Analysis of Histone Gene Expression in Adult Tissues of the Sea Urchins Strongylocentrotus purpuratus and Lytechinus pictus: Tissue-Specific Expression of Sperm Histone Genes

TOBY LIEBER, KAREN WEISSER, AND GEOFFREY CHILDS*

Departments of Genetics and Molecular Biology, Albert Einstein College of Medicine, Bronx, New York 10461

Received 18 November 1985/Accepted 10 April 1986

We analyzed the histone mRNA population found in several adult tissues of the sea urchin *Strongylocentrotus purpuratus* and in testis of *Lytechinus pictus*. Unique species of H1 and H2b mRNAs encoding the sperm-specific histone subtypes can be found exclusively in testis RNA. *S. purpuratus* contains two distinct testis-specific H1 transcripts, while *L. pictus* contains one such transcript. Each of these mRNAs is larger than either early or late embryonic H1 mRNAs. Other somatic adult tissues contain transcripts derived from members of the late embryonic H1 histone gene family. *S. purpuratus* contains one H2b transcript found exclusively in testis, while *L. pictus* contains two such H2b mRNAs. Similarly, in tissues other than testis, late H2b transcripts were found. While there is no sperm-specific H2a protein, a limited set of late histone H2a genes encoding primarily the H2a- β subtype is expressed in testis. The majority of the H2a protein found in diploid adult tissues is also the H2a- β subtype; however, the size of the H2a transcripts differs between testis and other tissues. We conclude that different members of the late H2a gene family are differentially expressed in embryos and adult tissues. We prepared and characterized cDNA clones encoding the sperm-specific H2b protein as well as the H2a- β protein found in testis.

Sea urchins express at least four distinct histone gene families in a temporal and tissue-specific manner. The early histone proteins, encoded by several hundred tandem repetitive genes (13, 21), are transiently expressed in the oocyte and up until the late blastula stage (6, 14). The late histone proteins (8, 33) are encoded by 5 to 10 genes that are not arranged as tandem repeats (7, 22). Late histone mRNAs are also expressed in the oocyte and early embryo; however, they accumulate until the gastrula stage of development, when they represent >98% of the total histone mRNA in the embryo (15). This program of differential histone gene expression in the sea urchin embryo results in the accumulation of structurally different subtypes of H1, H2a, and H2b proteins in the chromatin of the animal (6, 8, 27, 28, 33). The role of these different H1, H2a, and H2b subtypes remains unclear. While the early and late H3 and H4 genes encode different mRNAs (6, 14), the proteins they encode are identical (6, 14, 28, 32)

The sperm-specific H1 and H2b subtypes (36–38) and the cleavage stage H1, H2a, and H2b proteins (27, 30) are two other classes of developmentally regulated histone proteins. The sperm histones are expressed in a sex- and tissue-specificmanner only in the testis. Presumably they enable the sperm DNA to highly condense in the compact sperm pronucleus. Since the paternal genome is expressed by the two-cell stage, only 1.5 h following fertilization (20), there must be a mechanism to rapidly decondense and remodel the inactive sperm chromatin following fertilization (27, 30). Cleavage stage histones presumably have this role. In this regard, they resemble the replacement variant histones described in other organisms (3, 42, 44). Cleavage stage histones are made in occytes, where they accumulate as a pool of stored proteins (30).

While there is a considerable body of information about the transcriptional (6, 14, 23, 24, 41) and post-transcriptional

(24, 41) regulation of histone mRNAs in the early embryo, there are little or no data about which genes are utilized in the adult somatic and germ line cells. We report here the identity of the classes of histone mRNAs found in several somatic tissues and in testes that are actively producing sperm. Furthermore, we prepared cDNA clones encoding the testis H2b and H2a proteins from two different sea urchin species and demonstrate their tissue-specific expression.

MATERIALS AND METHODS

Preparation of RNA probes. RNA probes were prepared by using SP6 polymerase (New England BioLabs, Inc.) to transcribe off linearized templates as described by Melton et al. (25). A 100- μ Ci amount of [α -³²P]GTP, 400 Ci/mmol (Amersham Corp.), was included in each reaction in place of cold GTP. Following transcription, DNase treatment, and extraction of the reaction mixture with phenol-chloroform, the RNA was purified by ethanol precipitation in the presence of 0.7 M ammonium acetate.

The templates were the following: (i) For the H2a-specific template, a 900-base pair BamHI/EcoRI fragment of pSpL3. containing a Strongylocentrotus purpuratus late H2a gene (22), was subcloned into pSP65 (25). (ii) For the H2b-specific template, an 800-base pair BamHI/Bg/II insert of pSpL1, containing an S. purpuratus late H2b gene (22), was subcloned into the BamHI site of mp8 (26) from which a BamHI/EcoRI fragment was subcloned into pSP65. (iii) For the H3-specific template, a 270-base pair HpaII/HindIII fragment of pLpH3H4-1 (7) containing a late Lytechinus pictus H3 gene was subcloned into pSP64. (iv) For the H4-specific template, a 3.3-kilobase HindIII/BamHI fragment of a subclone of LpH3H4-21, containing an L. pictus late H4 gene, was subcloned into pSP65. (v) For the H1specific template, a 400-base pair EcoRI/HindIII fragment of Lp-1H1-1, containing a late L. pictus H1 gene (J. Knowles and G. Childs, unpublished data), was subcloned into pSP65.

^{*} Corresponding author.



FIG. 1. Analysis of the histone proteins found in adult tissues. Histone proteins isolated from (A) nuclei of the indicated S. *purpuratus* tissues and (B) sperm of both S. *purpuratus* and L. *pictus* were resolved on acid-urea Triton X-100 gels. Following electrophoresis the proteins were stained with amido black and photographed. The identity of each histone subtype is shown. SpH1 is the larger sperm-specific H1 subtype. H2b-1 and H2b-2 are the two larger sperm-specific H2b subtypes.

Gel electrophoresis. Histone mRNA species were resolved on 4.8% acrylamide–0.16% bisacrylamide–7 M urea gels (28 by 13 by 0.3 cm) as previously described (6). Electrophoresis was carried out at 150 V for 24 to 28 h.

