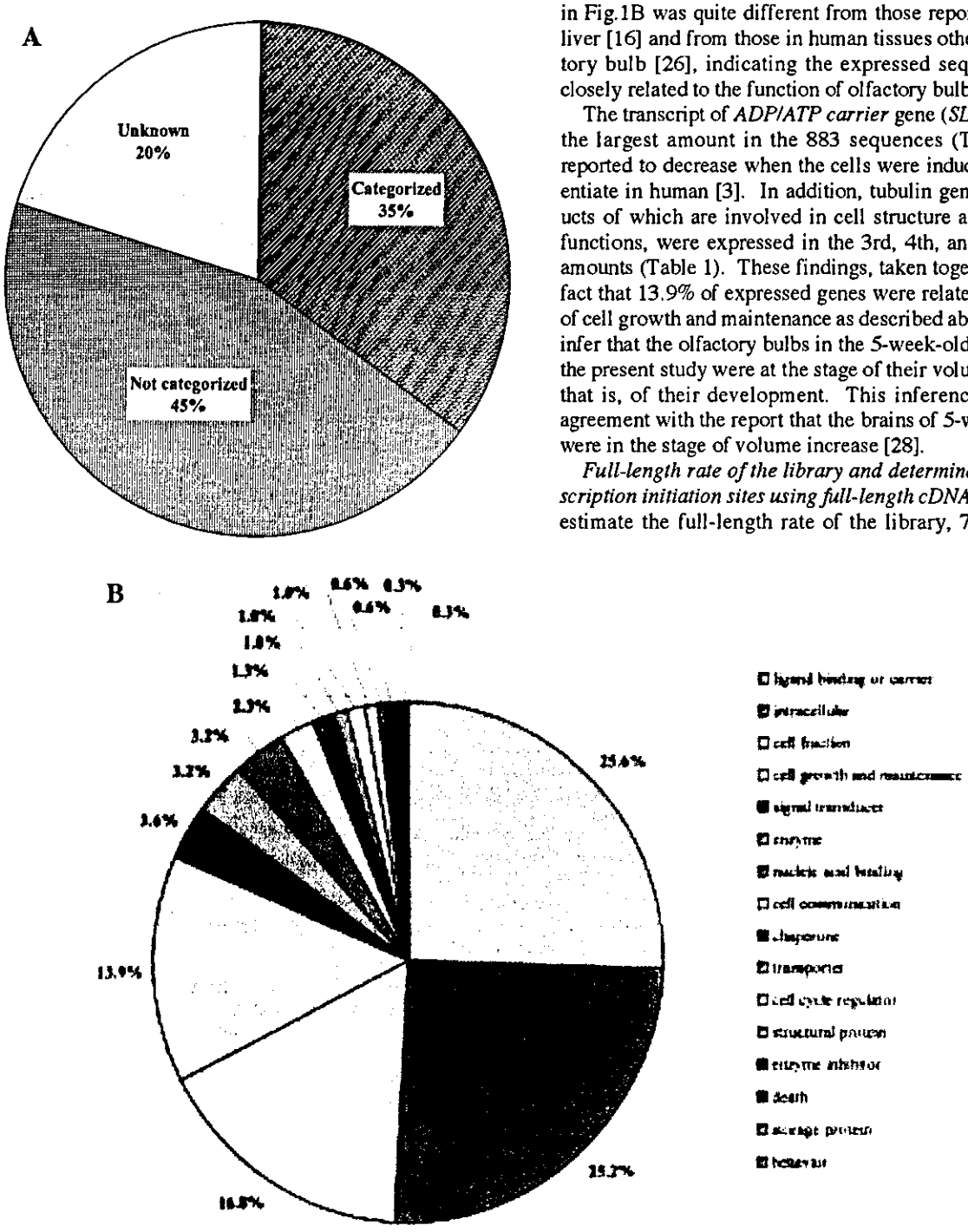
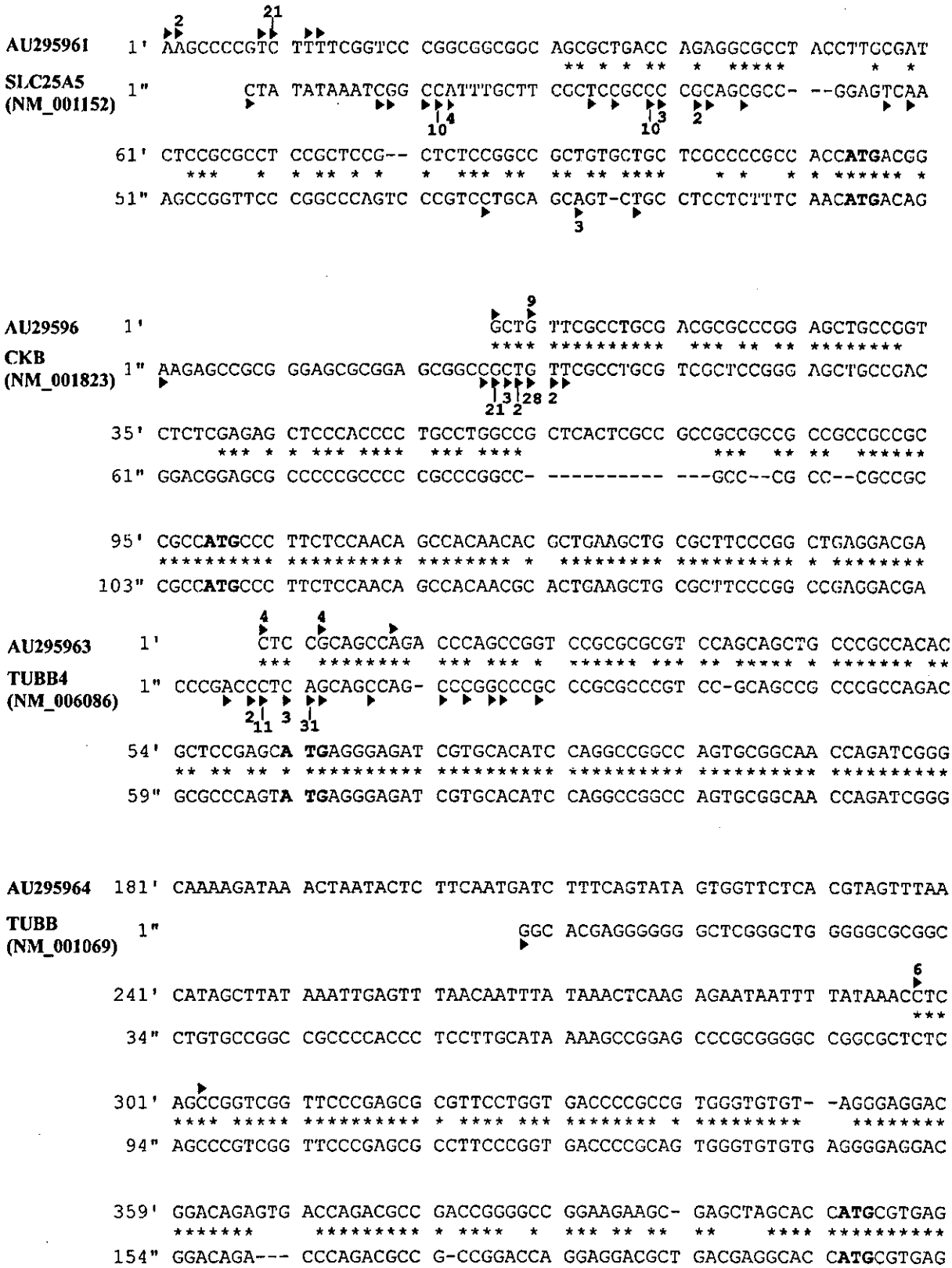


Fig. 1. Results of ontological analysis on 723 independent sequences expressed in the swine main and accessory olfactory bulbs using TIGR database. The 883 sequences obtained in the present study were clustered to 723 independent sequences, which were then classified into 3 groups, i.e., the functionally categorized group (Categorized), the functionally not-categorized group (Not categorized), in the unidentified group (Unknown) using TIGR database of Dec. 10, 2001 (http://tigrblast.tigr.org/tgi/) (Panel A). "Categorized" sequences were then further grouped by function using the preceding database (Panel B).

