Biochemical and Biophysical Research Communications **262,** 223–230 (1999)
Article ID bbrc.1999.1160, available online at http://www.idealibrary.com on IDEAL®

Molecular Cloning and Genomic Organization of Mouse Homologue of *Drosophila germ cell-less* and Its Expression in Germ Lineage Cells[1]

Tohru Kimura,*[,2] Kentaro Yomogida,† Naomi Iwai,* Yukiko Kato,* and Toru Nakano*

**Department of Molecular Cell Biology, †Department of Science for Laboratory Animal Experimentation, Research Institute for Microbial Diseases, Osaka University, 3-1, Yamada-oka, Suita-shi, Osaka 565-0871, Japan*

Received July 12, 1999

Primordial germ cells (PGCs) are founder cells of all gametes. A number of genes which control PGCs development have been identified in invertebrates, whereas such genes are by and large unelucidated in mammals. Here we describe cloning, genomic structure and expression of mouse homologue of *germ cell-less* (*gcl*) gene which is required for PGCs formation in *Drosophila*. The mouse gcl shows 34% identity compared with *Drosophila* gcl protein and contains BTB/POZ domain. The *gcl* gene consists of 14 exons and spans more than 50 kb. The CpG islands are found around exon 1 of the gene. Putative promoter region contains potential binding sites for various transcription factors. Northern blot analysis showed that its mRNA is highly expressed in adult testis with lower expression in ovary, ES (embryonic stem) cells, and various other organs. *In situ* hybridization analysis revealed strong expression of the *gcl* gene in the pachytene stage spermatocytes. The expression was also observed in post-migratory PGCs, but was not apparent in migratory and pre-migratory PGCs. Further studies including gene disruption analysis would provide an important insight into mammalian germ lineage development.

Key Words: **germ lineage; development; primordial germ cell; spermatogenesis.**

All gametes are known to arise from primordial germ cells (PGCs) (1, 2). In mouse, PGCs are set aside from other cell lineages during gastrulation and appear as a small population of cells near the base of allantois at embryonic day 7.5 (7.5E). They become incorporated into the hindgut endoderm and migrate to the genital ridges through dorsal mesentery by 10.5E. PGCs interact with each other and somatic cells in genital ridges to form sex cords which are the pre-structures of seminiferous tubules and ovarian follicles in the male and in the female, respectively. By 12.5E, sex differences become apparent: PGCs are aggregated in a striped pattern along the sex cords in the male gonads whereas they are in a dotted pattern in the female gonads (3). The genes involved in these processes are poorly elucidated.

In contrast, a large number of genes have been identified as regulators of PGCs development in *Drosophila melanogaster* and *Caenorhabditis elegans* (1). In *Drosophila*, PGCs, also called as pole cells, appear at the posterior pole of blastula. During cellularization, nuclei arriving at the posterior pole are directed to enter the germ line by the molecules stored in the posterior egg cytoplasm (the pole plasm). A large number of maternal effect genes are required for the pole cell formation. *germ cell-less* (*gcl*) is one of such genes and has several characteristics of germ cell determinants (4, 5). At first, gcl is specifically incorporated into pole cells. Second, its posterior localization requires the function of all the genes necessary for pole cell formation. Most importantly, the reduced *gcl* expression results in failure of pole cell formation and the overexpression of *gcl* results in transient increase of pole cells. Thus, *gcl* is believed to be one of the essential components during germ cell specification pathway.

In this study, we have cloned mouse homologue of *gcl* and determined its gene structure in order to gain more insights into the molecular basis of PGCs development in mammals. The mouse *gcl* was highly expressed in pachytene stage spermatocytes and PGCs in genital ridges of the male and the female, which im-

[1] Sequence data from this paper have been submitted in GenBank under Accession Nos. AF163665 and AF163666.

[2] To whom correspondence should be addressed. Fax: +81-6-6879-8362. E-mail: tkimura@biken.osaka-u.ac.jp.