RNA transfer. RNA gels were electrophoretically transferred to Nytran membranes (Schleicher & Schuell) in a Bio-Rad transblot apparatus. Transfer was at 4°C in TAE buffer (10 mM Tris, 5 mM sodium acetate, 0.5 mM EDTA, pH 7.8) for approximately 15 h at 300 mA. Following transfer, the Nytran membrane was soaked in $6 \times SSC$ (1× SSC is 150 mM NaCl plus 150 mM sodium citrate) and baked in a 65 to 80°C oven for 2 h as outlined by the manufacturer.

Primer extensions. The H2b-specific oligomer was synthesized on an Applied Biosystems automated DNA synthesizer. The oligomer was end labeled with T4 polynucleotide kinase (P-L Biochemicals) in the presence of a twofold molar excess of $[\gamma^{-32}P]ATP$ (3,000 Ci/mmol; Amersham). The end-labeled oligomer was separated from unincorporated label by DEAE-cellulose chromatography. The unincorporated triphosphates were eluted with 0.2 M NaCl, and the oligomer was eluted with 30% (vol/vol) triethylamine bicarbonate, pH 8.3. A 10-µg amount of each RNA sample was hybridized to a 10-fold molar excess of primer to histone mRNA found in total gastrula RNA. The hybridization was carried out in a 10-µl volume of 700 mM KCl–50 mM HEPES (N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid), pH 7–1 mM EDTA. The mixture was heated to 90°C for 3 min, allowed to come to room temperature for 6 h, and finally put at 4°C overnight.

The primers were extended with reverse transcriptase (Life Sciences) as outlined by Maniatis et al. (19). Actinomycin D (0.1 mg/ml) was included in each reaction. Reverse transcriptions were carried out for 0.5 h at 4°C, 1.5 h at room temperature, and finally, 2 h at 42°C. The reaction was stopped by addition of EDTA to 20 mM, and the RNA template was degraded by addition of NaOH to 0.5 M. Following neutralization with 0.5 M Tris (pH 7.5)–0.5 M acetic acid and ethanol precipitation, the single-stranded cDNA was electrophoresed on 5% acrylamide–7 M urea sequencing gels.

Preparation of RNA. RNA from embryos and eggs was isolated as described previously (6). Coelomocyte suspensions were made 25 mM EGTA [ethylene glycol-bis(β -aminoethyl ether)-N,N,N',N'-tetraacetic acid] and pelleted by centrifugation. RNA was isolated from cell pellets, using guanidine-HCl as for embryos. Testes, ovaries, intestines, and lantern tissue were frozen in liquid nitrogen upon removal from the animals. Muscle RNA was isolated by swirling the lantern tissue in guanidine-HCl. RNA from testes, ovaries, and intestines was prepared by homogenizing the tissues in guanidine-HCl, using a Brinkmann Polytron homogenizer.

All tissues with the exception of testis and ovary were collected from both male and female animals. These samples and ovary tissue had small quantities of contaminating eggs which contain high concentrations of stored early histone mRNAs (13, 21). This became most apparent with late H2b and with H3 and H4 probes and in control experiments with early histone gene probes. Since the histone proteins isolated from chromatin of adult tissues contain no detectable early histone subtypes (see Fig. 1) (L. Cohen, personal communication), the presence of early mRNA in adult tissues will not be considered relevant.

Hybridization conditions. Electroblot Northerns were prehybridized and hybridized as outlined by Schleicher & Schuell for Nytran membranes. Hybridizations were carried out at 50°C for 15 h. Five percent dextran sulfate, 10 mM dithiothreitol, and 10⁶ cpm of RNA probe per ml were included in the hybridizations. Blots were washed sequentially in 10× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–1% sodium dodecyl sulfate at room temperature, 1× SSC-1% sodium dodecyl sulfate at 37°C, 0.5× SSC at 65°C, and 0.5× SSC-1% sodium dodecyl sulfate–1% sodium pyrophosphate at 65°C.

Preparation of cDNA clones. Testis RNA was fractionated on 5% acrylamide-7 M urea gels as described above. The RNAs 350 to 650 nucleotides in size were excised and then eluted from the gel slice by the procedure of Gross et al. (10). A 1-µg amount of this RNA was annealed to 50 µg of random hexamers (P-L Biochemicals), and first-strand cDNA was made as described before (19). The second strand was synthesized by the procedure of Gubler and Hoffman (11). Escherichia coli DNA ligase and β-NAD were omitted from the reaction. The double-stranded cDNA (L. pictus) was then tailed and cloned into appropriately tailed Pst-cut pBR322 as described previously (11) or made blunt ended (S. purpuratus) with T4 DNA polymerase (New England BioLabs) and cloned into EcoRI-cut pUC9 vector following the addition of EcoRI linkers. The cDNA clones were transformed into E. coli RRI by the method of Hanahan (12). Clones hybridizing to the *Eco*RI insert of plasmid pSpL3 (22) were subjected to further analysis.

DNA sequence analysis. Testis-specific H2a and H2b

cDNA clones were subcloned into appropriately digested M13mp19 vector and sequenced by the dideoxy sequencing technique (34).

Triton-acid urea gels. Protein samples were resolved by Triton-acid urea gel electrophoresis as previously described (6, 8, 28).

RESULTS

Coding capacity of adult tissue RNAs. We determined the histone subtypes expressed in several adult tissues by analvsis of the histones extracted from chromatin. Histone proteins isolated from chromatin of several adult tissues and sperm were resolved on acid-urea Triton X-100 gels (Fig. 1). Several important facts can be inferred from the data displayed on these gels. First, the α subtypes (early histones) are absent from these adult tissues. Second, while embryos express equivalent quantities of three late H2a proteins, H2a- β , H2a- γ , and H2a- δ (6, 8, 28), each of the adult tissues contains predominantly the β form of H2a. Small quantities of H2a- δ are found in intestines and tubefeet and only in lower quantities or not at all in ovary or sperm. Only tubefeet contain significant amounts of H2a- γ . It is also clear that diploid tissues contain late forms of H2b. Finally, sperm normally contain a unique H1 and two novel H2b subtypes (36-38). L. pictus sperm histones contain both H2b-1 and H2b-2 proteins, while S. purpuratus sperm histones contain predominantly the H2b-2 subtype (Fig. 1B).