Table 2. Full-length of the 70 sequences randomly selected under the criteria described in Materials and Methods

Clone name/ accession number	Accession number	Gene	e-value	Length from CDS	Full/Not full
1-1_C07/AU295991	NM_001494.2	GDI2	0.0	+140	Not Full
1-1_F11/AU296073	NM_001069.1	TUBB	e^{-120}	+287	Not Full
10-1_F05/AU296444	NM_021999.1	Homo sapiens transmembrane protein BRI	e^{-130}	+92	Not Full
11-1_E05/AU296499	NM_000627.1	Human transforming growth factor-beta 1 binding protein	0.0	+804	Not Full
13-1_A06/AU296535	NM_022751.1	FLJ21610	0.0	+428	Not Full
13-1_A09/AU296538	NM_000592.3	Human complement component C4B	e^{-121}	+2905	Not Full
13-1_B03/AU296541	NM_006311.1	Homo sapiens nuclear receptor co-repressor N-CoR	e^{-134}	+6176	Not Full
14-1_C06/AU296617	NM_002199.2	IRF2	e^{-176}	+400	Not Full
14-1_D12/AU295987	NM_006703.1	NUDT3	0.0	+303	Not Full
15-1_A11/AU296665	NM_006460.1	Homo sapiens mRNA for HEXIM1 protein	e^{-126}	+61	Not Full
15-1_F09/AU296696	NM_014841.1	SNAP91	e^{-109}	+486	Not Full
7-1_F10/AU296250	U09609.1	Human p80HT (p80HT/NKFB-2)	e^{-158}	+1230	Not Full
8-1_H08/AU296338	NM_003442.1	ZNF143	e^{-156}	+3691	Not Full
9-1_D09/AU296372	NM_021189.1	NECL1	e^{-176}	+24	Not Full
1-1_G03/AU296076	NM_004581.1	RABGGTA	e^{-165}	-110	Full
1-1_G07/AU296077	NM_006400.2	DCTN2	0.0	-95	Full
10-1_F08/AU296445	NM_002861.1	PCYT2	0.0	-39	Full
11-1_E01/AU296497	NM_014392.1	D4S234E	0.0	-103	Full
13-1_B02/AU295998	NM_033296.1	PGR1	0.0	-63	Full
13-1_C06/AU296000	NM_006936.1	SMT3H1	e^{-101}	-114	Full
9-1_B04/AU295973	NM_014225.1	Human protein phosphatase 2A regulatory subunit alpha-isotype (alpha-PR65)	0.0	-18	Full
9-1_B06/AU296378	NM_000487.3	ARSA	e^{-134}	-138	Full
5-1_B05/AU296101	NM_004390.1	CTSH	e^{-162}	-66	Full
1-1_C04/AU296043	NM_000304.1	PMP22	e^{-143}	-63	Full
1-1_C08/AU295966	NM_006088.1	TUBB2	0.0	-90	Full
1-1_F12/AU296074	NM_006275.1	SFRS6	0.0	-138	Full
1-1_G05/AU295961	NM_001152.1	Human ADP/ATP carrier protein	0.0	-102	Full
1-1_G06/AU295995	NM_002629.1	PGAM1	0.0	-53	Full
1-1_G08/AU296078	NM_005276.1	Homo sapiens L-glycerol-3-phosphate	0.0	-51	Full
10-1_C01/AU295990	NM_005389.1	PCMT1	0.0	-178	Full
11-1_D05/AU296490	NM_004040.1	ARHC	e^{-113}	-380	Full
11-1_D07/AU296492	NM_080600.1	Human myelin-associated glycoprotein (MAG)	e^{-137}	-265	Full
11-1_E03/AU296014	NM_005007.1	Homo sapiens 10IHW 9006 IkBL protein (NFKBIL1)	0.0	-72	Full
11-1_E06/AU295962	NM_001823.1	Homo sapiens creatine kinase B	0.0	-89	Full
13-1_A07/AU296536	NM_003713.1	PPAP2B	e^{-103}	-391	Full
13-1_A10/AU296539	NM_022003.1	FXYD6	e^{-104}	-343	Full
13-1_A11/AU296540	NM_052950.1	Homo sapiens WD40- and FYVE-domain containing protein 2 (WDF2)	0.0	-352	Full
14-1_B12/AU296007	NM_001069.1	TUBB	0.0	-79	Full
14-1_C05/AU295969	NM_002866.1	Homo sapiens GTP-binding protein (RAB3A)	0.0	-118	Full
14-1_C08/AU295978	NM_057161.2	SORT1	0.0	-139	Full
14-1_E02/AU296630	NM_022171.1	TCTA	e^{-111}	-41	Full
15-1_A09/AU296664	NM_003779.2	B4GALT3	e^{-175}	-232	Full
15-1_A10/AU296013	NM_001628.1	AKR1B1	e^{-103}	-42	Full
15-1_F06/AU295963	NM_006086.1	TUBB4	0.0	-63	Full
15-1_F07/AU296695	NM_006628.1	Homo sapiens okadaic acid-inducible and cAMP-regulated phosphoprotein 19 (ARPP-19)	e^{-147}	-150	Full
15-1_F10/AU296697	NM_000046.1	Human arylsulfatase B (ASB)	e^{-110}	-41	Full
15-1_F11/AU295962	NM_001824.1	Homo sapiens creatine kinase B	0.0	-92	Full
15-1_F12/AU296017	NM_003795.1	SNX3	0.0	-212	Full
15-1_G03/AU296699	NM_006000.1	Tuba4	e^{-166}	-61	Full
15-1_G05/AU296701	NM_005047.1	Homo sapiens, Similar to proteasome (prosome, macropain) 26S subunit, non-ATPase, 5	0.0	-10	Full
5-1_B03/AU295970	NM_015710.1	GLTSCR2	0.0	-11	Full
6-1_H01/AU295978	NM_057161.2	PEAS	0.0	-139	Full
6-1_H06/AU295970	NM_015710.1	GLTSCR2	0.0	-10	Full
6-1_H07/AU296205	NM_005520.1	HNRPH1	0.0	-69	Full
6-1_H11/AU295963	NM_006086.1	TUBB4	0.0	-52	Full
7-1_B09/AU296007	NM_001069.1	TUBB	0.0	-78	Full
7-1_G03/AU295980	NM_006374.1	Homo sapiens mRNA for YSK1	e^{-134}	-226	Full
9-1_D11/AU295963	NM_006086.1	TUBB4	0.0	-64	Full
9-1_E05/AU296377	NM_002394.1	SLC3A2	e^{-115}	-135	Full
9-1_E06/AU296378	NM_000487.3	ARSA	e^{-134}	-139	Full
9-1_E07/AU296379	NM_016256.1	Homo sapiens N-acetylglucosamine-1-phosphodiester alpha-N-acetylglucosaminidase	e^{-122}	-18	Full
9-1_G03/AU295976	NM_001975.2	ENO2	0.0	-74	Full
9-1_G05/AU296016	NM_031992.1	Wbscr1	0.0	-33	Full
9-1_G07/AU296017	NM_003795.1	SNX3	0.0	-239	Full
5-1_A07/AU296093	NM_016142.1	Homo sapiens steroid dehydrogenase homolog	e^{-173}	-125	Full
5-1_A08/AU296094	NM_002568.1	Homo sapiens testis-specific poly(A)-binding protein (PABP)	0.0	-369	Full
5-1_A09/AU296095	XM_048295.3	Homo sapiens, clone MGC:23208 IMAGE:4872481	0.0	-453	Full
5-1_A10/AU295967	NM_033133.1	CNP	0.0	-90	Full
5-1_B02/AU296098	NM_004305.1	BIN1	0.0	-100	Full
5-1_B04/AU296100	NM_003016.1	Human splicing factor SC35	e^{-156}	-224	Full


