A CR1 element is embedded in a novel tandem repeat (*Hin*fI repeat) within the chicken genome

Juan Li and Frederick C. Leung

Abstract: Highly repetitive DNA sequences constitute a significant portion of most eukaryotic genomes, raising questions about their evolutionary origins and amplification dynamics. In this study, a novel chicken repetitive DNA family, the *Hin*fI repeat, was characterized. The basic repeating unit of this family displays a uniform length of 770 bp, which was defined by the recognition site of *Hin*fI. The *Hin*fI repeat was specifically localized in the pericentric region of chromosome 4 by fluorescence in situ hybridization and constitutes 0.51% of the chicken genome. Interestingly, a chicken repeat 1 (CR1) element has been identified within this basic repeating unit. Like other CR1 elements, this CR1 element also displays typical retrotransposition characteristics, including a highly conserved 3′ region and a badly truncated 5′ end. This direct evidence from sequence analysis, together with our Southern blot results, suggests that the *Hin*fI repeat may originate from a unique region containing a retrotransposed CR1 element.

Key words: satellite DNA, CR1 retrotransposon, *Hin*fI repeat, *Gallus gallus*.

Résumé : Les séquences très répétitives d'ADN constituent une part significative de la plupart des génomes eucaryotes, ce qui soulève des questions sur leurs origines évolutives et leurs dynamiques d'amplification. Dans ce travail, une nouvelle famille d'ADN répétitif chez le poulet, la répétition *Hin*fI, a été caractérisée. L'unité de base est d'une taille uniforme (770 pb) et elle est bordée de sites *Hin*f1. La répétition *Hin*f1 est située spécifiquement dans la région péricentromérique du chromosome 4, tel qu'observé par hybridation in situ en fluorescence, et constitue 0,51 % du génome du poulet. Fait à remarquer, un élément répété CR1 est logé au sein de l'unité de base. Comme d'autres éléments CR1, cet élément CR1 montre des caractéristiques typiques de la rétrotransposition dont une région en 3′ fortement conservée et une région en 5′ grandement tronquée. Ces évidences directes découlant de l'analyse de séquence ainsi que les résultats d'une analyse Southern suggèrent que la répétition *Hin*f1 pourrait tirer son origine d'une région unique contenant un élément CR1 rétrotransposé.

Mots clés : ADN satellite, rétrotransposon CR1, répétition *Hin*f1, *Gallus gallus*.

[Traduit par la Rédaction]

Introduction

Highly repetitive DNAs, or satellite DNAs, are generally believed to be involved in centromere condensation, chromosome rearrangement, and karyotype evolution (Miklos 1985). These DNAs are usually detected by restriction endonuclease digestion resulting in a ladder structure ranging from a monomer to a multimer of the repeat. With extreme diversity in sequence and copy number, highly repetitive DNAs are considered to be a quickly evolving part of the eukaryotic genome (Miklos 1985).

The rapid divergence of highly repetitive sequences raises fundamental questions concerning their evolutionary origins and amplification dynamics (Charlesworth et al. 1994). Tandemly arrayed repeats had been suggested to arise from smaller arrays through the process of amplification involving replication slippage and unequal crossing-over (Pech et al. 1979). Cross hybridization was found to occur between related species, lending support to the library hypothesis that related species share a library of conserved satellite DNA sequence (Fry and Salser 1977; Mestrovic et al. 1998; Pons et al. 2004). An alternative model proposed by Smith (1976) postulates that a nonfunctional sequence will experience crossing-over events, and the recombinant products will not be eliminated by selection. With continuous crossing-over events, a periodicity occurs in the sequence. This suggests that satellite DNAs may evolve from the wide spectrum of non-satellite DNA sequences.

According to this model, the evolutionary origin of non-satellite DNAs should also include the middle-repetitive sequences, which retrotranspose in an RNA-mediated pattern and constitute a large portion of the genome (Rogers 1985). Until recently, direct experimental evidence to support the above hypothesis remained limited. In *Drosophila virilis* species, the *pvB370* satDNAs had been found to show significant similarity to the long terminal repeats of the *pDv*

Received 28 January 2005. Accepted 27 July 2005. Published on the NRC Research Press Web site at http://genome.nrc.ca on 13 January 2006.

Corresponding Editor: T.E. Bureau.

J. Li and F.C. Leung.[1] Department of Zoology, The University of Hong Kong, Hong Kong SAR, China.

[1]Corresponding author (e-mail: fcleung@hkucc.hku.hk).

 doi: 10.1139/G05-090 © 2006 NRC Canada

transposon (Heikkinen et al. 1995), suggesting an evolutionary relationship between this satellite DNA family and this transposon. In whales and dolphins, the heterochromatic satellite DNAs were found to be directly evolved from the long interspersed nucleotide element (LINE)-like repeats, which experimentally supports the idea that retroelements may serve as the basic core of satellite DNA monomers (Kapitonov et al. 1998).

