Identification of a possible stem-cell-maintenance gene homologue in the unicellular eukaryote *Paramecium caudatum*

By Shinji OBARA,^{*,**)} Yoshinori IWATAKI,^{**)} and Kazuyuki MIKAMI^{**,†)}

(Communicated by Koichi HIWATASHI, M.J. A., May 12, 2000)

Abstract: In macronuclear DNA of the ciliated protozoan *Paramecium caudatum*, we identified a gene homologous to *aubergine/piwi* in *Drosophila* and *ZWILLE* in *Arabidopsis*, which have recently been shown to regulate the maintenance of stem cell in animals and plants. The DNA sequence was considered to be actually transcribed because it was also obtained from cDNA of vegetative phase cells by RACE-PCR. The gene named *pap* (*Paramecium aubergine/piwi* homologue) encodes 781 amino acids and contains three short introns (21bp, 22bp and 23bp) in the ORF. Forty four UAA codons and 8 UAG codons which code for glutamine in *Paramecium* were found in the ORF. This is the first finding of *piwi/ZWILLE* homologue in a protozoan. The gene was transcribed actively immediately after induction of conjugation, while the transcription was restricted amount during vegetative phase. The result suggests that the gene is integrated in the developmental process of conjugation.

Key words: *Paramecium*; unicellular eukaryote; stem cell; *aubergine/piwi*; *ZWILLE*; conjugation.

Introduction. In multicellular animals and plants, differentiated cells derive from stem cells. How stem cells remain undifferentiated and are induced to differentiate are important and intriguing subjects. Until recently, analyses on these subjects failed to reveal common molecular mechanisms among different organs and organisms. However, homologous genes that regulate stem cell maintenance have recently been shown in several animals and plants.¹⁾ In *Drosophila*, a gene called *piwi* was identified. In the *piwi* mutant, all of the germline stem cells differentiated, while in wild type cells germline stem cells divided to give rise to one germline stem cell and the other for differentiated progeny. This indicates that the *piwi* gene has a function to maintain cells in the undifferentiated state.²⁾ In *Arabidopsis*, the *ZWILLE* (ZLL) protein (also known as the PINHEAD protein) has a high level of sequence similarity to PIWI. The phenotype of the *zwille* mutant lead to a very similar conclusion regarding its function.³⁾ Moreover, the *piwi/ZLL* homologue, *prg-1*, has been reported in the *Caenorhabditis elegans*.²⁾ A homologue, *hiwi*, has also been identified in humans, though

its role in germ cell maintenance has not been demonstrated. However, no obvious homologue seemed to be seen in the totally sequenced genomes of *Saccharomyces cerevisiae* and bacteria.²⁾

Then, does such a homologue exist in a unicellular eukaryote, such as the ciliates? Ciliates differentiate two types of nuclei, the somatic macronucleus and the germinal micronucleus. The macronucleus can be considered homologous to somatic cells and the micronucleus homologous to germ cells of multicellular organisms. In the ciliate, *Paramecium caudatum*, the germinal micronucleus divides and gives rise to germinal micronuclei and somatic macronuclei at the sexual phase, conjugation.⁴⁾ The macronucleus is polygenomic and transcriptionally very active, while the micronucleus is diploid and almost inert. During conjugation, the micronucleus undergoes meiosis and produces two gametic nuclei. After reciprocal exchange of gametic nuclei between mating cells, a zygotic nucleus is formed by fusing two gametic nuclei in each cell of the mating pair. The zygote divides mitotically and nuclear products differentiate into macronuclear anlagen or remain as micronuclei. The differentiation of somatic nuclei and germinal nuclei occurs in a common cytoplasm. Therefore, the homologue, if is present in ciliates, must be very intriguing, whatever its function. In the present paper, a gene homologous to the *hiwi* gene in

^{*)} Biological Institute, Faculty of Science, Tohoku University, Aoba, Aramaki, Aoba-ku, Sendai 980-8578, Japan.

^{**) EEC, Miyagi University of Education, Aoba, Aramaki, Aoba-ku, Sendai 980-0845, Japan.}

^{†)} Correspondence to: K. Mikami.

human, *ZLL* in *Arabidopsis* and/or *piwi* and *aubergine* (*sting* as another name) in *Drosophila* was found in macronuclear DNA of *P. caudatum*. Features of the gene, named here *pap*, will be described.

