A consensus linkage map of the chicken genome

Citation for published version:

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Genome Research

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
A Consensus Linkage Map of the Chicken Genome

Martien A.M. Groenen,1,15 Hans H. Cheng,2 Nat Bumstead,3 Bernard F. Benkel,4 W. Elwood Briles,5 Terry Burke,6 Dave W. Burt,7 Lyman B. Crittenden,8 Jerry Dodgson,8 Jossi Hillel,9 Sue Lamont,10 Abel Ponce de Leon,11 Morris Soller,12 Hideaki Takahashi,13 and Alain Vignal14

1Animal Breeding and Genetics Group, Wageningen Institute of Animal Sciences, Wageningen University, 6709 PG Wageningen, The Netherlands; 2United States Department of Agriculture, Agricultural Research Service, Avian Disease and Oncology Laboratory, East Lansing, Michigan 48823 USA; 3Molecular Biology, Institute for Animal Health, Compton Laboratory, Compton, Berkshire RG20 7NN, UK; 4Agriculture and Agro-Food Canada, Center for Food and Animal Research, Ottawa K1A 0C6, Canada; 5Department of Biological Sciences, Northern Illinois University, DeKalb, Illinois 60115 USA; 6Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK; 7Department of Molecular Biology, Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, UK; 8Michigan State University, Department of Microbiology, East Lansing, Michigan 48823 USA; 9Department of Genetics, The Hebrew University of Jerusalem, Rehovot 76100, Israel; 10Department of Animal Science, Iowa State University, Ames, Iowa 50011-3150 USA; 11Department of Animal Science, University of Minnesota, St. Paul, Minnesota 55108 USA; 12Department of Genetics, Silberman Life Science Institute, The Hebrew University of Jerusalem, Jerusalem, 91904 Israel; 13Department of Genetic Resources 1, National Institute of Agrobiological Resources, Kannondai, Tsukuba 305-8602, Japan; 14Institut National de la Recherche Agronomique, Genetique Animale, Laboratoire de Genetique Cellulaire, 31326 Castanet-Tolosan, France

A consensus linkage map has been developed in the chicken that combines all of the genotyping data from the three available chicken mapping populations. Genotyping data were contributed by the laboratories that have been using the East Lansing and Compton reference populations and from the Animal Breeding and Genetics Group of the Wageningen University using the Wageningen/Euribrid population. The resulting linkage map of the chicken genome contains 1889 loci. A framework map is presented that contains 480 loci ordered on 50 linkage groups. Framework loci are defined as loci whose order relative to one another is supported by odds greater then 3. The possible positions of the remaining 1409 loci are indicated relative to these framework loci. The total map spans 3800 cM, which is considerably larger than previous estimates for the chicken genome. Furthermore, although the physical size of the chicken genome is threefold smaller then that of mammals, its genetic map is comparable in size to that of most mammals. The map contains 350 markers within expressed sequences, 235 of which represent identified genes or sequences that have significant sequence identity to known genes. This improves the contribution of the chicken linkage map to comparative gene mapping considerably and clearly shows the conservation of large syntenic regions between the human and chicken genomes. The compact physical size of the chicken genome, combined with the large size of its genetic map and the observed degree of conserved synteny, makes the chicken a valuable model organism in the genomics as well as the postgenomics era. The linkage maps, the two-point lod scores, and additional information about the loci are available at web sites in Wageningen (http://www.zod.wau.nl/vf/research/chicken/frame_chicken.html) and East Lansing (http://poultry.mph.msu.edu/).

The chicken is increasingly becoming of great interest as an intermediate evolutionary model organism, ideally placed between mammals and more distant vertebrates as the pufferfish and zebrafish. There are a number of different reasons for this increasing interest in the chicken genome. First, the genome size is only one-third that of mammals (Tiersch and Wachtel 1991) mainly because of its low amount of repetitive sequences and reduced intron sizes (Hughes and Hughes 1995). Furthermore, it has an interesting complex genomic structure with two chromosomal subtypes—macrochromosomes and microchromosomes (Bloom et al. 1993)—with the microchromosomes appearing to be somewhat more gene dense then the macrochromosomes, reaching densities comparable to that of the Fugu genome (McQueen et al. 1998; Clark et al. 1999). Second, the level of conserved synteny between chicken and humans appears to be very high (Burt et al. 1995; Hu et al. 1995; Klein et al. 1996; Jones et al. 1997; Groenen et al. 1999; Nanda et al. 1999). Third, the chicken is being studied intensively for genes affecting polygenic traits (quantitative trait loci or QTL), which drive international efforts toward detailed physical and linkage mapping in the chicken.

