The EMBO Journal Vol. 1 No. 1, pp.27–33, 1982

An unusual evolutionary behaviour of a sea urchin histone gene cluster

Meinrad Busslinger[1], Sandro Rusconi and Max L. Birnstiel*

Institut für Molekularbiologie II der Universität Zürich, Hönggerberg, CH-8093 Zürich, Switzerland

Communicated by M.L. Birnstiel
Received on 19 November 1981

DNA sequences of cloned histone coding sequences and spacers of sea urchin species that diverged long ago in evolution were compared. The highly repeated H4 and H3 genes active during early embryogenesis had evolved (in their silent sites) at a rate (0.5–0.6% base changes/Myr) similar to single-copy protein-coding genes and nearly as fast as spacer DNA (0.7% base changes/Myr) and unique DNA. Thus, evolution in the major histone genes conforms to a universal evolutionary clock based on the rate of base sequence change. By contrast, the H4 and H3 coding sequences and a non-transcribed spacer of the DNA clone h19 of *Psammechinus miliaris* show an exceptionally low rate of sequence evolution only 1/100 to 1/200 that predicted from the clock hypothesis. According to the classical model of gene inheritance, the h19 DNA sequences in the *Psammechinus* genome require unusual conservation mechanisms by selection at the level of the gene and spacer sequences. An alternative explanation could be recent horizontal gene transfer of a histone gene cluster from the very distantly related *Strongylocentrotus dröbachiensis* to the *P. miliaris* genome.

Key words: codon selection/spacer DNA/mutational drift/horizontal gene transfer

Introduction

The isolation of specific gene units (Birnstiel *et al.*, 1966; Birnstiel, 1967) has made it possible to look at the primary event of evolution: the changes in the base sequences of genes and spacers. Recent techniques of gene cloning and DNA sequencing have produced a wealth of data on the rate of evolution of protein-coding sequences (see below) and introns (van Ooyen *et al.*, 1979) but, curiously enough, no quantitative measurements have yet been reported for the rate of spacer evolution.

The DNA sequences coding for even relatively conservative proteins, e.g., haemoglobins, show rapid divergence by accumulation of "silent" base substitutions, especially for a pair of genes in the first 80 million years (Myr) or so (Perler *et al.*, 1980). Over longer periods the apparent rate of nucleic acid evolution decreases due to "saturation effects" (Perler *et al.*, 1980; Holmquist *et al.*, 1981), such as multiple base substitutions at the same site, back-mutation to the original base, and constraints due to special nucleic acid sequence motifs (e.g., exon-intron boundaries).

Rapid gene evolution by accumulation of mutations yielding synonymous codons was first predicted by Walker (1968) and King and Jukes (1969), and soon corroborated by RNA/DNA hybridization studies and RNA sequencing of sea urchin histone mRNA and histone gene sequences. A rate of $3-6 \times 10^{-9}$ substitutions/codon/year was obtained (Weinberg *et al.*, 1972; Birnstiel *et al.*, 1973; Grunstein *et al.*, 1973; Grunstein *et al.*, 1975). Histone genes (in their mutable base positions, see below) appeared to evolve as fast as the genes coding for the ephemeral fibrinopeptides and the bulk of the unique DNA (Birnstiel *et al.*, 1973).

Strong evidence suggests that a percentage of base changes resulting in synonymous codons are selected against (Fitch, 1980; Miyata and Hayashida, 1981; Holmquist and Pearl, 1980). But comparison of the DNA sequences of a great many different protein-coding genes shows that the overall rate of synonymous substitutions even over a large timespan has proceeded at a remarkably even rate of ∿0.5% base changes/Myr (Miyata *et al.*, 1981), i.e., the evolutionary clock of nucleic acid sequences (in their silent sites) seems to run at a similar rate in all structural genes analyzed to date.

We found that the rate of base substitution in the major histone gene families of five sea urchin species conforms to the evolutionary clock hypothesis. However, a DNA clone hl9 of *Psammechinus miliaris*, a minor histone gene repeat unit coding for all five types of histones, blatantly violates this rule.

Results

Evolution of the early embryonic sea urchin H3 and H4 genes by silent mutation

In all sea urchins analyzed to date the rapid cell division during embryogenesis is accompanied by high levels of histone mRNA synthesis, occurring on genes which are many hundred times repeated in the genome. They constitute the predominant family of histone genes present in these species (reviewed by Hentschel and Birnstiel, 1981): e.g., DNA clones h22 of *P. miliaris* (Clarkson *et al.*, 1976; Schaffner *et al.*, 1978); Sp2 and Sp17 of *Strongylocentrotus purpuratus* (Kedes *et al.*, 1975; Sures *et al.*, 1978); and Ph70 of *Paracentrotus lividus* (Spinelli *et al.*, 1979).

Strongylocentrotus dröbachiensis, a close relative of *S. purpuratus*, contains two histone gene maxi-families in a ratio of ∿2:1 (unpublished results) which were cloned as Nor1 and Nor5 (see **Materials and methods**). The sequences of these two clones resemble each other and DNA of clone Sp2/17 of *S. purpuratus* (see below); most probably they represent the early embryonic class of *S. dröbachiensis* histone genes.

Figure 1a and b shows the DNA sequences of clones h22, Sp2/17, Ph70, and Nor5. They code for identical H3 and H4 proteins, respectively, but differ by silent mutations, predominantly in the third base of the codons. The T1 oligonucleotides of the early blastula H4 mRNA of *Lytechinus* (Grunstein and Schedl, 1976) were matched, as far as possible with the H4 DNA sequences of the other sea urchins (Figure 1b).

The special codon composition of H3 and H4 structural genes shows that 25% and 26%, respectively, of all random base substitutions lead to synonymous codons in these genes. All other base substitutions, 3/4 of the total, were eliminated

[1]Present address: Laboratory of Gene Structure and Expression, National Institute for Medical Research, Mill Hill, London NW7 1AA, UK.

