

# Epigenetics Evolution and Replacement Histones: Evolutionary Changes at *Drosophila* H4r

Yoshinori Yamamoto<sup>1</sup>, Toru Watanabe<sup>1</sup>, Mayu Nakamura<sup>1</sup>, Norikazu Kakubayashi<sup>1,2</sup>, Yusuke Saito<sup>1</sup> and Yoshinori Matsuo<sup>1,4\*</sup>

<sup>1</sup>Laboratory of Adaptive Evolution, Faculty of Integrated Arts and Sciences, Minamijosanjima-cho 1-1, Tokushima University, Tokushima 770-8502, Japan

<sup>2</sup>Graduate School of Integrated Arts and Sciences, Minamijosanjima-cho 1-1, Tokushima University, Tokushima 770-8502, Japan

<sup>3</sup>Institute of Socio-Arts and Sciences, Tokushima University, Minamijosanjima-cho 1-1, Tokushima 770-8502, Japan

<sup>4</sup>Graduate School of Science and Technology, Tokushima University, Minamijosanjima-cho 2-1, Tokushima 770-8506, Japan

## Abstract

Histone 4 replacement (H4r) can replace replication-dependent H4 in *Drosophila*. To study the evolution of epigenetic mechanisms, the H4 and H4r genes from 14 *Drosophila* species were compared with regard to gene arrangement, codon bias and flanking sequences. Although the amino acid sequences of H4 and H4r are identical or nearly identical, the gene structures are quite different. The H4r gene is a single copy gene located 3R88C9 in *D. melanogaster* between punt and CEP78K, as it is in 11 closely related *Drosophila* species, but not in the three distantly related species. The H4r gene, unlike the H4 gene, has two introns and generates polyadenylated transcripts. The codon usage bias at particular sites differed between H4r and H4. The H4r gene had more GC pairs at 3<sup>rd</sup> codon position. Strongly conserved signal sequence was not found in the 5'-region or 3'-region of the H4r gene. These results suggested that the post transcriptional process such as modifying histone at or after translation will be important for replacing histones and remodeling the chromatin. The evolutionary changes that affect gene structure and codon usage might be a key step to develop epigenetic systems by replacement histones.

**Keywords:** Replacement histone; H4r; *Drosophila*; Epigenetics

## Introduction

Knowledge about histone variants and histone modifications has become very important for studying topics in cell biology such as gene expression, DNA replication, development, cell memory, and chromatin remodeling [1-3]. New information in these fields might be also helpful for understanding quantitative genetics and phenotypic evolution [4-10]. Gene structures of replication-dependent histones (RDHs) and replication-independent histones (RIHs) were studied in a broad range of the species [11-12]. However, detailed comparisons between these two types of histone genes with regard to genome organization, control region, and codon usage have not been published. The differences could be related to the function of a specific histone.

Histone 4 is a small (about 102 amino acids) and highly conserved protein [13]. H4 binds H3 to form the H3-H4 dimer, which is then assembled into the histone core of the nucleosome [13]. In *Drosophila*, there are some histone variants; for example, H3.3 is a H3 variant [14] but there is no such variant for H4. However, a replacement H4 (H4r) that replaces H4 has been reported [15]. These two types of histone 4, H4 and H4r, have identical amino acid sequences [15]. This is probably because the H4 is highly conservative and even a single amino acid substitution was not permitted during the evolution of the two types. This identity is why H4r is called a replacement histone not a histone variant.

In *Drosophila*, a single unit that contains five RDH genes (H1, H2A, H2B, H3, and H4) is repeated in tandem in a large gene cluster [16-23]. Therefore, each of the RDH genes exists in multiple copies (about 110 copies) [24-25]. Notably, the H4r gene structure differs substantially from the H4 gene structure; the H4r gene is a single copy gene with two introns that produces transcripts with a poly(A) tail [11,15,26]. H4 and H4r are expressed at different points in the cell cycle; however, it is not yet completely clear how or whether the obvious differences in gene structure relate to the specific function of each gene type.

Here, the H4r genes from 14 *Drosophila* species were examined to explore the functional significance of the H4r gene. The H4r genes from three *Drosophila* species (*D. mauritiana*, *D. erecta*, and *D. orena*) were newly cloned and sequenced. The genome arrangement, gene structure, and nucleotide sequences in the 5'-region, the coding

region, and the 3'-region of the H4r genes were compared among 14 *Drosophila* species to clarify the roles of these histone replacements. Analyzing the mode and rate of evolution of these replacement histones should illuminate their contributions to phenotypic changes in these organisms during evolution.

## Materials and Methods

### PCR and cloning

*Drosophila* strains from *D. mauritiana*, *D. erecta*, and *D. orena* were donated by Kyushu University. A DNA extraction kit (Sepa Gene Kit, Sanko Junyaku, Co., Ltd.) was used to extract genomic DNA from *Drosophila* larvae. PCR reactions were conducted with Takara Ex Taq as follows [27]: for *D. mauritiana* and *D. orena*, denaturation occurred at 94°C for 1 min, annealing at 54°C for 2 min, followed by polymerization at 70°C for 2 min with extension for 5 sec for 40 cycles. For *D. erecta*, the conditions for PCR were the same as above except that the annealing temperature was 53°C. The primers used for PCR were 5'-TTTGTGCGCAACGGG-3' (H4rF) and 5'-TGTGCTCCCTAAGC-3' (H4rR). The locations of primers for cloning of the H4r genes are shown in Figure 1. Plasmid vector PCR2.1 (Invitrogen) was used to clone each PCR product.