AU295961
SLC25A5
(NM_001152)
AU29596
CKB
(NM_001823)
AU295963
TUBB4
(NM_006086)
AU295964
TUBB
(NM_001069)

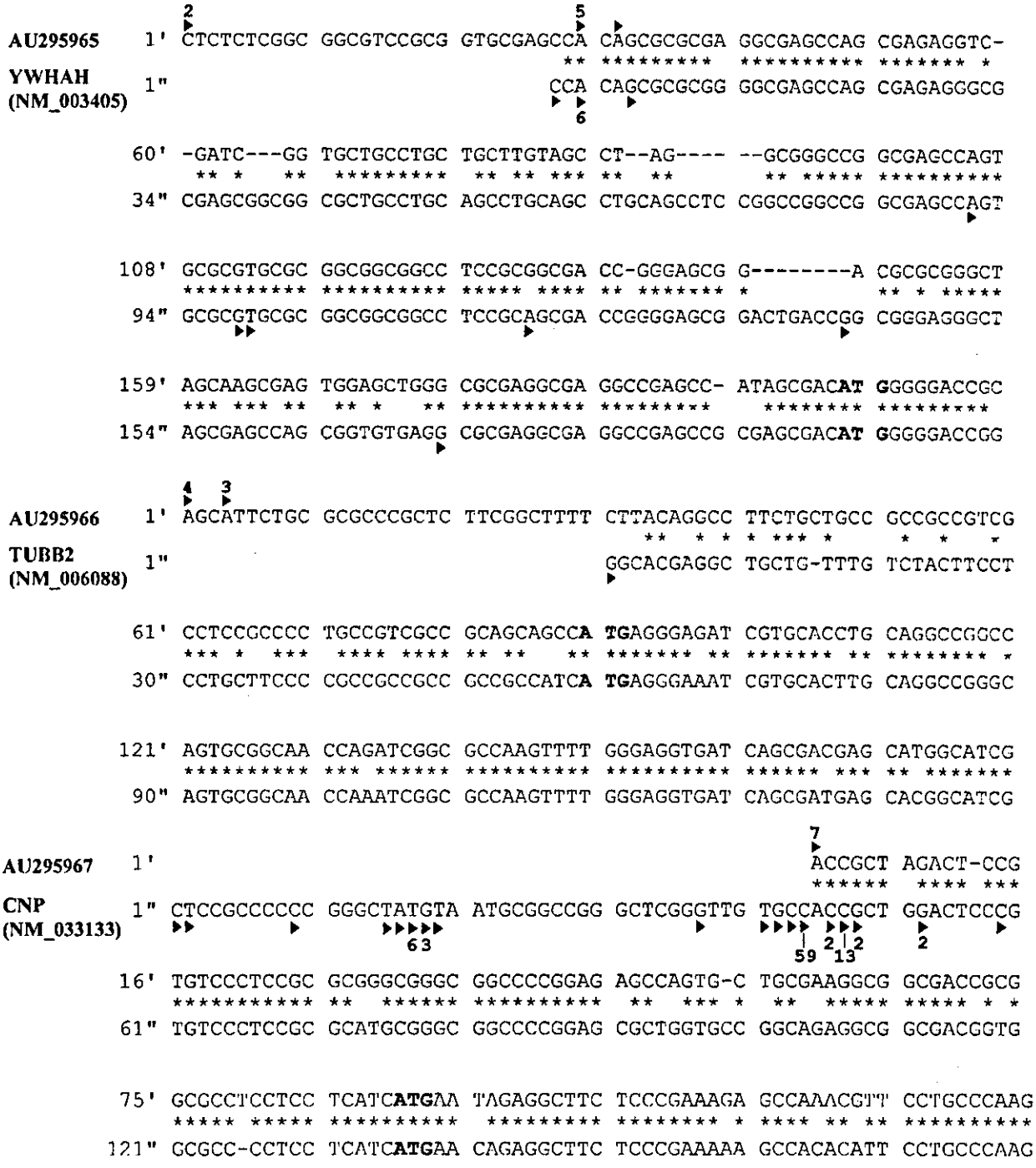


Fig. 2. Major transcriptional initiation sites for 7 genes, *SLC25A*, *CKB*, *TUBB4*, *TUBB*, *YWHAH*, *TUBB2*, and *CNP*. Sequences of a single gene appearing more than 6 times in the 883 sequences were selected (see Table 1). The 7 genes of the sequences thus selected corresponded to human genes *SLC25A*, *CKB*, *TUBB4*, *TUBB*, *YWHAH*, *TUBB2*, and *CNP*. The sequences were aligned for each gene to determine the major transcriptional initiation sites and the results of the alignment were presented in this figure. Since human major transcriptional initiation sites for *SLC25A*, *CKB*, *TUBB4*, *YWHAH*, and *CNP* genes were available (http://elmo.ims.u-tokyo.ac.jp/dbtss/), those sites were also presented on human sequences as a reference. For the remaining genes, the longest cDNA sequence of 5' upstream region registered in the NCBI was presented as a reference (http://www.ncbi.nlm.nih.gov/). Closed triangles indicate transcriptional initiation sites; numerals attached to the triangles indicate the number of transcriptional initiation, and the triangles without numerals present no information for the transcriptional initiation frequency.

selected under the conditions described in Materials and Methods were examined to determine whether those sequences contained the putative translation start site; 54 out of the 70 sequences contained the translation start site and its flanking 5' upstream sequences (Table 2), indicating that ca. 80% of cDNA clones in the present library were full-length sequences.