Chicken repeat 1 (CR1) belongs to a category of non-long-terminal-repeat (non-LTR) retrotransposons widely distributed among avian species (Chen et al. 1991). The consensus sequence of the CR1 retrotransposon spans 4556 bp and contains two long open reading frames (ORFs I and II) (Haas et al. 1997). It was suggested that ORF I encodes a nucleic acid binding protein. The characteristic reverse transcriptase (RT) was encoded by CR1 ORF II, which is responsible for the replication of the element. Propagation of the CR1 retrotransposon is likely via the mechanism of nick and prime; the reverse transcriptase recognized its mRNA and used it as a template for reverse transcription, priming from a nick in the chromosomal DNA (Luan et al. 1993). About 96 000 CR1 elements have been estimated to exist within the chicken genome (Wicker et al. 2005). The majority of elements have been found to concentrate at the G+C-rich regions (Olofsson and Bernardi 1983*a*). Since the G+C-rich region is also the region where most genes are concentrated, many CR1 repeats have been found to flank multiple genes. For example, the first CR1 element was found to flank the chicken *U1* RNA gene (Stumph et al. 1981). A cluster of CR1 repeats constitutes almost 16% of the chicken β-globin gene cluster (Reitman et al. 1993). Over 19 CR1 elements have been found to flank both the upstream and downstream regions of the chicken ovalbumin gene (Stumph et al. 1983). However, no CR1 element was reported to be located within the satellite region.

In this study, we will report a novel tandem repeat (the *Hin*fI repeat) identified in the chicken genome. Sequence analysis and Southern blot results suggest that this repeat may have evolved from a unique region containing a retrotransposed CR1 element, which provides experimental evidence for the model proposed by Smith (1976).

Materials and methods

Identification of the *Hin*fI repeat

To isolate the repetitive elements in the chicken genome, a chicken subgenomic library was constructed. Chicken genomic DNAs were completely digested by the restriction enzyme *Hae*III and subjected for electrophoresis in 0.8% *w*/*v* agarose gel. DNA fragments between 500 and 2000 bp were gel purified and ligated into a pBluescriptII SK (+) vector (Stratagene, La Jolla, Calif.). To screen for clones containing repetitive DNAs, chicken genomic DNAs were labeled randomly with digoxigenin using a Random Primed DNA labeling kit (Roche Diagnostics, Basel, Switzerland) according to the manufacturer's instructions. On the basis that DNAs with a higher copy number tend to give stronger hybridization signals, the clones with strong positive signals were considered potential clones containing repetitive DNAs. To confirm the highly repetitive nature of those strong positive clones, they were labeled with digoxigenin and hybridized with chicken genomic DNAs digested with different restriction enzymes. A clone of particular interest, Wp-6, was later confirmed to represent a novel repetitive family designated the *Hin*fI repeat.

Cloning and sequencing of multiple repeating units of the *Hin*fI repeat

To obtain the full length of the *Hin*fI repeating unit, a commercial chicken genomic library (adult Leghorn, male liver (EMBLSP6/T7 cloning vector, *Escherichia coli* K802 host strain)) purchased from Clontech (Mountain View, Calif.) was screened using the digoxigenin-labeled clone Wp-6 according to the manufacturer's instructions. Five positive λ phage clones were identified after a tertiary screening of 30 000 clones. The large genomic inserts (~8–15 kb) of the positive clones were released with restriction enzymes *Sst*I and *Bam*HI and further digested into smaller fragments (<10 kb). The genomic fragments containing multiple repeating units of the *Hin*fI repeat were identified by Southern blot analysis, gel purified, and subcloned into a pBluescriptII SK (+) vector. The positive clones were then sequenced in both directions with T3 and T7 sequencing primers on an ABI PRISM 310 genetic analyzer using the BigDye Terminator v. 3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, Calif.) according to the manufacturer's instructions.

Synthesis of digoxigenin-labeled DNA probes and Southern blot analysis

All probes used in the present study were digoxigenin-labeled using a PCR DIG Probe Synthesis Kit (Roche Diagnostics) according to the manufacturer's instructions. Briefly, the digoxigenin-11-dUTP concentration and extension time were adjusted after calculating the template GC value and insert length. Primers and free nucleotides were removed by Microcon 100 column filtration (Millipore, Amicon, Billerica, Mass.). The probe concentration of the Southern hybridization buffer was 30 ng/mL. Double-stranded DNA probes were denatured at 95 °C for 10 min before hybridization.