0006-291X/99 $30.00

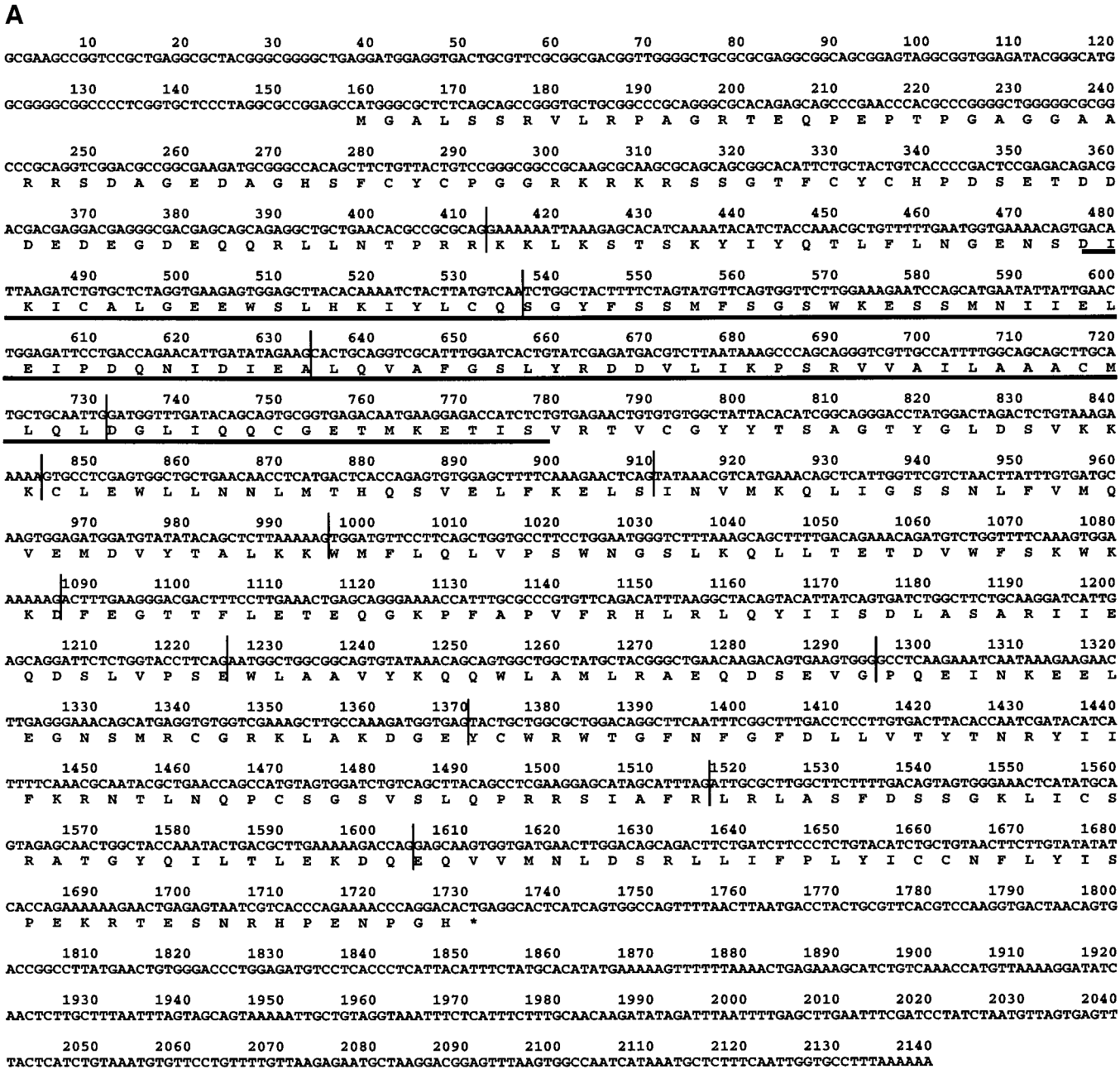

FIG. 1. (A) Nucleotide and amino acid sequences of mouse gcl. Sequences in EST clone span nucleotide number 305 to 1029. BTB/POZ domain is underlined. Exon/intron boundaries are shown by vertical lines. GenBank Accession No. of mouse gcl is AF163665. (B) Sequence alignment of mouse, *D. melanogaster*, and partial *C. elegans* gcl. BTB/POZ domain is underlined. GenBank Accession No. of *D. melanogaster* gcl is M97933. *C. elegans* gcl is found in the Cosmid clone (clone No. Y62E10, GenBank Accession No. AL031580).

plies its possible function in mammalian germ lineage development.

MATERIALS AND METHODS

Isolation of cDNA and gene for mouse gcl. The mouse EST (expressed sequence tags) clone (dbEST Id 641461) was purchased from Genome Systems Inc. (St. Louis, MO). The mouse adult testis cDNA library derived from B6 mice (6) was screened with the EST clone. The genomic fragments were obtained from two λ phage libraries containing 129J/Sv mice genome [(7), Stratagene, La Jolla, CA]. Sequences were determined using ABI PRIME 310 and 377 automated sequencers. The exons and introns were mapped by standard procedures.