Examination of the sequences of the 22 genes that appeared more than two times in the 883 sequences (Table 1) revealed that all were full-length sequences. Although the 5'-terminus of each sequence correspond to transcriptional initiation site of the gene, it is not possible to determine the major transcriptional initiation sites of each gene, if the number of the sequences for a given gene is limited. Therefore, the full-length sequences of a single gene appearing more than 6 times were aligned, showing no sequence polymorphism in the regions upstream of translation start sites. The 5' termini of the sequences were scored for identification of major transcriptional initiation sites of the respective genes. As a result shown in Fig. 2, the major transcriptional initiation sites of the 7 genes, which were orthologous to human genes, *SLC25A, CKB, TUBB4, TUBB, YWHAH, TUBB2,* and *CNP,* respectively, have been demonstrated. Since the transcription initiation sites of the 5 human genes, i.e., *SLC25A, CKB, TUBB4, YWHAH,* and *CNP* were shown in the website (http://elmo.ims.u-tokyo.ac.jp/dbtss/), the sites of those genes were presented in Fig. 2 as a reference. For the remaining 2 genes, *TUBB* and *TUBB2,* the longest sequence registered in the NCBI for each gene was presented in Fig. 2 as a reference (http://www.ncbi.nlm.nih.gov/). In *CKB, TUBB4, YWHAH,* and *CNP* genes presenting similarities in the upstream region of the translation start site as well as in the protein coding region between swine and human, the porcine major transcriptional initiation sites were demonstrated to be similar with those of human orthologs. This finding may suggest that a similar mechanism is operating for transcriptional regulation of those genes in swine and human. In contrast, the sequence of *SLC25A* between transcriptional initiation site and translation start site showed little similarity with the human counterpart, but a high similarity in protein coding regions. In addition, one major transcription initiation site was detected in porcine *SLC25A* gene, whereas two major transcription initiation sites were observed in the human gene (Fig. 2); the porcine initiation site was different from the human initiation sites in the distance from the translation start site. These facts may indicate that the expression of porcine *SLC25A* gene is controlled by a mechanism different from human.

In conclusion, the full-length cDNA library of porcine main and accessory olfactory bulbs constructed in the present study was demonstrated to contain virtually 80% full-length cDNAs. The sequence analysis of clones in the library provided not only information on the expressed genes in the porcine olfactory bulbs, but also on transcriptional initiation sites of genes. In our recent analysis, another type of *TUBB* transcript has been found in a full-length library of porcine hippocampus; this transcript was almost identical with the *TUBB* transcript of olfactory bulbs (present study) in protein coding sequence (similarity: 1,336 bp/1,338 bp or 443aa/445aa), but exhibited little similarity in the region between the transcriptional initiation sites and translation start site (unpublished data). When these findings were taken together, the extensive analysis of the full-length cDNA libraries would contribute to understand the function of the brain tissues including olfactory bulbs.

ACKNOWLEDGEMENT. This study was partly supported by a research grant from the Bio-oriented Technology Research Advancement Institution.

REFERENCES

1. Axel, R. 1995. The molecular logic of smell. *Sci. Am.* **273**: 154–159.
2. Bassuk, A.G. and Leiden, J.M. 1997. The role of Ets transcription factors in the development and function of the mammalian immune system. *Adv. Immunol.* **64**: 65–104.
3. Battini, R., Ferrari, S., Kaczmarek, L., Calabretta, B., Chen, S.T. and Baserga, R. 1987. Molecular cloning of a cDNA for a human ADP/ATP carrier which is growth-regulated. *J. Biol. Chem.* **262**: 4355–4359.
4. Broadwell, R.D. 1975. Olfactory relationships of the telencephalon and diencephalon in the rabbit. I. An autoradiographic study of the efferent connections of the main and accessory olfactory bulbs. *J. Comp. Neurol.* **163**: 329–345.
5. Buck, L.B. 1996. Information coding in the vertebrate olfactory system. *Annu. Rev. Neurosci.* **19**: 517–544.
6. Buck, L.B. 2000. The molecular architecture of odor and pheromone sensing in mammals. *Cell* **100**: 611–618.
7. Dulac, C. and Axel, R. 1995. A novel family of genes encoding putative pheromone receptors in mammals. *Cell* **83**: 195–206.
8. Hake, L.E. and Hecht, N.B. 1993. Utilization of an alternative transcription initiation site of somatic cytochrome c in the mouse produces a testis-specific cytochrome c mRNA. *J. Biol. Chem.* **268**: 4788–4797.
9. Halpern, M. 1987. The organization and function of the vomeronasal system. *Annu. Rev. Neurosci.* **10**: 325–362.
10. Hildebrand, J.G. and Shepherd, G.M. 1997. Mechanisms of olfactory discrimination: converging evidence for common principles across phyla. *Annu. Rev. Neurosci.* **20**: 595–631.
11. Kashiwadani, H., Sasaki, Y.F., Uchida, N. and Mori, K. 1999. Synchronized oscillatory discharges of mitral/tufted cells with different molecular receptive ranges in the rabbit olfactory bulb. *J. Neurophysiol.* **82**: 1786–1792.
12. Maruyama, K. and Sugano, S. 1994. Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene* **138**: 171–174.
13. Meredith, M. 1986. Vomeronasal organ removal before sexual experience impairs male hamster mating behavior. *Physiol. Behav.* **36**: 737–743.
14. Mori, K., Nagao, H. and Yoshihara, Y. 1999. The olfactory bulb: coding and processing of odor molecule information. *Science* **286**: 711–715.
15. Mori, K., von Campenhause, H. and Yoshihara, Y. 2000. Zonal organization of the mammalian main and accessory olfactory systems. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **355**: 1801–1812.

16. Ozawa, A., Yasue, H. and Hanazono, M. 1997. Analysis of porcine liver cDNA library and application for gene expression and genome mapping research. *Kenkyu J.* **20**: 20–25 (in Japanese).
17. Pinching, A.J. and Powell, T.P. 1971. The neuropil of the glomeruli of the olfactory bulb. *J. Cell. Sci.* **9**: 347–377.
18. Royet, J., Distel, H., Hudson, R. and Gervais, R. 1998. A re-estimation of the number of glomeruli and mitral cells in the olfactory bulb of rabbit. *Brain Res.* **788**: 35–42.
19. Sahara, Y., Kubota, T. and Ichikawa, M. 2001. Cellular localization of metabotropic glutamate receptors mGluR1, 2/3, 5 and 7 in the main and accessory olfactory bulb of the rat. *Neurosci. Lett.* **312**: 59–62.
20. Sam, M., Vora, S., Malnic, B., Ma, W., Novotny, M.V. and Buck, L.B. 2001. Neuropharmacology. Odorants may arouse instinctive behaviours. *Nature (Lond.)* **412**: 142.
21. Scalia, F. and Winans, S.S. 1975. The differential projections of the olfactory bulb and accessory olfactory bulb in mammals. *J. Comp. Neurol.* **161**: 31–55.
22. Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano, S. 1997. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. *Gene* **200**: 149–156.
23. Suzuki, Y., Ishihara, D., Sasaki, M., Nakagawa, H., Hata, H., Tsunoda, T., Watanabe, M., Komatsu, T., Ota, T., Isogai, T., Suyama, A. and Sugano, S. 2000. Statistical analysis of the 5' untranslated region of human mRNA using "Oligo-Capped" cDNA libraries. *Genomics* **64**: 286–297.
24. Suzuki, Y. and Sugano, S. 2001. Construction of full-length-enriched cDNA libraries. The oligo-capping method. *Methods. Mol. Biol.* **175**: 143–153.
25. Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakaki, Y., Nakamura, Y., Suyama, A. and Sugano, S. 2001. Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites. *EMBO Rep.* **2**: 388–393.
26. Suzuki, Y., Tsunoda, T., Sese, J., Taira, H., Mizushima-Sugano, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., Nakamura, Y., Suyama, A., Sakaki, Y., Morishita, S., Okubo, K. and Sugano, S. 2001. Identification and characterization of the potential promoter regions of 1031 kinds of human genes. *Genome Res.* **11**: 677–684.
27. Suzuki, Y., Yamashita, R., Nakai, K. and Sugano, S. 2002. DBTSS: DataBase of human Transcriptional Start Sites and full-length cDNAs. *Nucleic Acids Res.* **30**: 328–331.
28. van Eerdenburg, F.J., Poot, P., Molenaar, G.J., van Leeuwen, F.W. and Swaab, D.F. 1990. A vasopressin and oxytocin containing nucleus in the pig hypothalamus that shows neuronal changes during puberty. *J. Comp. Neurol.* **301**: 138–146.
29. Zou, Z., Horowitz, L.F., Montmayeur, J.P., Snapper, S. and Buck, L.B. 2001. Genetic tracing reveals a stereotyped sensory map in the olfactory cortex. *Nature (Lond.)* **414**: 173–179.