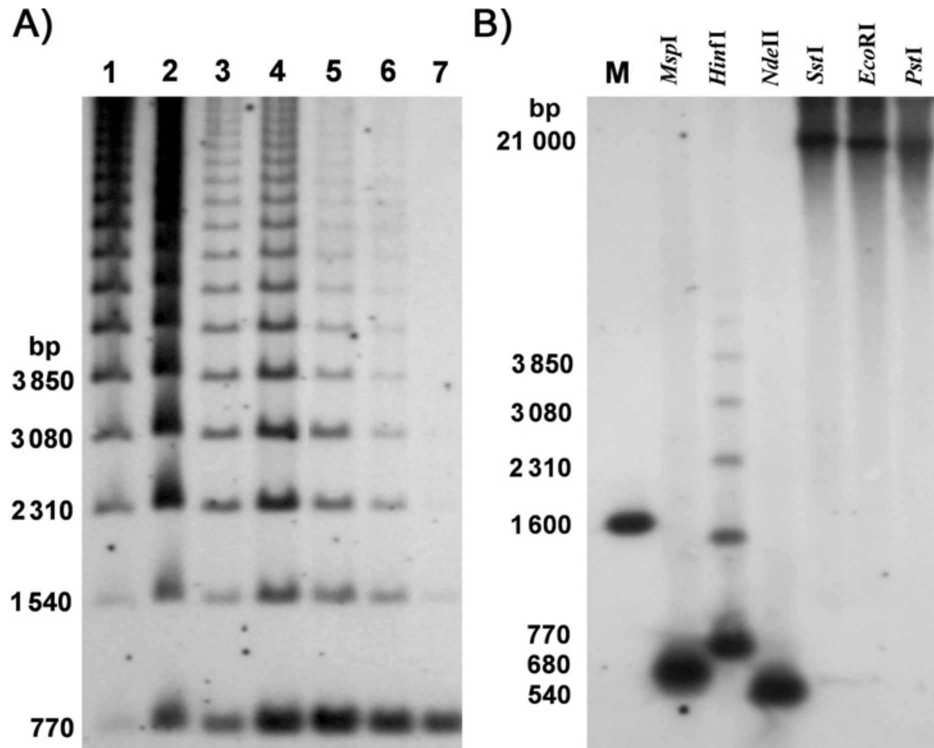
For Southern hybridization, genomic DNAs were digested with various restriction enzymes as indicated and subjected to agarose gel electrophoresis. After alkaline denaturation, DNAs were vacuum blotted and UV cross linked onto a positively charged nylon membrane (Roche Diagnostics). Hybridizations were carried out in 5× standard saline citrate (SSC), 0.1% SDS, and 1% blocking reagent (Roche Diagnostics) and washed under high stringency (0.5× SSC and 0.1% SDS at 68 °C). The digoxigenin-labeled probe was detected by alkaline phosphatase-conjugated anti-digoxigenin and CDP-star (Roche Diagnostics).

Determination of the *Hin*fI repeat copy number in chicken genome

Slot-blot analysis was performed to determine the genomic abundance of the *Hin*fI repeat within the chicken genome according to the previously described method of Wang et al. (2002). Briefly, a series of diluted chicken genomic DNAs (1000 ng, 500 ng, 250 ng, 125 ng, 62.5 ng) and *Hin*fI monomer units (20.4 ng, 10.2 ng, 5.1 ng, 2.55 ng, 1.275 ng) were slotted on the positively charged nylon membrane (Roche Diagnostics). The digoxigenin-labeled probe was

© 2006 NRC Canada

Fig. 1. Southern blot analyses of the *Hin*fI repeat in chicken genomic DNA. Hybridization probe is Wp-6 labeled with digoxigenin. M indicates the marker lane with the size shown in bp. (A) Chicken genomic DNAs were partially digested by *Hin*fI through control of time interval and run in 1.0% *w*/*v* agarose gel. Lane 1, 1 min; lane 2, 2 mins; lane 3, 3 mins; lane 4, 5 mins; lane 6, 10 mins; lane 7, 30 mins. The ladder pattern reveals tandem arrangement of the repeating units. (B) Chicken genomic DNAs were digested with restriction enzyme as indicated and run in 1.0% *w*/*v* agarose gel. Pronounced signals can be observed.

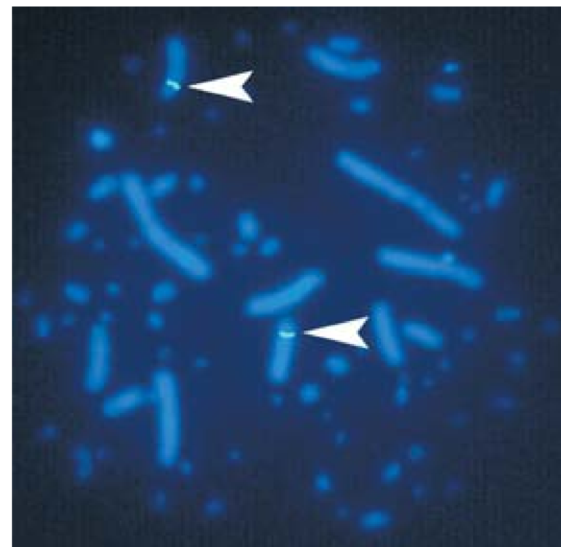


amplified from the non-CR1 region using primers *Hin*fI-F (5′-GCACGACGTACGACGTCT-3′) and *Hin*fI-R (5′-AGTGGGAGAAGAAGTTTT-3′). Hybridization was performed under the same conditions as Southern blot analysis. The intensity of hybridization signals was measured by densitometry. The relative amounts of *Hin*fI repeat within the chicken genome were determined using the calibration curve based on the signals of diluted *Hin*fI monomer units.