### Determination of nucleotide sequence

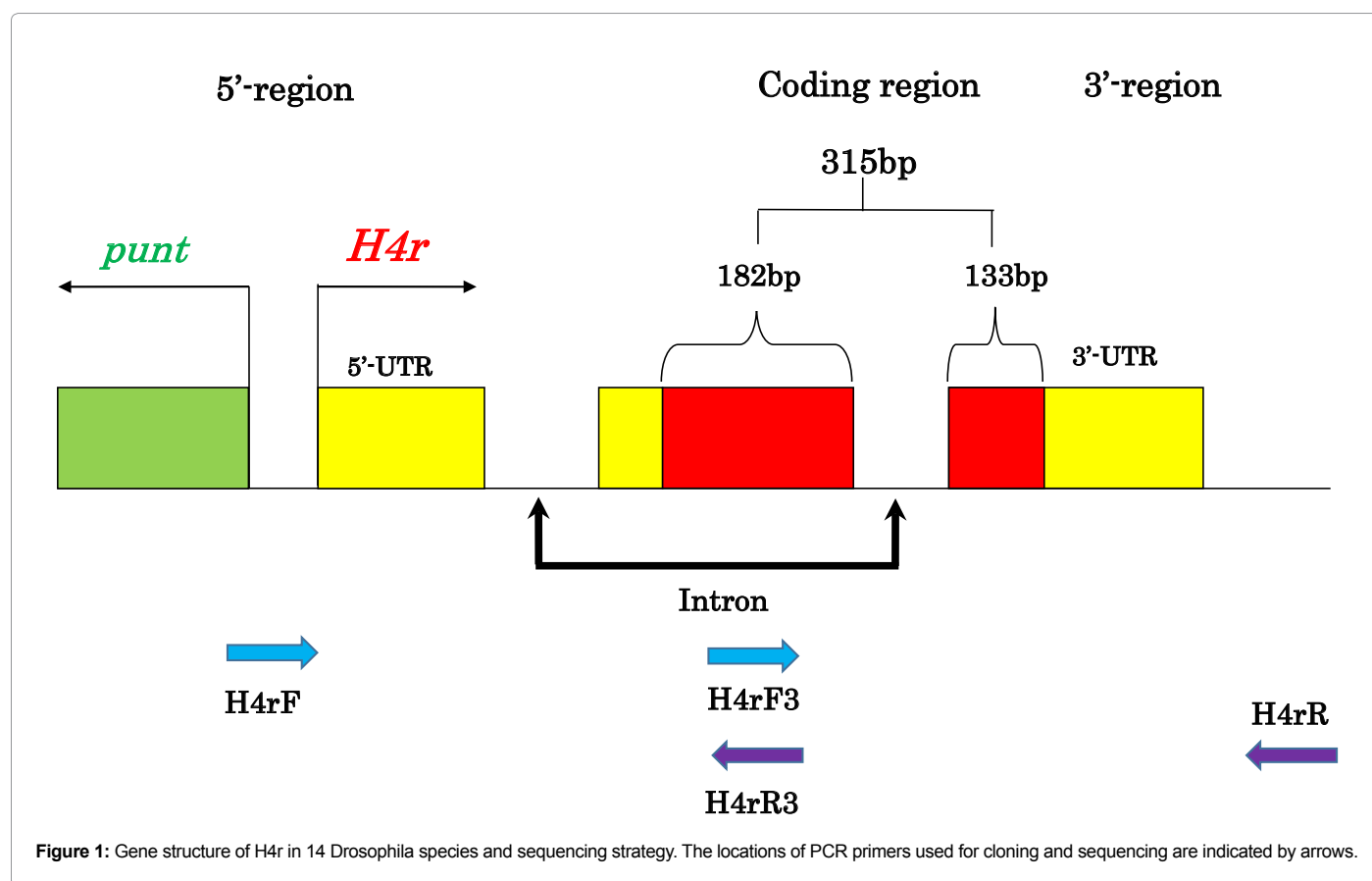
Nucleotide sequence was determined with a Dye Terminator sequence kit (Applied Biosystems) and an ABI 310 sequencer [28]. The sequencing strategy for the histone gene region is depicted

**\*Corresponding author:** Yoshinori Matsuo, Graduate School of Science and Technology, Tokushima University, Minamijosanjima-cho 2-1, Tokushima 770-8506, Japan; Tel: +81-88-656-7270; E-mail: [matsuo.yoshinori@tokushima-u.ac.jp](mailto:matsuo.yoshinori@tokushima-u.ac.jp)

**Received** June 15, 2016; **Accepted** July 20, 2016; **Published** July 26, 2016

**Citation:** Yamamoto Y, Watanabe T, Nakamura M, Kakubayashi N, Saito Y, et al. (2016) Epigenetics Evolution and Replacement Histones: Evolutionary Changes at *Drosophila* H4r. J Phylogenetics Evol Biol 4: 170. doi:10.4172/2329-9002.1000170

**Copyright:** © 2016 Yamamoto Y, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.



in Figure 1. The primers used for sequencing were 5'-GTTCT-CAAGCAGGT-3' (U-19), 5'-CAGGAAACAGCTATGAC-3' (Reverse), 5'-AAGCGTATCTCTGG-3' (H4rF3), and 5'-CCAGAGATACGCTT-3' (H4rR3).

## Data analysis

The nucleotide sequence of the replacement histone 4 gene (H4r) from each of three *Drosophila* species (*D. mauritiana*, *D. erecta*, and *D. orena*) has been deposited into the DNA Data Bank of Japan (DDBJ). The accession numbers for the H4r genes from *D. mauritiana*, *D. erecta*, and *D. orena* are LC127192, LC127193 and LC127194, respectively (Table 1). A second clone from *D. orena* was also sequenced, but not used for analysis; this clone differed from the other clone at only two nucleotide sites in the 3' region. Table 1 lists all 14 *Drosophila* species compared here; the DNA sequences of histone genes from the 11 other *Drosophila* species were obtained from either FlyBase (FB2015-01) or DDBJ (Table 1). Clustal W (ver 2.1) [29] available from DDBJ was used for multiple sequence alignment.

## Results

### Genome organization of H4r genes in fourteen *Drosophila* species

Genomic arrangement of each H4r gene from the 14 *Drosophila* species is shown diagrammatically in Figure 2. For 11 of the 14 species, the H4r gene was located between a punt gene and a Cep78K gene. The punt gene encodes a transforming growth factor-beta receptor with protein kinase activity. The Cep78K gene encodes a centrosomal protein of 78 KDa. In these 11 species, the transcriptional orientation of

the H4r gene is head-to-head relative to punt and tail-to-tail relative to Cep78K. In eight closely related species, the length of spacers between these three genes was highly similar. The spacer between punt and the H4r gene was only slightly longer in *D. pseudoobscura* and *D. persimilis* than in the other eight, but it was much longer in *D. willistoni*. For three more distantly related species (*D. mojavensis*, *D. virilis*, and *D. grimshawi*), punt could not be found upstream of H4r gene. In *D. virilis* and *D. grimshawi*, another gene containing a MADF (myb/SANT-like domain in Adf-1) domain was found upstream of the H4r gene. These findings suggested that a head-to-head pairing of the H4r and punt genes was not always necessary for proper expression of H4r; notably, head-to-head pairing is necessary for proper expression of the replication-dependent genes of H3 and H4 [30]. A Cep78 K gene was located downstream of H4r in all 14 species, and the intergenic spacing was similar length in all species. Because of the tail-to-tail arrangement, it was not clear whether a Cep78K gene positioned a short distance downstream of H4r is necessary for proper H4r expression.

### Comparative analysis of the coding regions for H4 and H4r genes in *Drosophila*

The amino acid sequences deduced from the H4 and H4r nucleotide sequences were aligned for all 14 *Drosophila* species (supplementary data). Of 28 histone 4 genes, 26 showed identical predicted amino acid sequences (consensus sequence) except for two variant H4 genes, one each in *D. sechelia* and in *D. willistoni* (supplementary data). In these two H4 variants, only one (*D. willistoni*) or two (*D. sechelia*) amino acid sites differed from the consensus sequence. Notably, the amino acid sequences of H4r homologs were highly conserved; there were no amino acid substitutions at any site among the 14 *Drosophila*

species. This finding suggested that the function of H4r gene was very important and was expected to be the same or similar to that of H4.

Although the amino acid sequences deduced from these genes were identical for all H4r homologs, the nucleotide sequences encoding these amino acids differed among all the genes; no two nucleotide sequences were identical. This means that there were cases in which the nucleotide sequence at synonymous sites differed. Codon usage bias for these genes is shown in Figure 3. For each histone gene, the total codon usage, the number of codon used summed over *Drosophila* species, was calculated. Although the difference of codon usage for the gene among species is known, a comparison for the gene is possible using the data from the same species. A bias in codon usage was observed for most amino acids, as has been previously reported for other genes [31-32]. However, any difference between two histone types in codon usage was the issue investigated here (Table 2). Although no significant difference in codon usage was observed between the two histone types for six amino acids (His, Asn, Pro, Ser, Tyr and Phe), highly significant differences were found for six other amino acids (Thr, Gly, Arg, Lys, Asp, and Gln). Additionally, moderate differences were found for five other amino acids (Leu, Ala, Val, Glu and Ile). To determine the distribution of codon bias, the codon usage at each amino acid site was listed for two histone 4 genes (Figure 4). At several specific sites, codon usage differed notably between these two histone genes. For example, there was an extreme difference of codon bias at two Lys sites, Lys5 ( $\chi^2_{d.f.=1}=24.3$ ,  $P<0.001$ ) and Lys20 ( $\chi^2_{d.f.=1}=17.4$ ,  $P<0.001$ ); however, no such difference was observed at other Lys sites. Similar site-specific codon bias was also evident for Arg, Gly and Leu. Whether these specific sites that exhibited biased codon usage are subject to histone modification is a matter of interest. A possible relationship between codon usage bias and histone modification [7,33] is indicated in Figure 4. Several sites that exhibit codon biased (e.g., Lys 5 and Lys 20) seemed to have

strong connections to histone modification. Thus the codon usage could be related to the functional difference between H4 and H4r genes.