*To whom reprint requests should be sent.

© IRL Press Limited, Oxford, England. 0261-4189/82/0101-0027$2.00/0

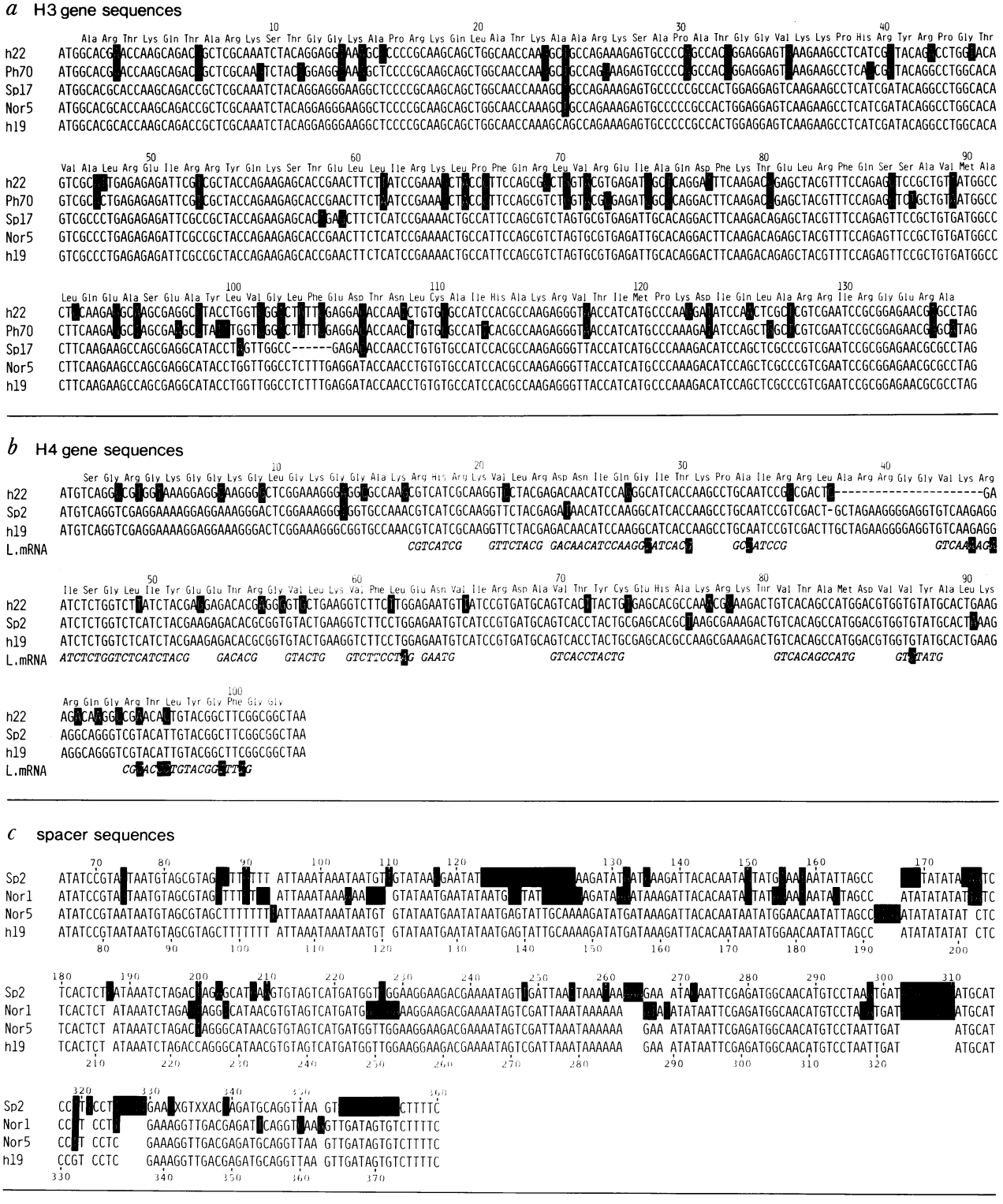

Fig. 1. A comparison of DNA of sea urchin H3 and H4 coding sequences and of spacer segments between the H2B and H3 gene. Single base substitutions and deletion/insertion events relative to the h19 sequence marked in black. Undetermined nucleotides --- in the gene; X in the spacer DNA. Position of the nucleotides downstream from the termination codon of the H2B gene given for the spacer sequence of clone Sp2 and h19 (above and below the line, respectively). Sequencing: *Paracentrotus* clone h19 see **Materials and methods**; *Psammechinus* clone h22, Schaffner *et al.* (1978); *Psammechinus* clone h19, Busslinger *et al.* (1980); h22 sequence in the H4 gene from codon 60 to 93, J.C. Irminger (unpublished results); H3 sequence of *S. purpuratus* clone Sp17 and spacer sequence of *S. purpuratus* clone Sp2, Sures *et al.* (1978); H4 clone Sp2 by Grunstein *et al.* (1981). T1 oligonucleotides of early embryonic H4 mRNA of *Lytechinus*, Grunstein *et al.* (1976), were computer matched as far as possible with the H4 DNA sequence.

by selection for the identical, conserved amino acid sequences of H3 and H4 which therefore represent the immutable nucleotides. To calculate base substitution as a percentage of mutable nucleotides, the observed percent divergence for the entire structural gene must be multiplied by a factor of ~4 (see **Materials and methods**). Table I gives observed percent divergence for all sites and for mutable bases in the silent sites for pairs of structural histone genes and Figure 2b shows the