Northern blot analysis. Total cellular RNA was isolated from 8 weeks old C57BL/6J mice using RNeasy Mini Kit (Qiagen, Valencia,

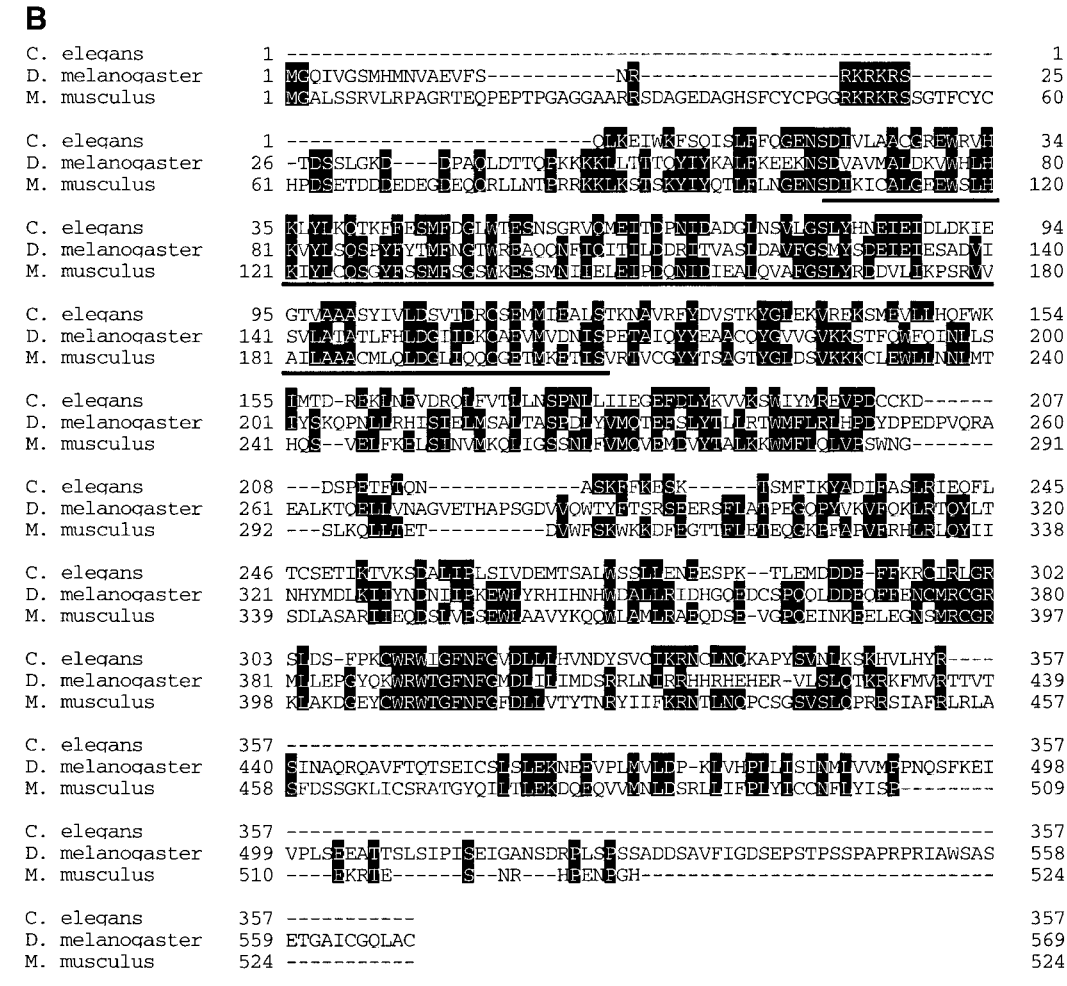

FIG. 1—*Continued*

CA). Ten micrograms of total RNA were subjected to Northern blot analysis. The 1.7 kb *Sma*I-*Eco*RV fragment was radiolabeled by BcaBest Labeling Kit (Takara, Japan) and used for the probe. Northern blot analysis was performed as described (8).

In situ hybridization analysis. The EST clone (dbEST Id 641461) was used to prepare the sense and antisense cRNA probes. Digoxigenin-11-UTP-labeled cRNA probes were prepared using DIG RNA Labeling Kit (Boeringer Mannheim, Mannheim, Germany). *In situ* hybridization was performed as described (9).

