Histone and histone gene compilation and alignment update

Dan Wells and Doris Brown

Department of Biology and Institute for Molecular Biology, University of Houston, Houston, TX 77204-5513, USA

INTRODUCTION

This collection of histones and histone genes is an update of an earlier compilation (Wells and McBride, 1989). As was done in the previous compilation, we have attempted to gather and align all histone sequences available in the GenBank and EMBL nucleic acid sequence databases, and in the protein data bases. This update includes 164 histone sequences that were either missed in the previous compilation or have only recently appeared in the data bases. This brings the total number of histones gathered and aligned to 412. For each histone class, the sequences are arranged into phylogenetically related groups (*vertebrates at the beginning and fungi at the end*). Table 1 is a quick reference guide to the sequences shown in figures 1–9 and refers the reader to the full citations listed at the end of the compilation. Table 1 also lists the accession numbers (AC#) of each sequence.

PROTEIN SEQUENCES

Data from direct amino acid sequence analysis along with data from translated gene sequences are presented for each of the histone subtypes (Figures 1–5). Consensus sequences were generated for each histone subgroup based on cumulative sequence information from this and the previous compilation. Numbering is based on the consensus sequence. Figure 1 shows the alignments for the H1 proteins. In this figure the alignments of H1 and H5/H1.0 variants are shown separately to emphasize the similarity of the vertebrate H5/H1.0 group. The (●) symbols immediately above the H5 consensus indicate amino acids shared in the two variant groups. For the H1 histones, only the central conserved hydrophobic region could be aligned with any degree of certainty. A dash (–) in a sequence indicates an unsequenced region. Blank spaces in the protein alignments indicate deletions with respect to the consensus. An asterisk (*) in the consensus indicates the absence of a consensus amino acid for that position. A gap in the consensus sequence indicates an insertion was introduced into at least one of the aligned sequences to maximize sequence similarity. A dot (·) below a consensus amino acid indicates identity with that amino acid.

GENE SEQUENCES

Histone gene alignments are displayed for H2A, H2B, H3 and H4. No alignment was attempted for the histone H1 genes due to their extremely diverged nature. In an effort to draw some functional conclusions from these alignments, genes known or presumed to be nonfunctional have been excluded from this compilation. In each coding region alignment the consensus (cos)

sequence is displayed on the top line followed by each individual histone gene sequence (Figures 6–9). As for the proteins, consensus sequences were generated based on cumulative sequence information from this and the previous compilation (Wells and McBride, 1989). The numbering begins at the first base of the consensus and only bases within the consensus are numbered. A dot (·) below a consensus base indicates identity with the consensus. A dash (–) in a sequence indicates an unsequenced region. Blank spaces within a sequence indicate a deletion with respect to the consensus was introduced to maximize the alignment. A gap in the consensus indicates an insertion was introduced into at least one of the sequences to maximize the alignment.

AVAILABILITY

The complete database is available on floppy disk for the Macintosh. Please send a blank disk and your complete mailing address and phone number with your request. Plans are also in progress to make the data base available through electronic mail. Send inquiries to me at DanW@jetson.uh.edu.

ACKNOWLEDGEMENTS

We would like to thank April Cook for her editorial assistance. This effort was supported in part by Grant GM38623 to DW from the National Institute of Health.

REFERENCES

1. Alonso, A., Breuer, B., Bouterfa, H. and Doenecke, D. (1988) EMBO J. 7, 3003–3008.
2. Artishevsky, A., Wooden, S., Sharma, A., Resendez, E., Jr. and Lee, A.S. (1987) Nature 328, 823–827.
3. Banfield, D.C.D., Honda, B.M. and Smith, M.J. (1988) J. Mol. Evol. 27, 36–44.
4. Birnstiel, M., Portmann, R., Busslinger, M., Schaffner, W., Probst, E. and Kressman, A. (1979) Alfred Benzon Symp. 13, 117–132.
5. Brandt, W.F., Andrade Rodrigues, J. and Von Holt, C. (1988) Eur. J. Biochem. 173, 547–554.
6. Brunk, C.F., Kahn, R.W. and Sadler, L.A. (1990) J. Mol. Evol. 30, 290–297.
7. Chalmers, M. and Wells, D. (1990) Nucl. Acids Res. 18, 3075.
8. Chaubet, N., Chaboute, M.-E., Philipps, G. and Gigot, C. (1987) Dev. Genet. 8, 461–473.
9. Cheng, G., Nandi, A., Clerk, S. and Skoultchi, A.I. (1989) Proc. Natl. Acad. Sci. U.S.A. 86, 7002–7006.
10. Cole, K.D., Kandala, J.C., Kremer, E. and Kistler, W.S. (1990) Gene 89, 265–269.