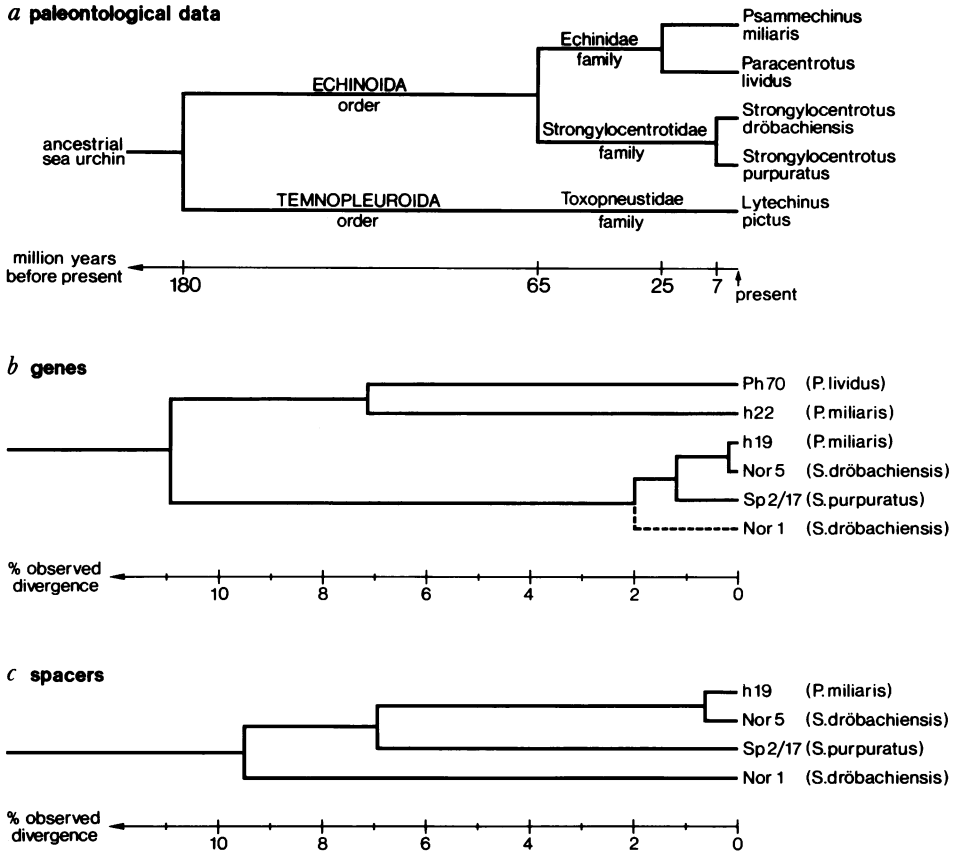

Fig. 2. Evolutionary tree for five sea urchin species and their early embryonic histone gene clusters. **a**, Phylogeny of five sea urchin species by paleontological classification of Durham (1966). **b**, and **c**, Evolutionary relationships between different sea urchin histone gene clusters based on sequence differences in the H3 and H4 gene **b** (see Table I) or in the spacer DNA sequence **c** (see Table II). Approximate percentage sequence divergence between the histone genes of DNA clone Nor1 and those of the other *Strongylocentrotus* DNA clones Nor5 and Sp2 part **b** extrapolated from the sequence differences in their spacer DNA (see part **c**), since no gene sequence of clone Nor1 has been determined.

resulting evolutionary relationship of histone DNA clones.

Divergence of the histone-coding sequences in relation to the time of speciation of the sea urchins

According to the paleontological classification shown in Figure 2a, *Lytechinus* belongs to the Temnopleuroida, whereas *Strongylocentrotus, Paracentrotus*, and *Psammechinus* are members of the Echinoida. These two orders diverged 180 Myr ago (Durham, 1966). The Echinidae, from which the Strongylocentrotidae separated, occurred in the Cretaceous period. Therefore, separation of these two families must have occurred around 65 Myr ago (Durham, 1966). *Paracentrotus* and *Psammechinus*, which are more closely related to one another, can be traced as separate species for at least 25 Myr (Durham, 1966), whereas *S. purpuratus* and *S. dröbachiensis* diverged ∿5 Myr ago (W. Durham, personal communication). The family of Strongylocentrotidae was restricted to the North Pacific until the Pliocene (∿5 Myr ago) when *S. dröbachiensis* migrated through the Arctic to the North Atlantic and Europe, whereas *Psammechinus* and *Paracentrotus* have always been North Atlantic species (W. Durham, personal communication).

Studies of the thermal stability of heteroduplexes formed between single copy DNAs of several sea urchin species support this phylogeny. DNA sequences in the Strongylocentrotidae appear to have diverged at 0.5% base changes/Myr in single-copy DNA (uncorrected for multiple hits, see below) implying a divergence time of 7 Myr for the two species *S. purpuratus* and *S. dröbachiensis* (Hall *et al.*, 1980). Heteroduplexes formed between *Psammechinus* single-copy tracer DNA and *Paracentrotus* driver DNA suggest a sequence divergence of ∿24–36% (Busslinger, 1981). At 0.5% base change/Myr these two species must have diverged ∿24–36 Myr ago.