Fluorescence in situ hybridization (FISH)

Fluorescence in situ hybridization was performed according to the method previously described (Wang et al. 2002). Briefly, chicken metaphase chromosome spreads were prepared from chicken lymphocyte cell line DT-40 following the standard cytogenetic procedures. After treatment with proteinase K and RNase A, the metaphase chromosomes were denatured in denaturation buffer (70% formamide, 2× SSC, 40 mmol/L phosphate buffer, pH 7.2) at 68 °C for exactly 2 min. The slides were then dehydrated and air dried. The digoxigenin-labeled probe was amplified from the non-CR1 region using primers *Hin*fI-F and Wp-R (5′-AGAAGATGTCCAACGTCGG-3′). Fifteen microlitres of probe (2 ng/μL) in hybridization mixture (50% formamide, 2× SSC, 500 μg salmon sperm DNA/mL, and 10% *w*/*v* dextran sulfate) were applied to the denatured chromosome for overnight hybridization in a 37 °C humid chamber. The slides were washed under stringent conditions (50% formamide, 2× SSC, 42 °C). About 25 μL fluorescein-conjugated anti-digoxigenin antibody (Roche Diagnostics) were applied and incubated at 37 °C for 30 min. After the final wash (PBS, pH 7.2), the slides were counterstained with 4′,6-diamidino-2-phenylindole (DAPI) and mounted with Vectashild (Vector Laboratories, Burlingame, Calif.). The slides were viewed under a Zeiss epifluorescent microscope (Carl Zeiss

Fig. 2. Localization of the *Hin*fI repeat on chicken mitotic chromosomes by FISH. The metaphase chromosomes were prepared from a chicken DT-40 cell line. The chromosomes were counterstained with DAPI and hybridization signals were indicated by arrowhead.



© 2006 NRC Canada

Fig. 3. Consensus nucleotide sequence of the *Hin*fI repeating unit. *Hin*fI sites flanking the repeating unit (upperlined) are indicated. The satellite CR1 region is double underlined. Forward (Wp-F, *Hin*fI-F) and reverse (Wp-R, *Hin*fI-R) PCR primers are indicated.

```
HinfI
GATTCTATGC TTCTCTCCAA AAGCACGACG TACGACGTCT TCTACGCAAA
                           HinfI-F
TAAACACGTT TGTCTCCGAG AAACATGGCA TTTTCAGCCA CAGAATACTG

TTCTCCTCCC AAACCACACG ATAGTTTCTC CCGCAACGKA CCACTTTCCT

CCCCAAAATA CAAACACGGC TTTCCCAGGA GACCACTTTT CTCCCCGACG
                            NdeII
TTGGACATCT TCTCCCCAAA ATCGGATCTT TCTCCCCACA AAATACCACT
   Wp-R
TCTCTACCCC AAATCCCATG CTTTCTCCCA CAAAAGACCA CTTCTCTCCA
                      Wp-F
CAGAATATCT GCTTTCTCCC CCACACGGAC AAATTCACCC CCCCAAGTAC

CACATTTTCC CCAGCAAAAT ACCACCTTTC TCCTGGGCGT AAGGTATTTC

CACCCCAAAA TACCACGGTT CTCCCCCAGA GAAAACTTCT TCTCCCACTT
                                      HinfI-R
AATACCACTT CCAGGACCTG AAGGGAGCAT CTAATCAGGA CGGAGAACGG

TTGCTTACGG GGCTGGAACT GAGGCAGGGG GCGTTCAGGT GAGRGAGTAC
                                        MspI
GARGACGTTT TTCACACCGA GGGGGGTGAC ACACTGGAAC CGGCTGCCCA

AGGAGGTTGG GGATGCCTCA TCCCTGGAGG CATTGGAGCC CAGGCTGGGT
                 MspI
GTGGCTCTGC GCAGCCCGGT CTACTGGTTG GCGACCCTGT GCAACTTAGC

CCGGCGGCGT TGAAACTCCG TGGTCCTTGA GGTCCTTTTC AACCCAGGCC

ATTCTGTCTT TCTGTGATCG GATTC
                NdeII HinfI
```

Inc., Germany). Images were captured and analyzed with CytoVision software (Applied Imaging Group, Rochester, N.Y.).

Results

Identification and characterization of the *Hin*fI repeat

A genomic DNA clone, designated as Wp-6 (GenBank accession No. AY491727), was identified as containing repetitive elements and sequencing analysis revealed that Wp-6 may represent a novel repetitive family. To investigate the repetitive nature of this novel repeat, Wp-6 was labeled with digoxigenin and hybridized to the digested genomic DNAs.

As shown in Fig. 1B, hybridization signals were detected at positions of 680 bp, 770 bp, and 540 bp when chicken genomic DNAs were digested by *Msp*I, *Hin*fI, and *Nde*II, respectively. Hybridization signals at the large molecular size were detected when DNAs were digested by *Sst*I, *Eco*RI, and *Pst*I, indicating the long extending pattern of this repetitive family. Its internal structural arrangement was further investigated by *Hin*fI partial digestion. Mild digestion by *Hin*fI produced a distribution of bands displaying a typical ladder structure (Fig. 1A). As many as 20 bands were identified in tandem and fragment length steadily reached the monomer length of 770 bp. Owing to the recognition site of *Hin*fI, this novel tandem repeat was hence named the *Hin*fI repeat.