Analysis of the histone mRNAs found in adult tissues. We identified testis-specific H1 and H2b transcripts and the other histone mRNAs that accumulate in adult tissues by northern analysis (1) of electrophoretically transferred 5% acrylamide-7 M urea gels. The analysis of the testis material was carried out with two different sea urchin species, S. purpuratus and L. pictus, whose last common ancestor lived over 60 million years ago (35). The probes used for this analysis are individual late histone gene fragments subcloned into vectors for synthesis of high-specific-activity RNA probes. While these probes show some bias for late histone mRNA, experiments with early histone gene probes (data not shown) and with a testis-specific probe (see below) with one exception failed to detect additional novel histone RNA species. The use of RNA probes enhanced the detection of very low levels of histone mRNAs in largely nondividing cells of muscle, coelomocyte, and intestine.

Histone H1 transcripts. Using cloned late H1 probes (Knowles and Childs, unpublished data), we could detect two large testis-specific mRNAs in S. purpuratus (Fig. 2) and a single testis-specific transcript larger than either early or late H1 mRNA in L. pictus (see Fig. 4). Only a single sperm H1 subtype has been reported. Its size, 248 amino acids (38), is larger than either early H1 (204 amino acids [17]) or late H1 (216 amino acids; Knowles and Childs, in preparation) protein. Further analysis of the coding capacity of these two testis-specific transcripts by isolation of cDNA clones will be necessary to resolve the question of how many different sperm H1 genes exist. Ovary, intestine, coelomocyte, and muscle all contain transcripts that comigrate with the embryonic late H1 mRNA. This is consistent with the immunological detection of late but not early H1 subtypes in intestine (29). The late H1 gene family therefore appears to be used by adult cells with the exception of those cells in the testis that are producing sperm.

Histone H2b transcripts. Both L. pictus and S. purpuratus testes contain H2b transcripts that are larger than either the early or late H2b mRNAs (Fig. 3A and 4). The larger size of

MOL. CELL. BIOL.



FIG. 2. Comparison of histone transcripts in eggs, embryos, and adult tissues of S. purpuratus. A 10-µg portion of RNA from the indicated sources was electrophoresed on 5% acrylamide-7 M urea gels, electrophoretically transferred to Nytran membranes, and hybridized to individual histone gene RNA probes as described in Materials and Methods. All samples hybridized to any given probe were from the same gel. Exposure times differ due to the different amounts of histone mRNAs in embryos and different tissues. E, Early histone mRNAs from tandem repetitive genes; L, late histone mRNAs from dispersed late histone genes; T, histone mRNAs encoding testis subtypes. See Materials and Methods for an explanation of the presence of some early-sized transcripts in tissues. Exposure times for indicated lanes were as follows: H1-late blastula and gastrula, 6 h, one intensifying screen, all others 15 h; H2a and H2b-15 h, with two intensifying screens; H3-blastula, gastrula, and testis, 6 h, with one screen, all others 15 h with one screen; H4-intestines, 15 h, with one screen, all others 6 h with one screen.

these mRNAs was expected because the sperm-specific H2b proteins are 21 to 23 amino acids longer than their other H2b counterparts (36, 37). *L. pictus* has two faint transcripts of



FIG. 3. Comparison of S. purpuratus embryonic and testis H2a and H2b transcripts. A 10- μ g portion of the indicated RNAs isolated from S. purpuratus was resolved on 5% acrylamide-7 M urea gels as described in Materials and Methods. The RNAs were transferred to Nytran membranes and hybridized with late H2b (A) and late H2a (B) RNA probes as outlined in the text. Marks indicate the two H2b transcripts in the testis 2 preparation seen better in shorter exposures. Exposure times: H2a, 5 days with one screen; H2b, 15 h with one screen.

higher molecular weight and three transcripts that comigrate with late H2b mRNA (Fig. 4). S. purpuratus contains a single predominant large H2b transcript in both testis RNA preparations (Fig. 3A). The testis 2 preparation contains, in addition to the single transcript found in the testis 1 preparation, an additional minor H2b transcript also larger than either early or late H2b mRNA (Fig. 3A). In contrast to L. pictus, S. purpuratus testis RNA contains little detectable late H2b transcripts. The late mRNAs seen in L. pictus are likely to be derived from diploid somatic cells in the testis and appear most abundant in this blot because the probe, a late H2b gene, is more homologous to itself than to the sperm-specific transcripts. Other experiments demonstrate that the large testis-specific transcripts make up the majority of the H2b mRNA in this tissue (see below).

Histone H2a transcripts. There are no known tissuespecific subtypes of the H2a, H3, and H4 proteins. In *S. purpuratus* the β subtype of H2a is the most abundant H2a protein in chromatin of most tissues (Fig. 1). Of six resolvable embryonic H2a mRNAs, only two are highly abundant in *S. purpuratus* testis RNA (Fig. 3B). Two testis RNA preparations each contain the third and fourth largest of the H2a mRNAs, while the testis 1 preparation also contains smaller amounts of the largest and smallest H2a mRNAs. It is likely that the third and fourth bands both encode an H2a- β protein. Busslinger and Barberis (4) have characterized testis H2a cDNA clones, and they all encode a protein H2a-3 shown to be the most abundant sperm H2a protein. H2a-3 of the sea urchin species *P. milaris* and H2a- β of *S. purpuratus* are probably equivalent (see below).

Interestingly, other S. purpuratus tissues contain quite different arrays of H2a mRNAs (Fig. 2 and 3B). In intestine, coelomocyte, and muscle, the largest of the six embryonic H2a mRNAs is the most abundant H2a transcript, while ovary contains contaminating early H2a mRNA and at least two other late H2a transcripts. It seems clear that either there are multiple H2a genes that encode β -subtype protein or a single β gene can encode different-sized transcripts that accumulate in a tissue-specific manner.