Materials and methods. *Cell culture and induction of conjugation.* The stock used was 27aG3s3 (mating type V) belonging to syngen 3 of *P. caudatum*. The amiconucleate clone of 27aG3s3 (27aG3s3amic) was obtained by sucking out the micronucleus with a microcapillary from a cell in the stationary phase of growth.^{5,6)} Culture and handling techniques for paramecia followed those of Hiwatashi.⁷⁾ Culture medium was fresh lettuce juice diluted with Dryl's salt solution⁸⁾ modified by the substitution of KH_2PO_4 for NaH_2PO_4 (pH 7.0) and inoculated with a non-pathogenic strain of *Klebsiella pneumoniae* as food organisms 1 or 2 days prior to being used. Cultivation was carried out at $25 \pm 1^\circ\text{C}$.

For northern hybridization analyses during conjugation, conjugation was induced by mixing mating-reactive cells of complementary types, stock KNZ5 (mating type V) and KNZ2 (mating type VI) belonging to syngen 3.

Preparation of DNA. Macronuclear DNA was isolated from 27aG3s3amic. Nuclear DNA was isolated by a modification of the method of Scott *et al.*⁹⁾ and Tsukii.¹⁰⁾ Cells were collected with a filter paper (Toyo, no. 1, Tokyo, Japan) and washed with distilled water. Cells packed by low-speed centrifugation were suspended in 5 ml of lysing solution A (1% acetic acid, 1% Nonidet P-40, 0.25 M sucrose). The cells were lysed by pipetting (about 10 times) with a pasteur pipette. To collect nuclei, the suspension was centrifuged at 500 g for 5 min. The precipitate was suspended again in 200 μl of distilled water, and squirted into 400 μl of lysing solution B (10 mM Tris-HCl pH 8.0, 50 mM sodium EDTA, 1% sodium dodecyl sulfate) at 65°C for 1 h. The lysate was treated with Proteinase K (MERCK, Darmstadt, Germany) (finally 100 $\mu\text{g}/\text{ml}$) at 37°C overnight. Nucleic acids were extracted from the lysate with phenol and phenol/chloroform, and precipitated with ethanol. The precipitate was resuspended in Tris-EDTA (pH8.0), treated with RNase A (finally 20 $\mu\text{g}/\text{ml}$) at 37°C for 1 h, extracted with phenol/chloroform, and finally precipitated with ethanol.

RNA extraction and cDNA synthesis. Total RNA was isolated using TORIZOL Reagent (GIBCO BRL, NY, U.S.A.) from the cells at vegetative phase. cDNA was synthesized with either Ready-To-Go T-Primed First Strand Kit (Amersham Pharmacia Biotech, Uppsala,

Table I. Primers used PCR amplification

Primer no.	Nucleotide sequences of the primers (5'-3')
1	GGIATHACIAARCCIGCIAT
2*	GCRTGYTCIGTRTAIGTIAC
3*	TCTCTTAACACCTCCTCTTCTGGC
4	TATCAGAGATGCCATCACATACAC
5*	GATTCCCTTTCTGATTTTCCAGCCAT
6	GGTAGGAGGCATACCGTGGGCTC
7*	CCATGCTTAGAGAAAATATCTACTCCACATATC
8	ATGCAAAAAGTAATGGTGAGAAGCAACTTCC
9*	TCAAATATAAAATAACACTTTTAAATTCATAAATCGGG
10	GGCCTGGAATGGAGTGCATATTTAG
11*	CACAATTTCCATAAGCATGTTACCAGG

*Antisense sequence.

Sweden) or SMART cDNA Synthesis Kit (CLONTECH, CA, U.S.A.).

Primers for the isolation of genes. Table I shows the nucleotide sequences of primers used for PCR amplifications. The degenerate primers (nos. 1 and 2) for PCR amplification of macronuclear DNA fragment of histone H4 gene were designed according to conserved regions of known histone H4. To isolate the flanking region of the amplified *Paramecium* histone H4 gene fragment, inverse PCR^{11,12)} was performed with circularized macronuclear DNA digested with *Hind*III. Primers (nos. 3 and 4) for inverse PCR were designed according to the sequence of PCR products obtained by using degenerate primers. Primers nos. 5 and 6 were used to confirm the DNA sequence of inverse PCR fragments.

The 5'-ends or 3'-ends of cDNAs were isolated by RACE-PCR method as follows. cDNAs created with SMART cDNA Synthesis Kit were blunt-ended, phosphorylated, and ligated into vector pT7Blue-3 digested with *Eco*RV. The ligation mixture was used as a template in PCR amplification. The fragments of 3'-end and 5'-end were obtained by PCR using the vector sequence primer and the gene specific sequence primers (nos. 6 and 7). Finally, the almost whole cDNA created with Ready-To-Go T-Primed First Strand Kit and macronuclear DNA coding sequence were isolated by PCR using primers (nos. 8 and 9), and sequenced with primers nos. 8 to 11.