The genomic abundance of the *Hin*fI repeat was estimated by slot-blot analysis (data not shown). Densitometric measurements indicate that the *Hin*fI repeat constitutes approximately 0.51% of the chicken genome. Assuming that the haploid chicken genome is 1.2×10^9 bp (Olofsson and Bernardi 1983*b*), 8000 copies of the *Hin*fI repeating unit were estimated to exist in the chicken genome.

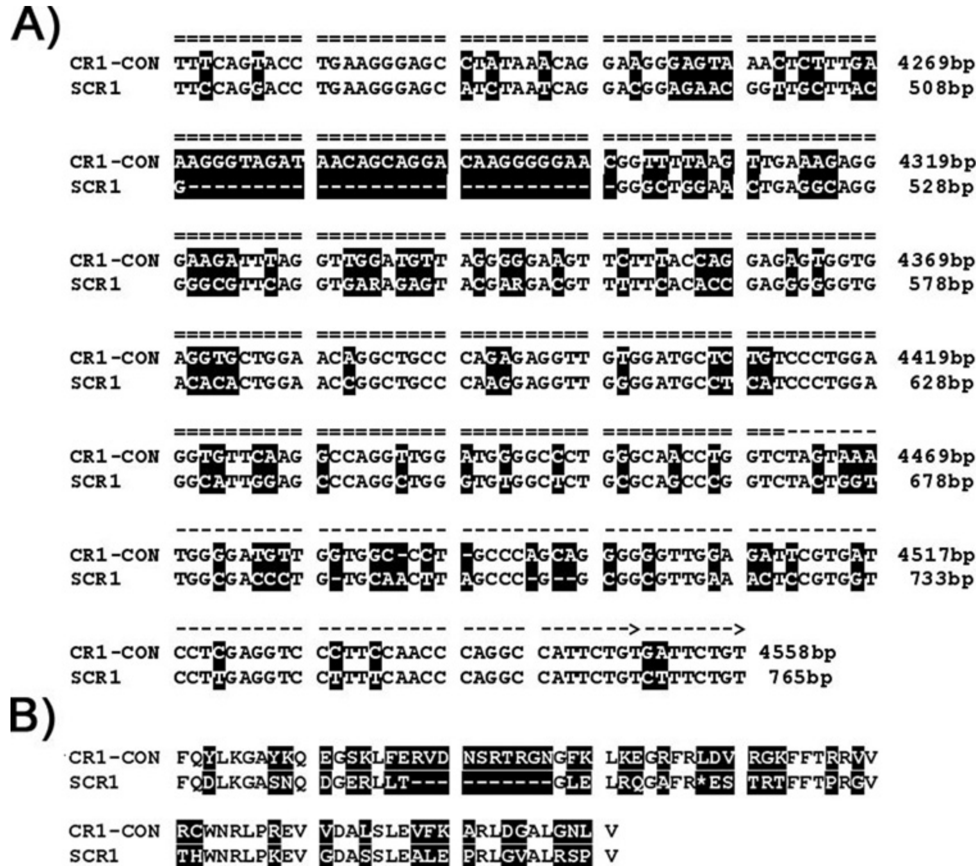
The localization of this *Hin*fI repeat was determined by in situ hybridization on chicken metaphase chromosomes. More than 30 metaphase cells were examined. Fluorescein signals were specifically detected in the pericentric regions of chromosome 4 (Fig. 2). No signal was detected at other locations. This specific chromosomal localization was further confirmed by performing a BLAST search in the chicken genome database (http://pre.ensembl.org/). The monomer unit sequence shares high sequence similarity (>98%) with one contig, 391.12, which is 1454 bp in length and contains one whole *Hin*fI monomer unit and one incomplete monomer unit. In line with our in situ hybridization result, this contig has been located on chromosome 4.

Sequence analysis of the *Hin*fI repeating unit

To characterize the basic repeating unit of the *Hin*fI repeat, a λ phage chicken genomic library was screened. Five

© 2006 NRC Canada

Fig. 4. Alignment of SCR1 and CR1 retrotransposons. (A) Aligned nucleotide sequence data between SCR1 and consensus CR1 retrotransposon (partial). Numbering refers to *Hin*fI monomer unit position (SCR1) and CR1 retrotransposon consensus sequence position (CR1-CON) (Haas et al. 1997) (U88211). The SCR1 element is located within the 3′ terminal region of the CR1 retrotransposon, including a partial coding region of CR1 ORF II (double upperlined) and 3′ UTR region (single upperlined). The characteristic terminal 8 bp direct repeats are indicated by an arrowhead. (B) Aligned amino acid sequence data deduced from SCR1 coding region and the terminal region (910 a.a. – 990 a.a.) of CR1 ORF II. Dashes indicate where gaps were introduced to optimize alignments. Non-identical nucleotides and amino acids were shaded in black.



clones containing multiple *Hin*fI repeating units were identified after screening of over 30 000 phage clones. The regions of interest were subcloned and partially sequenced (GenBank accession Nos. AY491728–AY491735).