Histone H3 and H4 transcripts. Adult tissues express a set of H3 and H4 transcripts that resemble late embryonic mRNAs in size and relative abundance (Fig. 2). Three size classes of late H3 and two size classes of late H4 transcripts are present in all tissues in about the same relative amounts. As discussed in Materials and Methods, ovary, coelomocyte, and muscle tissues were collected from both male and female animals. Very small quantities of contaminating eggs (which have large pools of stored early histone transcripts) are responsible for the early H3 and H4 RNAs seen in these preparations. This interpretation is supported by the absence of these transcripts in male testis and by the absence of early protein in chromatin (eggs have large reserves of mRNA in their abundant cytoplasm but only a single nucleus and therefore very little histone to contribute to chromatin preparations). The failure to clearly see late mRNAs for these genes in several tissues could also be due to the use of heterologous L. pictus probes to the S. purpuratus RNA samples.

Analysis of testis-specific H2b transcripts by primer extension analysis. The presence of testis-specific H2b transcripts was further studied by primer extension analysis. The amino acid sequences of *S. purpuratus* early and late H2b proteins have been inferred from DNA sequences (18, 40; R. Maxson, personal communication), and several sperm H2b proteins have been sequenced from other sea urchin species (36, 37). We prepared an oligonucleotide for primer extension analysis that would recognize each of these different H2b transcripts (Fig. 5B). The results of these experiments clearly indicate that testis RNA contains transcripts larger than early and late H2b mRNAs and that these extra sequences must be at the 5' end of the RNA (Fig. 5A). In *L. pictus*, two testis-specific products of 275 and 290 base pairs



FIG. 4. Comparison of (A) H1 and (B) H2b transcripts in embryos and testis of *L. pictus*. A 10- μ g portion of RNA from the indicated sources was prepared and hybridized with H1 and H2b RNA probes, exactly as described in the legends to Fig. 2 and 3. See legend to Fig. 2 for definition of T, L, and E.



FIG. 5. Analysis of H2b transcripts from embryos and adult tissues by primer extension. (A) The oligonucleotide depicted in (B) was labeled at its 5' end with polynucleotide kinase and used as a primer to assay the mRNAs from the indicated sources. The single-stranded cDNAs generated in the reactions were resolved on 5% acrylamide-7 M urea sequencing gels, using 5'-end-labeled pBR322 *Hinfl* or ϕ X174 *Hae*III digests as size standards. Testis 1 and 2 correspond to the same preparations used in the Northern blots shown in Fig. 4. Note the presence of two primer extension products in testis 2 sample. (B) The H2b-specific oligonucleotide sequence and the amino acid sequence it encodes on the H2b protein are shown. The amino acid residues this oligonucleotide spans in the early, late, and testis H2b-2 subtype H2b proteins are indicated. The asterisk denotes a mistake in the oligonucleotide produced, resulting in a mismatch at the 3' position. The oligonucleotide is therefore effectively a 13 mer. See legend to Fig. 2 for definition of T, E, and L.

can be visualized, while in S. purpuratus one major product of 285 base pairs can be seen along with a minor product (280 base pairs) in one of the two RNA preparations. The oligonucleotide spans amino acid residues 80 to 84 of the sperm H2b-2 protein (37) and residues 81 to 85 of the sperm H2b-1 protein (36). The predicted extension products for sperm H2b mRNA would be 251 to 254 nucleotides plus the length of the 5'-untranslated region. We therefore expect the leader region to be about 31 to 34 bases for S. purpuratus sperm H2b mRNA and 36 to 39 and 21 to 24 bases for the two L. pictus transcripts. Since the H2b oligonucleotide recognizes all H2b transcripts equally, the primer extension analysis reveals the true relative abundance of spermspecific transcripts to late transcripts found in testis tissue. Sperm-specific H2b mRNAs are present in much higher quantity than late mRNAs that presumably come from diploid somatic cells in the testis. The sizes of the early (blastula) and late (gastrula) H2b primer extension products coincide with the known sequences of the early and late H2b genes and the difference in size of early and late H2b mRNAs (6, 14).

The H2b oligonucleotide was also tested with several S. *purpuratus* adult tissues in addition to testis. A product 285 bases in length was never seen in any tissue other than testis or at any embryonic stage tested (Fig. 5). This result is consistent with the Northern blot analysis (Fig. 2) and further confirms the tissue specificity of the large testis H2b transcripts.

Preparation of testis-specific cDNA clones. Testis RNAs from *S. purpuratus* and *L. pictus* were separately resolved on preparative 5% acrylamide–7 M urea gels, and RNA the size of core histone mRNAs was excised and purified. Partially purified histone mRNA was used to construct cDNA libraries with random oligonucleotide primers (see Materials and Methods). Several independent H2b and H2a clones were selected from each library and their nucleotide sequence was determined.

Six independent S. purpuratus testis H2b clones were sequenced. The sequence of the longest cDNA insert is shown in Fig. 6. Each of the six clones that were sequenced encodes the sperm H2b-2 protein. The only difference in the deduced amino acid sequence among these six cDNAs was