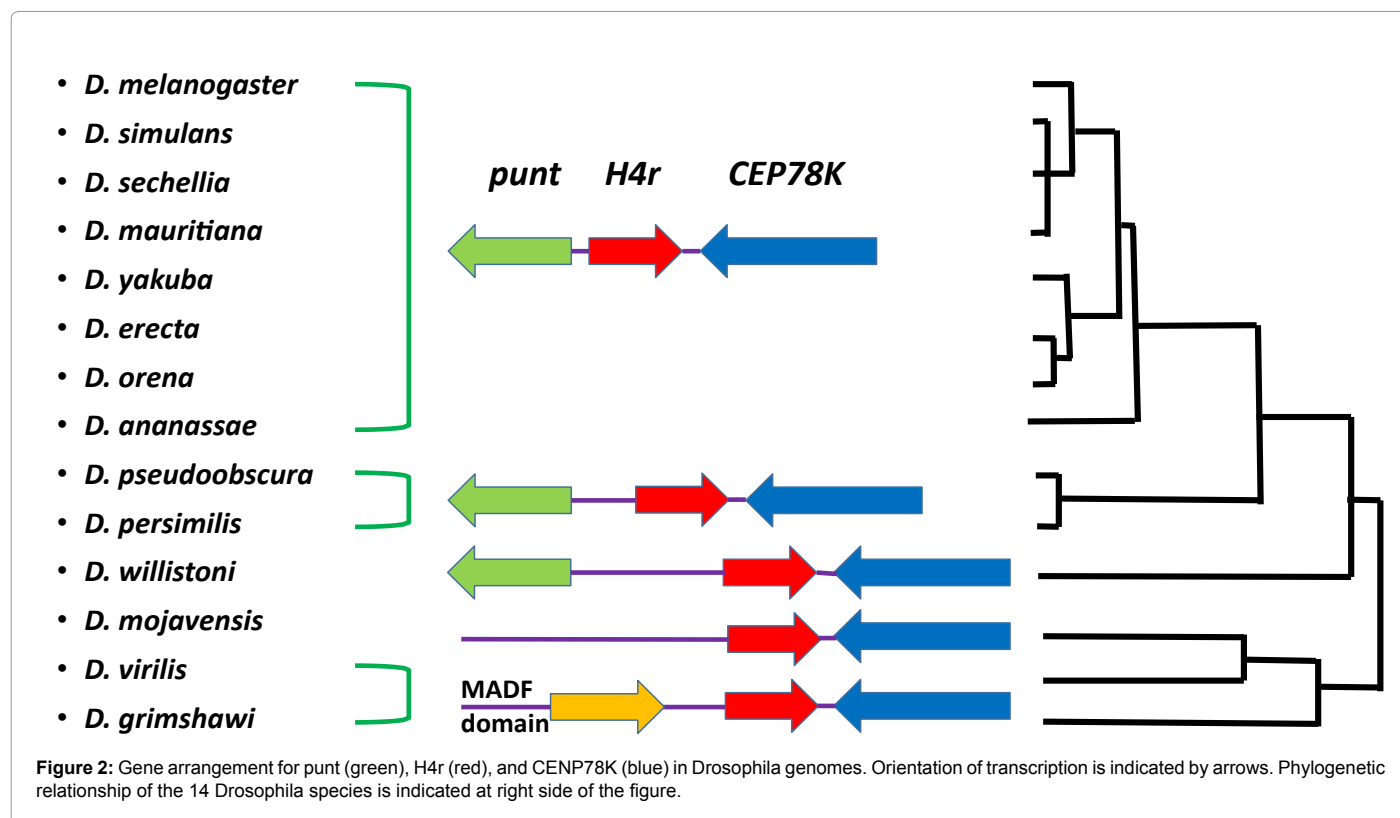
The GC content at the 3rd codon position in the histone 4 genes (H4 and H4r) for the 14 *Drosophila* species is depicted in Figure 5. Although variability in GC content among the species was observed, the GC content of the H4r gene was higher than that of the H4 gene in each species studied. This finding indicated that G or C was used more frequently at synonymous sites in the H4r gene than in the H4 gene.

### Conserved sequences in the 5'-region, the intron, and the 3'-region of H4r genes

Regions upstream of the H4r genes were compared among seven species in the melanogaster group (Figure 6). The spacer between the H4r and punt genes in each of these species was only 50 or 51bp long; in contrast, this spacer length between genes was not as strictly conserved relative to the more distantly related species (Figures 2 and 6). Notably, nine base pairs (AGGGCTGGT) upstream of the transcription start site for punt and seven base pair (ATACTAG) in the middle of spacer were relatively conserved among the melanogaster subgroup species, but not the more distantly related species. Thus in the 5'-region of the H4r gene, there seemed to be no signal sequence that has been conserved strongly as is the case for H3-H4 gene pair [23].

Additionally, the second half of the first intron was relatively conserved among the seven melanogaster subgroup species (Figure 7), although it was not conserved in the distantly related species.

Regions downstream of the H4r genes were also compared among the seven melanogaster subgroup species (Figure 8). The spacer between H4r gene and CEP78K gene was short (60 ~ 96 bp) and relatively constant length even in the distantly related species (Figure 2). However, no conserved sequences were found in this spacer region.



Species	H4 gene	H4r gene
<i>D. melanogaster</i>	X14215 (DDBJ)	CG3379 (FlyBase)
<i>D. simulans</i>	AB055959 (DDBJ)	GD18968 (FlyBase)
<i>D. sechellia</i>	AB105183 (DDBJ)	GM24174 (FlyBase)
<i>D. mauritiana</i>	AB105182 (DDBJ)	LC127192 (DDBJ)
<i>D. yakuba</i>	AB073635 (DDBJ)	GE24247 (FlyBase)
<i>D. erecta</i>	AB073634 (DDBJ)	LC127193 (DDBJ)
<i>D. oreana</i>	AB105181 (DDBJ)	LC127194 (DDBJ)
<i>D. ananassae</i>	GF19278 (FlyBase)	GF17447 (FlyBase)
<i>D. pseudoobscura</i>	AB249651 (DDBJ)	GA17414 (FlyBase)
<i>D. persimilis</i>	GL17799 (FlyBase)	GL24413 (FlyBase)
<i>D. willistoni</i>	GK10375 (FlyBase)	GK12810 (FlyBase)
<i>D. mojavensis</i>	GI13912 (FlyBase)	GI10680 (FlyBase)
<i>D. virilis</i>	GJ23527 (FlyBase)	GJ24352 (FlyBase)
<i>D. grimshawi</i>	GH23501 (FlyBase)	GH18656 (FlyBase)

**Table 1:** Analysis of the H4 and H4r genes in 14 Drosophila species. Nucleotide sequence data were obtained from FlyBase (annotation ID) or DDBJ (accession number).