Although the first genetic linkage map in chicken...
was published >60 years ago (Hutt 1936), it was not until the development of large numbers of molecular markers in the last decade that the generation of linkage maps in chicken increased. In chicken, three different linkage maps were developed using three different mapping populations. The first genetic map, based completely on DNA markers, was published by Bumstead and Palyga (1992). This map, based on the Compton (C) reference population, consisted solely of restriction fragment length polymorphism (RFLP) markers. The second genetic map to be published (Levin et al. 1993, 1994) was based on the East Lansing (EL) reference population and consisted primarily of RFLPs, random amplified polymorphic DNA (RAPD) markers, and chicken repeat element 1 (CRI) markers. Since then, both populations have been used to map a considerable number of microsatellite markers (Cheng et al. 1995; Crooijmans et al. 1997; Gibbs et al. 1997) and AFLP markers (Knorr et al. 1999) as well. The third map (Groenen et al. 1998, Herbergs et al. 1999) was based on a large F2 population and consisted solely of microsatellite and amplified fragment length polymorphism (AFLP) markers. Increasing marker densities and increased initiatives in physical mapping in chicken have necessitated the need of a single consensus linkage map in chicken. Because all three maps have many markers in common, this goal has become feasible for the large and intermediate-sized chromosomes.

In this paper we describe the integration of all available data of the three mapping populations, resulting in a consensus linkage map of the chicken genome comprised of 50 linkage groups, with a total of 1889 loci.

RESULTS AND DISCUSSION

Linkage Maps

The genotyping data from the three chicken mapping populations were combined and analyzed simultaneously using the CRIMAP linkage program. Contributions of genotyping data were made from laboratories that have been using the EL and C reference populations and from the Animal Breeding and Genetics Group of the Wageningen University using the Wageningen/Euribrid (WAU) population. A complicating factor for the integration of all maps in chicken is the fact that not all types of markers are evenly distributed over the macro- and microchromosomes, particularly the low abundance of microsatellites on the microchromosomes (Primmer et al. 1997). Consequently, many of the small linkage groups do not have a marker in common, making the integration impossible at present. Furthermore, linkage groups C15 and C20 had only one marker in common with the corresponding linkage groups in the WAU and EL data sets, and as a consequence, the other loci from C15 and C20 could not be positioned very precisely (Fig. 1, linkage groups E18C15W15 and E49C20W21).

The total number of different loci that have been typed on at least one of the three mapping populations was 2019. However, a relatively large proportion of these markers was either unlinked (95) or could not be positioned clearly on the linkage maps (35) and therefore were omitted from the final map shown in Figure 1. A large proportion of the omitted markers (89) are AFLP markers typed on the WAU population (Herbergs et al. 1999).

The resulting linkage map of the chicken genome contains 1889 loci (Table 1). The framework map contains 480 loci ordered on 50 linkage groups (Fig. 1). The possible positions of the remaining 1409 loci are indicated relative to these framework loci. When all linkage groups are taken into account, the total length of the linkage map is 4000 cM. However, it is expected that several of the smaller EL, C, and WAU linkage groups belong to the same chromosomes. If we correct for this fact, then the minimal length of the chicken consensus linkage map is ~3800 cM, which still is considerably larger than the previous estimates 2600–3000 cM for the chicken genome (Rodionov et al. 1992; Burt et al. 1995).

Although there are large differences in length between many of the EL (male) and the C linkage groups (female), these variations are most likely the result of differences between the lines used and typing errors in some of the RFLP markers in the C map. The differences in length between the male and female maps

<table>
<thead>
<tr>
<th>Table 1. Number and Type of Loci on Chicken Linkage Maps</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Loci</td>
</tr>
<tr>
<td>WAU</td>
</tr>
<tr>
<td>---------------</td>
</tr>
<tr>
<td>A. Loci</td>
</tr>
<tr>
<td>WAU</td>
</tr>
<tr>
<td>EL</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>Type I loci</td>
</tr>
<tr>
<td>Linkage groups</td>
</tr>
<tr>
<td>B. Markers</td>
</tr>
<tr>
<td>Microsatellite</td>
</tr>
<tr>
<td>Minisatellite</td>
</tr>
<tr>
<td>RFLP</td>
</tr>
<tr>
<td>AFLP</td>
</tr>
<tr>
<td>SSCP</td>
</tr>
<tr>
<td>ASO</td>
</tr>
<tr>
<td>RAPD</td>
</tr>
<tr>
<td>CR1</td>
</tr>
<tr>
<td>Classical</td>
</tr>
</tbody>
</table>

