CHICKEN HISTONE GENE ORGANISATION

A thesis submitted to the
University of Adelaide,
for the Degree of Doctor of Philosophy

by

RICHARD JAMES D'ANDREA, B.Sc. (Hons.)

Department of Biochemistry,
University of Adelaide,
Adelaide, South Australia

February, 1985

awarded 27-8-1985

CONTENTS

SUMMARY

The work discussed in this thesis involves analysis of the organisation of the core and H1 histone genes in the chicken.

1. DNA fragments specific for the coding regions of either the H_2A, H_2B, H3 or H4 genes were isolated from a previously characterised recombinant, λCHO1, and subcloned into pBR322. An H1 specific probe was provided by L.S. Coles (this laboratory).

2. A chicken genomic library was screened with ^{32}P-labelled core and H1 histone gene sequences and a bank of histone gene containing recombinants isolated.

Non-histone coding probes, from each end of the previously characterised recombinant λCHO1, were used to select overlapping clones λCHO3 and λCHO5.

3. Genomic clones, λCHO3 and λCHO5, and subclones derived from them were characterised by restriction enzyme and hybridisation analysis using the histone gene-specific probes. Fine restriction enzyme and Southern analysis allowed the number and position of each gene to be determined.

Further "chromosome crawling" with non-histone coding probes from each of these clones led to the isolation of λCHO7 (overlapping λCHO5) and confirmed the identification of other clones, already characterised, which extend λCHO3. This region, spanning 50 kb, contained thirteen histone genes and was covered by six λ-recombinants.

4. To aid in further analysis of the histone gene distribution in the chicken genome a cosmid genomic library (1 genome equivalent) was constructed in *E. coli* strain HB101. High molecular weight chicken erythrocyte DNA partially digested with *Sau*3AI, to an average size of 35–45 kb, was inserted into *Bam*HI digested cosmid arms prepared from

the plasmid pHC79. This library was screened using the histone gene-specific probes. Five different cosmid recombinants were isolated, three of which (6.3C, 2.1C and 5.1C) were fully characterised and are described in this thesis.

Cosmid 6.3C extended the region already characterised by 20 kb (eight further genes) giving a continuous genomic region of approximately 70 kb spanning twenty-one histone gene regions.

The genes in this region are arranged in clusters of 2-8. These clusters are separated by non-histone coding regions of 2.5-8 kb. It is clear from this data that no long-range histone gene repeat exists at this locus. However, the organisation of the chicken histone genes within this region is not totally random. H_2A genes are closely associated with H_2B genes and in these cases each gene pair is divergently transcribed. In addition, there is a non-random association of H1 genes with H_2A and H_2B genes. The fact that there are exceptions to these associations (i.e. H_2A occurs between H3 and H4 in λCH03) suggests that the expression of each individual gene is independent of the precise linkage.

Detailed restriction mapping of subclones containing individual gene clusters revealed regions containing symmetrical enzyme sites. In each case histone genes appear as an inverted duplication centered around H3 genes, viz. pCH8.4E, H4.H_2A.H3.H_2A.H4 and pCH11.5E, H1.H_2A.H_2B.H3.H3.H_2B.H_2A.H1. Analysis of sequence data from pCH8.4E indicated that in this case a 2.1 kb region containing H_2A and H4 genes is almost exactly duplicated in the reverse orientation. The boundaries of the duplication are characterised by a 10bp direct repeat at the H4 end and a closely related 10bp inverted repeat at the H_2A end. These repeats may represent the sites of recombination events which generated the symmetrical structure.

5. All the characterised recombinants with chromosomal, histone gene-containing inserts were probed for non-histone encoding regions using cDNA made from 5-day chick embryo poly (A) plus RNA as a probe. Only one cDNA positive region was detected and this corresponded to a region spanned by overlapping cosmid recombinants 2.1C and 5.1C. These recombinants also contain a single H3 gene (subsequently identified as a split H3 gene reported previously by Engel *et al*, 1982).

DNA from the histone-containing recombinants was also hybridised to *in vitro* labelled total chicken DNA and fragments containing repeat-sequence DNA were detected by autoradiography. Further hybridisation analysis with specific probes identified two major unrelated repeats interspersed with the histone genes. A third unknown repeated DNA sequence is represented only once in the regions analysed.

6. Work described in this thesis together with other results from this laboratory has allowed an extensive analysis of the chicken histone gene arrangement. In all, there are forty-two core and H1 histone genes detected in four independent regions, and it is likely that these account for the majority of the core and H1 histone genes in the chicken genome.