Amino Acid	$\chi^2$	d.f.
Thr	53.04***	3
Gly	33.24***	3
Arg	43.60***	5
Lys	17.96***	1
Leu	17.12**	5
Ala	11.61**	3
His	1.2	1
Val	12.82**	3
Asp	21.26***	1
Asn	4.08*	1
Ile	12.0**	2
Gln	21.32***	1
Pro	3.87	3
Ser	0.67	2 <sup>a</sup>
Tyr	0	1
Glu	10.32**	1
Phe	1.79	1

\*\*\*:  $P < 0.001$ ; \*\*:  $P < 0.01$ ; \*:  $P < 0.05$ ; <sup>a</sup>: The expected value which is smaller than one is combined with the next smallest value.

**Table 2:**  $\chi^2$  tests for codon bias of the H4 and H4r genes in 14 Drosophila species.

## Discussion

### Gene structure, control signal and function of H4r

In many cases, the gene pair encoding replication-dependent histones (e.g., H3-H4 or H2A-H2B in Drosophila) has a head-to-head orientation [16,24-25], and transcription of two genes is co-regulated [30]. In the closely related species of the melanogaster group, the H4r and punt genes were arranged in a head-to-head orientation and only about 50 bp apart. In the more distantly related species, however, the length of the spacer between two genes was not conserved. In *D. mojavensis*, *D. virilis*, and *D. grimshawi*, punt was not located near the H4r gene and was located in another region. This finding suggested that the H4r-punt pairing did not have some highly conserved functional significance. However, the possibility cannot be excluded that this pairing of genes has a function in the melanogaster group. CEP78K, which encodes a centriole protein, was located downstream of the H4r gene in all 14 species studied; however, the tail-to-tail orientation and the lack of a signal sequence element in the spacer suggested that these two genes were not co-regulated.

The gene structures of two types of histone gene were quite different from each other. A multiple-copy gene structure seemed to be favorable for producing a large amount of protein in a short period of time, while a single-copy structure seemed to be favorable for fine-tuning the expression level of the H4r gene. Additionally, an exon-intron structure allows for the production of multiple kinds of transcripts from a single transcription unit. The transcription termination signals are a hair-pin loop structure and a poly (A) signal for H4 gene and H4r gene, respectively [26,34-35]. A protein that binds to the hair-pin loop structure is part of transcription complex; this complex facilitates the coordination between histone mRNA transcription and DNA replication [30].

As the amino acid sequence of H4r was identical to that of most H4 copies; this identity is the reason that H4r is not considering a histone 'variant'. Because of the conserved protein, even a single amino acid substitution has not been tolerated during the evolution of H4 and H4r. Although the exact function of H4r is not clear, the strong conservation of this protein among the 14 Drosophila species suggested that this protein has an important role in chromatin remodeling. Therefore, H4r, like H3.3 and H2AvD, is expected to play a significant role in histone replacement. It seemed to be a key for chromatin remodeling that non modified or differently modified histones replace with the RDHs. Therefore the mechanisms which produce different control of expression will be important rather than the difference of primary structure.

In the region upstream of H4r gene, no transcriptional control signals were found. In addition, based on FlyBase data, the expression profile for the H4r gene in Drosophila (FB2015-01) indicated that the transcripts were observed at most developmental stages and at similar expression levels. Therefore, transcription may not be a critical step in the control of the H4r gene. Post-transcriptional or post-translational processes (e.g., histone modification) might be important for chromatin remodeling.

### Codon usage and histone modification

The codon usage of Drosophila histone genes has been investigated for many species [19,36-38]. Mechanisms of generating bias for codon usage such as selection-mutation balance, the effect of population size and bias for mutation pattern have been extensively studied [23,39-44]. The overall codon usage did not differ greatly between the two histone 4 genes, H4 and H4r. However, a site-by-site analysis of codon bias showed that some sites did exhibit substantial difference. Sites exhibiting moderate bias were not common. Thus, codon usage at particular sites did differ between H4 and H4r. The relationship between codon usage and histone modification was not clear, but codon bias is one characteristic difference between the two types of histone. It is generally recognized that histone modifications occur post-translationally [1,45]. However, some types of histone modification, e.g., methylation of H3 Lys 9, occur during translation [46]. This finding suggests the possibility that differences in codon usage could affect the efficiency of some histone modifications. This finding also means that the differences in codon usage between the two types of histone could have functional significance.

The differences of gene structure and codon usage observed between two types of histone genes may be relevant to the respective functions. Therefore, the evolutionary mechanism for developing