11. Domier, L.L., Rivard, J.J., Sabatini, L.M. and Blumenfeld, M. (1986) *J. Mol. Evol.* 23, 149–158.
12. Fecker, L., Ekblom, P., Kurkinen, M. and Ekblom, M. (1990) *Mol. Cell. Biol.* 10, 2848–2854.
13. Gantt, S. and Key, J.L. (1987) *Eur. J. Biochem.* 166, 119–125.
14. Gargiulo, G., Razvi, F., Ruberti, I., Mohr, I. and Worcel, A. (1985) *J. Mol. Biol.* 181, 333–349.
15. Gigot, C., Chaubet, N., Chaboute, M.-E., Ehling, M. and Philipps, G. (1987) *Plant Physiol. Biochem.* 25, 235–247.
16. Grimes, S., Weisz-Carrington, P., Daum, H., III., Smith, J., Green, L., Wright, K., Stein, G. and Stein, J. (1987) *Exp. Cell Res.* 173, 534–545.
17. Gruber, A., Streit, A., Reist, M., Benninger, P., Boehni, R. and Schuemperli, D. (1990) *Gene* 95, 303–304.
18. Harper, D.S. and Jahn, C.L. (1989) *Gene* 75, 93–107.
19. Hill, C.S. and Thomas, J.O. (1990) *Eur. J. Biochem.* 187, 145–153.
20. Howell, A.M., Cool, D., Hewitt, J., Ydenberg, B., Smith, M.J. and Honda, B.M. (1987) *J. Mol. Evol.* 25, 29–36.
21. Hraba Renevey, S. and Kress, M. (1989) *Nucl. Acids Res.* 17, 2449–2461.
22. Hurt, M.M., Chodchoy, N. and Marzluff, W.F. (1989) *Nucl. Acids Res.* 17, 8876.
23. Konrad, M. (1988) *J. Biol. Chem.* 263, 9468–9474.
24. Kosciessa, U. and Doenecke, D. (1989) *Nucl. Acids Res.* 17, 8861.
25. Kremer, H. and Hennig, W. (1990) *Nucl. Acids Res.* 18, 1573–1580.
26. Lieber, T., Angerer, L.M., Angerer, R.C. and Childs, G. (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85, 4123–4127.
27. Liu, T.J., Liu, L. and Marzluff, W.F. (1987) *Nucl. Acids Res.* 15, 3023–3039.
28. Mannironi, C., Bonner, W.M. and Hatch, C.L. (1989) *Nucl. Acids Res.* 17, 9113–9126.
29. Matsuo, Y. and Yamazaki, T. (1989) *Nucl. Acids Res.* 17, 225–238.
30. Maxson, R., Mohun, T., Gormezano, G. and Kedes, L. (1987) *Nucl. Acids Res.* 15, 10569–10582.
31. Meier, V.S., Boehni, R. and Schuemperli, D. (1989) *Nucl. Acids Res.* 17, 795.
32. Moss, S.B., Challoner, P.B. and Groudine, M. (1989) *Dev. Biol.* 133, 83–92.
33. Mueller, K., Lindauer, A., Bruederlein, M. and Schmitt, R. (1990) (unpublished from GenBank).
34. Palla, F., Casano, C., Albanese, I., Anello, L., Gianguzza, F., Dibernardo, M.G., Benura, C. and Spinelli, G. (1989) *Proc. Natl. Acad. Sci. U.S.A.* 86, 6033.
35. Porter, D., Brown, D. and Wells, D. (1990) DNA Sequence (in press).
36. Porter, D.C., Moy, G.W. and Vacquier, V.D. (1989) *Comp. Biochem. Physiol.* 92B, 381–384.
37. Rutledge, R.G., Neelin, J.M. and Seligy, V.L. (1988) *Gene* 70, 117–126.
38. Sellos, D., Krawetz, S.A. and Dixon, G.H. (1990) *Eur. J. Biochem.* 190, 21–29.
39. Smith, R.C., Dworkin-Rastl, E. and Dworkin, M.B. (1988) *Genes Dev.* 2, 1284–1295.
40. Spiker, S., Weisshaar, B., da Costa e Silva, O. and Halbrock, K. (1990) *Nucl. Acids Res.* 18, 5897.
41. Stark, M.J.R. and Milner, J.S. (1989) *Yeast* 5, 35–50.
42. Sturm, R.A., Dalton, S. and Wells, J.R.E. (1988) *Nucl. Acids Res.* 16, 8571–8586.
43. Taylor, J.D., Wellman, S.E. and Marzluff, W.F. (1986) *J. Mol. Evol.* 23, 242–249.
44. Thomas, G. and Padayatty, J.D. (1984) *Indian J. Biochem. Biophys.* 21, 1–6.
45. Tonjes, R. and Doenecke, D. (1985) *Gene* 39, 275–279.
46. Tonjes, R., Munk, K. and Doenecke, D. (1989) *J. Mol. Evol.* 28, 200–211.
47. Van Daal, A. and Elgin, S. (1989) (unpublished from GenBank).
48. Van Daal, A., Elgin, S. and Gorovsky, M.A. (1990) *J. Mol. Evol.* 30, 449–455.
49. Vanfleteren, J.R., Van Bun S.M. and Van Beeumen, J.J. (1987) *Comp. Biochem. Physiol.* 87B, 847–849.
50. Vanfleteren, J.R., Van Bun S.M. and Van Beeumen, J.J. (1988) *Biochem. J.* 255, 647–652.
51. Vanfleteren, J.R., Van Bun S.M. and Van Beeumen, J.J. (1989) *FEBS Lett.* 257, 233–237.
52. Vanfleteren, J.R., Van Bun S.M., De Baere, I. and Van Beeumen, J.J. (1990) *Biochem. J.* 265, 739–746.
53. Wefes, I. and Lipps, H.J. (1990) DNA Sequence 1, 25–32.
54. Wells, D. (1986) *Nucl. Acids Res.* 14 Supplement, r119–r149.
55. Wells, D. and McBride, C. (1989) *Nucl. Acids Res.* 17 Sequences Supplement, r311–r346.
56. Wilhelm, M.L. and Wilhelm, F.X. (1989) *J. Mol. Evol.* 28, 322–326.
57. Wolfe, S.A., Anderson, J.V., Grimes, S.R., Stein, G.S. and Stein, J.S. (1989) *Biochim. Biophys. Acta.* 1007, 140–150.
58. Wolfe, S.A., Anderson, J.V., Grimes, S.R., Stein, G.S. and Stein, J.S. (1989) (unpublished from GenBank).
59. Wu, S.C., Boegre, L., Vincze, E., Kiss, G.B. and Dudits, D. (1988) *Plant Mol. Biol.* 11, 641–649.
60. Wu, S.C., Gyoergye, J. and Dudits, D. (1989) *Nucl. Acids Res.* 17, 3057–3063.
61. Wu, S.C., Vegh, Z., Wang, X.-M., Tan, C.C. and Dudits, D. (1989) *Nucl. Acids Res.* 17, 3297.
62. Xu, H., Johnson, L. and Grunstein, M. (1990) *Mol. Cell. Biol.* 10, 2687–2694.
63. Yaguchi, M., Roy C., Dove M. and Seligy, V. (1977) *Biochem. Biophys. Res. Commun.* 76, 100–106.
64. Yang, Y.-S., Brown, D.T., Wellman, S.E. and Sittman, D.B. (1987) *J. Biol. Chem.* 262, 17118–17125.

Table 1. Sequence List

HISTONE H1

1. Mouse H1, AC # J03482 (Yang *et al.*, 1987)
2. Mouse H1-var.1, AC # M25365 (Cheng *et al.*, 1989)
3. Rat H1t, AC # M28409 (Wolfe *et al.*, 1989)
4. Rat H1d, AC # M31229 (Cole *et al.*, 1990)
5. *Xenopus* H1-B4, AC # M36655 (Smith *et al.*, 1988)
6. Sea Urchin H1d-Sp, AC # J03807 (Lieber *et al.*, 1988)
7. Sea Urchin H1-Sp, AC # P19376 (Porter *et al.*, 1989)
8. Sea Urchin H1-Pm, AC # V01144 (Birnstiel *et al.*, 1979)
9. *Drosophila* H1-Dh, AC # X17072 (Kremer & Hennig, 1990)
10. *Drosophila* H1-Dm, AC # X14215 (Matsuo & Yamazaki, 1989)
11. *Caenorhabditis* H1.1, AC # P10771 (Vanfleteren *et al.*, 1988)
12. *Caenorhabditis* H1.2, AC # P15796 (Vanfleteren *et al.*, 1990)
13. Pea H1, AC # X05636 (Gantt & Key, 1987)

HISTONE H5/H1.0

14. Mouse H1-0, AC # X13171 (Alonso *et al.*, 1988)
15. Pigeon H5, AC # P02260 (Yaguchi *et al.*, 1977)
16. *Xenopus* H5-A, AC # M22834 (Rutledge *et al.*, 1988)
17. *Xenopus* H5-B, AC # M22835 (Rutledge *et al.*, 1988)