Since sequences in the mutable bases often diverge quite widely (Table I), multiple hits must be frequent. To calculate the approximate total mutation rates, the observed percent divergence can be corrected for multiple substitutions (Salser, 1977; Perler *et al.*, 1980; and **Materials and methods**). The corrections do not exceed a factor of 2 (Table I) and are slightly too low since the incidence of base substitutions does not follow an exact Poisson distribution as assumed in these calculations (Holmquist and Pearl, 1980; Holmquist *et al.*, 1981). From the corrected percent divergence (Table I) and the times of speciation suggested by the paleontological record, the bases in the silent sites of the sea urchin histone genes are seen to evolve at 0.2–0.9% base changes/Myr. In the H4 gene of *Lytechinus* and *Psammechinus* the relatively low rate of 0.22%/Myr is most probably due to "saturation effects", because extremely long timespans are involved (see **Introduction**); also, only the more conserved oligonucleotides can be aligned with the *Psammechinus* sequence and so the data are biased. The rate of 0.92%/Myr for *Psammechinus* and *Paracentrotus* could be too high due to the inaccuracy of the fossil date (see above). Despite these minor differences, our values compare very favorably with those for the evolution of globin (0.51%/Myr) and insulin (0.7%/Myr) genes (Perler *et al.*, 1980; Miyata *et al.*, 1980).

Table I. Evolution of sea urchin histone genes.

Compared histone coding sequences		Observed percentage divergence for all sites		Observed percentage divergence for mutable bases in the silent sites		Corrected percentage divergence for mutable bases in the silent sites		Rate of silent substitutions
		H3	H4	H3	H4	H3	H4	
h22 ⟷ Ph70								
P. miliaris	*P. lividus*	7.2%		28.7%		46%		0.92%/Myr
h22 ⟷ Sp2/17								
P. miliaris	*S. purpuratus*	11.4%	10.8%	45.4%	41.7%	70%	69%	0.54%/Myr
h22 ⟷ Nor5								
P. miliaris	*S. dröbachiensis*	11.1%		44.5%		70%		0.54%/Myr
h22 ⟷ H4 mRNA								
P. miliaris	*L. pictus*		12.5%		45.7%		78%	0.22%/Myr
Sp17 ⟷ Nor5								
S. purpuratus	*S. dröbachiensis*	1.0%		4.0%		6%		0.60%/Myr
h19 ⟷ Sp2/17								
P. miliaris	*S. purpuratus*	1.3%	1.3%	5.0%	5.1%	7%	7%	(0.05%/Myr)
h19 ⟷ Nor5								
P. miliaris	*S. dröbachiensis*	0.2%		1.0%		1%		(0.008%/Myr)

Comparative analysis of sea urchin histone genes based on Figures 1a and b. Incompletely sequenced codons were omitted from the evaluation of sequence differences. Observed and corrected percentage divergences for mutable bases in the silent sites in **Materials and methods**. Rate of synonymous substitution is the ratio of the corrected percentage divergence to the corresponding time of speciation (Figure 2a).

By heteroduplex melting Hall *et al.* (1980) determined that the unique DNAs of *S. purpuratus* and *S. franciscanus*, and of *S. purpuratus* and *S. dröbachiensis*, show a median sequence divergence of 21% and 7.7%, respectively (Hall *et al.*, 1980). After correction for multi-hit events (assuming that all bases are mutable), we obtained a rate of sequence evolution for sea urchin single-copy DNA of 0.6–0.74%/Myr.

From our sequence and rate comparisons: firstly, all genes considered can be assembled into an evolutionary tree (Figure 2b) paralleling the phylogeny of the species and superimposable on it (Figure 2a); secondly, the mutation rate of synonymous substitutions in the major, early embryonic histone genes of five sea urchin species is in the same range as the rate of sequence evolution of unique genes and only slightly less than that of single-copy DNA.

DNA clone h19 of P. miliaris is strikingly similar to a major histone DNA repeat of S. dröbachiensis

The relatively rare DNA clone h19, represented in *Psammechinus* by 3–5 copies/haploid genome, was initially selected from a collection of lambda-histone DNA recombinants on the basis of its unusual restriction pattern, with a view to studying the structure and expression of histone gene variants distinguishable from those encoded by the predominant clone h22 (Birnstiel *et al.*, 1979). The DNA sequences of clone h19, when transposed into protein sequences (Busslinger *et al.*, 1980), were found to code potentially for H2A and H2B proteins differing from those predicted from the h22 DNA sequences (Schaffner *et al.*, 1978). The amino acid sequences encoded by the DNA clone h19 closely resembled those encoded by DNA clone Sp2/17 of the Pacific species *S. purpuratus* (Busslinger *et al.*, 1980) which suggested that heterodox evolutionary mechanisms might be at work. More surprising still, the DNA sequences of genes and spacers of DNA clones h19 of *P. miliaris* and Sp2/17 of *S. purpuratus* (Sures *et al.*, 1978) were closely similar (Busslinger *et al.*, 1980). This was a totally unexpected finding since the two sea urchin species diverged from each other long ago, as confirmed by the very large sequence divergence obtained in DNA reannealing experiments between the unique DNAs of these two species (Busslinger, 1981). We have now discovered that the H3 DNA of the abundant Nor5 DNA of the European species *S. dröbachiensis* and the rare h19 clone of *Psammechinus* are identical apart from a single base substitution (A-T, codon 24; see Figure 1a).

To determine whether the extraordinary evolutionary behaviour of DNA clone h19 was confined to structural genes, we investigated the spacer DNA lying between the H2B and the H3 gene in the DNA clones h19 and Nor5, and homologous DNA sections of clones Nor1 and Sp2. The sequences (Figure 1c) which are ∿300 bases long, map 3′ to the terminal palindromic sequence typical for histone genes (Busslinger *et al.*, 1979; Hentschel and Birnstiel, 1981), well clear of the mRNA coding sequences. These non-transcribed spacer DNA sequences of the *P. miliaris* clone h19 are unambiguously homologous with the sequences of the *Strongylocentrotus* clones Sp2, Nor1, and Nor5, whereas any matching between spacer partners from other species (not shown) is arbitrary due to the high degree of sequence divergence in the spacer DNA.