RESULTS AND DISCUSSION

Isolation and sequence of the mouse gcl cDNA. To clone mouse *gcl*, we compared *D. melanogaster* gcl amino acid sequence to conceptual translation of ESTs in the GenBank by use of the National Center for Biotechnology Information's BLAST search program (10, 11), and found highly homologous sequences in mouse, rat and human ESTs. Using one of mouse ESTs (dbEST Id 641461), adult testis cDNA library was screened to isolate the full-length cDNA clone. Deduced amino acid sequences from the cDNA indicate that it contains open reading frame which encodes a protein of 524 amino acids (Fig. 1A). Thirty-four percent of amino acids are identical to *D. melanogaster* gcl, but N- and C-terminal ends are relatively diverged (Fig. 1B). Database search using full-length mouse gcl revealed that it contains BTB/POZ domain. BTB/POZ domains are utilized as protein-protein interaction interfaces and found in transcriptional repressors and actin-binding proteins (12–14). Since *D. melanogaster* gcl is localized in nucleus (5), it might be involved in transcriptional repression in germ line specification pathway. From this point of view, it is noteworthy that *PIE-1*, essential gene for germ line specification in *C. elegans*, acts as a repressor for genes expressing in somatic cell lineage (15, 16). Searching of *C. elegans* database showed that *C. elegans* genome also contain *gcl* sequences (Fig. 1B). Identity of partial *C. elegans* gcl to mouse and *D. melanogaster* gcl are 34 and 30%, respectively. Examination of phylogenetic distances by both UPGMA method and NJ method predicted that

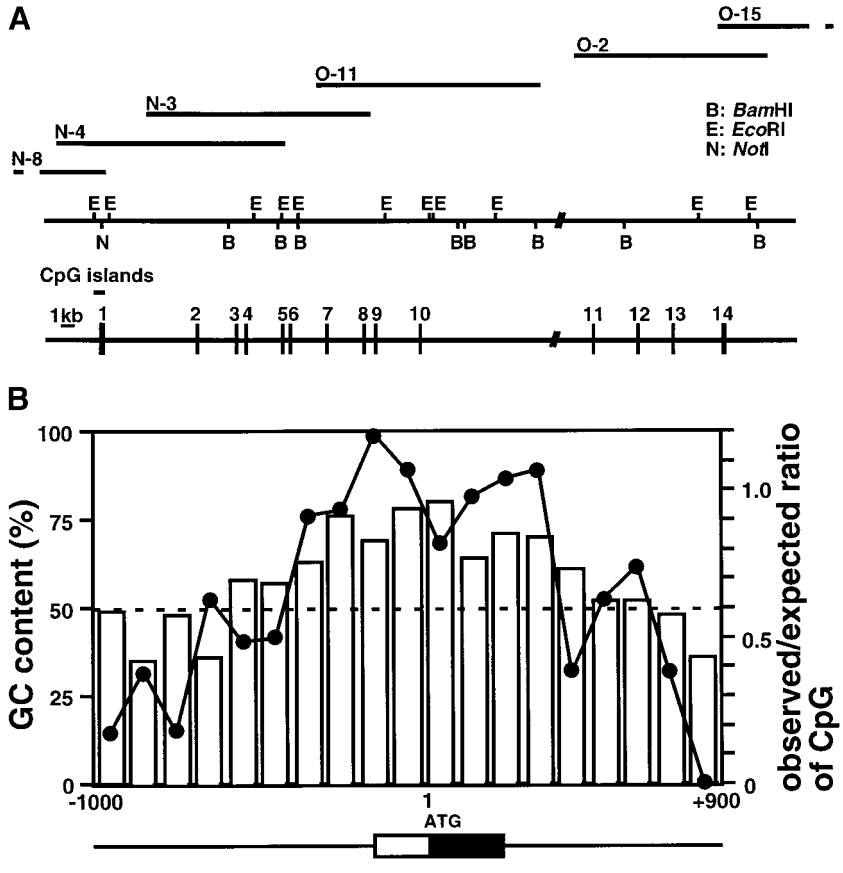

FIG. 2. (A) Genomic organization of mouse *gcl* gene. Position of exons (black boxes) and partial restriction enzyme map are shown. The top six lines indicate the λ phage clones (N-3, N-4, N-8, O-2, O-11 and O-15). (B) The CpG islands. GC contents (open column) and observed/expected ratio of CpG (closed circles) were calculated as described with a window of 100 nucleotides (17). Nucleotide number 1 corresponds to A of initiation codon. The CpG islands have been defined by the nucleotide sequences which contain more than 50% GC contents and more than 0.6 of observed/expected ratio of CpG, thus showing the border by dotted line. Schematic genomic structure is shown below the graph (thin line, 5′-upstream region; open bar, 5′-untranslated region; closed bar, open reading frame). (C) Sequences of putative promoter region and exon 1. CpG is shown in outlined letters. The potential binding sites for transcription factor and its orientation are shown by the arrow below the sequences. Gray arrowhead points to 5′ end of cDNA clone which contains the most 5′ portion. GenBank/EMBL Accession No. of this sequence is AF163666.

mouse gcl is closer to *C. elegans* gcl than to *D. melanogaster* gcl. Homologous sequences were not found in *S. cerevisiae* genome. These results indicate that *gcl* is evolutionarily conserved gene among various multicellular animals.