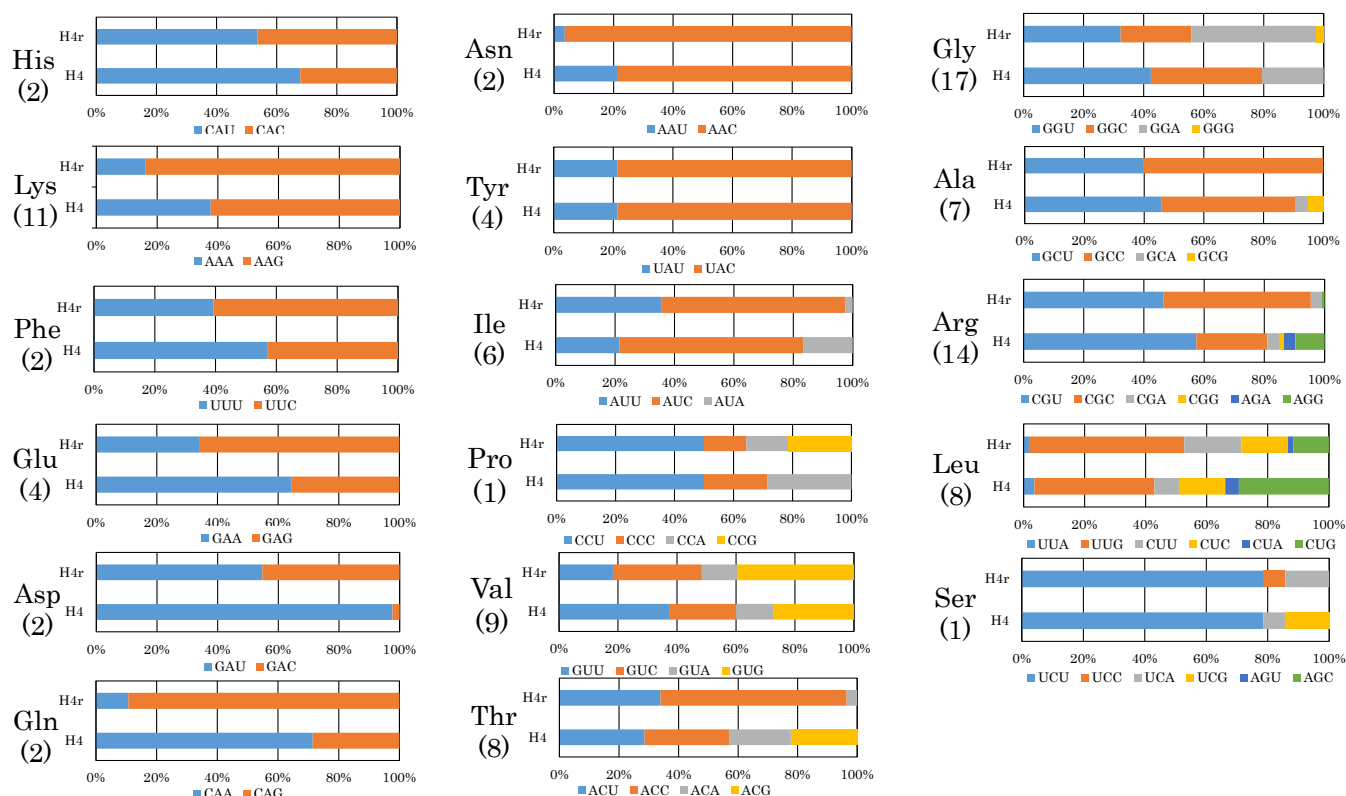
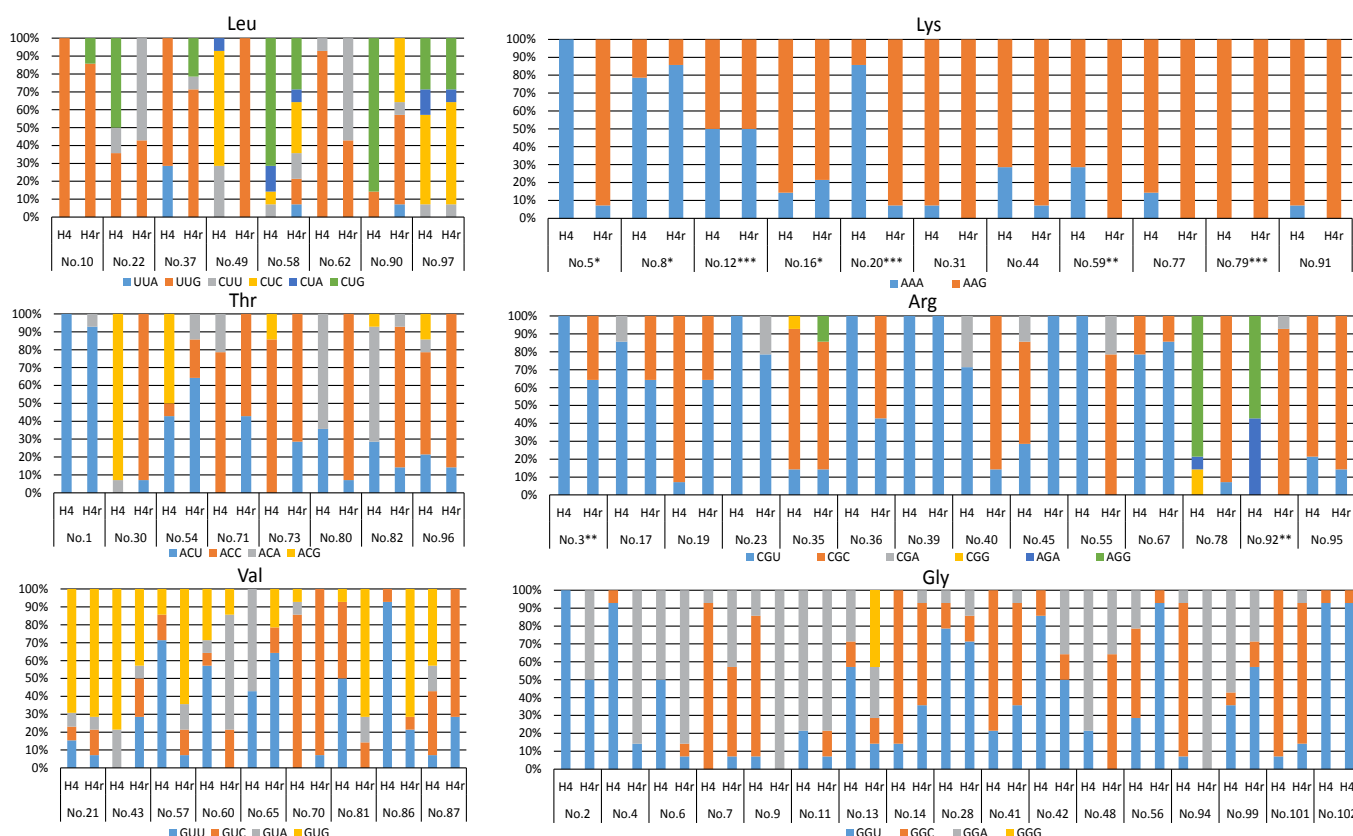
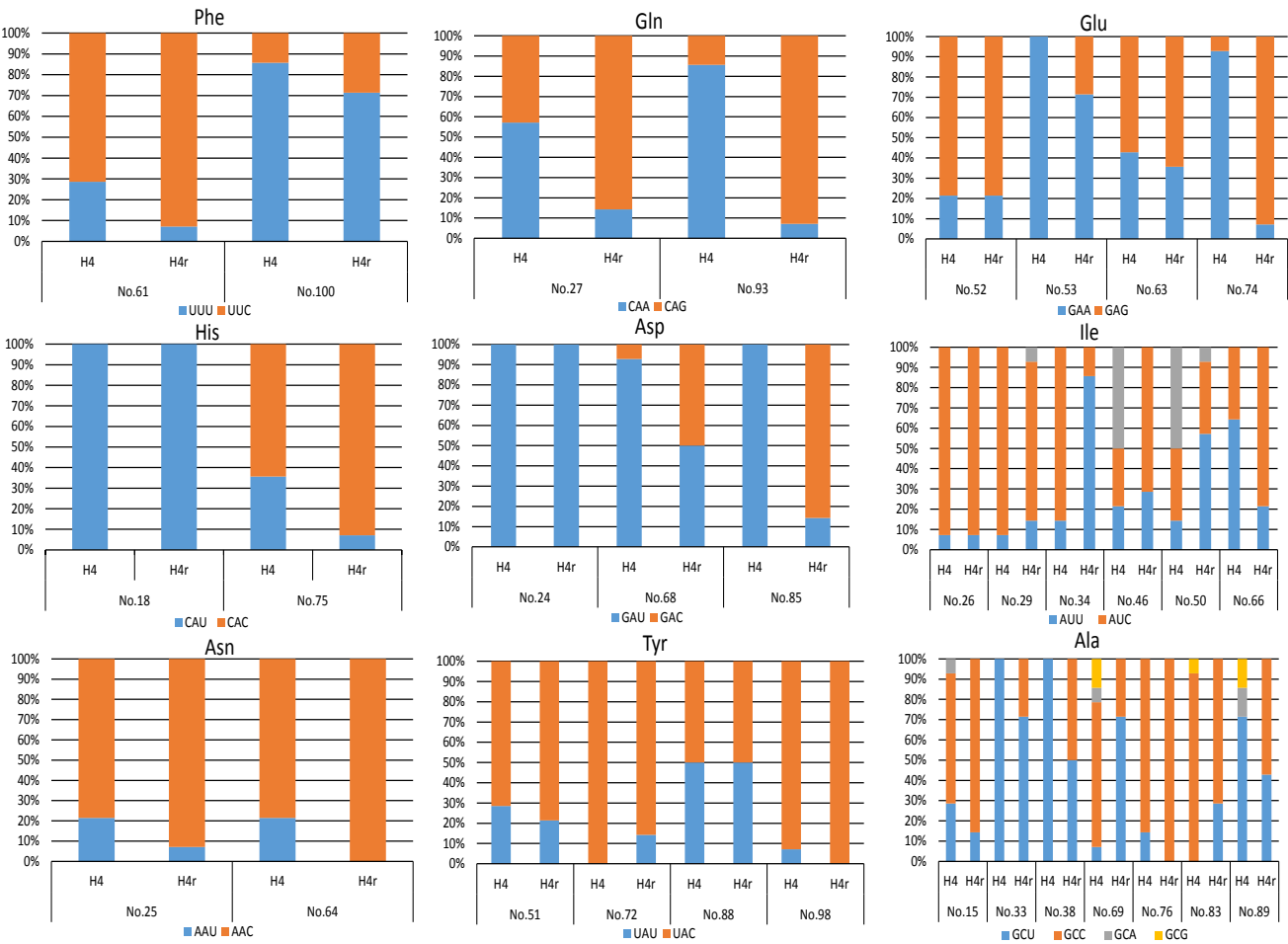