HISTONE H2A

18. Human H2A.X, AC # X14850 (Mannironi *et al.*, 1989)
19. Mouse H2A.1-53, AC # M33988 (Gruber *et al.*, 1990)
20. Mouse H2A.1, AC # X37736 (Fecker *et al.*, 1990)
21. Mouse H2A-614, AC # M16148 (Hurt *et al.*, 1989)
22. Mouse H2A, AC # X16495 (Kosciessa & Doenecke, 1989)
23. Mouse H2A-291A, AC # X05862, Y00117 (Liu *et al.*, 1987)
24. Chicken H2A-CH07, AC # X07763, X07758 (Sturm *et al.*, 1988)
25. Chicken H2A-pCH22.0B, AC # X07762, X07757 (Sturm *et al.*, 1988)
26. Chicken H2A-pCH11.0EL, AC # X07764, X07759 (Sturm *et al.*, 1988)
27. Chicken H2A-pCH11.0ER, AC # X07765, X07760 (Sturm *et al.*, 1988)
28. Chicken H2A-CH1-10, AC # X07766, X07761 (Sturm *et al.*, 1988)
29. Duck H2A, AC # X14730 (Tonjes *et al.*, 1989)
30. Sea Urchin H2A-L3-Sp, AC # X06642 (Maxson *et al.*, 1987)
31. Sea Urchin H2A-Pl, AC # M25281 (Palla *et al.*, 1989)
32. Sea Urchin H2A-Pm, AC # V01144 (Birnstiel *et al.*, 1979)
33. *Drosophila* H2A-Dh, AC # X17072 (Kremer & Hennig, 1990)
34. *Drosophila* H2A-Dm-1, AC # X14215 (Matsuo & Yamazaki, 1989)
35. *Drosophila* H2A-Dm-2, AC # X14215 (Matsuo & Yamazaki, 1989)
36. *Drosophila* H2A-Dv, AC # M32467 (Domier *et al.*, 1986)
37. *Drosophila* H2AvD-Dm, AC # X15549 (Van Daal & Elgin, 1989)
38. *Platynereis* H2A, AC # X53330 (Sellos *et al.*, 1990)
39. Parsley H2A, AC # X53831 (Spiker *et al.*, 1990)
40. Rice H2A, AC # M35379 (Thomas & Padayatty, 1984)
41. *Volvox* H2A-III, AC # M31921 (Mueller *et al.*, 1990)
42. *Volvox* H2A-IV, AC # M31922 (Mueller *et al.*, 1990)
43. *Tetrahymena* H2Av, AC # X15548 (Van Daal *et al.*, 1990)
44. Yeast H2A-1, AC # M18455 (Konrad, 1988)

HISTONE H2B

45. Mouse H2B, AC # M25487 (Moss *et al.*, 1989)
46. Mouse H2B-291A, AC # X05862, Y00117 (Liu *et al.*, 1987)
47. Mouse H2B-291B, AC # X05863, Y00117 (Liu *et al.*, 1987)
48. Chicken H2B-CH07, AC # X07763, X07758 (Sturm *et al.*, 1988)
49. Chicken H2B-pCH22.0B, AC # X07762, X07757 (Sturm *et al.*, 1988)
50. Chicken H2B-pCH11.0EL, AC # X07764, X07759 (Sturm *et al.*, 1988)
51. Chicken H2B-pCH11.0ER, AC # X07765, X07760 (Sturm *et al.*, 1988)
52. Chicken H2B-CH1-10, AC # X07766, X07761 (Sturm *et al.*, 1988)
53. Duck H2B, AC # X14731 (Tonjes *et al.*, 1989)
54. Sea Urchin H2B-L1-Sp, AC # X06640 (Maxson *et al.*, 1987)
55. Sea Urchin H2B-L3-Sp, AC # X06642 (Maxson *et al.*, 1987)
56. Sea Urchin H2B-L4-Sp, AC # X06643 (Maxson *et al.*, 1987)
57. Sea Urchin H2B-Pm, AC # V01144 (Birnstiel *et al.*, 1979)
58. Sea Urchin H2b.1-Ee, AC # P13281 (Hill & Thomas, 1990)
59. Sea Urchin H2b.2-Ee, AC # P13282 (Hill & Thomas, 1990)
60. Sea Urchin H2b.3-Ee, AC # P13283 (Hill & Thomas, 1990)
61. Sea Star H2B, AC # X05619, X05620 (Howell *et al.*, 1987)
62. *Drosophila* H2B-Dh, AC # X17072 (Kremer & Hennig, 1990)

63. *Drosophila* H2B-Dm, AC # X14215 (Matsuo & Yamazaki, 1989)
64. *Drosophila* H2B-Dv, AC # M32468 (Domier *et al.*, 1986)
65. *Platynereis* H2B, AC # X53330 (Sellos *et al.*, 1990)
66. *Caenorhabditis* H2B.2, AC # P04255 (Vanfleteren *et al.*, 1989)
67. Wheat H2B, AC # P05621 (Brandt *et al.*, 1988)
68. *Volvox* H2B-III, AC # M31921 (Mueller *et al.*, 1990)
69. *Volvox* H2B-IV, AC # M31922 (Mueller *et al.*, 1990)
70. Yeast H2B-1, AC # M37743 (Xu *et al.*, 1990)

HISTONE H3

71. Mouse H3, AC # X16148 (Hurt *et al.*, 1989)
72. Mouse H3.1-221, AC # M32460 (Taylor *et al.*, 1986)
73. Mouse H3.2-221, AC # M32459 (Taylor *et al.*, 1986)
74. Mouse H3.1, AC # X16496 (Kosciessa & Doenecke, 1989)
75. Mouse H3.2, AC # M33989 (Gruber *et al.*, 1990)
76. Mouse H3.3, AC # X13605 (Hraba Renevey & Kress, 1989)
77. Rabbit H3.3, AC # X51897 (Chalmers & Wells, 1990)
78. Hamster H3.2, AC # M28265 (Artishevsky *et al.*, 1987)
79. Duck H3-1, AC # X14732 (Tonjes *et al.*, 1989)
80. Duck H3-2, AC # X14732 (Tonjes *et al.*, 1989)
81. Duck H3.4, AC # M14396 (Tonjes & Doenecke, 1985)
82. Sea Urchin H3-Pl, AC # M25281 (Palla *et al.*, 1989)
83. Sea Urchin H3-Pm, AC # V01144 (Birnstiel *et al.*, 1979)
84. Sea Star H3-Pb, AC # X07504 (Banfield *et al.*, 1988)
85. Sea Star H3-Di, AC # X07505 (Banfield *et al.*, 1988)
86. *Drosophila* H3-Dh, AC # X17072 (Kremer & Hennig, 1990)
87. *Drosophila* H3-Dm, AC # X14215 (Matsuo & Yamazaki, 1989)
88. *Drosophila* H3-Dv, AC # M33982 (Domier *et al.*, 1986)
89. Copepod H3-Tc, AC # X52393 (Porter *et al.*, 1990)
90. *Platynereis* H3, AC # X53330 (Sellos *et al.*, 1990)
91. Alfalfa H3-1.1, AC # M35867 (Wu *et al.*, 1988)
92. Alfalfa H3-pH3c-1, AC # X13674 (Wu *et al.*, 1989)
93. Alfalfa H3-pH3c-3, AC # X13675 (Wu *et al.*, 1989)
94. Alfalfa H3-pH3c-11, AC # X13676 (Wu *et al.*, 1989)
95. Maize H3, AC # M35388 (Chaubet *et al.*, 1987)
96. Maize H3, AC # M36658 (Gigot *et al.*, 1987)
97. Rice H3-H3R11, AC # X13678 (Wu *et al.*, 1989)
98. Rice H3-H3R-21, AC # X13680 (Wu *et al.*, 1989)
99. *Tetrahymena* H3II-Ta, AC # X17126 (Brunk *et al.*, 1990)
100. *Tetrahymena* H3II-Tau, AC # X17127 (Brunk *et al.*, 1990)
101. *Tetrahymena* H3II-Tb, AC # X17128 (Brunk *et al.*, 1990)
102. *Tetrahymena* H3II-Tc, AC # X17129 (Brunk *et al.*, 1990)
103. *Tetrahymena* H3II-Tca, AC # X17130 (Brunk *et al.*, 1990)
104. *Tetrahymena* H3II-Tcu, AC # X17131 (Brunk *et al.*, 1990)
105. *Tetrahymena* H3II-Te, AC # X17125 (Brunk *et al.*, 1990)
106. *Tetrahymena* H3II-Tf, AC # X17132 (Brunk *et al.*, 1990)
107. *Tetrahymena* H3II-Th, AC # X17133 (Brunk *et al.*, 1990)
108. *Tetrahymena* H3II-Tl, AC # X17134 (Brunk *et al.*, 1990)
109. *Tetrahymena* H3II-Tm, AC # X17135 (Brunk *et al.*, 1990)
110. *Tetrahymena* H3II-Tmi, AC # X17136 (Brunk *et al.*, 1990)
111. *Tetrahymena* H3II-Tn, AC # X17137 (Brunk *et al.*, 1990)
112. *Tetrahymena* H3II-Tni, AC # X17138 (Brunk *et al.*, 1990)
113. *Tetrahymena* H3II-Tp, AC # X17139 (Brunk *et al.*, 1990)
114. *Tetrahymena* H3II-Tpa, AC # X17140 (Brunk *et al.*, 1990)
115. *Tetrahymena* H3II-Tpi, AC # X17143 (Brunk *et al.*, 1990)
116. *Tetrahymena* H3II-Tpy, AC # X17141 (Brunk *et al.*, 1990)
117. *Tetrahymena* H3II-Tr, AC # X17144 (Brunk *et al.*, 1990)
118. *Tetrahymena* H3II-Ts, AC # X17142 (Brunk *et al.*, 1990)
119. *Tetrahymena* H3II-Tt, AC # X17145 (Brunk *et al.*, 1990)
120. *Cluyveromyces* H3, AC # X14230 (Stark & Milner, 1989)