Genomic organization of the mouse gcl gene. We determined genomic structure of *gcl* gene. The mouse *gcl* gene comprises 14 exons, spanning more than 50 kb (Fig. 2A). The sequences around exon/intron boundaries were matched to consensus sequences of splicing donors and acceptors (Table 1). The CpG islands were identified around exon 1 including putative promoter region; GC contents were more than 50% and observed/expected ratio of CpG was above 0.6, which satisfies the criteria for CpG islands (Fig. 2B (17)). TFSEARCH program predicted that putative promoter region contains potential binding sites for various transcription factors such as Sp1, Sry, CREB (cAMP responsive element binding proteins), Octamer binding proteins and STAT (signal transducers and activators of transcription) [Fig. 2C (18, 19)]. Since some of them are known to play important roles in germ cell development (20–22), further promoter analysis would be required to examine the importance of these elements for *gcl* expression.

Expression of gcl mRNA in adult mice and embryonic gonad. Expression of *gcl* mRNA in various adult organs was examined by Northern blot analysis. As shown in Fig. 3A, expression was most prominent in testis where two transcripts, approximately 3 kb and 4 kb, were equally expressed. In contrast, in other organs including ovary, *gcl* expression level was much lower and the size of major transcripts was 4 kb. Low level expression was also observed in D3 ES (embryonic