DNA sequencing. Amplification products were isolated by electrophoresis and were purified with GENECLAN III (BIO101, CA, U.S.A.). The purified DNA was sequenced directly or indirectly after cloning into the plasmid vector pGEM-T Easy (Promega, WI, U.S.A.). The DNA sequencing was performed with either SILVER SEQUENCE DNA Sequencing System

Fig. 1. Heterogeneous composition of the macronuclear DNA upstream of the histone H4 gene; (a) The longer PCR product was composed of 5'-[region A] + [region B] + [region C] + [the beginning of histone H4 coding]-3'. The shorter product contained region A and C but lacked the part of region B. λ , λ HindIII digests; Ma, PCR products of macronuclear DNA with the primer nos. 6 and 5. (b) Position of the region A in the expected *hiwi* homologue on macronuclear DNA.

(Promega) or ABI PRISM 310 Genetic Analyzer (PE Biosystems, CA, U.S.A.).

Sequence analysis. DNA and amino acids (aa) sequences were analyzed using the GENETYX-MAC 9.0 (Software Development Co., Ltd., Tokyo, Japan). Homology search was performed with Gapped BLAST.¹³⁾ Comparing translated nucleotide sequences of all reading frames with a protein sequence in database was performed with the BLASTX program. Comparison of the aa sequence with a protein sequence database was performed with the BLASTP program through the worldwide web (WWW) server of the National Center for Biotechnology Information (NCBI; Washington, DC, U.S.A.). Multiple aa sequence alignment was performed with CLUSTAL W¹⁴⁾ through the WWW server of the DNA Data Bank of Japan (DDBJ; Mishima, Japan).

Northern blot analysis. Total RNA was isolated using the TORIZOL Reagent from cells at certain times before and after mating. The RNA (15 μ g/lane) was electrophoresed on a 1% agarose-formaldehyde denaturing gel and subsequently transferred onto Hybond-N⁺ membrane (Amasham Pharmacia Biotech). The blots on the membrane were hybridized with the mixture of probes for the *pap* and the glyceraldehyde-3-dehydrogenase (*gapdh*) and detected with AlkPhos DIRECT system for chemiluminescence (Amasham Pharmacia Biotech). The probe for *gapdh* was designed according to the macronuclear DNA sequences of *gapdh* in *P. caudatum* (Obara and Watanabe, to be published). The hybridization and washing were performed at 50 °C.

Results and discussion. *Isolation and sequencing of the upstream region of macronuclear histone H4 gene.* The macronuclear DNA of an ami-

cronucleate stock 27aG3s3 was amplified with two primers. Primer no. 5 was an antisense sequence that is present in the coding region of histone H4 and primer no. 6 was a sense sequence from a version upstream of the gene (Table I). PCR with the two primers produced two kinds of DNA fragments that were different in length (Fig. 1). The longer fragment was 3.3kbp and the shorter was 1.1kbp. The longer one was roughly composed of 5'-[region A] + [region B] + [region C] + [the beginning of histone H4 coding]-3'. The shorter one contained region A and C but lacked the part of region B. The region A had the same sequence in both of the fragments. The region B that was present only in the longer fragment showed a homology in aa sequence with a human protein serine/threonine kinase *stk2* (GenBank, accession no. L20321; expect value, $2e-29$). Region C did not show any homology with known proteins.

We noticed that region A had homology in aa sequence with the human *HIWI* protein according to BLASTX search (Table II). This was quite interesting, because the human gene *hiwi* has been identified as *piwi* homologue. The group of homologues includes the *hiwi* gene in human, *aubergine*, *piwi* in *Drosophila*, *ZLL* in *Arabidopsis* and/or *prg-1* in *C. elegans*.

