Original Article

Cytogenetic and Genome Research

Cytogenet Genome Res
DOI: 10.1159/000314923

Accepted: January 13, 2010
by J. Smith
Published online: July 3, 2010

Precise Centromere Positioning on Chicken Chromosome 3

A. Zlotina[a] S. Galkina[a] A. Krasikova[a] R.P.M.A. Crooijmans[b] M.A.M. Groenen[b]
E. Gaginskaya[a] S. Deryusheva[a]

[a] Saint-Petersburg State University, Saint-Petersburg, Russia; [b] Animal Breeding and Genomics Centre, Wageningen University, Wageningen, The Netherlands

Key Words
Centromere positioning · Chicken · CNM · FISH · GGA3 · Immunostaining · Lampbrush chromosomes

Abstract
Despite the progress of the chicken *(Gallus gallus)* genome sequencing project, the centromeric sequences of most macrochromosomes remain unknown. This makes it difficult to determine centromere positions in the genome sequence assembly. Using giant lampbrush chromosomes from growing oocytes, we analyzed in detail the pericentromeric region of chicken chromosome 3. Without knowing the DNA sequence, the centromeres at the lampbrush stage are detectable by immunostaining with antibodies against cohesin subunits. Immunostaining for cohesin followed by FISH with 23 BAC clones, covering the region from 0 to 23 Mb on chicken chromosome 3 (GGA3), allowed us to map the GGA3 centromere between BAC clones WAG38P15 and WAG54M22 located at position 2.3 and 2.5 Mb, respectively. This corresponds to the gap between 2 supercontigs at the 2.4-Mb position in the current GGA3 sequence assembly (build 2.1). Furthermore, we have determined that the current putative centromeric gap at position 11.6–13.1 Mb corresponds in fact to a long cluster of tandem chicken erythrocyte nuclear membrane repeats (CNM).

Genome sequencing of the red jungle fowl *(Gallus gallus)* was completed at 6.6× coverage in 2004 [International Chicken Genome Sequencing Consortium, 2004] followed by a new draft genome assembly (Gallus_gallus-2.1) in 2006 (http://www.ncbi.nlm.nih.gov). Although this new build improved the genome sequence assembly, the supercontigs on which this genome sequence is based are still interrupted by numerous gaps. Furthermore, many contigs still remain unanchored to specific chromosomes and instead are combined on chromosome 'ChrUn'. On the current assembly, centromeres are indicated by 1.5-Mb gaps on macrochromosomes and by 0.5-Mb gaps on microchromosomes.

Notwithstanding many attempts to clone centromeric repeats from the chicken genome, only 4 repetitive sequences have been described so far as centromere associated sequences. These repeats are the chicken erythrocyte nuclear membrane (CNM) repeat found in pericentromeric regions of microchromosomes and macrochromosomes 6, 9 and W [Matzke et al., 1990; Wang et al., 2002; Krasikova et al., 2006], the *Xho*I repeat near the centromeric region of chromosome W [Solovei et al., 1998; Krasikova et al., 2006], the *Hinf*I repeat concentrated in the pericentromeric region of chromosome 4 [Li and Leung, 2006], and the partially inverted tandem repeat (PIR) specific for the centromere of chromosome

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

1424–8581/10/0000–0000$26.00/0

Accessible online at:
www.karger.com/cgr

Svetlana Galkina
Saint-Petersburg State University
Oranienbaumskoe sh., 2
Saint-Petersburg, 198504 (Russia)
Tel. +7 812 450 7311, Fax +7 812 450 7310, E-Mail spbchromas@gmail.com