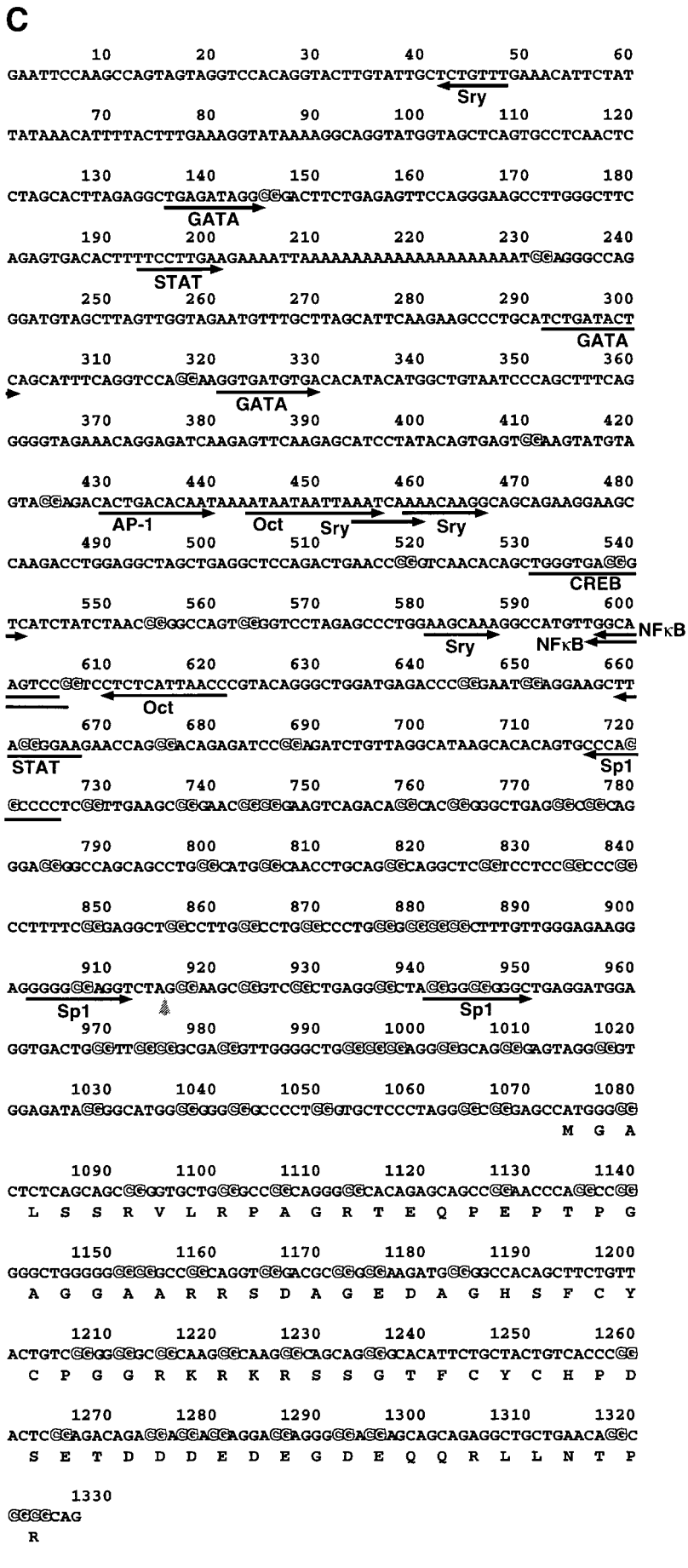

FIG. 2—*Continued*

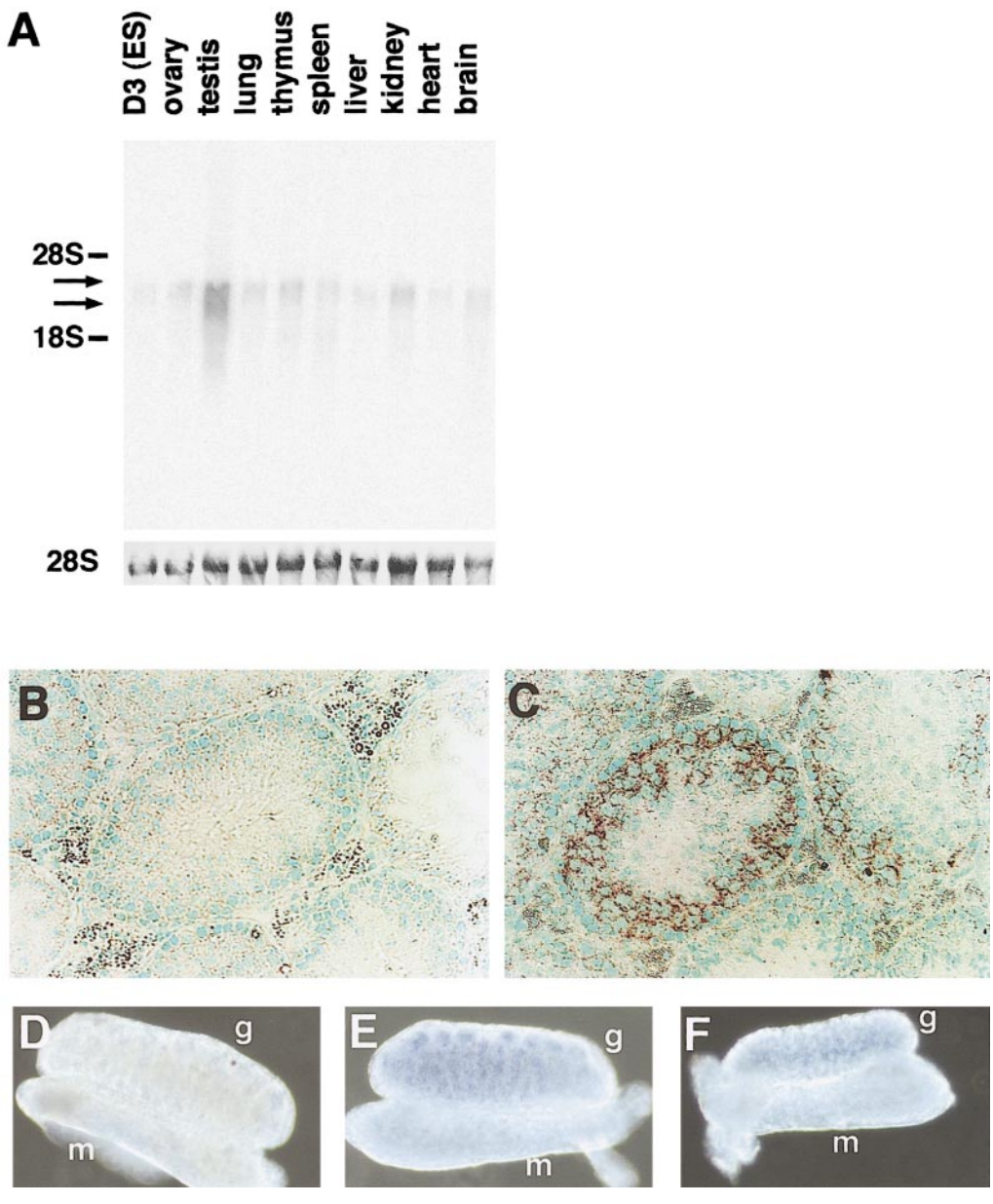

FIG. 3. Expression of mouse *gcl* mRNA. (A) Northern blot analysis of adult mice organs and ES cells. Ten micrograms of total RNA were subjected to Northern blot analysis. Two mRNA species are indicated by arrows. To show the quantity and integrity of RNA, 28S rRNA was visualized by staining the same filter with methylene blue. (B, C) *In situ* hybridization of sections of adult mice testis. Signals were identified in pachytene stage spermatocytes when using antisense probe (C) but not sense probe (B). Nuclei were counterstained with methylgreen. (D to F) Whole mount *in situ* hybridization of 13.5E male (D, E) and female (F) gonads using sense (D) and antisense (E, F) probe. Striped signals and dotted signals were observed in the male (E) and female (F) gonads, respectively, which corresponds to the localization of PGCs in gonads. g, gonad; m, mesonephros.

stem) cells (Fig. 3A). *In situ* hybridization of sections of testis revealed that *gcl* is highly expressed in primary spermatocytes, especially in pachytene stage spermatocytes, but not in spermatogonia and spermatids (Fig. 3C). We also examined expression in PGCs by whole mount *in situ* hybridization. PGCs-specific signal was not observed during pre-migratory and migratory stages (data not shown). However, in the post-migratory stages, strong signals were visible in PGCs in the gonads. Characteristic striped signals and dotted signals were observed in the male and female gonads, respectively (Figs. 3E, 3F), which corresponds to the localization of PGCs in gonads.

Early stages of germ line specification processes would be rather different between invertebrates and mammals. For instance, germ cell determinants such

TABLE 1

Exon/Intron Boundaries of Mouse *gcl* Gene

	412 413	
(exon 1)	GCCGCGCAGgtacggaacg.....ttcttaacagGAAAAAATT	(exon 2)
	536 537	
(exon 2)	TTATGTCAAgtgagtattt.....ttttttccagTCTGGCTAC	(exon 3)
	633 634	
(exon 3)	ATATAGAAGgtatcacctc.....ttgctttcagCACTGCAGG	(exon 4)
	731 732	
(exon 4)	CTGCAATTGgtaagtagtg.....tagtttacagGATGGTTTG	(exon 5)
	844 845	
(exon 5)	AAAGAAAAAgtgagccacc.....tggaccccagGTGCCTCGA	(exon 6)