The spacer DNA segment of Nor5 (*S. dröbachiensis*) and Sp2 (*S. purpuratus*) differ by 6.9% base substitutions (Table II). Correcting for multiple hits gave a rate of 0.7% base substitution/Myr in spacer DNA. The h19 and Nor5 non-transcribed spacer DNAs differ by only two base substitutions out of 308 bases (0.6%) and by two small insertions and are therefore virtually identical. As *Psammechinus* and *Strongylocentrotus* diverged 65 Myr ago this gives a remarkably low rate of spacer evolution of 0.004% base changes/Myr. This high degree of sequence conservation characterizes not only the H3 and the non-transcribed spacer segment discussed here but also most, if not all, of the 6.7 kb histone gene clusters, since to date both h19 and Nor5 have closely similar restriction maps. Since these two sea urchin species have evolved for a very long time, the DNA sequences

Table II. Evolution of a spacer DNA sequence between the H2B and H3 gene.

Compared histone DNA	Number of substitutions	Number of deletions	Number of substitutions + deletions	Average length of compared DNA sequences (bp)	Percentage divergence (substitutions) only	Total percentage divergence (substitutions + deletions)
Sp2 ⟷ Nor1	32	15	47	301	10.6%	15.6%
Sp2 ⟷ Nor5	21	13	34	304	6.9%	11.2%
Sp2 ⟷ h19	23	12	35	301	7.6%	11.6%
Nor5 ⟷ h19	2	2	4	309	0.6%	1.3%
Nor1 ⟷ h19	29	9	38	306	9.5%	12.4%
Nor5 ⟷ Nor1	27	10	37	308	8.8%	12.0%

Comparative analysis of spacer sequences based on the DNA sequence in Figure 1c. Total divergence, last column, obtained by attributing the same weight to each insertion/deletion event as to a single base substitution.

in h19 and Nor5 seem to have escaped the continuous "random walk" of DNA sequences, which suggests that either the recovery of the DNA clone h19 from *P. miliaris* was an artifact or its presence in the *P. miliaris* genome is due to very unusual evolutionary mechanisms.

Clone h19 is a genuine component of the P. miliaris genome

That the isolation of DNA clone h19 from *P. miliaris* resulted from a contamination of the DNA preparation by gonadal tissue of *S. dröbachiensis*, was ruled out by a second cloning experiment using DNA of a single *P. miliaris* individual, when h19-type DNA clones were found three times, together with 130 h22-type clones of the predominant gene family (Rusconi, 1979). Previously, Southern blots using different restriction enzymes had revealed the standard h22 maxi-family in this very same individual and hence clearly identified this individual as *P. miliaris*. One of these three fresh h19′ DNA clones, all of which differed from the original h19 by a small deletion in the spacer DNA between the H3 and the H2A gene (Rusconi, 1979), was further characterized by sequencing the spacer DNA lying between the H2B and the H3 gene; this proved identical with the original DNA clone h19 (data not shown).

The presence of h19 DNA in *Psammechinus* cannot be due to contaminating *S. dröbachiensis* sperm, at very low concentration in the *Psammechinus* sperm population, for two reasons: firstly, *P. miliaris* spawns in July and *S. dröbachiensis* in January; secondly, the second collection of *P. miliaris* was gathered from the Isle of Cumbrae (W. Scotland) where *S. dröbachiensis* does not occur (W. Finleyson, personal communication) although *P. miliaris* and *S. dröbachiensis* do share the same habitat on the Norwegian coast line. Furthermore, h19 could not have been cloned as a ubiquitous laboratory contaminant because h19 and Nor5 are not identical. Moreover, *S. dröbachiensis* sea urchins were kept in our laboratory only after work on DNA clone h19 had been completed.

That DNA clone h19 had been cloned from parasitic organisms inhabiting both *Psammechinus* and *Strongylocentrotus* sperm and gonads is also ruled out since this DNA represents the major repeat unit in *S. dröbachiensis*. Moreover, the arrangement of histone genes differs dramatically between species of different taxonomic orders (Hentschel and Birnstiel, 1981) and yet DNA clone h19 has all the hallmarks of a typical sea urchin histone clone. Also the h19 DNA clone has been isolated four times from *P. miliaris* DNA and each time a DNA clone has been obtained which closely resembles Nor5, the representative of the less abundant histone gene cluster of *S. dröbachiensis*, as if Nor5 had become separated from Nor1 by "subcloning" into *P. miliaris* from which it can now be isolated as h19.

Discussion

Evolution of spacers and genes in repeated histone gene families

The histone gene clusters active during the cleavage stages of sea urchin embryos are classic examples of the parallel (Birnstiel and Grunstein, 1972; Weinberg *et al.*, 1972) or tandem (Brown *et al.*, 1972) evolution of multi-gene families (Smith, 1973). Although within species tandem repeats are highly homogeneous in sequence, between species the sequences of structural histone genes are known to evolve rapidly (Weinberg *et al.*, 1972; Grunstein *et al.*, 1973). Sequencing of representatives of these multi-gene families now reveals their sequence evolution.

Evolutionary trees based on silent mutations (predominantly within the third base of the codons) of the histone DNA sequences and on the fossil record seem to be superimposable (compare Figures 2a and b) and compatible with results from reannealing experiments on single-copy DNA (Hall *et al.*, 1980; Busslinger, 1981). We demonstrate that each of these histone gene families evolves by a random walk of their mutable bases, at a rate closely similar to that of unique genes and single-copy DNA, apparently undisturbed by any homogenization or "correction processes" which act on repeated genes within the species. The hypothetical "master sequences" must be subject to continuing mutational drift.

The rapid sequence change in genes and spacers allows analysis of the evolutionary relationship even between recently diverging histone gene clusters. Thus, within *Strongylocentrotus*, the divergence of Nor1 and the progenitor of Nor5 and Sp2/17 appears to predate speciation of *S. purpuratus* and *S. dröbachiensis* (see Figure 1c). The base substitution of 6% between the structural genes of *S. purpuratus* DNA clones Sp2/17 and *S. dröbachiensis* DNA clone Nor5 (Table I) suggests a time of divergence of ∿6 Myr and agrees well with previous paleontological and biochemical data (Durham, 1966; Hall *et al.*, 1980).