Journal of Reproduction and Development, Vol. 49, No. 2, 2003

—Original—

Trophinin is Expressed in the Porcine Endometrium During the Estrous Cycle

Saeko NAKANO[1,4], Hisashi KISHI[1], Hidehiko OGAWA[3], Hiroshi YASUE[1], Akira OKANO[2] and Kiyoshi OKUDA[4]

[1]Genome Research Group, National Institute of Agrobiological Sciences, Tukuba, Ibaraki 305–8602, [2]Department of Animal Breeding and Reproduction, National Institute of Livestock and Grassland Sciences, Tsukuba Norin-kenkyu-danchi P.O. Box 5, Ibaraki 305–0901, [3]Department of Bioscience, Faculty of Applied Bioscience, Tokyo University of Agriculture, Tokyo 156–8502 and [4]Laboratory of Reproductive Endocrinology, Graduate school of Natural Science and Technology, Okayama University, Okayama 700–8530, Japan

Abstract. We investigated endometrial expression of trophinin mRNA and protein, homophilic cell adhesion molecules, during the estrous cycle of gilts. An immunopositive reaction for trophinin was observed in the luminal and glandular epithelia of the endometrium at all stages of the estrous cycle, but not in endometrial stromal cells or the myometrium. A partial coding sequence of porcine trophinin was similar to sequences in humans and mice, with homologies of 75% and 70%, respectively. As in humans and mice, the trophinin gene is expressed in the endometrium. Trophinin, however, is expressed in the endometrium of the pig throughout the estrous cycle, higher expression levels were observed at some points of the luteal phase, as in humans. These findings suggest that regulation of trophinin gene expression in the pig is different from that in mice, but similar to that in humans. Furthermore, the present results suggest that the pig might be a suitable model for studying the physiological importance of trophinin in early pregnancy in humans.

Key words: Trophinin, Semi-quantitative assay, Endometrium, Gilt

(J. Reprod. Dev. 49: 127–134, 2003)

Adhesion molecules have an important role in attachment of the embryo to the maternal uterine epithelium. Among the adhesion molecules, integrins are perhaps the best characterized. They are involved in both cell to cell and cell to ECM interactions including attachment of the conceptus to the epithelium [2, 6, 9, 14, 25]. Some integrins are present on the endometrial epithelium of humans [1, 12, 13, 15, 16, 23, 24], rodents [19] and pigs [3, 4], and also certain integrins are constitutively expressed at high levels on the uterine luminal epithelium throughout the estrous cycle [3, 4]. Another adhesion molecule whose role is not well understood is trophinin. Trophinin was initially identified as an intrinsic membrane protein that mediated apical cell adhesion between trophoblastic teratocarcinoma HT-H and endometrial adenocarcinoma SNG-M cells [7].

In humans, trophinin is not expressed in the endometrium during the proliferation phase, whereas in primates [7, 22] and mice [21], high levels of trophinin transcripts and proteins are detected in the trophoblasts and the endometrial epithelial cells at the implantation site of the endometrium during early pregnancy.

Accepted for publication: November 27, 2002
A. Okano (aokano@affrc.go.jp)

Experiments with ovariectomized mice have shown that estrogen, but not progesterone, is required for trophinin expression in the uterus [21]. Implanting blastocysts has no effect on trophinin expression in the endometrial epithelial cells of the mouse, although strong trophinin expression is found in the endometrial epithelial cells at the implantation site [18, 21]. Recently, trophinin expression was detected in the luteal phase endometrium of both normal and infertile women with peak expression in the mid-luteal phase [27].

Together, these findings indicate that trophinin plays an important role in implantation in the mammalian species so far examined, although different species may have different mechanisms or other factors by which the trophinin gene is regulated. For the study of porcine trophinin, we have initiated experiments to characterize whether or not the porcine trophinin is expressed during the estrous cycle. In the present study, we attempted to sequence the mRNA of porcine trophinin, and to observe the pattern of trophinin distribution in the endometrium throughout the estrous cycle, using an immunohistochemical technique. Semi-quantitative expression levels of porcine trophinin were also determined using the polymerase chain reaction (PCR) during the estrous cycle.

Materials and Methods

Animals

Sexually mature crossbred (Landrace × white Yorkshire × Deurock) gilts raised at the National Institute of Livestock and Grassland Sciences were used after being checked for two consecutive estrous cycles. The day when the animals exhibited behavioral estrous activity was designated as day 0 of the estrous cycle.

Sample collection

Each group of animals was sacrificed on days 4, 8, 12, 15, 18 and 20 of the cycle and day 15 of pregnancy. The uterus was removed immediately after electric stunning and exsanguination. A part of endometrial tissue from one of the uterine horns was mechanically collected and stored at –80 C until extraction of total RNA. The other uterine horn was cut to a length of 3–5 cm and embedded into O.C.T. compound (SAKURA Finetechnical, Tokyo, Japan) in liquid nitrogen vapor, and stored at –80 C until the immunohistochemical study.

Immunohistochemistry

Cryosections (6–8 μm thickness) were made with a Cryostat (HM 500; Micron Laboratories, Walldorf, Germany) and mounted on 3-aminopropyl triethoxysilane-coated slides (Matsunami, Osaka, Japan). The sections were fixed in acetone at –80 C for 5 min and dried in air. They were rinsed in 0.1 M PBS, and were incubated with 3% H_2O_2 to quench endogenous peroxidase activity. After soaking in blocking solution (1% BSA in PBS) at room temperature for 30 min, the sections were incubated at 4 C overnight with the primary monoclonal antibody against human trophinin (clone 3–11, mouse IgM; [22]), kindly provided by Dr. M. N. Fukuda, The Burnham Institute, La Jolla, CA, USA. After washing with 0.1 M PBS, the sections were incubated with biotinylated goat anti-mouse IgM secondary antibody (Anti-IgM Mouse Goat-poly-Biotin; Chemicon, Temecula, CA, USA) at room temperature for 1 h and washed with 0.1 M PBS. Positive reactions were visualized using an ABC-peroxidase-staining kit (Vector Laboratories, Elite, CA, USA) diluted in blocking solution at 37 C for 1 h. After 3 washes with 0.1 M PBS for 5 min each, the sections were incubated with 2% avidin-biotin complex (Vector) at 37 C for 30 min. The sections were then reacted with 0.5% 3, 3′-diaminobenzidine tetrahydrochloride (Sigma Chemical, St Louis, MO, USA) and 0.01% H_2O_2 to visualize trophinin. All sections were counterstained with hematoxylin.