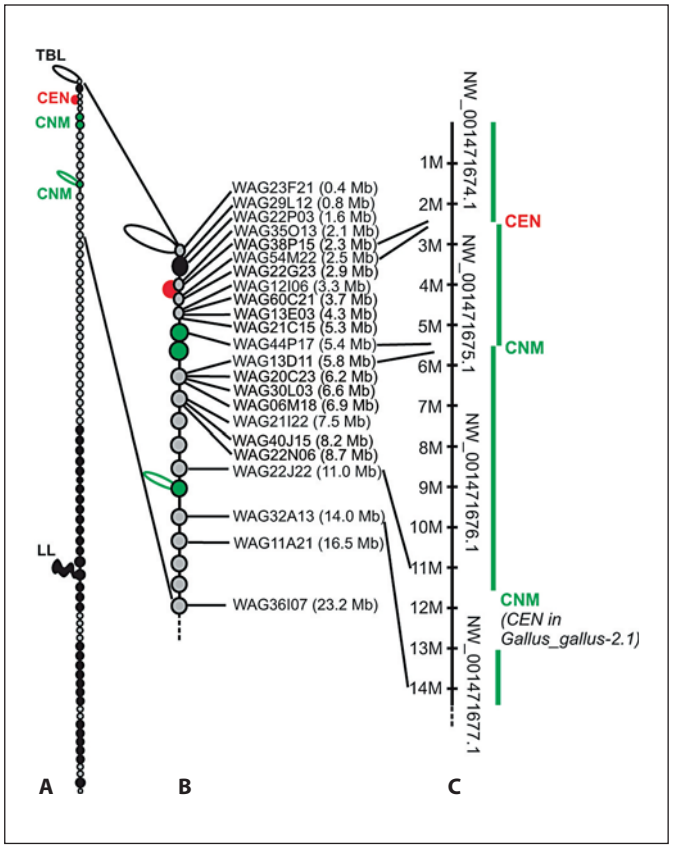

Fig. 1. Localization of BAC clones on chicken chromosome 3. **A** Cytological map of chicken LBC3. The red circle indicates the centromere-specific cohesin-enriched structure (CEN). Chromomeres and loops bearing CNM repeats are shown in green. Chromosome landmarks: LL, lumpy loop; TBL, telomere bow-like loop. **B** Enlarged region with mapped BAC clones. Positions of BAC-linked markers in the sequence assembly of GGA3 are indicated. **C** Scheme of contigs for the 0–14-Mb region of GGA3 (retrieved from http://www.ncbi.nih.gov). The position of the centromere (CEN) and CNM repeat arrays are indicated.

8 [Wang et al., 2002; Krasikova et al., 2006]. Even in the case of known centromeric repeats these might not represent reliable centromeric markers. In fact, pericentromeric repeat CNM has been shown to be concentrated both at centromeric regions of microchromosomes and at subtelomeric regions of the long arms of the same chromosomes [Krasikova et al., 2006; Deryusheva et al., 2007].

Recently, we have shown that centromeres are marked by cohesin-enriched structures on chromosomes at the lampbrush stage [Krasikova et al., 2006]. Lampbrush chromosomes (LBCs) are highly extended bivalents that can be found in the growing oocytes of many animals including birds [reviewed in Gaginskaya et al., 2009]. Being extremely long, more than 30 times longer than the corresponding mitotic metaphase chromosomes, LBCs represent a very promising system for high resolution physical gene mapping [Galkina et al., 2006]. In our previous studies we took advantage of using LBCs to map centromeres in chicken and Japanese quail [Krasikova et al., 2006; Deryusheva et al., 2007]. We showed that in most cases predicted centromeric regions in the chicken genome sequence assembly perfectly match with the position of centromeric cohesin-enriched structures [Krasikova et al., 2006] except for chromosome 3. The position of the centromere detected by immunostaining with antibodies against cohesin subunits was very close to the loop bearing BAC clone WAG29L12 (0.8 Mb; see fig. 1), which does not correspond to the gap at position 11.6–13.1 Mb annotated as the centromere on GGA3. We assume that the annotated centromeric gap in the current GGA3 sequence assembly is a non-centromeric cluster of CNM repeats mapped on chicken chromosome 3 [Krasikova et al., 2006]. The question which gap in the sequence assembly of GGA3 actually represents the centromeric DNA repeat therefore still remained open. Here we focus on a detailed analysis of the pericentromeric region of GGA3 to answer this question.

Materials and Methods

Twenty-three chicken BAC-clones from the Wageningen chicken BAC library [Crooijmans et al., 2000; http://www.bioinformatics.nl/gbrowse/cgi-bin/gbrowse/] were selected from the 0–23-Mb region of GGA3 for FISH mapping on chicken LBC3. The BACs cover the region of the first 9 Mb at 0.1–0.5-Mb intervals, and flank in the closest proximity the annotated centromeric gap at position 11.6–13.1 Mb and 2 other gaps in the region on GGA3 (fig. 1).

Chicken LBCs were isolated manually from growing oocytes of 0.5–1.5 mm in diameter according to standard protocols. The detailed description of LBC preparation is available at http://www.exeter.ac.uk/projects/lampbrush/. FISH with BAC clones on LBCs was performed as described [Galkina et al., 2006]. Oligonucleotide probes specific for CNM repeats [Krasikova et al., 2006; Deryusheva et al., 2007] were additionally used to assign the clusters of this repeat on chicken chromosome 3 to the specific positions in the GGA3 sequence assembly. Before FISH some preparations were stained with an antibody against STAG2, one of the cohesin subunits, as previously described [Krasikova et al., 2006] to map the centromere on LBC3.