The degree of divergence between the histone spacers of sea urchin species is in good accord with the evolutionary clock of DNA sequences. For point mutations only, which clearly predominate over deletion/insertions within the spacer seg-

ment (see Figure 1c), a sequence divergence of 6.9% is obtained for *S. dröbachiensis* and *S. purpuratus* DNA spacer sequences (Nor5 vs. sp2). In evaluating the sequence divergence of spacer DNA, the significance of deletions which are ~50% as frequent as point mutations, is unclear. If a deletion mutation is equal to a single base change the value of divergence is increased to 11% (see Table II). Both these values agree well with 7.7% for single-copy sequence divergence shown by DNA reassociation experiments between these two species (Hall *et al.*, 1980).

As evolutionary distance between spacer pairs increases, they appear to be increasingly randomized by small deletions/insertions, while at the same time preserving their overall lengths (Kedes, 1979), so that any sequence alignment becomes arbitrary. The pattern of evolution in histone spacers differs distinctly from that of ribosomal spacers of *Xenopus laevis*. In the sea urchin, histone gene cluster spacers show little sequence repetition, and the sequence change appears to be more gradual, whereas sequence repetition and saltation of simple DNA sequences are the main characteristic of *X. laevis* ribosomal spacer DNA (Boseley *et al.*, 1979). Thus, two extreme possible modes of spacer evolution are revealed.

Classical explanations for the low rate of h19 DNA sequence evolution

DNA sequences are not random; for instance, dinucleotide frequencies are distinctly skewed in many prokaryotic and eukaryotic DNAs, the dinucleotide CpG, being, for example, dramatically under-represented (Russel *et al.*, 1976). Strong selection of sub-classes of the degenerate codons is evident in many genes, including those of histone (Elfstratiadis *et al.*, 1977; Schaffner *et al.*, 1978; Sures *et al.*, 1978). Also, codon context may be important and may regulate mRNA translation in both prokaryotes and eukaryotes (Bossi and Roth, 1980). In structural gene sequences there may be selection against motifs which might be erroneously interpreted as regulatory signals. There is, indeed, a bias towards the use of codons differing by more than one nucleotide from the termination triplets (Shaw *et al.*, 1977). Specific mRNA sequences *per se* may be selected for their secondary structure or for their ability to bind proteins involved in messenger transport or packaging.

Measurements of the rate of divergence for many structural genes show such constraints, which nevertheless do not prevent gene sequences from evolving at a rapid pace, although not as fast as single-copy DNA, most of which is thought not to consist of gene sequences (Crick, 1971).

If h19 has been part of the Echinidae genome since this family diverged from the Strongylocentrotidae (see Figure 2a), then virtually any base change must be lethal for this particular DNA. The above mechanisms cannot adequately explain such a high degree of selection acting over a combined evolutionary timespan of 130 Myr. It is even less obvious how similar constraints could apply to the non-transcribed spacer sequences (e.g., Figure 1c). Moreover, the extraordinary low rate of sequence evolution of the DNA clone h19 cannot be ascribed to the repetitive nature of histone genes *per se* because, as argued above, the major histone clones evolve at a speed consistent with the evolutionary DNA clock hypothesis and no additional constraints can be invoked on the basis of corrective mechanisms for repetitive genes.

The similarity between DNA clones h19, Nor5, and Sp2/17 cannot result from convergent evolution by a random process. Furthermore, the conservation-by-selection model requires a very strange temporal pattern of evolution in the three different histone DNA clusters. Figures 2b and c show that the evolution of Nor5 and Sp2/17 conforms to the accepted rate of DNA evolution. These two clones show a sequence divergence consistent with the combined evolutionary distance of ~10 Myr, during which these genes and spacers behave like any other structural genes. The evolution of the h19 DNA sequences does not fit this pattern; thus, the progenitors to both Nor5 and h19 DNA sequences would have had to evolve extremely slowly after separation of the Echinidae and Strongylocentrotidae. However, after speciation of *S. dröbachiensis* and *S. purpuratus*, leading to the appearance of Nor5 and Sp2/17 DNA, base substitutions in *S. purpuratus* gene clusters would have had to increase to about twice the standard rate of the clock of DNA sequences, while an exceptionally low rate of sequence change would have had to persist in Nor5 and h19.

A heterodox evolutionary pattern of h19 DNA sequences

Since h19 appears to be a genuine component of the *P. miliaris* genome and classical concepts fail to explain its presence in that genome, we must consider the possibility of a horizontal gene transfer of an h19-like gene cluster from *S. dröbachiensis* to *P. miliaris* within the last million years or so. The extent and significance of a horizontal gene transfer can be assessed by searching for unusual interspecific similarities between genes against the background of a diversity of other genes (Wilson *et al.*, 1977). Clones h19 and Nor5 clearly show such unusual similarity compared with the sequence divergence of other histone genes and with that of unique DNA.

Recently, cloning and DNA sequencing techniques have revealed a surprisingly high incidence of structural gene mobility within prokaryotic and eukaryotic species, due to illegitimate recombination, e.g., I.S. sequences and transposons in bacteria, and Ty and *copia* elements in yeast and in the fruit fly (reviewed by Calos and Miller, 1980). There is even tenuous evidence for transposable elements in *S. purpuratus* (Childs *et al.*, 1981). There are many suggestive findings that elements potentially capable of promoting the mobility of genes exist and, there is irrefutable evidence for the presence of jumping genes. Can such genes jump across the species barriers of eukaryotes, as has been thought likely in bacteria (Iida *et al.*, 1981)?