RNA extraction and cDNA synthesis

The tissues were mechanically crushed to a fine powder with a mincer (Microtech Nition Co. Ltd., Chiba, Japan) and homogenized with a homogenizer (Microtech Nition) in ISOGEN (Nippongene Co. Ltd., Tokyo, Japan) to prepare total RNA. One microgram of total RNA was used to generate single strand cDNAs in 15 μl of reaction mixture using a First-strand cDNA Synthesis Kit (Amersham Pharmacia Biotech Inc., Buckinghamshire, UK), and stored at –20 C until RT-PCR.

Identification of partial cDNA of porcine trophinin

We designed two primers, F3 (5′- GCA GTG AAG CCA GCA TTA GC -3′) and 889 (5′- CTG GTG CTG GGT CCA TCA CAA AAA C -3′), based on

the human sequence [7] (EMBL/GenBank/DDBJ accession number U04811). The cDNA fragments were amplified with these primers using an Advantage® cDNA PCR Kit (CLONTECH Laboratories. Inc., Palo Alto, CA, USA). The PCR reaction was carried out using 1 μl of the reverse transcription solution in a total volume of 50 μl. The mixture was heated to 94 C for 2 min and amplified for 35 cycles (94 C for 30 sec; 58 C for 30 sec; 72 C for 90 sec) with a final extension at 72 C for 4 min using a PTC-200 DNA Engine (MJ Research Inc., Waltham, MA, USA). The PCR product was subjected to agarose gel electrophoresis, visualized by ethidium bromide staining, extracted from the gel using a Quantum Prep® Freeze 'N Squeeze DNA Gel Extraction Spin Column (Bio-Rad Laboratories, Hercules, CA, USA), subcloned into pGEM®-T Easy vector (Promega, Madison, WI, USA), and sequenced.

Semi-quantitative porcine trophinin gene expression

Total RNA was extracted from each sample and reverse transcribed as described above. The PCR reaction was carried out using 1 μl each of reverse transcription solution and 20 pmol each of the primers F2 and R2 (Fig. 2a), newly designed from a partial sequence of porcine trophinin in a total volume of 20 μl. The mixture was heated to 94 C for 2 min and amplified for 32 cycles (94 C for 30 sec; 58 C for 30 sec; 68 C for 60 sec) with a final extension at 68 C for 4 min. For each PCR reaction, the number of cycles used was optimized so that the amplification process was carried out within the exponential (linear) range. Porcine β-actin was used as a relative standard for the determination of semi-quantitative porcine trophinin gene expression. The PCR reaction for the β-actin was carried out using 1 μl of the reverse transcription solution in a total volume of 20 μl. The mixture was heated to 94 C for 2 min and amplified for 19 cycles (94 C for 30 sec; 62 C for 30 sec; 72 C for 60 sec) with a final extension at 72 C for 4 min. The primers used in the PCR for porcine β-actin were described previously [26]. For each PCR reaction, the number of cycles used was optimized so that the amplification process was carried out within the exponential (linear) range.

Aliquots of the PCR reaction products (5 μl) with 1 μl 6 × loading dye (Promega) were electrophoresed in a 1.5% agarose gel (Nacalai Tesque, Kyoto, Japan) containing ethidium bromide (Nacalai) in a constant 100 V field. The ethidium bromide-stained gels were evaluated with a UV transilluminator, and photographed by Printgraph-CX® (ATTO Co., Tokyo, Japan). The integrity of the individual bands was analyzed by computerized densitometry using the Lanes and Spot Analyzer® program (ATTO). The level of porcine trophinin gene expression in each sample was expressed as a percentage of the expression of porcine β-actin.

Statistics

Values obtained from semi-quantitative PCR assays were shown as the mean ± S.E.M. of three animals from each group (except day 15; n=4). One-way analysis of variance was carried out and the significance of difference between means was determined by Duncan's multiple range test [20].

Results

Localization of trophinin in the uterus during the estrous cycle

An immunopositive reaction was specifically found in the luminal and glandular epithelial cells of the endometrium of the early pregnant pig (Fig. 1a, b) as was found in humans [7] and mice [21]. Therefore, the antibody used in the present study appears to recognize porcine trophinin. An immunopositive reaction for trophinin was found in the luminal and glandular epithelial cells of the endometrium on days 8 (luteal phase) and 20 (follicle phase) of the estrous cycle (Fig. 1d and e; Fig. 1f and g). No positive reaction was found in the endometrial stromal cells (Fig. 1d and f) or the cells of the myometrium (Fig. 1b). The distribution of positive reactions for trophinin was not different among the sections obtained from other days of the estrous cycle (data not shown).

Partial coding sequence of porcine trophinin mRNA

We generated a PCR product of 819 base pairs from porcine uterine endometrium extracts. The sequence of this product is shown in Fig. 2a and has been submitted to EMBL/GenBank/DDBJ under accession number AB081725. Porcine trophinin has 75.7% homology with human trophinin (EMBL/GenBank/DDBJ accession number U04811) and 69.8% homology with mouse trophinin (accession number AF241244) (Fig. 2b). Most of the putative