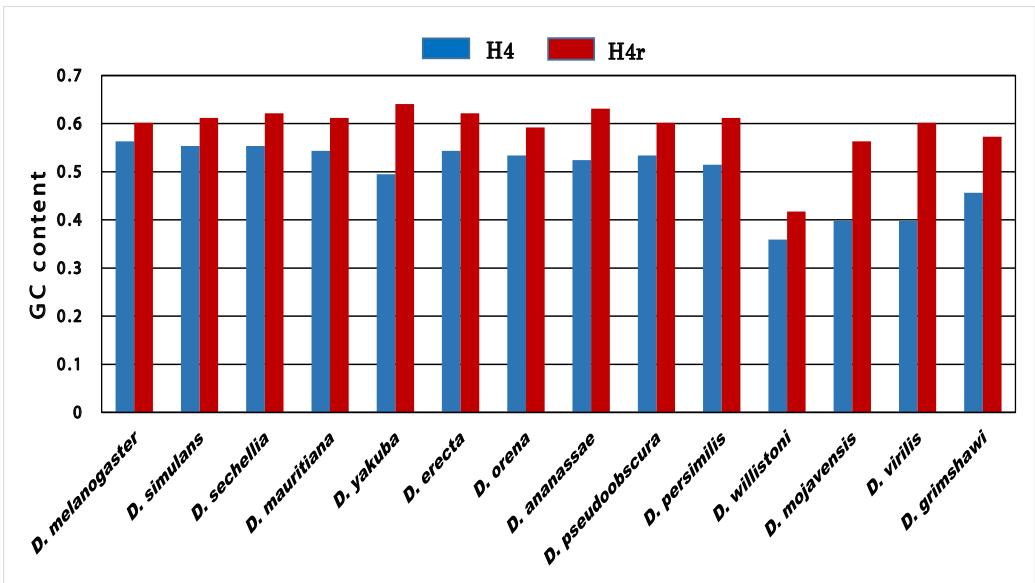
Figure 3: Codon usage of histone genes for H4 and H4r in Drosophila. Numbers within the parenthesis indicate the number of constituent amino acids in H4 and H4r.



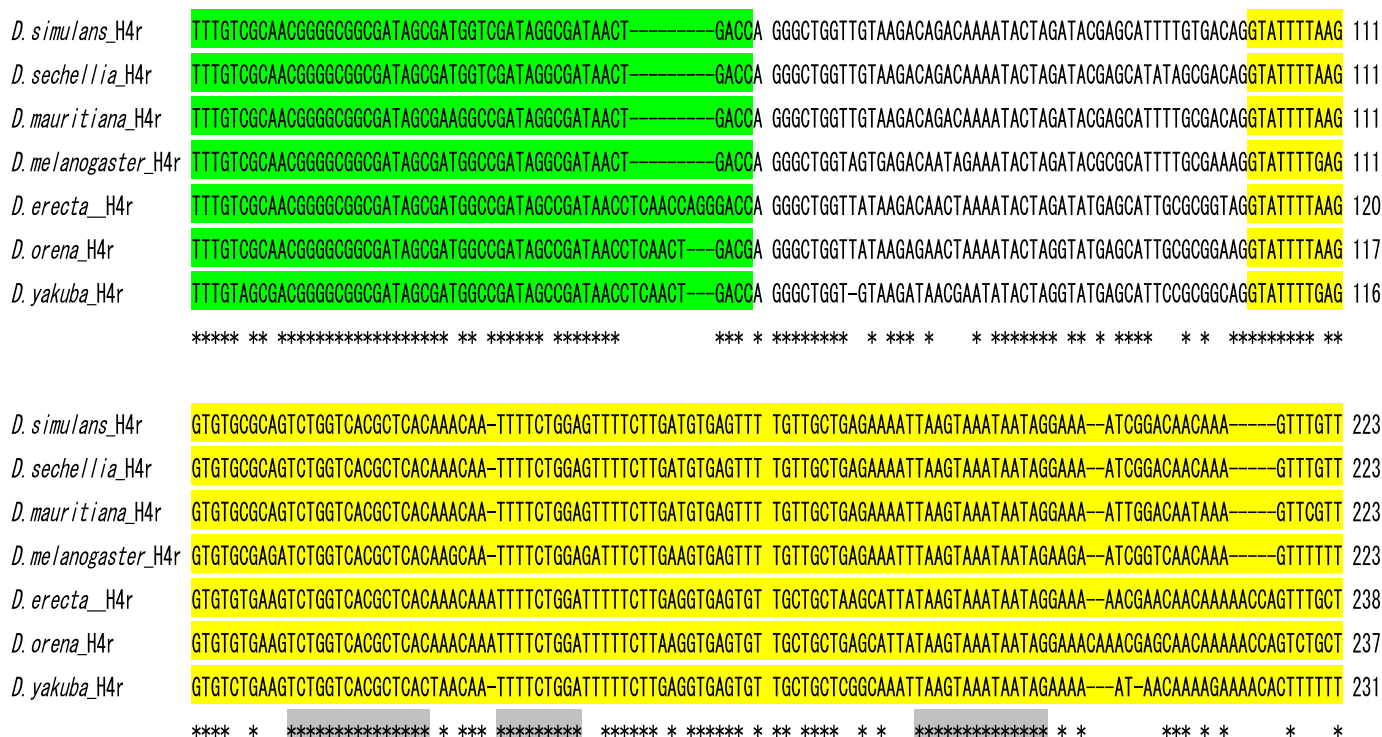




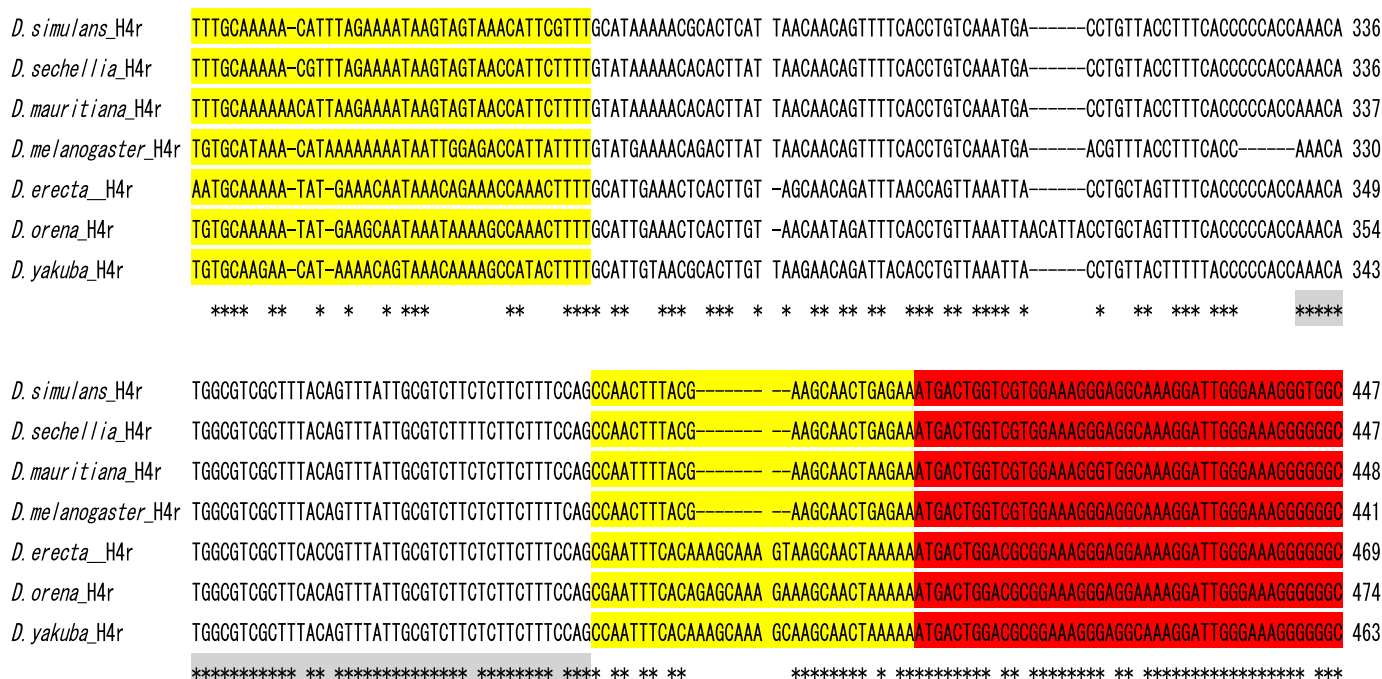
**Figure 4:** Codon usage at each amino acid site of the respective histone 4 genes in Drosophila. The locations of histone modifications are indicated as follows: \*(Acetylation mark), \*(Methylation mark) and \*(Acetylation and methylation mark) after the position number.