The retroviruses of animal cells may act as vectors for a horizontal gene transfer. These viruses cross species boundaries (Benveniste and Todaro, 1974, 1976) and can mobilize and extricate defined host genes (reviewed by Bishop, 1981). Viruses have been found in close association with sea urchin sperm (J. Wooley, personal communication). There is no information on retroviruses of the sea urchin, let alone their species specificity, but viral gene mobilization is an exciting possibility.

Gene transfer might also be by species hybridization. However, this seems unlikely since the unique DNA sequences of the haploid *S. dröbachiensis* chromosome set cannot now be detected within the *Psammechinus* genome (Busslinger, 1981); also, *S. dröbachiensis* and *P. miliaris* spawn at different seasons and sea urchin interspecific hybrids are not viable in later developmental stages.

The horizontal gene transfer may well be a genuine, albeit rare occurrence. It may not be limited to repeated genes although scrutiny of other repeated gene families might help

reveal similar unusual events. Irrespective of whether the classical theory of selection of DNA sequences or horizontal gene transfer ultimately explain the h19 DNA in *P. miliaris*, our work calls for new mechanisms and concepts. The theory of horizontal gene transfer is a salutary challenge to received views of a totally coherent evolution and orderly transmission of genes in eukaryotes.

Materials and methods

Isolation of recombinant histone DNA clones

Sperm DNA isolated from a single *P. miliaris* individual (from the Isle of Cumbrae, Scotland) was digested with the restriction enzyme HindIII. The resulting DNA fragments were fractionated according to size by agarose gel electrophoresis. DNA molecules of 6–7 kb length were eluted from agarose and then ligated into the HindIII site of the vector pBR322. This ligated DNA was used to transfect *Escherichia coli* HB101. Ampicillin-resistant colonies were screened for histone DNA inserts by hybridization with nick-translated h19 or h22 DNA. Novel histone DNA clones of the h19 and h22 type were ultimately identified by comparison of their restriction pattern with that of the original DNA clones h19 and h22, previously isolated by Clarkson *et al.* (1976). The DNA clones Nor1 and Nor5 were isolated from sperm DNA of a single *S. dröbachiensis* individual (from the Biological Station, Dröbak, Norway) as outlined above, with the difference that the size fractionation prior to the ligation step was omitted and that the vector pAT153 was used instead of pBR322.

DNA sequencing

All DNA sequences were determined according to the procedure of Maxam and Gilbert (1977). The H3 gene was localized on the DNA clone Ph70 (Spinelli *et al.*, 1979) by double digestion with the restriction enzymes PvuII and BstEII, both of which cut all sea urchin H3 DNAs at conserved positions, i.e., PvuII at codon 19/20 and BstEII at codon 117/118 (Figure 1a). The H3 DNA sequence of DNA clone Ph70 and that of DNA clone Nor5 was determined from each of these two sites in both directions. The spacer sequences of the DNA clones h19, Nor1, and Nor5 were determined from a single XbaI site (present in the h19 sequence at position 217–222, Figure 1c). The spacer DNA sequence of clone h19 was confirmed by sequencing from a TaqI site (270–273) and an AluI site (95–98) towards the H2B gene.

Evaluation of the sequence data

Incompletely sequenced codons were omitted from our comparative analysis of the structural histone gene sequences. The observed percentage sequence divergence, based only on the mutable nucleotides of the H3 and H4 genes, was calculated by categorizing all silent sites by toleration of one, two, or three different nucleotide substitutions. Assuming that transitions and transversions are equally probable we scored each of these silent sites as 1/3, 2/3, and 1, respectively. The total number of mutable bases was determined by summing over all silent sites using the above factors. The observed percentage divergence for mutable bases is defined as the ratio of the observed number of nucleotide substitutions to the total number of mutable bases in the two compared genes. Assuming a Poisson-distributed frequency of mutational events, the observed percentage divergence can be corrected for multiple substitutions at the same nucleotide site (Perler *et al.*, 1980; Salser, 1977). The corrected percentage divergence for mutable bases in the histone genes was calculated according to the method described in Perler *et al.* (1980).

Acknowledgements

We wish to thank Dr. R. Flavell for the opportunity to finish this work in his laboratory, Dr. R. Portmann for his advice throughout this work, Dr. C. Spadafora for supplying us with *P. lividus* DNA, F. Ochsenbein for graphical work, and S. Oberholzer for typing the manuscript. We are grateful to Dr. M. Chipchase for critical reading of the manuscript and for suggesting revisions. This work was supported by a grant of the State of Zürich and by the Swiss National Research Council, Grant No. 3.743.70.