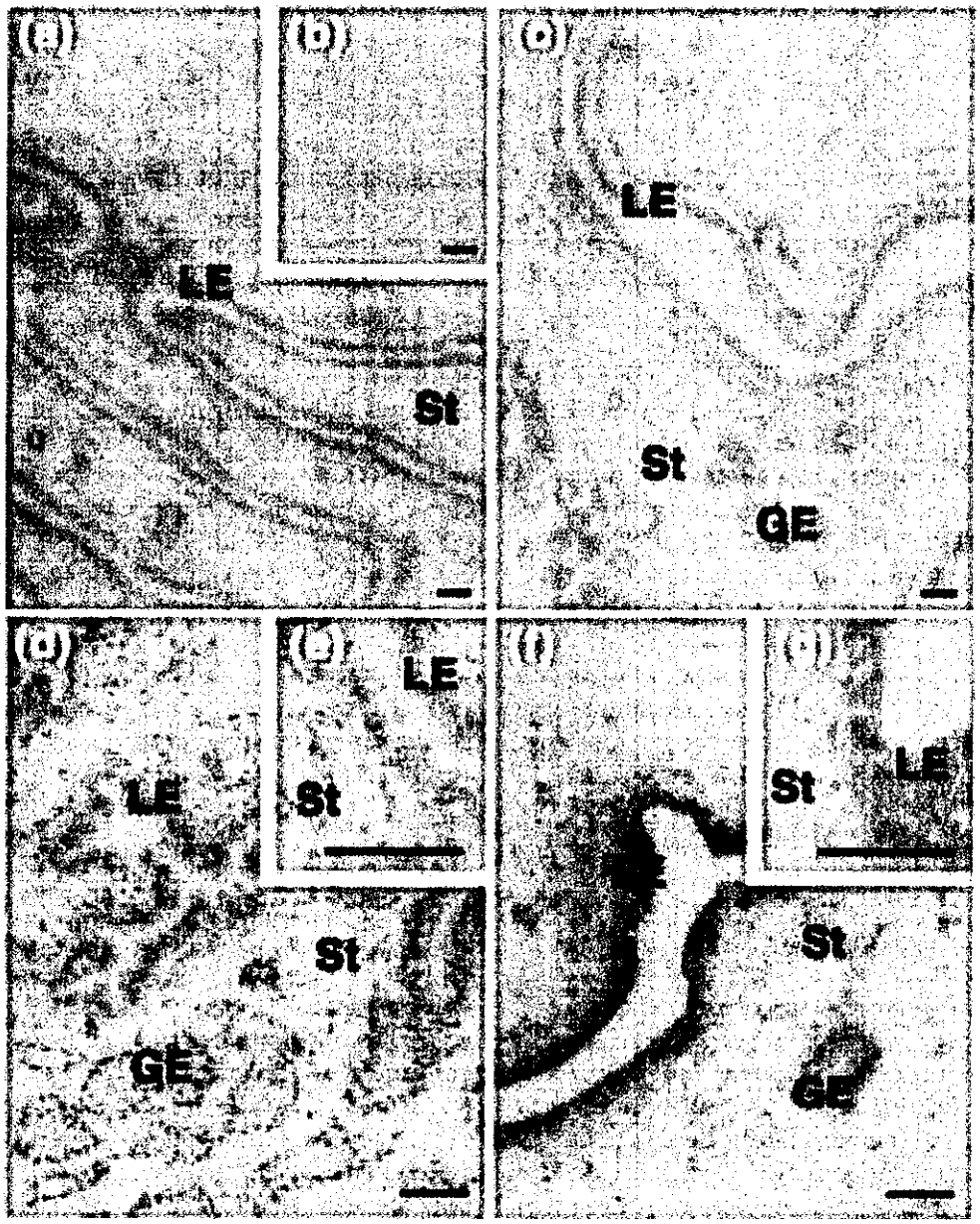


Fig. 1. Localization of porcine trophinin in the uterus during the estrous cycle and pregnancy in the pig. Photomicrographs taken in early pregnancy (day 15, a, b, c), the luteal phase (day 8 of estrous cycle, d, e) and follicular phase (day 20 of estrous cycle, f, g) are shown. Micrographs 'a' and 'b' show the endometrium and myometrium, respectively, taken from the same section. The tissue sections were incubated with (a, b, d, e, f, g) or without (c) a monoclonal antibody against human trophinin. Micrographs shown in 'a' 'b' and 'c' were taken at magnification of × 100. Micrographs shown in 'd' and 'f', and 'e' and 'g' were taken at magnifications of x 200 and × 400, respectively. The solid bar in each panel represents 100 μm. GE, glandular epithelium; LE, luminal epithelium; St, stroma.

porcine trophinin polypeptide consists of tandem decapeptide repeats that start with phenylalanine (Fig. 2c).

Semi-quantitative porcine trophinin gene expression during estrous cycle

Porcine trophinin gene was expressed throughout the estrous cycle, although the level of

a

```
1   TTTGGTGGCACACCTTGTACCAGTGCCAACTTCGCAGGTGGAGTTAGCCCTAGTTTCAGT  60
1   F  G  G  T  P  C  T  S  A  N  F  A  G  G  V  S  P  S  F  S    20

61  GGTCCACTGAACACCAGTACCAGTTTCAGTGGTGGAGCCAGCTCTGGTTTTGGAGGGACA 120
21  G  P  L  N  T  S  T  S  F  S  G  G  A  S  S  G  F  G  G  T    40

121 CTCAGCGCCACTGCTGGCTTCAATGATGCACTCATTACAACCACCAGCTTTGGCAGTACA 180
41  L  S  A  T  A  G  F  N  D  A  L  I  T  T  T  S  F  G  S  T    60

181 CTCGGCACCAGTGCAGTCTTTAGTGGTGCACTTAGCACCAGCACTGCCTTTGGTGGCACA 240
61  L  G  T  S  A  V  F  S  G  A  L  S  T  S  T  A  F  G  G  T    80

241 CTCAGCACTAGTGTCTACTTTGGTGGTTCTCCTAGTTCCAGTGCCAGCTTTGGTGGCACA 300
81  L  S  T  S  V  Y  F  G  G  S  P  S  S  S  A  S  F  G  G  T   100

301 CTCAGTACCAGTATATGCTTCAGTGGTTCTCCTAGCACCAGCACTGGTTTTGATGGTGTA 360
101 L  S  T  S  I  C  F  S  G  S  P  S  T  S  T  G  F  D  G  V   120

361 TTCAGTACCAGTGTCTCCTTTGGTGGCTCTTCCAGCACCAGCACTGACTTTAGTGGTACG 420
121 F  S  T  S  V  S  F  G  G  S  S  S  T  S  T  D  F  S  G  T   140

421 CTAAGCACCAGCGTCTGCTTTGGTGATTCACTCAGCACTAGTTCCAGCTTTGGTGGTGCA 480
141 L  S  T  S  V  C  F  G  D  S  L  S  T  S  S  S  F  G  G  A   160

481 CTCAGCACCAGTGTTGGCTTTGGTGGTGCACTCAGTACCAGTGTTGGCTTTGGTGGTGCA 540
161 L  S  T  S  V  G  F  G  G  A  L  S  T  S  V  G  F  G  G  A   180

541 CTCAACACCAGTGCTGGTTTTAGCAGTGCTGTCAGCATTAACACTGGCTTCAACAGTGTA 600
181 L  N  T  S  A  G  F  S  S  A  V  S  I  N  T  G  F  N  S  V   200

601 CCCAGCACCAACTCTGGCTTTAGCAGTGTGTTCAGCACCAGTGCTGACTTCAGTGGGGCA 660
201 P  S  T  N  S  G  F  S  S  V  F  S  T  S  A  D  F  S  G  A   220

661 CTTAACACCACTACTGACTTTGGCAGTGCTTCTAGCACCAGCATTGGCTTTGGTGGAGCC 720
221 L  N  T  T  T  D  F  G  S  A  S  S  T  S  T  G  F  G  G  A   240

721 CCCAGCACCAGCTTCTGCTTTGGCAGTTTGTCTAACACCAACCTATGCTTTGGCGGCCCT 780
241 P  S  T  S  F  C  F  G  S  L  S  N  T  N  L  C  F  G  G  P   260

781 CCTAGCACCAACACCTGCTTTAGTGGTCCTACCAATGCC                       819
261 P  S  T  N  T  C  F  S  G  P  T  N  A                       273
```