Α

S. purpuratus testis H2b-2 cDMA sequence

10203853TCAGCTTCAT CTTAACCGCA ACA ACA ATG CCA AGA AGT CCG TCC AAG ACT AGC CCG AGG
HET Pro Arg Ser Pro Ser Lys Thr Ser Pro ArgAAG GGT AGC CCT CGC CGA GAG GGT AGC CCT AGT CGC AAG GCT AGC CCC AAG
Sly Ser Pro Arg Arg Gly Ser Pro Ser Arg Lys Ala Ser Pro Lys Arg Gly113
CGC AAG GGA GCA AAG CGA GCT GGA AAG GGA AGG GGT CGT CGT AGG AAT GTC GTC AGG
Gly Lys Gly Ala Lys Arg Ala Gly Lys Gly Gly Arg Arg Arg Asn Val Val Arg113
CGT CGC CGC CGT CGC CGA GAG AGG AGC TAC GGC ATC TAC AAA GTG CTG AGG
Arg Arg Arg Arg Arg Glu Ser Tyr Gly Ile Tyr Ile Tyr Lys Val Leu Lys128
CGT CGC CGC CGA CAC TC GGC ATC TC CGC GGA ATG TCT GTG AGG GTT CAC CCC GAC ACT GGC ATC TCC AGG CGT GGA ATG TCT GTG AAG
Arg Arg Arg Arg Arg Glu Ser Tyr Gly Ile Tyr Ile Tyr Lys Val Leu LysCAG GTT CAC CCC GAC AAT TC GGC ATC TCC CGT GGA ATG TCT GTG ATG AAC AGC
Gln Val His Pro Asp Thr Gly Ile Ser Ser Arg Gly MET Ser Val MET Aan SerTTC GTC AAT GAT ATC TTC GGG AGG ATC GCC GGA GAG ATC CAG AGG GCT TCC CGT CTG ACT ACT
AAT GAT ATC TTC GGG AGG ATC GCC GGA GAG ATC CAG ACT GCT GTT CGC CTT
Ala Asn Asp Ile Phe Gly Arg Ile Ala Gly Glu Ala Ser Arg Leu Thr Arg383
GCC AAC CGA GGA GAG AGG ATC AGC AGC GCG GAG ATC CAG ACT GCT GTT CGC CTT
Ala Asn Arg Arg Ser Thr Ile Ser Ser Arg Glu Ile Gln Thr Ala Val Arg Leu
Ser Glu Gly Thr Lys Ala Ser383
GCC CTC CCA GGA GAA AGG GAA CG ACC AAG CAC CGC TG TAA AACC AGG GCT GTA CAAG AGG GCC GTG
Leu Leu Pro Gly Glu Leu Ala Lys His Ala Val Ser Glu Gly Thr Lys Ala Val

В

L. pictus testis H2b cDNA sequence

 15
 30
 45

 ATT TAC AGG GTG CTG AAG CAG GTC CAC CCA GAC ACT GGC ATC TCG AGC CGT GCA
 11e Tyr Arg Val Leu Lys Gin Val His Pro Asp Thr Gly Ile Ser Ser Arg Ala

 60
 75
 90
 105

 ATT TC CGC ATG ATG AAT AGC TTC GTC AAC GAT GTC TTC GAG CGG ATC GCT GGG GAG
 105

 MET Ser Val MET Asn Ser Phe Val Asn Asp Val Phe Glu Arg Ile Ala Gly Glu
 105

 GCT TCT CGC TTG TGT CAA GCC AAC CGA CGT CGC ACC ATC TCC AGC CGG GAG ATC
 135

 Ala Ser Arg Leu Cys Gln Ala Asn Arg Arg Arg Thr Ile Ser Ser Arg Glu Ile
 165

 165
 180
 195
 210

 CAG ACT GCT GTT CGT CTA CTT CTC CCA GGA GAA CTG GCC AAU CAT GCT GTA TCC
 110
 210

 165
 180
 195
 210

 CAG ACT GCT GTT CGT CTA CTT CTC CCA GGA GAA CTG GCC AAU CAT GCT GTA TCC
 191
 108

 165
 225
 240
 255
 268
 278

 GAG GGA ACG AAG GCC GTG ACC AAT TAT ACC ACC TCC CGC TAA ATTTCACCAG CTTTCGTTTA
 228
 298
 298

 GCCTTTAGCT GGAAAATAAC CAAATGGCT
 298
 298
 298
 298
 298

an arginine at position 26 instead of a lysine (Fig. 6). This difference is due to a single base change in one of the six clones that could have resulted from either polymorphism in the testis RNA (made from several individuals) or reverse transcriptase errors. Three amino acid sequence changes were found between the *S. purpuratus* H2b-2 protein and the equivalent protein from a different species, *P. angulosus*. The *P. angulosus* protein has an alanine at position 5 (instead of serine), an arginine at position 43 (instead of asparagine), and a lysine at position 46 (instead of an arginine). These differences most likely reflect the evolutionary distance between these two species. All six of our cDNAs were >99% homologous with each other. These results support the conclusion that *S. purpuratus* testis contains one major sperm H2b protein (Fig. 1B) and only a single major H2b transcript (Fig. 3A and 5) encoding the testis H2b-2 subtype. The cDNA clone shown encodes a leader sequence of 23 nucleotides, eight bases shorter than a full-length transcript predicted by the primer extension analysis. An additional 20 to 30 bases of 3'-untranslated sequence including the highly conserved 3'-terminal palindromic sequence (5, 13) are also not included in the H2b cDNA clone.

The sequence of an H2b cDNA from the *L. pictus* clone bank was also determined (Fig. 6). This insert extends from amino acid residue 57 (relative to testis H2b-2 protein) to a portion of the 3'-untranslated region. The protein encoded by this cDNA is typical of H2b proteins but contains several changes. Perhaps this is due to the much greater evolutionary distance between *L. pictus* and most other sea urchin С

S. purpuratus testis H2a-B cDNA sequence

D

L. pictus testis H2m cDNA sequence

FIG. 6. Nucleotide sequences of testis H2b and H2a cDNA clones from *S. purpuratus* and *L. pictus*. The nucleotide sequences of representative cDNA clones encoding testis H2b or H2a proteins along with the deduced amino acid sequences are shown. The only amino acid change seen among the six *S. purpuratus* H2b cDNAs sequenced results from an A-to-G transition at nucleotide position 103, resulting in an arginine residue replacing the lysine.

species studied (35). A genomic clone exactly corresponding to this cDNA clone has been selected, and the pentapeptide repeats at the N terminus of the protein coding exon showed it to be an H2b-2 gene (Z. C. Lai and G. Childs, manuscript in preparation). The amino acid sequence differences seen are all in variable positions on H2b-2 proteins. The L. pictus H2b-2 protein contains a cysteine at position 99 and asparagine at position 100 (Fig. 7). Other sperm H2b-2 proteins contain threonine, alanine, or serine at position 99 and threonine or histidine at position 100. The cysteine residue is unusual and has not been seen in other sperm H2b proteins. This residue may play a role in cross-linking of nucleosomes and in the condensation of sperm chromatin. While the analysis of the L. pictus genomic clone showed it to encode a sperm H2b protein, we tested the specificity of the cDNA by Northern analysis, probing individual adult tissues and embryonic RNA. The sperm H2b cDNA hybridizes almost exclusively with a single testis mRNA and only poorly with *S. purpuratus* early H2b mRNA and poorly or not at all with late H2b transcripts (Fig. 8). This result further demonstrates the tissue-specific accumulation of these sperm H2b mRNAs.