	910 911	
(exon 6)	AGAACTCAGgtatgtgaca.....ctttttacagTATAAACGT	(exon 7)
	995 996	
(exon 7)	CTTAAAAAGgtattgatgt.....ttctaactagTGGATGTTC	(exon 8)
	1086 1087	
(exon 8)	GGAAAAAAGgtacattgaa.....tcttaaatagACTTTGAAG	(exon 9)
	1224 1225	
(exon 9)	TACCTTCAGgtaagaaaac.....tctgttttagAATGGCTGG	(exon 10)
	1294 1295	
(exon 10)	TGAAGTGGGgtgagtatgc.....atcattttagGCCTCAAGA	(exon 11)
	1370 1371	
(exon 11)	GATGGTGAGgtaggtctgg.....tctctcgtagTACTGCTGG	(exon 12)
	1516 1517	
(exon 12)	AGCGTTTAGgtaggatggc.....ttaattctagATTGCGCTT	(exon 13)
	1604 1605	
(exon 13)	AAAGACCAGgtacgtgtgc.....tgtcttgtagGAGCAAGTG	(exon 14)

Note. Capital letters and lower cases indicate exon and intron sequences, respectively. A at the nucleotide number 1511 is changed to G in 129J/Sv genome, but amino acid is not changed by this polymorphism.

as pole plasm are not stored in mammalian egg cytoplasm (1, 2). Instead of germ cell determinants, signals from extraembryonic ectoderm are considered to direct a small population of cells in extraembryonic mesoderm to enter the germ line (1, 23). It has been postulated that mouse homologues of germ cell determinants in *Drosophila* might play a role in later processes of germ cell specification. Consistent with this hypothesis, mouse *gcl* is expressed in post-migratory stage but not in pre-migratory and migratory stages. Similar expression pattern has been reported in mouse homologue of *vasa* gene which is also required for pole cell formation in *Drosophila* (24). The mouse *gcl* is also highly expressed in pachytene stage spermatocytes, again like mouse *vasa* homologue, suggesting the implication in spermatogenesis.

Recently, other group independently cloned mouse *gcl* as a binding protein of the DP-3 component of E2F transcription factors which regulate G1/S cell cycle transition (25). It is possible that gcl might have some roles other than germ cell formation. Here we describe the *gcl* expression in mouse testis and embryonic gonads. At present, it is open question whether its potential role in cell cycle regulation could be involved in germ cell development. Gene disruption analysis is now underway to evaluate the role of mouse *gcl* in the mammalian germ lineage development.