The DNA fragment amplified from cDNA and its features. To know whether the *hiwi* homologue in *P. caudatum* is transcribed or not, we tried to amplify cDNA of the amicronucleate cells at vegetative phase by 3'-RACE using with the primer no. 6 (Table I). We obtained two different length DNA fragments. They have almost identical DNA sequence with each other from the position of primer no. 6 to the putative stop codon. The difference in the length is due to the length

Fig. 3. Map of the 3 introns in the *pap* gene of macronuclear DNA. The exons are shown as open boxes and introns are shown as black lines. 5' and 3'-UTR are shown as gray lines. Oligonucleotide primers used for PCR amplification or for sequencing are shown with arrows, number are primer no. in Table I. Arrowhead: poly (A) addition site.

code UAA indicates glutamine in *P. caudatum*.¹⁹⁾ The present result shows that UAG in *P. caudatum* codes for an aa, probably glutamine as was previously shown in *P. primaurelia*²⁰⁾ and in *P. tetraurelia*.²¹⁾

Homology search. A BLAST search with the BLASTP program revealed that the aa sequence deduced from nucleotide sequence shares statistically significant homology with AUBERGINE and PIWI. According to the BLASTP search, the aa sequence from nos. 4 to 756 has 24.7% homology with AUBERGINE (expect value, 3e-64). The sequence of aa nos. 4 to 781 showed 24.2% homology with PIWI (expect value, 2e-60). The sequence of aa from nos. 238 to 755 has 29.6% homology with HIWI (expect value, 5e-56). The homology with PRG-1 showed 25.6% (expect value, 2e-51) on the sequence of aa from nos. 115 to 750. And the homology with PRG-2 showed 27.2% (expect value, 6e-48) on the sequence of aa from nos. 203 to 750. The results clearly demonstrate homology to the *aubergine*, *piwi*, *hiwi*, *prg-1* and *prg-2* gene family (Fig. 2). Therefore, we named this gene *pap* (*Paramecium aubergine/piwi* homologue). The gene family including *piwi*, *hiwi*, *ZLL* and others seems to regulate the maintenance of stem cells in animals and plants.¹⁾

Isolation and sequencing the *pap* gene of the macronuclear genome. The entire sequence of the *pap* gene of the macronuclear genome was determined. The sequence has been deposited in the DDBJ/EMBL/GenBank databases under the accession no. AB035972. The *pap* gene contains three introns, that is, intron I (21bp), intron II (22bp) and intron III (23bp) (Fig. 3). All of them are very short as was found in the hemoglobin gene of *P. caudatum*¹⁹⁾ and the β -tubulin gene of *P. tetraurelia*.²²⁾ These introns show eukaryotic consensus sequences of GT/AG at the splicing junctions.²³⁾

The transcriptional activity of *pap* during conjugation. To know the change of transcriptional level of PAP mRNA during sexual phase, that is, conjugation, RNA was isolated from cells at various times after mixing

mating-reactive cells of complementary types. The northern hybridization revealed that the PAP mRNA was little detected at stationary phase but increased drastically soon after induction of conjugation, while the mRNA of GAPDH was actively transcribed before conjugation and inclined to decrease as the process of conjugation advanced. As shown in Fig. 4, mating reactive cells at vegetative phase did not show apparent band on PAP mRNA, while a strong band appeared on GAPDH mRNA. Soon after mating reaction, however, a strong band appeared on the PAP. The expression continued up to 12 hours after mixing mating reactive cell of complementary types though the amount inclined to decrease as the conjugation process advanced. In contrast to the pattern of PAP, GAPDH expressed actively during vegetative phase and decreased gradually during conjugation. The results show that the *pap* gene was transcribed very actively at the beginning of conjugation.

Discussion on the prospective function of PAP. Stem cells can reproduce themselves as well as generate differentiated progeny.^{1),24)} The ability to remain in an undifferentiated state or to generate one or more differentiated cell types as occasion demands is a unique role of stem cells in multicellular organisms. In plants such as *Arabidopsis*, stem cells generate most of their adult structures. From two populations of stem cells known as 'meristems', located at opposite ends of the embryo, are formed the shoot and root systems. How does a stem cell remain undifferentiated, and how does it regulate the production of differentiated progeny? In *Drosophila*, Wilson *et al.*¹⁶⁾ reported that *aubergine* enhanced *osker* mRNA translation in the ovary. The *osker* protein is known to be a component of the pole granule. The gene *piwi* has recently been shown to be required for self renewing division of germ-line stem cells (GSCs) during oogenesis in *Drosophila*.²⁾ In ciliates, each cell has somatic macronuclei and germ micronuclei. The differentiation of the somatic macronuclei from the germ micronuclei occurs during conjugation. DNA sequence of the homologue *pap* was obtained not only from

Fig. 4. Changes in the amount of PAP mRNA during conjugation. Northern blots of the RNA were hybridized with probes for the *pap* and *gapdh*. st2 : Cells on the 2nd day of stationary phase culture, st3 : Cells on the 3rd day of stationary phase culture. 0.5 hours after mixing coincide in time with the cell stage of mating reaction and formation of mating clump, 4 hours with the initiation of meiotic prophase of meiosis I, 8 hours with the metaphase of meiosis I, 12 hours with formation of gametic nuclei and fusion of gametic nuclei, 16 hours with the critical stage of nuclear differentiation, 20 and 24 hours with the developing stage of somatic macronuclei.