HISTONE H4

121. Mouse H4-c12, AC # X13235 (Meier *et al.*, 1989)
122. Mouse H4, AC # X13236 (Meier *et al.*, 1989)
123. Rat H4t, AC # M28409 (Wolfe *et al.*, 1989)
124. Rat H4-pPS7, AC # M24020 (Grimes *et al.*, 1987)
125. Rat H4, AC # X13554 (Wolfe *et al.*, 1989)
126. Duck H4-1, AC # X14732 (Tonjes *et al.*, 1989)
127. Duck H4-2, AC # X14732 (Tonjes *et al.*, 1989)
128. *Xenopus* H4-1-pX1H4-1, AC # M23776, M23777 (Gargiulo *et al.*, 1985)
129. Sea Urchin H4-L1-Sp, AC # X06639 (Maxson *et al.*, 1987)
130. Sea Urchin H4-L2-Sp, AC # X06641 (Maxson *et al.*, 1987)

131. Sea Urchin H4-Pm, AC # V01144 (Birnstiel *et al.*, 1979)
132. *Drosophila* H4-Dh, AC # X17072 (Kremer & Hennig, 1990)
133. *Drosophila* H4-Dm, AC # X14215 (Matsuo & Yamazaki, 1989)
134. *Drosophila* H4-Dv, AC # M32469 (Domier *et al.*, 1986)
135. Copepod H4-Tc, AC # X52393 (Porter *et al.*, 1990)
136. *Platynereis* H4, AC # X53330 (Sellos *et al.*, 1990)
137. *Caenorhabditis* H4, AC # P18678 (Vanfleteren *et al.*, 1987)
138. Maize H4-C13, AC # M36659 (Gigot *et al.*, 1987)
139. *Tetrahymena* H4II-Ta, AC # X17126 (Brunk *et al.*, 1990)
140. *Tetrahymena* H4II-Tau, AC # X17127 (Brunk *et al.*, 1990)
141. *Tetrahymena* H4II-Tb, AC # X17128 (Brunk *et al.*, 1990)
142. *Tetrahymena* H4II-Tc, AC # X17129 (Brunk *et al.*, 1990)
143. *Tetrahymena* H4II-Tca, AC # X17130 (Brunk *et al.*, 1990)
144. *Tetrahymena* H4II-Tcu, AC # X17131 (Brunk *et al.*, 1990)
145. *Tetrahymena* H4II-Te, AC # X17125 (Brunk *et al.*, 1990)
146. *Tetrahymena* H4II-Tf, AC # X17132 (Brunk *et al.*, 1990)
147. *Tetrahymena* H4II-Th, AC # X17133 (Brunk *et al.*, 1990)
148. *Tetrahymena* H4II-Tl, AC # X17134 (Brunk *et al.*, 1990)
149. *Tetrahymena* H4II-Tm, AC # X17135 (Brunk *et al.*, 1990)
150. *Tetrahymena* H4II-Tmi, AC # X17136 (Brunk *et al.*, 1990)
151. *Tetrahymena* H4II-Tn, AC # X17137 (Brunk *et al.*, 1990)
152. *Tetrahymena* H4II-Tni, AC # X17138 (Brunk *et al.*, 1990)
153. *Tetrahymena* H4II-Tp, AC # X17139 (Brunk *et al.*, 1990)
154. *Tetrahymena* H4II-Tpa, AC # X17140 (Brunk *et al.*, 1990)
155. *Tetrahymena* H4II-Tpi, AC # X17143 (Brunk *et al.*, 1990)
156. *Tetrahymena* H4II-Tpy, AC # X17141 (Brunk *et al.*, 1990)
157. *Tetrahymena* H4II-Tr, AC # X17144 (Brunk *et al.*, 1990)
158. *Tetrahymena* H4II-Ts, AC # X17142 (Brunk *et al.*, 1990)
159. *Tetrahymena* H4II-Tt, AC # X17145 (Brunk *et al.*, 1990)
160. *Stylonychia* H4-OM, AC # M24411 (Harper & Jahn, 1989)
161. *Stylonychia* H4-CH4g, AC # X16019 (Wefes & Lipps, 1990)
162. *Stylonychia* H4-CH4k, AC # X16018 (Wefes & Lipps, 1990)
163. *Physarum* H4-1, AC # X15142 (Wilhelm & Wilhelm, 1989)
164. *Physarum* H4-2, AC # X15141 (Wilhelm & Wilhelm, 1989)