Results and Discussion

The combination of immunostaining for centromeric cohesin-enriched structures with high-resolution multi-color FISH on chicken LBCs allowed us to narrow down the position of the centromere on chromosome 3 to a region between BACs WAG38P15 and WAG54M22 (figs. 1 and 2A). These BACs are assigned to position 2.3 and 2.5 Mb, respectively. This indicates that the elusive GGA3 centromeric repeats should be located in a 0.2-Mb region between these 2 BACs. In the current GGA3 sequence assembly, a gap does exist in this region at position 2.4 Mb which we argue is the location of the centromere.

When the actual position of the centromere is located at 2.4 Mb, the question arises: what sequences are located at the current putative centromeric gap at position 11.6–13.1 Mb and the third gap in the region of interest, the gap at position 5.6 Mb? Gaps in the sequence assembly are usually the result of a sequencing break at tandem repeat clusters. In the sequence assembly of GGA3, a few units of CNM repeat can be found near position 11.5 Mb. Earlier we have detected 2 large blocks of this repeat on LBC3 by FISH with CNM-specific probes [Krasikova et al., 2006]. To verify whether the 2 CNM arrays detectable by FISH are located at the positions corresponding to these 2 gaps we performed FISH with CNM-specific oligoprobes followed by repeated FISH with BAC clones assigned to genomic positions closely surrounding the gaps. Proximal CNM-positive FISH signals are mapped at the 2 chromomeres between BACs WAG44P17 and WAG13D11 at positions 5.4 and 5.8 Mb, respectively (fig. 1). This FISH pattern confirms our assumption that CNM repeats are clustered at position 5.6 Mb, where a gap exists in the current sequence assembly of GGA3. The other CNM array was revealed in long transcription units on lateral loops arising from the chromomere flanked by BACs WAG22J22 and WAG32A13 (figs. 1, 2B); these BACs are assigned to positions 11.0 and 14.0 Mb, respectively. Together with the fact that CNM units are present in the contig at the position around 11.5 Mb, this supports the idea that the current centromeric gap at position 11.6–13.1 Mb consists of CNM repeat sequences.

The pattern of CNM hybridization on GGA3 reminded us of chicken microchromosomes with CNM repeats at both termini [Krasikova et al., 2006; Deryusheva et al., 2007]. Indeed, H4 gene-rich isochores, which are typical for the majority of chicken microchromosomes, concentrate in the region flanked by CNM clusters on GGA3

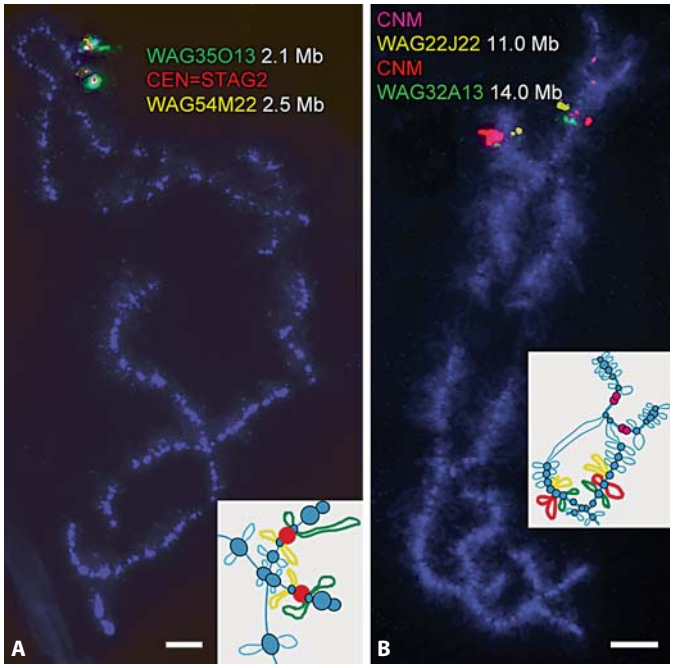

Fig. 2. Centromere positioning on chicken chromosome 3. **A** Immunofluorescent detection of the centromere followed by FISH with BAC clones WAG35O13 (green) and WAG54M22 (yellow) on LBC3. Centromeric cohesin-enriched structures (red) localize between these BACs . The schematic drawing shows the distribution of these signals (inset). Scale bar = 5 μm. **B** FISH with CNM-specific oligoprobes (red and pink) and BAC clones WAG22J22 (yellow) and WAG32A13 (green). The schematic drawing shows the distribution of these FISH signals (inset). Chromosomes are counterstained with DAPI (blue). Scale bar = 10 μm.