A total of 18 repeating units were obtained with more than 95% sequence identity. A consensus sequence of 770 bp from sequence alignment was identified (Fig. 3), which agreed with the estimation from Southern blot analysis. Two *Msp*I restriction sites have been identified at positions of 590 bp and 666 bp, generating a band of 700 bp and consistent with Southern blot analysis. Similarly, 2 *Nde*II restriction sites were located at 224 bp and 765 bp, generating the bands of 540 bp and 230 bp. Since hybridization probe Wp-6 was located within the 540 bp region, the band at 230 bp was not detected in Southern blot analysis (Fig. 1B). In addition, the recognition sites of *Sst*I, *Pst*I, and *Eco*RI were not found within the repeating unit.

A striking sequence similarity was observed among the regions of the *Hin*fI repeating unit (459–765 bp) and CR1 element. This CR1-like element is 307 bp in length and contains a partial coding region of the CR1 ORF II and the complete 3′ untranslated region (UTR) of the CR1 retrotransposon (Fig. 4). Deduced amino acid sequence within the coding region shares 51% sequence identity with the terminal region (910–990 amino acids) of the CR1 retrotransposon ORF II (990 amino acids) (Haas et al. 1997). As shown in Fig. 4, this CR1-like element also shares high sequence similarity (73%) with consensus CR1 retrotransposon within its 3′ UTR region. The characteristic 8 bp direct repeats ((CATTCTGT)(CTTTTCTGT)) can be identified, revealing the hallmark of the CR1 retrotransposon (Silva and Burch 1989). We therefore named this CR1-like element identified within the *Hin*fI repeating unit the satellite CR1 (SCR1) element.

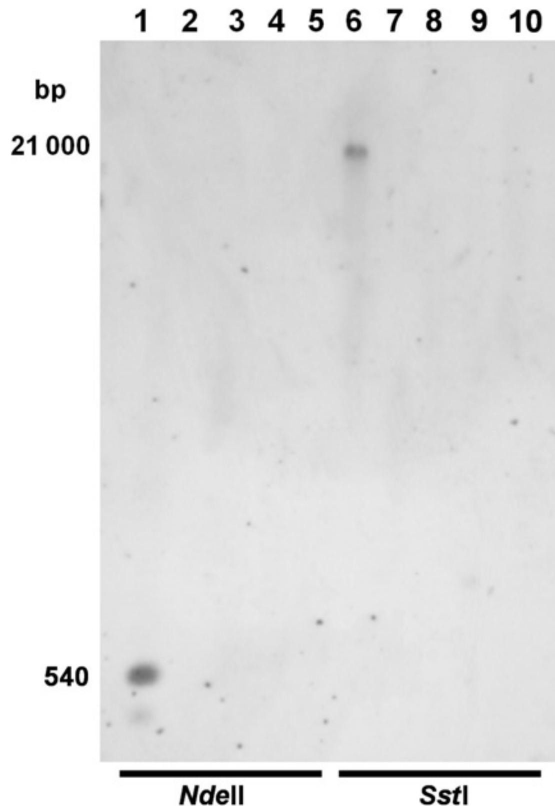
The remaining 463 bp of the *Hin*fI repeat is defined as the non-CR1 region. Sharing low sequence similarity with the CR1 element, this region is 50% G+C rich and contains no internal substructure. Within the non-CR1 region, no specific sequence has been found to flank the satellite CR1 element. The duplication of the CR1 retrotransposon integration site is also not detectable within this region (Silva and Burch 1989).

Species distribution of the *Hin*fI repeat

To investigate the species distribution of the *Hin*fI repeat, 5 avian species (Chinese Francolin (*Francolinus pintadeanus*), Pigeon (*Columba livia*), Pheasant (*Phusianus colchicus*),

© 2006 NRC Canada

Fig. 5. Species distribution of the *Hin*fI repeat by Southern blot analysis. Genomic DNAs (1.5 μg/lane) from White Leghorn chicken (lanes 1 and 6), Chinese Francolin (lanes 2 and 7), turkey (lanes 3 and 8), ring necked pheasant (lanes 4 and 9), and pigeon (lane 5 and 10) were digested by *Nde*II and *Sst*I and run on 0.8% *w*/*v* agarose gels. The hybridization probe was amplified from the non-CR1 region. Comigration of the size ladder is marked.

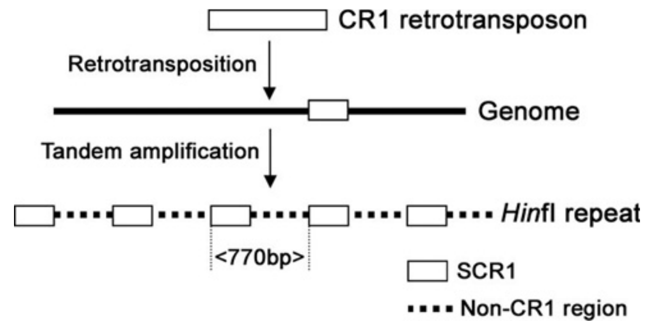


Turkey (*Melegris gallopavo*), and White Leghorn chicken (*Gallus gallus*)) were included in this study. Owing to the wide distribution of the CR1 retrotransposon among avian species (Chen et al. 1991), the probe amplified from the non-CR1 region was investigated. As shown in Fig. 5, hybridization signals were only detected in the White Leghorn chickens, indicating the species-specific distribution of the *Hin*fI repeat. Owing to the great variety of breeds and populations of chicken, we further examined the existence of the *Hin*fI repeat within another chicken breed (Shekki chicken) and the red jungle fowl, the assumed progenitor of all domestic breeds (Delany 2004). Similar results (data not shown) were obtained that confirmed the *Hin*fI repeat as a repetitive family specific to *Gallus gallus*.