Figure 1. Histone H1 Proteins Alignments

	20	40	60	80	100	120
COS	20	40	60	80	100	120
18	SRSGAGLOPVRVRLIRKGN	YARRGACAPYLAUVVET	TALEILAGNAARDKKTR	LIPRHIOJAIRNDEEINK	LGVTIPGGVTIPHOAVLP	PKRTEKHHKAGK
19	T.....	H.....	S.....	R.....	S.....	SATVGP
20	C.....	S.....	M.....	R.....	S.....	APSGGKKATOASOEY
21	T.....	S.....	M.....	R.....	S.....	S.....
22	T.....	S.....	M.....	R.....	S.....	S.....
23	T.....	S.....	M.....	R.....	S.....	S.....
24	T.....	S.....	M.....	R.....	S.....	S.....
25	T.....	S.....	M.....	R.....	S.....	S.....
26	T.....	S.....	M.....	R.....	S.....	S.....
27	T.....	S.....	M.....	R.....	S.....	S.....
28	V.....	M.....	M.....	HEF.....	K.....	DS.....S.....
29	A.....	F.....	Q.....	V.....	S.....	S.....AS.....
30	GA.....	P.....	K.....G.....M.....	S.....	V.....	G.....SS.....
31	S.....T.....	F.....	K.....G.....M.....	S.....	V.....	S.....KA.....
32	VKG.....	I.....	M.....A.....V.....	S.....	S.....	S.....KA.....
33	VKG.....	I.....	M.....A.....V.....	S.....	S.....	S.....KA.....
34	VKG.....	I.....	M.....A.....V.....	S.....	S.....	S.....KA.....
35	VKG.....	I.....	M.....A.....V.....	S.....	S.....	S.....KA.....
36	VKG.....	I.....	M.....A.....V.....	S.....	S.....	S.....KA.....
37	A.GKAGKDSGR...	AV.....A.....	I.....H.....K.....S.....T.....	G.....DS.....	I.....H.....K.....S.....	I.....G.....E.....T.....VODPQR
38	K.....S.....	I.....H.....K.....S.....T.....	G.....DS.....	I.....H.....K.....S.....	I.....G.....E.....T.....VODPQR	GNVILSOAY
39	ETA.KA.K.FGG.NGPRKKSVT..VK...	IG.Y.K.R.Q.T.....	A.....V.....	L.....V.....	G.....A.....F.....H.....	NP.....AERAAKEP
40	HTK.L..F-RKISIPAGI.....	I.....Y.....K.....K.....	I.....	N.....V.....	I.....	G.....D.....S.....H.....SKGKGEEAA
41	A....KT.GK.	AV.....K.....	I.....	N.....V.....	I.....	G.....D.....S.....H.....SKGKGEEAA
42	A....KT.GK.	AV.....K.....	I.....	N.....V.....	I.....	G.....D.....S.....H.....SKGKGEEAA
43	AGGKGGK.K.GK..VGAKNKKTPO..YK	I.....F.....K.....G.....V.....	T.....A.....	G.....DI.....	I.....H.....K.....S.....	G.....HST.NRSSAKTAEPR
44	-----ASO....AK....T.....	R.....	Q.....I.....S.....	T.....	D.....	N.....HON.....SA.AT..SOEL

Figure 2. Histone H2A Protein Alignments

	20	40	60	80	100	120
COS	20	40	60	80	100	120
45	PEPAKSAPAKSGKKAATKOTKGGKRRSRKESYSIIYKVK	KOVHEDTGTSSKAGINSEVNDIEPFIAGEASRLAHYKRS	STTSREIOTAVRLMLPGLAKHAVSEGTKAVTKYTSK			
46	---VR.V.....	A.....D.....	V.....V.....	S.....		ILWNKEYIPLSF
47	A.....D.....	V.....V.....	S.....			
48	T.....	A.....D.....	V.....V.....	S.....		
49	D.....	K.....	R.....	D.....	S.....	
50	D.....	R.....	T.....	V.....		
51	D.....	R.....	T.....	V.....		
52	D.....	K.....	R.....	D.....	S.....	
53	D.....	K.....	R.....	D.....	S.....	
54	D.....	K.....	R.....	D.....	S.....	
55	AQ.G.....KAPRPSG.....	R.R.....G.....	R.S.....V.....	A.....K.....	V.....V.....	S.....
56	AOA.G.....KAPRPSG.....	R.R.....G.....	R.S.....V.....	A.....K.....	V.....V.....	S.....
57	-----HASRATD.....PKR.....	G.....	R.....S.....	V.....	Q.....K.....S.....	V.....
58	PT.....					
59	PSQSPTRKSPTRK.POKGK.A.R.....					
60	PKSPKSPRKGSPRKA.PKRG.....					
61	PRSPKSPKKGSPRKA.PKRGG.A.R.CKGGRRTVV.....					
62	PPKASGKGK.....					
63	PKTSKGAAKKAGKQAQ.NIT.ND..K.K.....A.....		S.....	A.....		T.....
64	PKTSKGAAKKAGKQAQ.NIT.TD.K.K.....A.....		S.....	A.....		
65	A.....S.....L.....		S.....			
66	PKVASKGAKKAASKA.AARS.E.K.R.R.....	S.....	S.....	A.....		
67	PPK.SA.GAK.AAK.V.KP.D.....KA.....	V.....R.....	S.....	V.....	S.....	I.....
68	PKPKPAENKVEKAAETPAGKKPKA.KRLP.GRTASKEAGEG..RGRKGSRAKGV.T.K..F.....	I.....	S.....	I.....	KL.....SAK.R.KP.....	S.....V.....
69	APKKEEPASQAAPAEVKAAPKAPKAPKKK.KKAP.KKVA.EPAGGEDG..SK..A.VAKS.T.KL.....	S.....	I.....	KV.T.....K.SR.....	KP.V.....	F.....
70	APKKEEPASQAAPAEVKAAPKAPKAPKKK.KKAP.KKSA.EPAGDAEGD.KK..A.VAKS.T.KL.....	S.....	I.....	KV.T.....K.SR.....	KP.V.....	F.....
						R.....S.....TOA

Figure 3. Histone H2B Protein Alignments

	20	40	60	80	100	120
71	ARTKOTARKSTGGKAPRKOLATKAARKSAPATGGVKKPHRYRPGTVAIREIRYOKSTELLIRKIPFORLVR	IAODEKTD	IRFOSSAVMAIOEASEAYLVGLFEDTNLC	AIHAKRVTIMPKD	LOLAR	RIGERA
72						
73						
74						
75						
76		S.....	A.IG.....			
77		S.....	A.IG.....			
78						
79						
80						
81		C.....	H.H.....	C.....		S.....
82				E.....		T.....
83				E.....		
84				E.....		
85				E.....		
86						
87						I.....
88						
89						
90						
91		F.....	K.....	V.S.....	A.....	
92		F.....	K.....	F.....	S.....	A.....
93		F.....	K.....		S.....	A.....
94			K.....	H.L.....	A.....	
95		F.....	K.....		A.....	A.....
96		F.....	K.....		A.....	A.....
97		F.....	K.....		A.....	A.....
98		F.....	K.....		A.....	A.....
99	A.....	S.....	I.....			
100	A.....	S.....	I.....			
101	A.....	S.....	I.....			
102	A.....	S.....	I.....			
103	A.....	S.....	I.....			
104	A.....	S.....	I.....			
105	A.....	S.....	I.....			
106	A.....	S.....	I.....			
107	A.....	S.....	I.....			
108	A.....	S.....	I.....			
109	A.....	S.....	I.....			
110	A.....	S.....	I.....			
111	A.....	S.....	I.....			
112	A.....	S.....	I.....			
113	A.....	S.....	I.....			
114	A.....	S.....	I.....			
115	A.....	S.....	I.....			
116	A.....	S.....	I.....			
117	A.....	S.....	I.....			
118	A.....	S.....	I.....			
119	A.....	S.....	I.....			
120		S.....	K.....	F.....	IG.....SV.....S.....	A.....OK.....K.....L.....S.....