(A) The total number of different loci analyzed on the Wageningen (WAU), East Lansing (EL), and Compton (C) linkage maps, as well as the number of loci that are shared among the different maps. (B) The different types of markers on the different chicken linkage maps. Included are loci located on the consensus linkage map, as in Fig. 1.
Figure 1 (See pages 139–142.) Framework consensus linkage map of the chicken genome. The framework loci (loci whose order relative to one another is supported by odds >3, Keats et al. 1991) have been ordered and their position is indicated by the number at left. The possible location of the loci whose order is not supported by odds >3 is indicated by an error bar. Loci that have been mapped cytogenetically are underlined; those known to represent expressed sequences (identified genes and ESTs) are shown in boldface type.
based on the WAU population generally are small with an overall difference between male and female maps of only 1.15%. Because of the much larger number of informative meioses within the WAU population, the framework of the consensus map is mainly built up by microsatellites typed on this population. Therefore, for the consensus map, the differences seen between the size of the male and female maps are similar as those described for the WAU map (Groenen et al. 1998).

Discrepancies Between the Different Maps

Discrepancies for seven loci were observed in map locations in the original maps from the three populations used. (1) G6PD has been mapped on E1 and CW, and (2) EIF4A2 has been mapped on E36 and C3. As these are both RFLPs, it is possible that different bands were scored in each population. With regard to EIF4A2, it is noteworthy that its location within the EL data is supported by the comparative mapping data (Fig. 2). In the EL data EIF4A2 maps to E36, close to the SNON and TPRC genes. In human, all three genes are located on the q arm of chromosome 3. However, given the discrepancies in the data, these two loci have not been included on the map. (3) LEI0144 has been mapped on chromosome 4 (WAU) and chromosome Z (EL); (4) MCW0066 has been mapped on chromosome 2 (C), E30 and W10; (5) MCW0166 has been mapped on chromosome 2 (WAU) and chromosome 4 (EL). (6) The BAT8 gene was mapped by a single laboratory (Spike and Lamont 1995) to the end of chromosome 4 (EL) and chromosome 16, the chromosome to which the major histocompatibility complex (MHC) class I and II genes have been mapped. Finally, (7) The α-tubulin gene (TUBA) was mapped to C15 and to E22. Because this gene has been mapped by two different laboratories using different methods, it is likely that two different loci of the α-tubulin gene family have been mapped. These two genes have been included in the map as TUBAAa and TUBAb, respectively.

To resolve the discrepancies between the microsatellites, these markers were retyped. Typing of LEI0144 on the EL population by M. Groenen and coworkers, (unpubl.) showed that this gene also mapped to chromosome 4, which is in agreement with the WAU data. Similarly, retyping of this marker by T. Burke and coworkers, (unpubl.) confirmed their previous results of LEI0144 mapping to the Z chromosome. A possible explanation for this discrepancy is that the typings are done by the two groups using different primers [Crooijmans et al. (1997) and Gibbs et al. (1997), respectively]. This marker has therefore been included on both locations on the map as LEI0144a and LEI0144b. The new typing results showed clearly that the previous assignment of MCW0066 to C chromosome 2 and of MCW0166 to EL chromosome 4 are not correct and that these two markers are on linkage group E30C14W10 and chromosome 2, respectively.

In addition, although there were no conflicting data between them, the mapping results of two loci for the three different populations were rather unexpected or somewhat unlikely. Marker LEI0192 appeared to be unlinked in the EL population, whereas it is mapped to chromosome 6 using the WAU population and to linkage group C21 using the C population. Similarly, marker HUJ0006 has been mapped to chromosome 6 on the WAU map and it is unlinked in the other two maps. Moreover, both markers are linked tightly to each other and located at the end of chromosome 6 of the WAU map, whereas they are unlinked in both of the other maps. Flipping of the typing phase of HUJ0006 in the original MAPMANAGER file of the EL population results in linkage to several markers on chromosome 6 but not to LEI0192. Flipping of the phase had no effect in these markers in the C data set. Physical mapping of a BAC clone containing LEI0192 confirmed its location at the end of chromosome 6 (V. Fillon and A. Vignal, pers. comm.).

Two additional discrepancies are observed between the consensus linkage map and the cytogenetic map. The CYP19 gene was mapped on linkage group E029C09W09, whereas it was mapped to chromosome 1 on the cytogenetic map (Tereba et al. 1991). Comparative mapping data support the location of this gene on linkage group E29C09W09. On the linkage map the MAX gene maps to chromosome 4, whereas cytogenetically it was mapped to chromosome 5p (Nanda et al. 1997). In this case, the comparative mapping data support the cytogenetic location of this gene on chromosome 5, where several other human genes have been mapped that are located on chromosome 14q.