Discussion

In the present study, a novel highly repetitive family, the *Hin*fI repeat, was characterized from the chicken genome. Its basic repeating unit is 770 bp in length and sequence analysis reveals that a CR1 element is located within the repeating unit. This CR1 element is 307 bp in length, belonging to the category of short CR1 elements (smaller than 400 bp), which is also the main category (>98%) existing within the chicken genome (Vandergon and Reitman 1994). Such short length is due to the special retrotransposition mechanism employed by the CR1 retrotransposon in which only the 3′ fraction of the retrotransposon RNA is inserted in the chromosomal DNA, resulting in truncated 5′ ends (Luan et al. 1993). This typical 5′ truncation character is also shared by the SCR1 element. Starting from the 3′ end, it only shares the extended 5′ length up to the terminal region of CR1 ORF II (Burch et al. 1993; Haas et al. 1997). Unlike most LINEs, the CR1 retrotransposon does not contain the common 3′ poly(A) tail, but instead ends in 8 bp repeats arranged in a direct pattern (Silva and Burch 1989). Such a highly homologous 3′ UTR is especially important for efficient retrotransposition of the CR1 retrotransposon (Haas et al. 1997). The sharing of an 8 bp direct repeat by this SCR1 element strongly suggests it is the progeny of a CR1 retrotransposon.

Fig. 6. Proposed evolution model of the *Hin*fI repeat.



Fluorescein signals were specifically detected on the pericentric region of chromosome 4, revealing that this SCR1 element is located in a region generally believed to be A+T rich (Miklos 1985). The present finding is the first report showing a progeny of CR1 retrotransposon localizing to the pericentric region. Hence, retrotransposition of the CR1 retrotransposon within the chicken genome is much wider than we expected. The estimation of 8000 SCR1 copies by slot-blot analysis reveals that these elements may constitute 8.3% of 96 000 CR1 elements within the chicken genome (Wicker et al. 2005). This finding broadens our knowledge of CR1 distribution. In addition to propagating in a reverse trasncriptase dependent manner and dispersing randomly flanking multiple genes (Reitman et al. 1993; Stumph et al. 1981; Stumph et al. 1983), the CR1 elements may also accumulate within satellite regions owing to the amplification of *Hin*fI monomer units.

Identification of a retrotransposed CR1 element located within the basic repeating unit of a *Hin*fI repeat points us to an interesting issue on the evolutionary origin of the *Hin*fI repeat, i.e., whether the 770 bp monomer repeat evolves from a unique region containing the retrotransposed CR1 element or the CR1 element is just an occasional insertion into the amplified monomer region. In present study, 2 probes were employed to investigate the evolutionary origin. Southern blot revealed that smear signals were detected in species related to chicken, such as the Ring Necked Pheasant, Chinese Francolin, and Turkey, when the whole repeating unit was used as the hybridization probe (data not shown). However, pronounced signals were detected only in White Leghorn chicken and red jungle fowl when the same blot was stripped and rehybridized with the probe amplified from the

© 2006 NRC Canada

non-CR1 region (Fig. 5). These results suggest that the *Hin*fI repeat probably evolves from a unique region containing a retrotransposed CR1 element during the *Gallus gallus* speciation (proposed mode, Fig. 6). The retrotransposed element is located within a basic repeating unit of satellite DNA sequences, and a similar case has been previously described in the human genome (Frommer et al. 1984). An *Alu* element (300 bp) was found to be located within a 2.47 kb repeat. Since the non-*Alu* region in the basic repeating unit has also been detected in gorillas and chimpanzees (Cooke et al. 1982), the 2.47 kb repeat is unlikely to have evolved directly from a unique region containing the retrotransposed *Alu* element. The insertion of an *Alu* element is a later incident compared with the amplification of the 2.47 kb repeat (Frommer et al. 1984).

Present findings support the model proposed by Smith (1976), which predicts that satellite DNAs may expand from a wide spectrum of non-satellite DNAs. The retrotransposed elements may occasionally be fixed and further amplified into satellite DNAs owing to continual unequal crossing over. The supporting evidence has been elucidated from *Drosophila* (Heikkinen et al. 1995) and mammals (Kapitonov et al. 1998). The present finding that the *Hin*fI repeat may evolve from a unique region containing a retrotransposed CR1 element may serve as one more piece of empirical evidence from the chicken genome.

References

Burch, J.B., Davis, D.L., and Haas, N.B. 1993. Chicken repeat 1 elements contain a *pol*-like open reading frame and belong to the non-long terminal repeat class of retrotransposons. Proc. Natl. Acad. Sci. U.S.A. **90**: 8199–8203.

Charlesworth, B., Sniegowski, P., and Stephan, W. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. Nature (London), **371**: 215–220.