Figure 4. Histone H3 Protein Alignments

	20	40	60	80	100
QOS	SRGKGKGLKGGAKRH	RKVLRDNTGITKPAIRRLARGGVKRI	SGLIYEETRGVVKFLENVIRDAVYVTEHAKRKT	VTAMDVVYALKROGRTLYGEGG	
121					
122					
123					
124					
125					
126					
127					
128					
129					
130					
131					
132	T.....				
133	T.....				
134	T.....				
135	T.....				
136					
137					
138					
139	VG.....	M.V.....	SKRSNKAS.E.....		
140	AG.....	M.V.....	SKRSNKAS.E.....		
141	AG.....	M.V.....	SKRSNKAS.E.....		
142	AG.....	M.V.....	SKRSNKAS.E.....		
143	VG.....	M.V.....	SKRSNKAS.E.....		
144	AG.....	M.V.....	SKRSNKAS.E.....		
145	AG.....	M.V.....	SKRSNKAS.E.....		
146	AG.....	M.V.....	SKRSNKAS.E.....		
147	VG.....	M.V.....	SKRSNKAS.E.....		
148	AG.....	M.V.....	SKRSNKAS.E.....		
149	AG.....	M.V.....	SKRSNKAS.E.V.....		
150	AG.....	M.V.....	SKRSNKAS.E.....		
151	VG.....	M.V.....	SKRSNKAS.E.....		
152	VG.....	M.V.....	SKRSNKAS.E.....		
153	PS.....	M.V.....	A..SNKAS.E.....		
154	VG.....	M.V.....	SKRSNKAS.E.....		
155	AG.....	M.V.....	SKRSNKAS.E.....		
156	VG.....	M.V.....	SKRSNKAS.E.....		
157	AG.....	M.V.....	SKRSNKAS.E.....		
158	VG.....	M.V.....	SKRSNKAS.E.....		
159	AG.....	M.V.....	SKRSNKAS.E.....		
160	A...V...Y.V.....	TK.S.KET.M.....			
161	A...V...Y.V.....	TK.S.KET.M.....			
162	A...V...Y.V.....	TK.S.KET.M.....			
163					
164					

Figure 5. Histone H4 Protein Alignments

Fig 6. (cont.)

```

220      240      260      280      300
COS  CACGCTGCGCGCACACAGAGACACCGGCATCATCCCGCCGACCTGCGAGTGGCCATCGCCACGAGAGAGCTGAAACAAGCTGCTGGGGGGGTGACCATCGCTCAGGCT
18  ...T.G.....G.A.....C.....C.....C.....C.....G.....C.....A
19  ...T.G.....G.....G.....C.....C.....C.....C.....C.....G.....C.....C
20  ...G.....G.....G.....G.....C.....C.....C.....C.....C.....G.....C.....C
21  ...G.....G.....G.....G.....C.....C.....C.....C.....C.....G.....G.....G
22  ...C.....C.....A.....A.....T.....T.....TAAAC.....C.A.T.....C
23  ...G.....G.....G.....C.....C.....C.....C.....C.....C.....G.....G
24  ...G.....G.....G.....C.....C.....C.....C.....C.....C.....C.....C
28  ...G.....G.....G.....ATGATTC.....AA.....AA.....C.....G.....C
29  ...G.....G.....G.....C.....C.....C.....C.....C.....C.....C.....C
30  ...C.....C.....C.....T.....T.....TG.....T.....CA.T.A.C.....C.....A
31  ...A.....A.....AT.GA.G.....T.....T.....T.A.C.T.AG.A.T.....A.A.C.....T.T.A.T.....A
32  ...AT.TA.G.....A.....T.....T.A.C.TG.G.T.T.T.A.A.C.....T.T.A.C.....A
33  ...T.T.A.....T.....A.TA.A.T.....T.....AT.....T.....T.T.A.....TTCA.C.C.T.T.C.....C
35  ...T.....T.T.....TA.A.T.T.G.T.T.....A.....T.....T.A.....CTC.C.C.A.T.A.A...
36  ...T.....A.....A.....A.....A.....A.....A.....A.....A.....A.....A
37  ...AT.GAAG...TTG.AGT.AA.T...CT.T...T.A.....C.....T...GGA.....G..GC...AT.AA..CA.....GCT..C
38  ...A.....C.....T.....A.G.....A.A.TT...T.....T.....T.....AT.....CTC...T.C.....C.....A
39  ...A.....AA.G.....GA.A.....T.T.A.TT...T.T...AG.GA.G...T.....GGA..A..T...CT..A.T..AT.T...T...
41  ...A.....T.....A.....A.....G.....A.....TA.C..T.....T.....T.....GG.....T.....AT.....GTCC..C
42  ...G.....G.....A.....G.....G.....A.C.....C.....C.....GG.....AC.....T.....GTCC..C
43  ...T.T...T.TAAG..TTT...AGTC.GAA.A...CT.T.T.T...TCT...T.TA.AGT..T.A..AT.AG.T.TT  T..AT.AA..CT.....T..GCT...
44  ...T.T...TA.G..T.....A.A..T.T..AA.A..TT...AT...T...A.A..T.T.C..AT...TAAC..T...T.C..A...

320      340      360      380
COS  GGTGTCCTGCCCAACATCCAGGCGGTGCTGCTGCCCAAGAGACCGACACGACACCAACACCAAGGCCAAGGGGAAG
18  ...C.....C.....AG.GCC.C.GTGGG.C.G....C.CCTCGGGCGGCAAGAGGCCACCCAGGCCTCCAGGAGTACTAA
19  ...A.T..GC...A..CAAGTC...A.CGG.....GAG..A.C.GTGCAGG.TC.AC..C.....GGCAACGTCATTCTGTGCGAGGCCTACTAA
20  ...C.....C.....G.GCC.CA.....G.GCC.CA.....A..ATAA
21  ...C.....G.....G.....G.GCC.CAT...G.....C...TGA
22  ...T.....T.A...C.....T.G.GTC.CA.....C...TAA
23  ...C.....C.....T.....G.GCC.CA.....TGA
29  ...G.G.....A.....GCC...T.AA.C...TAA
30  ...C.....A.T.....TC.....GC.T.....TAA
31  ...A.....T.....T.....ATCA.G.T.A
33  ...GT...T.....A.A..TT.T.....A.....GG.TT.A
35  ...C.GT...T.T.A.....T.T...T.....G.....GG..T.A
37  ...A.T..GC...A..CAAGTC...A.CGG.....GAG..A.C.GTGCAGG.TC.AC..C.....GGCAACGTCATTCTGTGCGAGGCCTACTAA
38  ...A.T.....T.....T.....AC.....GCG.AAT.A
39  ...TT...A.....A.CC.A..T.T...A.....T.CTG.....AG.TGCTAAG..CCA...TCTCCATCTAAGGCCCGGAAAGTCACCAAGAAGGCTTAA
41  ...G.....A.....T.....T.....T.A.GGTTGG..AGGC.AGG...CTGCCCTAA
42  ...G.....T.C.....T.....T.A.GGCGG..AGGC.AGG...CTGCCCTAA
43  ...A.T..TC.....TAAA.CT..CT..GGT...C.CT.TACT..A...GATCTAGTGTAA..CTGCTGAACCTCGTTGA
44  ...TT...A.....TCAAACT..T...A.....T.T.C.....GCT.....TTCTCAAG.ATTATAA

```


Figure 7. (cont.)