Finally, LEI0229 was mapped to both the Z chromosome (EL) and the W chromosome (C). The most likely explanation is that LEI0229 maps to one of the pseudoautosomal regions of the chicken Z chromosome (Fridolfsson et al. 1998).

Anchoring Linkage Groups to Chromosomes

In addition to the integration of the three linkage maps, eventually these maps will have to be integrated with the physical map in chicken. However, the integration of the physical and genetic maps presents considerable difficulties and has proceeded at a slower pace, as a result (Morisson et al. 1998). The chicken karyotype is composed of 2n = 78 chromosomes which, according to their size, are classified as macro- and microchromosomes (Bloom et al. 1993). Due to the presence of microchromosomes in chicken, a standard karyotype could only be established for the eight large macrochromosomes and the two sex chromo-
Figure 2  Comparative mapping results among chicken, man, and mouse. The order of the loci is according to the linkage map shown in Fig. 1. The second column in each linkage group shows the location of the loci on the human cytogenetic map according to Genome Data Base (http://www.gdb.org/); the third column shows the map location in the mouse. Blocks of conserved synteny between chicken and man and between chicken and mouse are shaded.
Many markers on the consensus map have been mapped cytogenetically as well, allowing the integration of the linkage map with the physical map. These loci are underlined in Figure 1. Because only a standard karyotype has been established for the macrochromosomes, only the linkage groups of these larger chromosomes could be assigned to their corresponding chromosomes. For historical reasons, the microchromosome containing the MHC is named chromosome 16.

To enable identification of the microchromosomes, a set of large insert clones is being developed that can be used as tags in two-color fluorescence in situ hybridization (Fillon et al. 1998). Polymorphic markers have been developed for many of these large insert clones (Morisson et al. 1998; P.A. Thomson and T. Burke, unpubl.), which will allow the assignment of the linkage groups to the corresponding microchromosomes as well. Furthermore, additional large insert BAC clones have been isolated using microsatellites from the small linkage groups, which provides additional probes for the identification of the microchromosomes (R. Crooijmans, V. Fillon, M. Groenen, and A. Vignal, unpubl.).

Comparative Gene Mapping
Genetic markers within or adjacent to known genes have been classified as type I markers (O’Brien 1991). The inclusion of type I markers on the linkage map makes it possible to access the mapping information that is available in densely mapped species such as humans and mice. Currently, the consensus map contains 350 markers within expressed sequences, 235 of which represent identified genes or sequences that have significant sequence identity to known genes. These loci are shown in boldface type in Figure 1. The orthologs of 204 of these 235 genes have also been mapped in human (Fig. 2). The comparative mapping data based on the consensus linkage map show a considerable amount of chromosomal conservation retained between man and chicken during evolution. This is in sharp contrast with the comparative mapping data between chicken and mouse, in which the amount of chromosomal conservation is considerably lower. Similar results are obtained for the comparisons between the genomes of different mammals, indicating that there have been extensive rearrangements during the evolution of the mouse genome and at a much higher rate than in birds or the other mammals (Andersson et al. 1996). Based only on the linkage data presented in this paper, at least 87 different chromosomal regions can be identified between man and chicken (Fig. 2). For many of these chromosomal regions, physically mapped genes provide additional evidence for the observed conservation of linkage (Burt et al. 1995; Andersson et al. 1996; Nanda et al. 1999; Burt et al. 1999). Furthermore, the physically mapped genes have identified additional conserved regions between the genomes of chicken and man. Based on this number of conserved chromosomal regions, it has been calculated that the number of autosomal conserved segments shared between the chicken and human genomes is probably <100 (Burt et al. 1999). This level of conservation of synteny between chicken and human, in combination with the threefold more compact genome of the chicken, makes it an excellent evolutionary model organism in addition to Fugu, mouse, and rat. This is particularly true for the microchromosomes, which appear to be somewhat more gene dense than the macrochromosomes, thereby reaching gene densities close to that of Fugu (Angrist 1998; McQueen et al. 1998; Clarke et al. 1999). Furthermore, the higher level of conservation of genome organization (Gilley et al. 1997; Reboul et al. 1999) and the easy accessibility of the chicken as an experimental animal in studies regarding complex polygenic traits are additional features favoring the chicken over other models such as Fugu.