Chen, Z.Q., Ritzel, R.G., Lin, C.C., and Hodgetts, R.B. 1991. Sequence conservation in avian CR1: an interspersed repetitive DNA family evolving under functional constraints. Proc. Natl. Acad. Sci. U.S.A. **88**: 5814–5818.

Cooke, H.J., Schmidtke, J., and Gosden, J.R. 1982. Characterisation of a human Y chromosome repeated sequence and related sequences in higher primates. Chromosoma (Berlin), **87**: 491–502.

Delany, M.E. 2004. Genetic variants for chick biology research: from breeds to mutants. Mech. Dev. **121**: 1169–1177.

Frommer, M., Prosser, J., and Vincent, P.C. 1984. Human satellite I sequences include a male specific 2.47 kb tandemly repeated unit containing one Alu family member per repeat. Nucleic Acids Res. **12**: 2887–2900.

Fry, K., and Salser, W. 1977. Nucleotide sequences of HS-alpha satellite DNA from kangaroo rat Dipodomys ordii and characterization of similar sequences in other rodents. Cell, **12**: 1069–1084.

Haas, N.B., Grabowski, J.M., Sivitz, A.B., and Burch, J.B. 1997. Chicken repeat 1 (CR1) elements, which define an ancient family of vertebrate non-LTR retrotransposons, contain two closely spaced open reading frames. Gene, **197**: 305–309.

Heikkinen, E., Launonen, V., Muller, E., and Bachmann, L. 1995. The pvB370 *Bam*HI satellite DNA family of the *Drosophila virilis* group and its evolutionary relation to mobile dispersed genetic *pDv* elements. J. Mol. Evol. **41**: 604–614.

Kapitonov, V.V., Holmquist, G.P., and Jurka, J. 1998. L1 repeat is a basic unit of heterochromatin satellites in cetaceans. Mol. Biol. Evol. **15**: 611–612.

Luan, D.D., Korman, M.H., Jakubczak, J.L., and Eickbush, T.H. 1993. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell, **72**: 595–605.

Mestrovic, N., Plohl, M., Mravinac, B., and Ugarkovic, D. 1998. Evolution of satellite DNAs from the genus *Palorus* — experimental evidence for the “library” hypothesis. Mol. Biol. Evol. **15**: 1062–1068.

Miklos, G.L.G. 1985. Localized highly repetitive DNA sequences in vertebrate and invertebrate genomes. *In* Molecular evolutionary genetics. *Edited by* R.J. MacIntyre. Plenum Press, New York, N.Y. pp. 241–321.

Olofsson, B., and Bernardi, G. 1983*a*. The distribution of CR1, and Alu-like family of interspersed repeats, in the chicken genome. Biochim. Biophys. Acta, **740**: 339–341.

Olofsson, B., and Bernardi, G. 1983*b*. Organization of nucleotide sequences in the chicken genome. Eur. J. Biochem. **130**: 241–245.

Pech, M., Streeck, R.E., and Zachau, H.G. 1979. Patchwork structure of a bovine satellite DNA. Cell, **18**: 883–893.

Pons, J., Bruvo, B., Petitpierre, E., Plohl, M., Ugarkovic, D., and Juan, C. 2004. Complex structural features of satellite DNA sequences in the genus *Pimelia* (Coleoptera: Tenebrionidae): random differential amplification from a common “satellite DNA library”. Heredity, **92**: 418–427.

Reitman, M., Grasso, J.A., Blumenthal, R., and Lewit, P. 1993. Primary sequence, evolution, and repetitive elements of the *Gallus gallus* (chicken) beta-globin cluster. Genomics, **18**: 616–626.

Rogers, J.H. 1985. The origin and evolution of retroposons. Int. Rev. Cytol. **93**: 187–279.

Silva, R., and Burch, J.B. 1989. Evidence that chicken CR1 elements represent a novel family of retroposons. Mol. Cell Biol. **9**: 3563–3566.

Smith, G.P. 1976. Evolution of repeated DNA sequences by unequal crossover. Science (Washington, D.C.), **191**: 528–535.

Stumph, W.E., Kristo, P., Tsai, M.J., and O’Malley, B.W. 1981. A chicken middle-repetitive DNA sequence which shares homology with mammalian ubiquitous repeats. Nucleic Acids Res. **9**: 5383–5397.

Stumph, W.E., Baez, M., Lawson, G.M., Tsai, M.J., and O’Malley, B.W. 1983. Higher-order structural determinants for expression of the ovalbumin gene family. Ciba Found. Symp. **98**: 80–95.

Vandergon, T.L., and Reitman, M. 1994. Evolution of chicken repeat 1 (CR1) elements: evidence for ancient subfamilies and multiple progenitors. Mol. Biol. Evol. **11**: 886–898.

Wang, X., Li, J., and Leung, F.C. 2002. Partially inverted tandem repeat isolated from pericentric region of chicken chromosome 8. Chromosome Res. **10**: 73–82.

Wicker, T., Robertson, J.S., Schulze, S.R., Feltus, F.A., Magrini, V., Morrison, J.A., Mardis, E.R., Wilson, R.K., Peterson, D.G., Paterson, A.H., and Ivarie, R. 2005. The repetitive landscape of the chicken genome. Genome Res. **15**: 126–136.

© 2006 NRC Canada