b

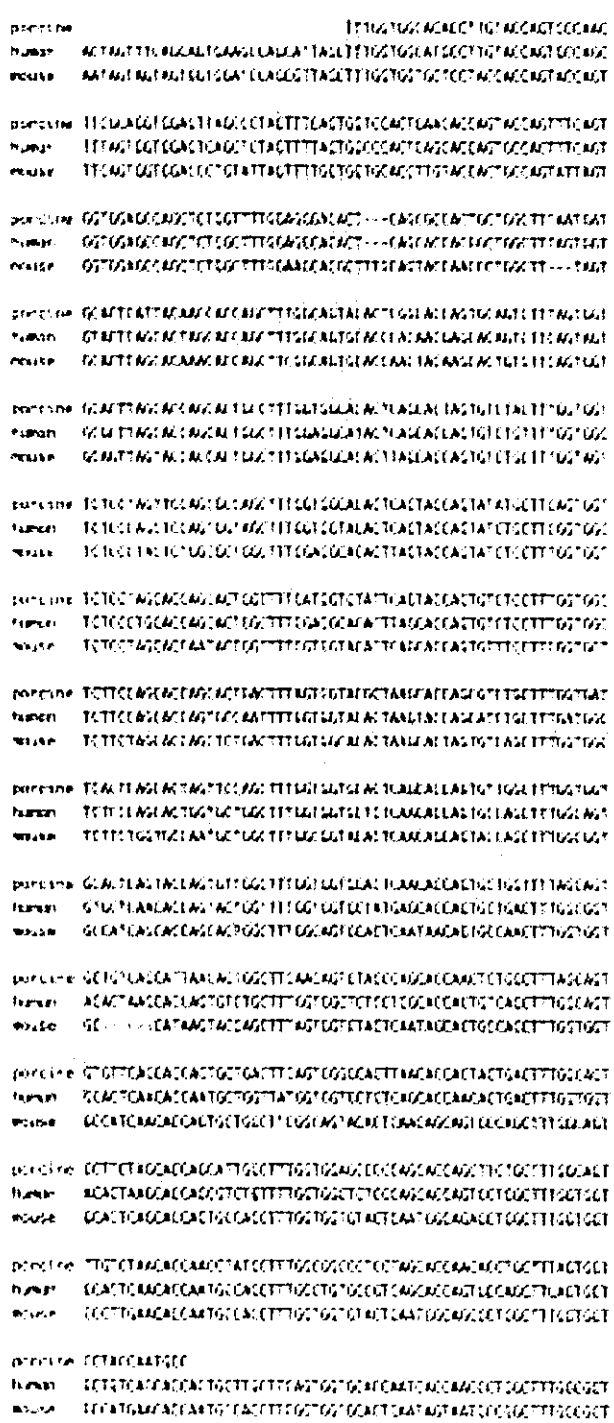


Fig. 2 Structure of trophinin. Partial CDS of porcine trophinin cDNA clone and deduced amino acid sequence (a). The underlines in panel 'a' represent the primers using semi-quantitative PCR assay. Alignment of trophinin mRNA in pigs, humans and mice (b). Conserved sequences are highlighted with a dark gray background. Tandem decapeptide repeats in trophinin (c). This sequence data is available from EMBL/GenBank/DDBJ under accession number AB081725.

its expression appeared to fluctuate. The expression level tended to be lower around the time of ovulation than at other times of the cycle. Significant ($p<0.05$) differences were observed between days 8, 12,18 and day 20 (Fig. 3).

Discussion

A partial coding sequence of porcine trophinin gene obtained in the present study showed that

```
C      1 FGGTPCTSAN
          FAGGVSPS
          FSGPLNTSTS
          FSGGASSG
          FGGTLSATAG
          FNDALITTTS
          FGSTLGTSAV
          FSGALSTSTA
          FGGTLSTSVY
          FGGSPSSSAS
          FGGTLSTSIC
          FSGSPSTSTG
          FDGVFSTSVS
          FGGSSSTSTD
          FSGTLSTSVC
          FGDSLSTSSS
          FGGALSTSVG
          FGGALSTSVG
          FGGALNTSAG
          FSSAVSINTG
          FNSVPSTNSG
          FSSVFSTSAD
          FSGALNTTTD
          FGSASSTSIG
          FGGAPSTSFC
          FGSLSNTNLC
          FGGPPSTNTC
          FSGPTNA   273
```

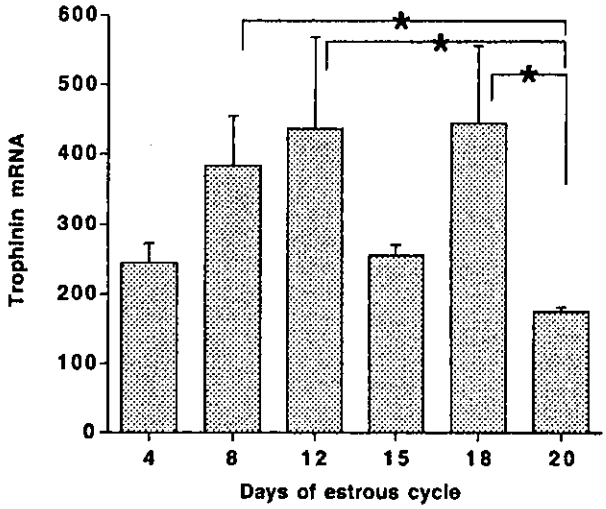


Fig. 3 Levels of porcine trophinin mRNA expression during the estrous cycle. Each value was calculated from the percentage of the expression level of porcine β-actin by triplicate or quadruplicate experiments and represents the mean ± S.E.M. (*$P<0.05$).

most parts of the peptide product consist of tandem decapeptide repeats, as is the case with human [7] and mouse [21] trophinins. These findings indicate that the mRNA reported here is a partial sequence of porcine trophinin, and that the mRNA is actually expressed in the porcine endometrium.

The immunohistochemical and semi-quantitative PCR results indicate that trophinin is expressed in the porcine uterus throughout the estrous cycle. However, positive immunostaining was not found in the endometrial stromal cells or the myometrium on any day of the cycle. This finding suggests that trophinin is not expressed ubiquitously. It has been well demonstrated that corpora lutea are active and progesterone levels in the plasma increase on days 8–12 of the estrous cycle in the pig [8, 17]. The latter two studies also showed that developing follicles mature, and estrogen levels increase on days 19–20 of the cycle. The present semi-quantitative PCR results show that there were significant differences between days 8, 12, 18, but not 15, and day 20. A previous study with mice showed that estrogen is important for inducing mouse trophinin mRNA expression [21]. Therefore, it appears that the induction of trophinin mRNA expression would be different between pigs and mice. On the other hand, trophinin is strongly expressed in the endometrial epithelium during the implantation window and the early secretory phase in humans [7]. Recently, Wang *et al.* (2002) demonstrated that trophinin immunoexpression was detected in the mid-luteal phase in women [27]. Taken together, regulation of trophinin gene expression in pigs might be similar to that in primates, although trophinin mRNA and protein in the endometrium were detected even in the follicular phase in the pig.

We are unable to explain why expression of the trophinin gene continued throughout the estrous cycle of the pig. In the human menstrual cycle, there is a progestational proliferation phase during which decidua are formed [5]. In the mouse uterus, progesterone suppresses cell proliferation during the estrous cycle while estrogen stimulates it [11]. Although morphological changes are observed in the endometrium of the pig during the estrous cycle [10], these changes are minor when compared with the changes in humans and mice. The difference between the pattern of trophinin expression in pigs and the patterns of expression in humans and mice could be due to differences in morphological changes in the endometrium during the estrous cycle among species.

In conclusion, this is the first report of a partial coding sequence of porcine trophinin and trophinin expression in the luminal and glandular epithelial cells of the uterus in the pig throughout the estrous cycle. Trophinin gene expression tended to be