Although the number of loci that are available for comparative mapping are still too limited to draw detailed conclusions, it is noteworthy that several of the small linkage groups in chicken, which most likely represent different microchromosomes, seem to represent, in almost their entirety, large fragments of specific human chromosomes (e.g., E29C09W09, E21E31C25W12, E48C28W13W27, E41W17, E54, E49C20W21, and chromosome 7).

Future Directions
The current map contains 801 microsatellite markers, which are the markers of choice for whole genome scans. However, the marker density is only sufficiently high for the macrochromosomes and a subset of the microchromosomes. Therefore, many more microsatellites are still needed to obtain (near) complete genome coverage in these kinds of studies. The integration of all the linkage maps and the cytological map in chicken is the first necessary step toward achieving this goal by identifying those regions that are particularly devoid of microsatellite markers. The major drawback, however, is the relatively low abundance of microsatellites on many of the microchromosomes (Primmer et al. 1997). Currently, increasing efforts are being put into the development of physical maps for several regions of the chicken genome, for example, Chromosome 16 (N. Bumstead, unpubl.) and linkage groups E29C09W09 and E53C34W16 (R. Crooijmans and M. Groenen, unpubl.). This has become feasible through an increased number of loci on the linkage map and because of the development of publicly available chicken YAC (Toye et al. 1997) and BAC (R.P.M.A. unpubl.).
Crooijmans, R. Acar, R.J.M. Dijkhof, J.J. van der Poel, and M.A.M. Groenen, in prep.; J. Dodgson, unpubl.) libraries. It is to be expected that physical maps eventually will become available for all chicken chromosomes. This in turn will make the targeted development of microsatellite markers possible for those regions that currently lack any such markers, thereby allowing the characterization of these regions in QTL studies as well.

METHODS

Mapping Populations

The three mapping populations have been described in detail previously. Briefly, the EL population (Crittenden et al. 1993) consists of 52 BC1 animals derived from a backcross between a partially inbred jungle fowl line and a highly inbred white leghorn line. The C population (Bumstead and Palyga 1992) consists of 56 BC1 animals derived from a backcross between two inbred white leghorn lines that differed in their disease resistance. The WAU population (Groenen et al. 1998) consists of 456 F2 animals from a cross between two broiler dam lines originating from the white Plymouth Rock breed.

Markers

A detailed description of all individual loci, including their references and the number of informative meioses, is available at the web site of the Animal Breeding and Genetics Group in Wageningen (http://www.zod.wau.nl/vf/research/chicken/frame_chicken.html) and East Lansing (http://poultry.mph.msu.edu/)

Linkage Analysis

For each of the linkage groups, the genotyping data of the three populations were combined into a single file. To analyze the genotyping data of the backcross populations, together with data from the WAU population, the genotypes for these two populations were recoded as either being 1:1 (homozygous) or 1:2 (heterozygous). The combined data therefore consisted of 12 individual families, 1 EL, 1C, and 10 W. Linkage analysis was performed using CRIMAP version 2.4 (Green et al. 1990). Initially, a two-point linkage analysis was performed in which all markers were analyzed against each other. Tables containing all two-point lod scores for all markers are available at the web site of the Animal Breeding and Genetics Group in Wageningen. When possible, markers that had been typed on all three maps were used to start building the map using the CRIMAP-BUILD option. Finally, the order of the framework loci was checked using the CRIMAP-flips5 function.

ACKNOWLEDGMENTS

We thank our many colleagues that have contributed to the mapping of the loci that are described in this paper, including R. Acar, S. Ambady, R. Bergé, D. Burke, R. Crooijmans, D. Dawson, E. Dufour, R. Dijkhof, M. Gibbs, O. Hanotte, L. Heltemes, J. Herbergs, B. Van Hest, J. Hillel, H. Khatib, C. Knorr, I. Levin, S. McConnell, C. Moran, M. Morisson, R. Okimoto, J. Palya, F. Pitel, B.J. Sheldon, E.J. Smith, M. Siwek, C. Spike, S. Suchyta, P. Thomson, A. Toye, R.J. Vallejo, T. Veenendaal, B.J. van Hest, and A. Wardle. The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES


Green, P., K. Falls, and S. Crooks. 1990. Documentation for CRI-MAP, version 2.4 Washington School of Medicine, St. Louis, MO.


Herbergs, J., M. Siwek, R.P.M.A. Crooijmans, J.J. van der Poel, and M.A.M. Groenen. 1999. Multicolour fluorescent detection and

Received September 22, 1999; accepted in revised form November 16, 1999.