macronuclear DNA but also from cDNA of vegetative cells. The evidence suggests that the gene is actually transcribed during vegetative phase. However, the result of northern hybridization revealed that the PAP mRNA was little detected at stationary phase but increased drastically soon after induction of conjugation (Fig. 4). This means that the genetic information of *pap* gene is especially required for the developmental process of conjugation in *P. caudatum*. According to the result of Fig. 4, the apparent band of PAP mRNA appeared during conjugation. After mixing the complementary types, conjugation is induced and sequential events occurs: meiosis (4 to 8 hours after the mixing), formation of gametic nuclei and nuclear fusion between gametic (male and female) nuclei (12 to 13 hours after), mating pair separation (about 14 hours after) and the determination of the nuclear differentiation (about 16 hours after). The function of *pap* seems to have some relationship with these micronuclear events. It remains to be solved what function the *pap* gene has on the differentiation of cells and/or nuclei. Any way, the presence of this homologue in *Paramecium* provides an unique opportunity to investigate the role of the gene family within unicellular organisms.

Acknowledgements. We are grateful to Dr. J. Forney (Purdue University) and Dr. E. Cole (St. Olaf College) for helpful suggestions and critical reading the manuscript. We also thanks to Dr. A. Yasui, Dr. T. Watanabe (Tohoku University) and Dr. H. Endo (Kanazawa University) for helpful suggestions. The research was supported by a Grant-in-Aid for Scientific Research (C) (No. 08680785) from the Ministry of Education, Science, Sports and Culture of Japan.

References

- 1) Benfey, P. N. (1999) *Curr. Biol.* **9**, 171-172.
- 2) Cox, D. N., Chao, A., Baker, J., Chang, L., Qiao, D., and Lin, H. (1998) *Genes Dev.* **12**, 3715-3727.
- 3) Mouessian, B., Schoof, H., Haecker, A., Jürgens, G., and Laux, T. (1998) *EMBO J.* **17**, 1799-1809.
- 4) Hiwatashi, K., and Mikami, K. (1989) *Inter. Rev. Cytol.* **114**, 1-19.
- 5) Koizumi, S. (1974) *Exp. Cell Res.* **88**, 74-78.
- 6) Mikami, K. (1979) *J. Exp. Zool.* **208**, 121-128.
- 7) Hiwatashi, K. (1968) *Genetics* **58**, 373-386.
- 8) Dryl, S. (1959) *J. Protozool.* **6** (Suppl.), 25.
- 9) Scott, J. M., Mikami, K., Leeck, C. L., and Forney, J. D. (1994) *Mol. Cell. Biol.* **14**, 2479-2484.
- 10) Tsukii, Y. (1996) *Europ. J. Protistol.* **100**, 165-169.
- 11) Ochman, H., Gerber, A. S., and Fartl, D. L. (1988) *Genetics* **120**, 621-624.
- 12) Triglia, T., Peterson, M., and Kenp, D. J. (1988) *Nucleic Acids Res.* **16**, 8186.
- 13) 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.
- 14) Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22**, 4679-4680.
- 15) Schüpbach, T., and Wieschaus, E. (1991) *Genetics* **129**, 1119-1136.
- 16) Wilson, J. E., Connell, J. E., and Macdonald, P. M. (1996) *Development* **122**, 1631-1639.
- 17) Schmidt, A., Palumbo, G., Bozzetti, M. P., Tritto, P., Pimpinelli, S., and Schäfer, U. (1999) *Genetics* **151**, 749-760.
- 18) Lin, H., and Spradling, A. C. (1997) *Development* **124**, 2463-2476.
- 19) Yamauchi, K., Ochiai, T., and Usuki, I. (1992) *Biochim. Biophys. Acta* **1171**, 81-87.
- 20) Caron, F., and Meyer, E. (1985) *Nature* **314**, 185-188.
- 21) Preer, J. R., Jr., Preer, L. B., Rudman, M. B., and Barnett, A. J. (1985) *Nature* **314**, 188-190.
- 22) Dupuis, P. (1992) *EMBO J.* **11**, 3713-3719.
- 23) Breathnach, R., Benoist, C., O'Hare, K., Gannon, F., and Chambon, P. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4853-4857.
- 24) Morrison, S. J., Shah, N. M., and Anderson, D. J. (1997) *Cell* **88**, 2872-2898.