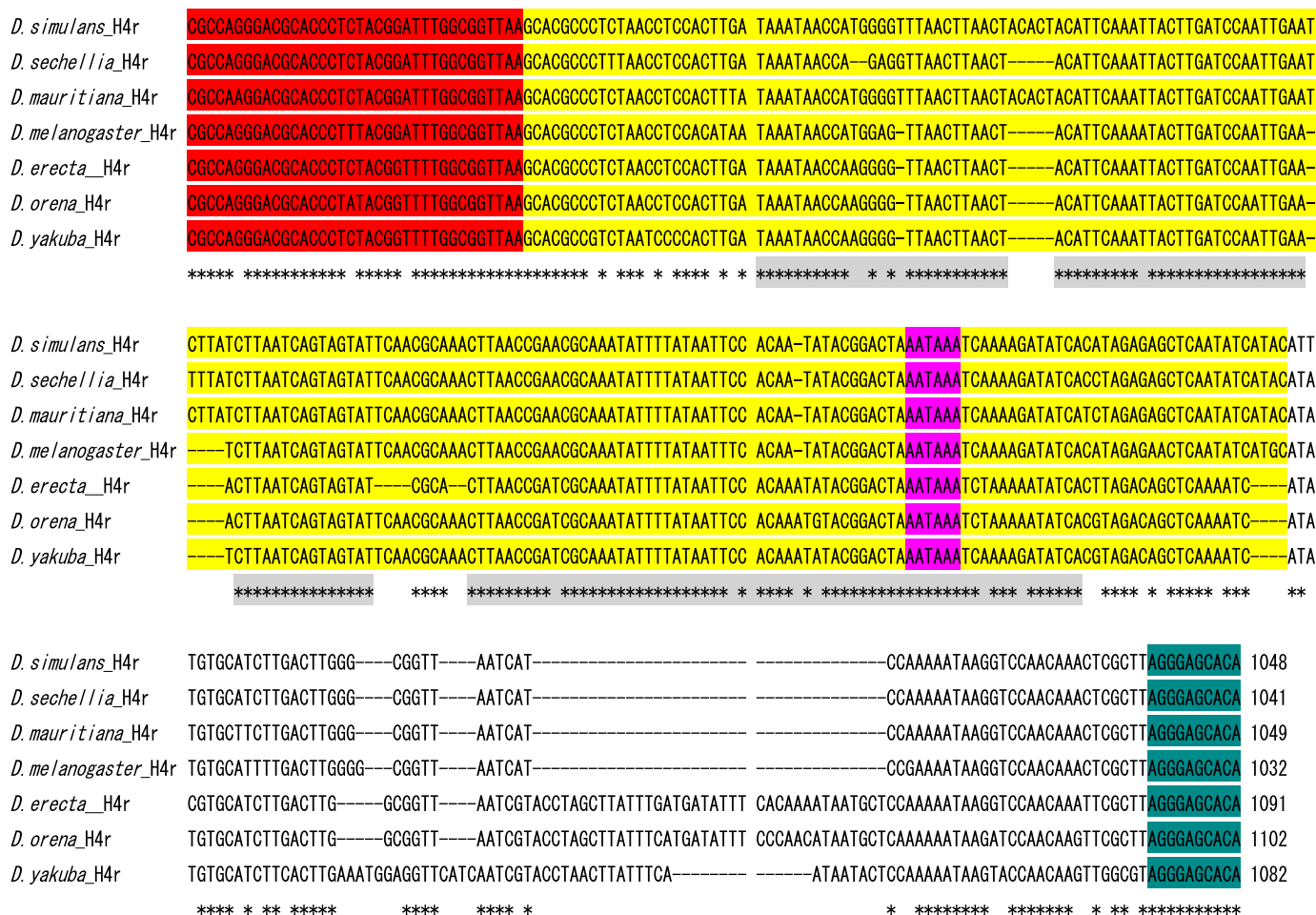
**Figure 5:** GC content at 3<sup>rd</sup> codon positions of the H4 (blue) and H4r (red) genes in Drosophila.



**Figure 6:** Nucleotide sequence comparison of the 5' flanking region of H4r genes in Drosophila. Nucleotide sequences of the melanogaster group are aligned. Asterisk under the sequence indicates identical nucleotides at the respective sites in seven Drosophila species. The regions are shown by the same colors as indicated in Figure 1.



**Figure 7:** Nucleotide sequence comparison of the first intron in H4r genes in seven Drosophila species. Nucleotide sequences of the melanogaster group are aligned. Asterisk under the sequence indicates identical nucleotides at the respective sites in seven Drosophila species. The regions are shown by the same colors as indicated in Figure 1.



**Figure 8:** Nucleotide sequence comparison of the 3' flanking region of *H4r* genes in seven *Drosophila*. Nucleotide sequences of the melanogaster group are aligned. Asterisk under the sequence indicates identical nucleotides at the respective sites in seven *Drosophila* species. The regions are shown by the same colors as indicated in Figure 1.

diverse gene structures and codon bias might be important for the histone replacement system and chromatin remodeling. To understand the evolution of histone replacement system, other histones variant, such as variants of H3 and H2A, should be also investigated.

## Conclusions

In this study, the replacement histone 4 genes from *Drosophila* are analyzed to investigate the functional differentiation and conservation. The results suggested that the post-transcriptional and post-translational processes such as histone modifications are important for replacing histones and remodeling chromatin. Evolutionary mechanisms that affect gene structure and codon usage might be important in the emergence of epigenetic systems that depend on replacement histones.

## Supplementary Data

Additional Figure S1 is available.