References

Benveniste,R.E., and Todaro,G.J. (1974) *Nature,* **252**, 456-459.
Benveniste,R.E.; and Todaro,G.J. (1976) *Nature,* **261**, 101-108.
Birnstiel,M.L., Wallace,H., Sirlin,J.L., and Fischberg,M. (1966) in Vincent, W.S., and Miller,O.L. (eds.), *Proceedings of the International Symposium on the Nucleolus – Its Structure and Function,* National Cancer Inst. Monogr., **23**, 431-444.
Birnstiel,M.L. (1967) in de Renck,A.V.S., and Knight,J. (eds.), *Ciba Foundation Symposium on Cell Differentiation,* A.D. Churchill Ltd., London, pp. 178-191.
Birnstiel,M.L., and Grunstein,M. (1972) *FEBS Symp.,* **23**, 349-365.
Birnstiel,M.L., Weinberg,E.S., and Pardue,M.L. (1973) in Hamkalo, and Papaconstantinou (eds.), *Molecular Cytogenetics,* Plenum Press, NY, pp. 75-93.
Birnstiel,M.L., Portmann,R., Busslinger,M., Schaffner,W., Probst,E., and Kressmann,A. (1979) in Engberg, J., Klenow,H., and Leick,V. (eds.), *Specific Eukaryotic Genes,* Alfred Benzon Symp., **13**, Munksgaard, pp. 117-132.
Bishop,J.M. (1981) *Cell,* **23**, 5-6.
Boseley,P., Moss,T., Mächler,M., Portmann,R., and Birnstiel,M.L. (1979) *Cell,* **17**, 19-32.
Bossi,L., and Roth,J.R. (1980) *Nature,* **286**, 123-127.
Brown,D.D., Wensink,P.C., and Jordan,T. (1972) *J. Mol. Biol.,* **63**, 57-73.
Busslinger,M., Portmann,R., and Birnstiel,M.L. (1979) *Nucleic Acids Res.,* **6**, 2997-3008.
Busslinger,M., Portmann,R., Irminger,J.C., and Birnstiel,M.L. (1980) *Nucleic Acids Res.,* **8**, 957-977.
Busslinger,M. (1981) Thesis, University of Zürich.
Calos,M.P., and Miller,J.H. (1980) *Cell,* **20**, 579-595.
Childs,G., Maxon,R., Cohn,R.H., and Kedes,L.H. (1981) *Cell,* **23**, 651-663.
Clarkson,S.G., Smith,H.O., Schaffner,W., Gross,K.W., and Birnstiel,M.L. (1976) *Nucleic Acids Res.,* **3**, 2617-2632.
Crick,F. (1971) *Nature,* **234**, 25-27.
Durham,J.W. (1966) in Moore,R.C. (ed.), *Treatise on Invertebrate Paleontology, Part U, Echinodermata 3,* University of Kansas Press, Lawrence, Kansas, pp. 282, 374, 427, 431-435.
Efstratiadis,A., Kafatos,F.C., and Maniatis,T. (1977) *Cell,* **10**, 571-585.
Fitch,W.M. (1980) *J. Mol. Biol.,* **16**, 153-209.
Grunstein,M., Schedl,P., and Kedes,L. (1973) in Hamkalo, and Papaconstantinou (eds.), *Molecular Cytogenetics,* Plenum Press, NY, pp. 115-123.
Grunstein,M., Schedl,P., and Kedes,L. (1975) *J. Mol. Biol.,* **104**, 351-369.
Grunstein,M., and Schedl,P. (1976) *J. Mol. Biol.,* **104**, 323-349.
Grunstein,M., Diamond,K.E., Koppel,E., and Grunstein,J. (1981) *Biochemistry (Wash.),* **20**, 1216-1223.
Hall,T.J., Grula,J.W., Davidson,E.H., and Britten,R.J. (1980) *J. Mol. Evol.,* **16**, 95-110.
Hentschel,C.C., and Birnstiel,M.L. (1981) *Cell,* **25**, 301-313.
Holmquist,R., and Pearl,D. (1980) *J. Mol. Evol.,* **16**, 211-267.
Holmquist,R., Pearl,D., and Jukes,T.H. (1981) in Goodman, M. (ed.), *Macromolecular Sequences in Systematics and Evolutionary Biology,* Plenum Press, in press.
Iida,S., Meyer,J., and Arber,W. (1980) *Cold Spring Harbor Symp. Quant. Biol.,* **45**, 27-39.
Kedes,L.H., Chang,A.C.Y., Housman,D., and Cohen,S.N. (1975) *Nature,* **255**, 533-538.
Kedes,L.H. (1979) *Annu. Rev. Biochem.,* **48**, 837-870.
King,J.L., and Jukes,T. (1969) *Science (Wash.),* **164**, 788.
Miyata,T., Yasunaga,T., and Mishida,T. (1980) *Proc. Natl. Acad. Sci. USA,* **77**, 7328-7332.
Miyata,T., and Hayashida,H. (1981) *Proc. Natl. Acad. Sci. USA,* **78**, 5739-5743.
Maxam,A.M., and Gilbert,W. (1977) *Proc. Natl. Acad. USA,* **74**, 560-564.
Perler,F., Efstratiadis,A., Lomedico,P., Gilbert,W., Kolodner,R., and Dodgson,J. (1980) *Cell,* **20**, 555-566.
Rusconi,S. (1979) Diploma Thesis, University of Zürich.
Russell,G.J., Walker,P.M.B., Elton,R.A., and Subak-Sharpe,J.H. (1976) *J. Mol. Biol.,* **108**, 1-23.
Salser,W. (1977) *Cold Spring Harbor Symp. Quant. Biol.,* **42**, 985-1002.
Schaffner,W., Kunz,G., Daetwyler,H., Telford,J., Smith,H.O., and Birnstiel,M.L. (1978) *Cell,* **14**, 655-671.
Shaw,R.F., Bloom,R.W., and Bowman,J.E. (1977) *J. Mol. Evol.,* **9**, 225-230.
Smith,G.P. (1973) *Cold Spring Harbor Symp. Quant. Biol.,* **38**, 507-513.
Spinelli,G., Gianguzza,F., Casano,C., Acierno,P., and Burckhardt,J. (1979) *Nucleic Acids Res.,* **6**, 545-560.
Sures,I., Lowry,J., and Kedes,L.H. (1978) *Cell,* **15**, 1033-1044.
van Ooyen,A., van den Berg,J., Mantei,N., and Weissmann,C. (1979) *Science (Wash.),* **206**, 337-344.
Walker,P.M.B. (1968) *Nature,* **219**, 228-232.
Weinberg,E.S., Birnstiel,M.L., Purdom,I.F., and Williamson,R. (1972) *Nature,* **240**, 225-228.
Wilson,A.C., Carlson,S.S., and White,T. (1977) *Annu. Rev. Biochem.,* **46**, 573-639.