We also compared the structure of several H2a cDNA clones from *L. pictus* and *S. purpuratus* (Fig. 6). Three *S. purpuratus* clones that we sequenced each encode part of an H2a protein identical in sequence with H2a proteins found in sperm of *P. angulosus* and *P. milaris* (4, 39). Comparison of this sequence and a *P. milaris* sperm H2a-3 protein encoding cDNA (4) showed them to be identical in predicted amino acid sequence and 95% conserved on the nucleotide level. We conclude that the H2a-3 and - β subtypes of H2a are equivalent. We directly tested the observation that H2a- β



FIG. 7. Comparison of sea urchin H2a and H2b amino acid sequences. The amino acid sequences of H2a (A) and H2b (B) testis proteins from *S. purpuratus* and *L. pictus* deduced from the cDNA sequences are compared with previously determined *S. purpuratus* early H2a and H2b sequences (encoding α subtypes). The repeating pentapeptides at the N terminus of the H2b proteins are underlined. The one-letter amino acid code is as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His: I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; Y, Tyr.

mRNA was a different size in testis compared with other tissues (Fig. 2 and 3B) by stringent hybridization of the H2a- β cDNA to a Northern blot containing RNA from tissues and embryos. The testis H2a- β cDNA preferentially hybridizes with the same testis mRNAs as the late H2a probe does but not with the larger H2a- β mRNA that is most abundant in late embryos and other tissues (Fig. 9). The testis H2a- β probe also hybridizes with an additional mRNA (not seen with other H2a probes) in the testis 2 RNA preparation that is larger than any of the previously identified H2a mRNA species. This mRNA has been seen in varying quantities in different testis RNA preparations but not in any of the other tissues tested (data not shown).

The L. pictus H2a cDNA encodes a minor H2a subtype differing from H2a- β . The predicted amino acid sequence of this sperm H2a protein differs in a minimum of three residues from the S. purpuratus protein and extends to a termination codon predicting a protein of 123 amino acid residues. The H2a-3 protein is 125 residues in length, with the extra residues residing at the C terminus of the protein (39). The nucleotide sequences of the H2a cDNAs from the two species that we have analyzed differ by 12%. This is the same degree of diversion found between the S. purpuratus sperm H2a- β cDNA clone (Fig. 6) and an S. purpuratus late H2a gene on clone L3 that encodes a different but unidentified subtype (24; Maxson, personal communication). As expected, based on the predicted amio acid sequences (Fig. 7) (4), the *S. purpuratus* and *P. milaris* sperm H2a genes are more closely related to one another on the nucleotide level (95%) than to other late histone genes found within the same genome (88%).

DISCUSSION

We have characterized the histone mRNAs found in adult tissues of the sea urchin. The transcripts we have identified are derived with two exceptions from the late histone gene family. We conclude that the late histone genes are active at all stages of the sea urchin life cycle. These genes are transcribed in oocytes. The mRNAs are stored in mature eggs, and they continue to be synthesized at all stages of embryogenesis (15). The term late histone genes is a misnomer and, perhaps as with the 5S rRNA genes (43), the term somatic genes is a more accurate name.

All members of the late histone gene family are not coordinately expressed in tissues. Indeed, we have found a considerable amount of developmental and tissue-specific choices in the accumulation of some H2a family members. The β subtype of H2a is the most abundant H2a protein found in all tissues examined (Fig. 1). The H2a-B protein appears to be encoded by more than one size class of mRNA (Fig. 2, 3, and 9). Cell-free translation of late embryonic H2a mRNAs excised from acrylamide gels demonstrated that the H2a-B mRNA is the largest of the late H2a mRNA species (6). This large H2a transcript also predominates in all adult tissues except testis, where two smaller and one larger H2a transcripts are most abundant (Fig. 2, 3, and 9). Based on the sequence of three independently isolated H2a cDNA clones and our cell-free translation experiments (data not shown). these two testis H2a mRNAs (bands 4 and 5, Fig. 3) encode the H2a-3 sperm protein (H2a-B). Indeed, stringent hybridization of the testis H2a- β cDNA insert to Northern blots confirmed this hypothesis (Fig. 9). There are most likely several H2a-B genes encoding these different-sized transcripts.

We have also demonstrated that late H3 and H4 genes are active in the adult tissues tested. These same genes are utilized along with late H1, H2a, and H2b genes in embryos, ovary, intestine, muscle, and coelomocytes, yet in testis their synthesis is coordinated instead with testis-specific H1 and H2b genes and a limited subset of late H2a genes. The most likely explanation for this observation is that each of the five classes of histone proteins can be independently expressed in response to different subset-specific transcription factors produced in embryos and tissues.

The synthesis of mRNAs encoding sperm-specific H1 and H2b proteins is under strict tissue-specific control. We could not detect the corresponding mRNAs in any tissue or embryonic stage except the testis. The sperm-specific protein H2b-2 is the major H2b subtype expressed by S.



FIG. 8. Analysis of mRNA species homologous with *L. pictus* testis H2b cDNA probe. A 10- μ g portion of RNA isolated from the indicated sources was resolved on 5% acrylamide-7 M urea gels and transferred to Nytran membranes as outlined in Materials and Methods. The riboprobe hybridized to these filters was constructed by subcloning a *Pst-HincII* fragment of the *L. pictus* H2b cDNA clone into pSP65. The RNA probe was prepared and hybridized to the filters, and the filters were processed as described in the text. (A) *S. purpuratus* RNA; (B) *L. pictus* RNAs. T, Mobility of sperm-specific H2b mRNA; E, mobility of early H2b mRNA; L, mobility of late H2b mRNAs.