## Acknowledgement

This research was supported by a Grant-in-Aid for Scientific Research to Y. M. from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

## References

1. Grunstein M (1997) Histone acetylation in chromatin structure and transcription. *Nature* 389: 349-352.
2. Kouzarides T (2007) Chromatin modifications and their function. *Cell* 128: 693-705.
3. Mito Y, Henikoff JG, Henikoff S (2005) Genome-scale profiling of histone H3.3 replacement patterns. *Nat Genet* 37: 1090-1097.
4. Feng S, Jacobsen SE, Reik W (2010) Epigenetic reprogramming in plant and animal development. *Science* 330: 622-627.
5. Rothbart SB, Strahl BD (2014) Interpreting the language of histone and DNA modifications. *Biochim Biophys Acta* 1839: 627-643.
6. Weber CM, Henikoff S (2014) Histone variants: dynamic punctuation in transcription. *Genes Dev* 28: 672-682.
7. Elliott GO, Murphy KJ, Hayes JJ, Thiriet C (2013) Replication-independent nucleosome exchange is enhanced by local and specific acetylation of histone H4. *Nucleic Acids Res* 41: 2228-2238.
8. Riddle NC, Minoda A, Kharchenko PV, Alekseyenko AA, Schwartz YB, et al. (2011) Plasticity in patterns of histone modifications and chromosomal proteins in *Drosophila* heterochromatin. *Genome Res* 21: 147-163.
9. Wang CI, Alekseyenko AA, LeRoy G, Elia AEH, Gorchakov AA, et al. (2013) ChIP-mass spectrometry captures protein interactions and modified histones associated with dosage compensation in *Drosophila*. *Nat Struct Mol Biol* 20: 202-209.
10. Maze I, Noh KM, Soshnev AA, Allis CD (2014) Every amino acid matters: essential contributions of histone variants to mammalian development and disease. *Nat Rev Genet* 15: 259-271.
11. Isenberg I (1979) Histones. *Annu Rev Biochem* 48: 159-191.
12. Wells DE (1986) Compilation analysis of histones and histone genes. *Nucleic Acids Res* 14: r119-r149.



13. Kornberg RD (1977) Structure of chromatin. *Annu Rev Biochem* 46: 931-954.
14. Fretzin S, Allan BD, Angela van Daal, Elgin SCR (1991) A *Drosophila melanogaster* H3.3 cDNA encodes a histone variant identical with the vertebrate H3.3. *Gene* 107: 341-342.
15. Akhmanova A, Miedema K, Hennig W (1996) Identification and characterization of the *Drosophila* histone H4 replacement gene. *FEBS Lett* 388: 219-222.
16. Matsuo Y, Yamazaki T (1989) tRNA derived insertion element in histone gene repeating unit of *Drosophila melanogaster*. *Nucleic Acids Res* 17: 225-238.
17. Matsuo Y, Yamazaki T (1989) Nucleotide variation and divergence in the histone multigene family in *Drosophila melanogaster*. *Genetics* 122: 87-97.
18. Kremer H, Hennig W (1990) Isolation and characterization of a *Drosophila hydei* histone DNA repeat unit. *Nucleic Acids Res* 18: 1573-1580.
19. Fitch DHA, Strausbaugh LD (1993) Low codon bias and higher rates of synonymous substitution in *Drosophila hydei* and *D. melanogaster* histone genes. *Mol Biol Evol* 10: 397-413.
20. Tsunemoto K, Matsuo Y (2001) Molecular evolutionary analysis of a histone gene repeating unit from *Drosophila simulans*. *Genes Genet Syst* 76: 355-361.
21. Kakita M, Shimizu T, Emoto M, Nagai M, Matsuo Y, et al. (2003) Divergence and heterogeneity of the histone gene repeating units in the *Drosophila melanogaster* species subgroup. *Genes Genet Syst* 78: 383-389.
22. Nagoda N, Fukuda A, Nakashima Y, Matsuo Y (2005) Molecular characterization and evolution of the repeating units of histone genes in *Drosophila americana*: coexistence of quartet and quintet units in a genome. *Insect Mol Biol* 14:713-717.
23. Matsuo Y (2006) Genomic structure and evolution of the histone gene family in *Drosophila*. *Current Topics in Genetics* 2:1-14.
24. Pardue ML, Kedes LH, Weinberg ES, Birnstiel ML (1977) Localization of sequences coding for histone messenger RNA in the chromosomes of *Drosophila melanogaster*. *Chromosoma* 63: 135-151.
25. Lifton RP, Goldberg ML, Karp RW, Hogness DS (1977) The organization of the histone genes in *Drosophila melanogaster*: functional and evolutionary implications. *Proc Cold Spring Harbor Symp Quant Biol* 42: 1047-1051.
26. Marzluff WF, Wagner EJ, Duronio RJ (2008) Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat Rev Genet* 9: 843-854.
27. Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, et al. (1985) Enzymatic amplification of  $\beta$ -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230: 1350-1354.
28. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74: 5463-5467.
29. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673-4680.
30. Salzler HR, Tatomer DC, Malek PY, McDaniel SL, Orlando AN, et al. (2013) A sequence in the *Drosophila* H3-H4 promoter triggers 1 Histone Locus Body assembly and biosynthesis of replication-coupled histone mRNAs. *Dev Cell* 24: 623-634.
31. Shields DC, Sharp PM, Higgins DG, Wright W (1988) "Silent" sites in *Drosophila* genes are not neutral: evidence of selection among synonymous codons. *Mol Biol Evol* 5: 704-716.
32. Sharp PM, Lloyd AT (1993) Codon usage. In 'An atlas of *Drosophila* genes' by G. Maroni. Oxford University Press, New York.
33. Li KK, Luo C, Wang D, Jiang H, Zheng YG (2012) Chemical and biochemical approaches in the study of histone methylation and demethylation. *Med Res Rev* 32: 815-867.
34. Schumperli D (1986) Cell-cycle regulation of histone gene expression. *Cell* 45: 471-472.
35. Sabath I, Skrajna A, Yang XC, Dadlez M, Marzluff WF, et al. (2013) 3'-End processing of histone pre-mRNAs in *Drosophila*: U7 snRNP is associated with FLASH and polyadenylation factors. *RNA* 19: 1726-1744.
36. Matsuo Y (2000) Evolutionary change of codon usage for the histone gene family in *Drosophila melanogaster* and *Drosophila hydei*. *Mol Phylogenet Evol* 15: 283-291.
37. Matsuo Y (2000) Molecular evolution of the histone 3 multigene family in the *Drosophila melanogaster* species subgroup. *Mol Phylogenet Evol* 16: 339-343.
38. Matsuo Y (2003) Evolution of the GC content of the histone 3 gene in seven *Drosophila* species. *Genes Genet Syst* 78: 309-318.
39. Bernardi G, Bernardi G (1986) Compositional constraints and genome evolution. *J Mol Evol* 24: 1-11.
40. Moriyama EN, Hartl DL (1993) Codon usage bias and base composition of nuclear genes in *Drosophila*. *Genetics* 134: 847-858.
41. Akashi H (1995) Inferring weak selection from patterns of polymorphism and divergence at "silent" sites in *Drosophila* DNA. *Genetics* 139: 1067-1076.
42. Powell JR, Moriyama EN (1997) Evolution of codon usage bias in *Drosophila*. *Proc Natl Acad Sci U S A* 94: 7784-7790.
43. Poh YP, Ting CT, Fu HW, Langley CH, Begun DJ (2012) Population genomic analysis of base composition evolution in *Drosophila melanogaster*. *Genome Biol Evol* 4: 1245-1255.
44. Lawrie DS, Messer PW, Hershberg R, Petrov DA (2013) Strong purifying selection at synonymous sites in *D. melanogaster*. *PLoS Genet* 9: e1003527.
45. Jenuwein T, Allis CD (2001) Translating the histone code. *Science* 293: 1074-1080.
46. Rivera C, Saavedra F, Alvarez F, Diaz-Celis C, Ugalde V, et al. (2015) Methylation of histone H3 lysine 9 occurs during translation. *Nucleic Acids Res* 43: 9097-9106.