	160	180	200	220	240	260	280
COS	GGCATCTCCTCAAGCCATGGCCATCATGA	CTCTTGGTCAAGCAGCATCTTCGGAGCGCAT	CGCCGGAGGCTTCCCGCTGGCCCACTACACAA	AGGCTCCACCATCACCTCCCGGAGATCCAG			
45							
46							
47							
48							
49							
52							
53							
54							
55							
56							
62							
63							
64							
65							
68							
69							

	300	320	340	360	380
COS	ACCGCCCTGCCGCTGCTGCTGCCCGGAGAGCT	GGCCAGCAGCCCGCTCTCCGAGGGCACCAGGCTGT	CACCAAGTACACCACCTCCCAAGTAA		
45					
46					
47					
48					
52					
53					
54					
55					
56					
61					
62					
63					
65					
68					
69					
70					

	20	40	60	80	100	120
105	ATGGCTCTACTAAGCAGACCGCTCGCAAGTCCACTGGTGTAAAGCCCGCGCAAGCTCGCGCCGACCGGCGGTGAAGAAGCCCGGATACCGGCGCGGCAC					
71	C.....G.....	C.C.C.C.....G.....	G.....	AG.....	G.....	C.....
72	C.....G.....	T.C.C.C.C.....G.....	G.....	AG.....	G.....	T.C.....
73	G.....	C.....C.....G.....	G.....	AG.....	A.T.....	C.....
74	C.....	G.....G.....G.....A.G.....	T.T.A.AAG.T.A.....	T.....	T.....	T.....
75	C.A.C.....	A.G.....C.....G.....A.....	G.T.G.AAG.G.C.T.T.....	G.....	T.C.....	T.....
76	A.....	T.C.....A.G.....A.A.A.G.....	T.A.A.....	AGT.G.C.T.T.T.A.G.....	A.T.T.T.A.....	T.....
77	C.....	C.C.C.C.....G.....	G.....	AG.....	G.....	G.....
78	G.C.G.....	G.G.C.C.G.....G.....	A.G.....	AG.G.....	G.....	C.....
79	G.C.G.....	G.G.T.....G.....G.....	A.G.....	AG.G.....	G.....	C.....
80	G.C.G.....	G.G.T.....G.....G.....	A.G.....	AG.G.....	G.....	C.....
81	A.A.C.....	G.G.T.....G.....G.....	A.G.....	AG.G.....	G.....	C.....
82	A.A.C.....	G.G.T.....G.....G.....	A.G.....	AG.G.....	G.....	C.....
83	A.A.C.....	G.G.T.....G.....G.....	A.G.....	AG.G.....	G.....	C.....
84	A.A.A.....	G.....G.....G.....	T.....	A.AGT.....	T.....	A.A.A.....
85	C.C.....	A.A.A.....AG.....C.C.G.A.....	T.....	A.G.A.A.AGT.A.....	T.....	C.....
86	A.....	A.A.A.....A.C.....C.....G.....	T.....	A.A.T.AGT.A.A.T.....	T.C.....	A.T.C.T.C.T.....
87	C.....	A.T.....A.G.....A.....G.....	T.....	AGT.T.A.....	A.T.....	C.T.C.T.A.....
88	A.A.....	T.....T.A.....C.A.G.....	T.....	A.G.....	T.C.....	A.T.C.T.A.....
89	A.A.....	C.T.....C.A.G.....	T.....	A.G.....	T.C.....	A.T.C.T.A.....
90	A.A.....	C.T.....C.A.G.....	T.....	A.G.....	T.C.....	A.T.C.T.A.....
91	A.A.....	C.T.....C.A.G.....	T.....	A.G.....	T.C.....	A.T.C.T.A.....
92	A.A.....	C.T.....C.A.G.....	T.....	A.G.....	T.C.....	A.T.C.T.A.....
93	A.A.....	C.T.....C.A.G.....	T.....	A.G.....	T.C.....	A.T.C.T.A.....
94	A.A.....	C.T.....C.A.G.....	T.....	A.G.....	T.C.....	A.T.C.T.A.....
95	C.C.G.....	G.G.....G.C.C.C.....G.....	G.....	G.G.....	G.....	C.T.C.G.....
96	C.C.G.....	G.G.....G.C.C.C.....G.....	G.....	G.G.....	G.....	C.T.C.G.....
97	C.C.G.....	G.G.....G.C.C.C.....G.....	G.....	G.G.....	G.....	C.T.C.G.....
98	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
99	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
100	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
101	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
102	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
103	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
104	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
105	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
106	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
107	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
108	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
109	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
110	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
111	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
112	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
113	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
114	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
115	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
116	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
117	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
118	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
119	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....
120	A.A.....	A.A.A.....A.A.A.....A.A.A.....	A.A.A.....	T.T.A.A.....	T.T.A.T.....	A.T.....

Figure 8. Histone H3 Coding Sequence Alignments.

COS	20	40	60	80	100
121	ATGCTGGACGTGTAAAGGTGGAAGGCTCTGGGTAAAGGCGCGCGCAAGGACAC		AGGAAAGGTTCTCCGGGATACATCCAAGGTATCCCAAGCCCG		
122	C.G...T.....C.T.G.A.....T.....C.....		C.T.....C.....G.C.....G.C.....T.....		
123	T.C.C.G.A.....C.....C.A.T.....T.....C.T		C.C.A.G.G.C.C.....G.C.....G.C.....A.....		
124	T.....T.....G.T.G.A.T.....T.....T.T		C.C.A.C.G.C.C.....G.C.....G.C.....A.....		
125	A.G...C.C.....C.....A.T.....T.....T.T		C.C.....A.....G.C.....G.C.....T.....		
126	C.CA.A.C.G.C.G.G.C.C.....G.....G.....		C.C.....G.G.C.C.....G.C.....G.....		
127	C.CA.A.C.G.C.G.G.C.C.....G.....G.....		C.C.....GT.G.C.C.....G.C.....G.....		
128	C.C.....A.....AT.....G.A.A.....G.....		G.....		
129	C.....A.....G.A.C.A.....A.....T.T		C.C.....T.G.T.....A.....A.....		
130	A.....A.....G.A.C.A.....A.....T.T		C.C.....CT.G.T.....G.A.....A.....		
131	-----A.C.A.....T.....A.T.T		C.C.....A.A.C.....C.....T.....		
132	A.....T.....C.....CT.....A.A.T.A.....T.T		C.C.A.GT.G.T.....T.....G.....G.....A.....		
133	A.....T.....A.C.....CT.....A.....T.T		C.C.A.G.G.T.....T.....G.....G.....T.....		
134	A.....T.....C.....CT.....A.....T.T		C.C.A.GT.G.T.....T.....G.C.....G.A.....		
135	A.....A.C.....A.C.....T.....C.A.A.....T.T		C.C.A.GT.G.T.....T.....T.....G.....G.....		
136	G.G.C.A.G.C.G.G.G.....A.....G.....T.....C.....		A.T.T.....T.....G.....		
137	G.G.C.A.G.C.G.G.G.....A.....G.....T.....C.....		C.C.....C.....C.....C.....G.C.T.G.....		
138	ATG.TCG.....A.....T.....A.....T.....A.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
139	ATG.CCG.....A.....T.....A.....T.....A.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
140	ATG.CCG.....T.....A.....A.....A.T.....A.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
141	ATG.CCG.....T.....A.....A.....A.T.....A.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
142	ATG.CCG.....T.....A.....A.....A.T.....A.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
143	ATG.TCG.....A.....T.....A.....T.....A.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
144	ATG.CCG.....T.....A.....T.....A.....T.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
145	ATG.CCG.....T.....A.....T.....A.....T.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
146	ATG.CCG.....T.....A.....T.....A.....T.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
147	ATG.TCG.....A.....T.....A.....T.....A.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
148	ATG.CCG.....A.....T.....A.....T.....A.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
149	ATG.CCG.....T.....A.....T.....A.....T.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
150	ATG.CCG.....T.....A.....T.....A.....T.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
151	ATG.TCG.....A.....T.....A.....T.....A.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
152	ATG.TCG.....A.....T.....A.....T.....A.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
153	ATGCCCTC.....T.....A.....T.....T.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
154	ATG.TCG.....A.....T.....A.....T.....A.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
155	ATG.TCG.....A.....T.....A.....T.....A.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
156	ATG.TCG.....A.....T.....A.....T.....A.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
157	ATG.CCG.....T.....A.....T.....A.....T.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
158	ATG.TCG.....A.....T.....A.....T.....A.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
159	ATG.CCG.....T.....A.....T.....A.....T.....TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		TCC.A.GATCAA.AA.C.TCT.TG.....T.T.....		
160	G.C...A.....T.....G.ATAC.A.A.TT.T.....A.....TCA.....AA.....G.C.....ATG.A.....		ACC.A.....TCA.....AA.....G.C.....ATG.A.....		
161	G.C...A.....T.....TAC.....A.T.....A.....ACC.A.....TCA.....AA.....A.CA.....ATG.C.....		ACC.A.....TCA.....AA.....A.CA.....ATG.C.....		
162	G.C...A.....T.....ATAC.A.....TT.T.....A.....ACC.A.....ATCA.....AA.....G.C.....ATG.C.....		ACC.A.....ATCA.....AA.....G.C.....ATG.C.....		
163	G.C...A.....T.....A.C.G.A.C.C.....A.....A.....G.....T.....G.....T.....		G.....T.....G.....T.....G.....T.....		
164	C.G.A.C.....A.T.A.....A.....A.....A.....A.....A.....A.....A.....A.....A.....		A.G.A.A.....A.....A.....A.....A.....A.....A.....A.....A.....A.....		