FIG. 9. Analysis of transcripts homologous with testis H2a- β cDNA. A 10- μ g amount of the indicated *S. purpuratus* RNA preparations was resolved on 5% acrylamide-7 M urea gels and processed as described in the legends to the previous figures. The blot was hybridized to the isolated insert of the *S. purpuratus* testis H2a cDNA clone (Fig. 6) which had been labeled by nick translation. The dots indicate the mobility of the six embryonic late H2a mRNA transcripts seen in Fig. 3B. The large transcript most abundant in the testis 2 preparation is not detectable with the late H2a probe (Fig. 3b), using exposure times shown.

purpuratus. Three different pieces of information support this conclusion. First, only one major sperm H2b protein could be detected on acid-urea Triton gels (Fig. 1B). Second, only one major-sized mRNA or primer extension product was seen (Fig. 5). Since each testis RNA preparation was prepared from a pool of several individuals, this finding cannot be attributed to variation in individuals. Finally, the sequences of six independent cDNA clones were determined and each encodes an H2b-2 protein. This protein contains multiple copies of a pentapeptide repeat characteristic of the sperm H2b-2 protein at its N terminus (Fig. 7). Testis RNA preparation 2 (Fig. 3 and 5) also contained a small quantity of an additional H2b transcript that may correspond to an H2b-1 protein. Analysis of the coding and flanking regions of genomic clones encoding these two proteins may help to elucidate why S. purpuratus does not express normal quantities of H2b-1 protein. The testis-specific H2b genes are each unique sequences (Lai and Childs, in preparation), so mutations affecting expression of the H2b-1 gene could more readily spread in the population. Obviously, copious quantities of this subtype are not essential for proper condensation of sperm DNA.

In contrast, *L. pictus* expresses two distinct size classes of sperm-specific H2b mRNAs (Fig. 4 and 5). It is likely that these two transcripts encode both H2b-1 and H2b-2 sperm subtypes visible in Fig. 1, though we have not examined enough cDNA clones to conclusively demonstrate this. Based on analysis of cDNA clones, *P. milaris*, like *L. pictus*,

Vol. 6, 1986

expresses a more typical pattern of sperm H2b subtypes (4). This species contains similar quantities of both H2b-1 and H2b-2 subtypes.

The results presented in this study further support the idea that members of the histone multigene family are differentially expressed in a temporal fashion in embryos and in adult tissues. The isolation of sperm-specific subtype H2b genes expands the array of available substrates for the identification and isolation of regulatory factors responsible for the observed stage and tissue specificity of histone gene expression.

ACKNOWLEDGMENTS

We thank Rob Maxson for providing us with the DNA sequences of *S. purpuratus* late H2a and H2b genes used in some of our sequence comparisons. We are also grateful to Scott Hawley whose suggestions improved this manuscript.

This work was supported by a Public Health Service grant from the National Institutes of Health. Toby Lieber is supported by a Public Health Service predoctoral training grant. Computer resources used to carry out some of our studies were provided by the BIONET National Computer Resource for Molecular Biology whose funding is provided by Public Health Service grant 1 U41 RR-01685-02 from the National Institutes of Health.

LITERATURE CITED

- 1. Alwine, J. C., D. J. Kemp, and G. R. Stark. 1977. Method for detecting RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization to DNA probes. Proc. Natl. Acad. Sci. USA 74:5350-5354.
- Bonner, W. M., and A. L. Laskey. 1974. A film detection method for tritium labeled proteins and nucleic acids in polyacrylamide gels. Eur. J. Biochem. 46:83–88.
- Brush, D., J. B. Dodgson, O. K. R. Choi, P. W. Stevens, and J. D. Engel. 1985. Replacement variant histone genes contain intervening sequences. Mol. Cell. Biol. 5:1307-1317.
- 4. Busslinger, M., and A. Barberis. 1985. Synthesis of sperm and late histone cDNAs of the sea urchin with a primer complementary to the conserved 3' terminal palindrome: evidence for tissue-specific and more general histone gene variants. Proc. Natl. Acad. Sci. USA 82:5676-5680.
- 5. Busslinger, M., R. Portman, and M. Birnstiel. 1979. A regulatory sequence near the 3' end of sea urchin histone genes. Nucleic Acids Res. 6:2997-3008.
- Childs, G., R. Maxson, and L. H. Kedes. 1979. Histone gene expression during sea urchin embryogenesis: isolation and characterization of early and late messenger RNAs of Strongylocentrotus purpuratus by gene-specific hybridization and template activity. Dev. Biol. 73:153-173.
- Childs, G., C. Nocente-McGrath, T. Lieber, C. Holt, and J. Knowles. 1982. Sea urchin (L. pictus) late stage H3 and H4 genes: characterization and mapping of a clustered but nontandemly linked multigene family. Cell 31:383–393.
- 8. Cohen, L., K. M. Newrock, and A. Zweidler. 1975. Stagespecific switches in histone synthesis during embryogenesis in the sea urchin. Science 190:194–197.
- Erickson, A. H., and G. Blobel. 1983. Cell-free translation of messenger RNA in a wheat germ system. Methods Enzymol. 96:38-50.
- Gross, K., E. Probst, W. Schaffner, and M. Birnstiel. 1976. Molecular analysis of the histone gene cluster of Psammechinus milaris. I. Fractionation and identification of five individual histone mRNAs. Cell 8:455–470.
- 11. Gubler, U., and B. J. Hoffman. 1983. A simple and very efficient method for generating cDNA libraries. Gene 25:263–269.
- 12. Hanahan, D. 1983. Studies on transformation of Escherichia coli. J. Mol. Biol. 166:557-580.
- 13. Hentschel, C. C., and M. L. Birnstiel. 1981. The organization and expression of histone gene families. Cell 25:301-313.
- 14. Hieter, P. A., M. Hendricks, K. Hemminki, and E. S. Weinberg.

1979. Histone gene switch in the sea urchin embryo: Identification of late embryonic histone mRNAs and the controls of their synthesis. Biochemistry **18**:2707–2715.