ACKNOWLEDGMENTS

We thank Drs. H. Nojima, H. Tanaka, Y. Takihara, D. Tomotsune, S. Nishiguchi, K. Deguchi, K. Abe, and M. Okabe for reagents and advise. We also thank Ms. A. Mizokami for excellent secretary work. The sequence of clone Y62E10 has not yet been finished and was provided by the *C. elegans* Sequencing Consortium in draft form, and may contain errors. This work was supported in part by grants from Research for the Future Program of Japan Society for the Promotion of Science (JSPS-RFTF98L01101), and the Ministry of Education, Science, Sports and Culture.

REFERENCES

1. Wylie, C. (1999) *Cell* **96,** 165–174.
2. McLaren, A. (1999) *Genes Dev.* **13,** 373–376.
3. Rugh, R. (1990) The Mouse: Its Reproduction and Development, Oxford University Press, Oxford.
4. Jongens, T. A., Hay, B., Jan, L. Y., and Jan, Y. N. (1992) *Cell* **70,** 569–584.
5. Jongens, T. A., Ackerman, L. D., Swedlow, J. R., Jan, L. Y., and Jan, Y. N. (1994) *Genes Dev.* **8,** 2123–2136.
6. Tanaka, H., Yoshimura, Y., Nishina, Y., Nozaki, M., Nojima, H., and Nishimune, Y. (1994) *FEBS Lett.* **355,** 4–10.
7. Ikawa, M., Wada, I., Kominami, K., Watanabe, D., Toshimori, K., Nishimune, Y., and Okabe, M. (1997) *Nature* **387,** 607–611.
8. Kimura, T., Nakayama, K., Penninger, J., Kitagawa, M., Harada, H., Matsuyama, T., Tanaka, N., Kamijo, R., Vilcek, J., Mak, T. W., *et al.* (1994) *Science* **264,** 1921–1924.

9. Wilkinson, D. G. (19992) *In Situ* Hybridization: A Practical Approach, Oxford University Press, Oxford.
10. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410.
11. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402.
12. Albagli, O., Dhordain, P., Deweindt, C., Lecocq, G., and Leprince, D. (1995) *Cell Growth Differ.* **6,** 1193–1198.
13. Dhordain, P., Albagli, O., Lin, R. J., Ansieau, S., Quief, S., Leutz, A., Kerckaert, J. P., Evans, R. M., and Leprince, D. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 10762–10767.
14. Robinson, D. N., and Cooley, L. (1997) *J. Cell Biol.* **138,** 799–810.
15. Seydoux, G., Mello, C. C., Pettitt, J., Wood, W. B., Priess, J. R., and Fire, A. (1996) *Nature* **382,** 713–716.
16. Batchelder, C., Dunn, M. A., Choy, B., Suh, Y., Cassie, C., Shim, E. Y., Shin, T. H., Mello, C., Seydoux, G., and Blackwell, T. K. (1999) *Genes Dev.* **13,** 202–212.
17. Gardiner-Garden, M., and Frommer, M. (1987) *J. Mol. Biol.* **196,** 261–282.
18. Heinemeyer, T., Wingender, E., Reuter, I., Hermjakob, H., Kel, A. E., Kel, O. V., Ignatieva, E. V., Ananko, E. A., Podkolodnaya, O. A., Kolpakov, F. A., Podkolodny, N. L., and Kolchanov, N. A. (1998) *Nucleic Acids Res.* **26,** 362–367.
19. Akiyama, Y. http://www.rwcp.or.jp/papia/.
20. Sassone-Corsi, P. (1998) *J. Mol. Med.* **76,** 811–817.
21. Pesce, M., Gross, M. K., and Scholer, H. R. (1998) *BioEssays* **20,** 722–732.
22. Parker, K. L., Achedl, A., and Aschimmer, B. P. (1999) *Annu. Rev. Physiol.* **61,** 417–433.
23. Lawson, K. A., Dunn, N. R., Roelen, B. A., Zeinstra, L. M., Davis, A. M., Wright, C. V., Korving, J. P., and Hogan, B. L. (1999) *Genes Dev.* **13,** 424–436.
24. Fujiwara, Y., Komiya, T., Kawabata, H., Sato, M., Fujimoto, H., Furusawa, M., and Noce, T. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 12258–12262.
25. de la Luna, S., Allen, K. E., Mason, S. L., and La Thangue, N. B. (1999) *EMBO J.* **18,** 212–228.