[Andreozzi et al., 2001]. The equivalence of chicken microchromosomes to gene-rich T-blocks in mammals was hypothesized earlier [Andreozzi et al., 2001]. We analyzed the distribution of 54 human orthologs for genes localized in the region of GGA3 mentioned above. In human, these genes split into 3 groups on HSA2 which correlates with 3 gene-rich blocks on 2p13.3–p23.1 [Saccone et al., 1999]. These correlations suggest a microchromosomal origin of the region between the gaps at positions 5.6 Mb and 11.6–13.1 Mb on GGA3. The sequences at position 11.6–13.1 Mb on GGA3 therefore presumably represent an inactivated centromere of a fused microchromosome that showed some centromere-specific features, when analyzed previously.

In conclusion, we have shown again that giant chromosomes in the lampbrush form represent a powerful system for integration of precise physical gene mapping and genome sequence data, including chromosome re-

gions with unknown DNA sequences. Our present data provide strong evidence for centromere location at position 2.4 Mb on GGA3 in the current sequence assembly (build 2.1). In the next version of the chicken genome assembly, the centromere of chromosome 3 should therefore be moved to the gap at position 2.4 Mb. The gaps at positions 5.6 Mb and 11.6–13.1 Mb should be preserved as they represent long clusters of tandem repeats.

Acknowledgements

This work was supported by the Russian Foundation for Basic Research (project #08-04-00493). A.K. was supported by a personal grant from the President of Russia (# MK-1677.2009.4). We used the equipment of Core Facility 'CHROMAS' (Saint-Petersburg State University).

References

Andreozzi L, Federico C, Motta S, Saccone S, Sazanova AL, et al: Compositional mapping of chicken chromosomes and identification of the gene-richest regions. Chromosome Res 9:521–532 (2001).

Crooijmans RPMA, Vrebalov J, Dijkhof RJ, van der Poel JJ, Groenen MAM: Two-dimensional screening of the Wageningen chicken BAC library. Mamm Genome 11:360–363 (2000).

Deryusheva S, Krasikova A, Kulikova T, Gaginskaya E: Tandem 41-bp repeats in chicken and Japanese quail genomes: FISH mapping and transcription analysis on lampbrush chromosomes. Chromosoma 116:519–530 (2007).

Gaginskaya E, Kulikova T, Krasikova A: Avian lampbrush chromosomes: a powerful tool for exploration of genome expression. Cytogenet Genome Res 124:251–267 (2009).

Galkina S, Deryusheva S, Fillon V, Vignal A, Crooijmans R, et al: FISH on avian lampbrush chromosomes produces higher resolution gene mapping. Genetica 128:241–251 (2006).

International Chicken Genome Sequencing Consortium: Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432:695–716 (2004).

Krasikova A, Deryusheva S, Galkina S, Kurganova A, Evteev A, Gaginskaya E: On the positions of centromeres in chicken lampbrush chromosomes. Chromosome Res 14: 777–789 (2006).

Li J, Leung FC: A CR1 element is embedded in a novel tandem repeat (*Hin*fI repeat) within the chicken genome. Genome 49:97–103 (2006).

Matzke MA, Varga F, Berger H, Schernthaner J, Schweizer D, et al: A 41–42 bp tandemly repeated sequence isolated from nuclear envelopes of chicken erythrocytes is located predominantly on microchromosomes. Chromosoma 99:131–137 (1990).

Saccone S, Federico C, Solovei I, Croquette MF, Della Valle G, Bernardi G: Identification of the gene-richest bands in human prometaphase chromosomes. Chromosome Res 7: 379–386 (1999).

Solovei I, Ogawa A, Naito M, Mizuno S, Macgregor H: Specific chromomeres on the chicken W lampbrush chromosome contain specific repetitive DNA sequence families. Chromosome Res 6:323–327 (1998).

Wang X, Li J, Leung FC: Partially inverted tandem repeat isolated from pericentric region of chicken chromosome 8. Chromosome Res 10:73–82 (2002).