- Knowles, J. A., and G. Childs. 1984. Temporal expression of late histone messenger RNA in the sea urchin Lytechinus pictus. Proc. Natl. Acad. Sci. USA 81:2411-2415.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) 227:680-685.
- Levy, S., I. Sures, and L. H. Kedes. 1982. The nucleotide and amino acid sequence of a gene for H1 histone that interacts with euchromatin. The early embryonic H1 gene of the sea urchin Strongylocentrotus purpuratus. J. Biol. Chem. 257:9438-9443.
- Liebermann, D., B. H. Liebermann, J. Weinthal, G. Childs, R. Maxson, A. Mauron, S. N. Cohen, and L. H. Kedes. 1983. An unusual transposon with long terminal inverted repeats in the sea urchin Strongylocentrotus purpuratus. Nature (London) 306:342-347.
- 19. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual, p. 131. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Maxson, R. E., and J. C. Egrie. 1980. Expression of maternal and paternal histone genes during early cleavage stages of the echinoderm hybrid Strongylocentrotus purpuratus × Lytechinus pictus. Dev. Biol. 74:335–342.
- Maxson, R. E., T. J. Mohon, R. Cohn, and L. Kedes. 1983. Expression and organization of histone genes. Annu. Rev. Genet. 17:239-277.
- Maxson, R. E., T. J. Mohon, G. Gormezano, G. Childs, and L. H. Kedes. 1983. Distinct organizations and patterns of expression of early and late gene sets in the sea urchin. Nature (London) 301:120-125.
- Maxson, R. E., and F. H. Wilt. 1981. The rate of synthesis of histone mRNA during the development of sea urchin embryos (Strongylocentrotus purpuratus). Dev. Biol. 83:380–386.
- Maxson, R. E., and F. H. Wilt. 1982. Accumulation of the early histone messenger RNAs during the development of Strongylocentrotus purpuratus. Dev. Biol. 94:435-440.
- Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green. 1984. Efficient in vitro synthesis of biologically active RNA and RNA probes from plasmids containing a bacteriophage Sp6 promoter. Nucleic Acids Res. 12:7034-7056.
- Messing, J., and J. Vieira. 1982. A new pair of M13 vectors for selecting either DNA strand of double digest restriction fragments. Gene 19:269–276.
- Newrock, K. M., C. R. Alfgame, R. V. Nardi, and L. H. Cohen. 1977. Histone changes during chromatin remodeling in embryogenesis. Cold Spring Harbor Symp. Quant. Biol. 42:421-431.
- Newrock, K. M., L. H. Cohen, M. B. Hendricks, R. J. Donnelly, and E. S. Weinberg. 1978. Stage-specific mRNAS coding for subtypes of H2a and H2b histones in the sea urchin embryo. Cell 14:327-336.
- 29. Pehrson, J. R., and L. H. Cohen. 1984. Embryonal histone H1 subtypes of the sea urchin Strongylocentrotus purpuratus: purification, characterization, and immunological comparison with H1 subtypes of the adult. Biochemistry 23:6761–6764.
- 30. Poccia, D., J. Salik, and G. Krystal. 1981. Transitions in histone variants of the male pronucleus following fertilization and evidence for a maternal store of cleavage-stage histones in the sea urchin egg. Dev. Biol. 82:287–296.
- 31. Roberts, B. E., and B. M. Patterson. 1973. Efficient translation of tobacco mosaic virus RNA and rabbit globin 9S RNA in a cell-free system from commercial wheat germ. Proc. Natl. Acad. Sci. USA 70:2330-2334.
- 32. Roberts, S. B., K. Weisser, and G. Childs. 1984. Sequence comparisons of non-allelic late histone genes and their early stage counterparts. J. Mol. Biol. 174:647-662.
- Ruderman, J. V., and P. R. Gross. 1974. Histones and histone synthesis in sea urchin development. Dev. Biol. 36:286–298.
- 34. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci.

USA 74:5463-5467.

- Smith, A. B. 1981. Implications of lantern morphology for the phylogeny of post-paleozoic echinoids. Paleontology 24: 779-801.
- 36. Strickland, M., W. N. Strickland, W. F. Brandt, and C. Von Holt. 1977. The complete amino acid sequence of histone H2b₍₁₎ from sperm of the sea urchin Parechinus angulosus. Eur. J. Biochem. 77:263-275.
- Strickland, W. N., M. Strickland, W. Brandt, and C. Von Holt. 1977. The complete amino acid sequence of histone H2b₍₂₎ from sperm of the sea urchin Parechinus angulosus. Eur. J. Biochem. 77:277-286.
- 38. Strickland, W. N., M. Strickland, W. F. Brandt, C. Von Holt, and B. W. Liebold. 1980. The primary structure of histone H1 from sperm of the sea urchin Parenchinus angulosus. Sequence of the C-terminal CNBr peptide and the entire primary structure. Eur. J. Biochem. 104:567-578.
- 39. Strickland, W. N., M. Strickland, P. C. de Groot, and C. Von Holt. 1980. The primary structure of histone H2a from the sperm

cell of the sea urchin Parechinus angulosus. Eur. J. Biochem. 109:151-158.

- 40. Sures, I., J. Lowry, and L. H. Kedes. 1978. The DNA sequence of sea urchin (S. purpuratus) H2a, H2b, and H3 coding and spacer regions. Cell 15:1033–1044.
- Weinberg, E. S., M. B. Hendricks, K. Hemminki, P. E. Kuwabara, and L. A. Farrelly. 1983. Timing and rates of synthesis of early histone mRNA in the embryo of Strongylocentrotus purpuratus. Dev. Biol. 98:117–129.
- 42. Wells, D., and L. Kedes. 1985. Structure of a human histone cDNA: evidence that basally expressed histone genes have intervening sequences and encode polyadenylated mRNAs. Proc. Natl. Acad. Sci. USA 82:2834-2838.
- Wormington, W. M., and D. D. Brown. 1983. Onset of 5S RNA gene regulation during Xenopus embryogenesis. Dev. Biol. 99:248-257.
- Wu, R. S., and W. M. Bonner. 1981. Separation of basal histone synthesis from S-phase histone synthesis in dividing cells. Cell 27:321-330.