Figure 9. Histone H4 Coding Sequence Alignments

Figure 9. (cont.)

	120	140	160	180	200
120	CCATCCGGGAC	TGGGAGGTGGTCAAGCGATCTCCGGCCTCATCTACGAGGAGACCGCGGTGTGTGAAGTCTTCCTGGAGAACGTCAT			
121	C.....	C..G..A..G.....	G.....	G.....	G.....
122	C.....	C..A..A..G.....	G.....	G.....	G.....
123	C.....	C..A..A..G.....	G.....	G.....	G.....
124	C.....	C..A..A..G.....	G.....	G.....	G.....
125	G..TT.....	C..C.....T.....G..T..T.....	T.....C..C.....T.....T..A..T..G.....	G.....	G.....
126	G..C.....	C..C..C.....G..G.....	G.....C.....C.....G.....	G.....	G.....
127	G..C.....	C..C..C.....G..G.....	G.....C.....C.....G.....	G.....	G.....
128	T..T..T.....	A..T..A..A.....A..T..AGT..T.....A.....T..A..T..C.....T.....	T.....	T.....	T.....
129	T..T..T.....	A..T..A..A.....A..T..AG..T.....A.....T..A..C..C.....T.....	T.....	T.....	T.....
130	T..T..T.....	A..AA..G..A..T.....A..G.....T..T.....A.....A.....A.....T.....	T.....	T.....	T.....
131	T..TT..A.....	T..T.....C..G.....A.....T..A..T.....T..A..A..T..C..T..T.....TT.....A..T..T.....	T.....	T.....	T.....
132	T.....TT.....	C..T.....C..G.....A..T..A.....A.....A..G..T..C..T.....T..T.....A.....	T.....	T.....	T.....
133	T..A..T.....	T..T.....G.....A.....T..T..G.....T.....A..A..C.....C.....T.....T.....	T.....	T.....	T.....
134	T..T..T.....	C..CA.....A.....T.....A.....T..A.....A..A..A.....C.....T.....	T.....	T.....	T.....
135	G.....GA..A..G..C..C..G.....G..G.....G.....C.....C.....A.....C.....	T.....	T.....	T.....	T.....
136	A..A..AA.....C..CA..AA.....A.....A..A.....AG..TCT.....T.....A..AAAC..C..C..GAAG.....C.....G.....	T.....	T.....	T.....	T.....
137	A..A..AA.....C..CA..AA.....A.....A..A.....ATC.....T.....A..AAAC..T..C..GATCA.....C.....G.....	T.....	T.....	T.....	T.....
138	A..A..AA.....C..A..AA.....A.....A.....ATC.....T.....A..AAAC..C..A..GATCA.....C.....T.....	T.....	T.....	T.....	T.....
139	T.....A..T.....	C..T.....C..T.....G.....T..AG..AA..AC.....T..A..C.....AC.....T.....	T.....	T.....	T.....
140	T..T..A..T.....	C..T.....C..T.....G.....T..AG..AAGCG.....A..T..A.....T.....AC.....T..A.....	T.....	T.....	T.....
141	C.....	G.....	G.....	G.....	G.....
142	C.....	G.....	G.....	G.....	G.....
143	C.....	G.....	G.....	G.....	G.....
144	C.....	G.....	G.....	G.....	G.....
145	C.....	G.....	G.....	G.....	G.....
146	C.....	G.....	G.....	G.....	G.....
147	C.....	G.....	G.....	G.....	G.....
148	T.....TG.....	A..G.....C.....C.....G.....A.....A..T.....A.....	T.....	T.....	A.....
149	T.....TG.....	A..G.....C.....C.....G.....A.....A..T.....A.....	T.....	T.....	A.....
150	T.....T.....	A.....A.....A.....A.....A.....T.....T.....T.....A.....A.....A.....	T.....	T.....	A.....
151	T.....T.....	A.....A.....A.....A.....A.....T.....T.....T.....A.....A.....A.....	T.....	T.....	A.....
152	T.....T.....	A.....A.....A.....A.....A.....T.....T.....T.....A.....A.....A.....	T.....	T.....	A.....
153	T.....T.....	A.....A.....A.....A.....A.....T.....T.....T.....A.....A.....A.....	T.....	T.....	A.....
154	T.....T.....	A.....A.....A.....A.....A.....T.....T.....T.....A.....A.....A.....	T.....	T.....	A.....
155	T.....T.....	A.....A.....A.....A.....A.....T.....T.....T.....A.....A.....A.....	T.....	T.....	A.....
156	T.....T.....	A.....A.....A.....A.....A.....T.....T.....T.....A.....A.....A.....	T.....	T.....	A.....
157	T.....T.....	A.....A.....A.....A.....A.....T.....T.....T.....A.....A.....A.....	T.....	T.....	A.....
158	T.....T.....	A.....A.....A.....A.....A.....T.....T.....T.....A.....A.....A.....	T.....	T.....	A.....
159	T.....T.....	A.....A.....A.....A.....A.....T.....T.....T.....A.....A.....A.....	T.....	T.....	A.....
160	T.....T.....	A.....A.....A.....A.....A.....T.....T.....T.....A.....A.....A.....	T.....	T.....	A.....
161	T.....T.....	A.....A.....A.....A.....A.....T.....T.....T.....A.....A.....A.....	T.....	T.....	A.....
162	T.....T.....	A.....A.....A.....A.....A.....T.....T.....T.....A.....A.....A.....	T.....	T.....	A.....
163	T.....T.....	A.....A.....A.....A.....A.....T.....T.....T.....A.....A.....A.....	T.....	T.....	A.....
164	T.....T.....	A.....A.....A.....A.....A.....T.....T.....T.....A.....A.....A.....	T.....	T.....	A.....