



# THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Chicken genome

**Citation for published version:**

Burt, DW 2005, 'Chicken genome: current status and future opportunities' *Genome Research*, vol 15, no. 12, pp. 1692-8., 10.1101/gr.4141805

**Digital Object Identifier (DOI):**

[10.1101/gr.4141805](https://doi.org/10.1101/gr.4141805)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher final version (usually the publisher pdf)

**Published In:**

Genome Research

**Publisher Rights Statement:**

©2005 by Cold Spring Harbor Laboratory Press

**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](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Chicken genome: Current status and future opportunities

David W. Burt

Department of Genomics and Genetics, Roslin Institute (Edinburgh), Midlothian EH25 9PS, United Kingdom

The chicken genome sequence is important for several reasons. First, the chicken shared a common ancestor with mammals ~310 million years ago (Mya) at a phylogenetic distance not previously covered by other genome sequences. It therefore fills a gap in our knowledge and understanding of the evolution and conservation of genes, regulatory sequences, genomes, and karyotypes. The chicken is also a major source of protein in the world, with billions of birds used in meat and egg production each year. It is the first livestock species to be sequenced and so leads the way for others. The sequence and the 2.8 million genetic polymorphisms defined in a parallel project are expected to benefit agriculture and cast new light on animal domestication. Also, as the first bird to be sequenced, it is a model for the 9600 avian species thought to exist today. Many of the features of the chicken genome and its biology make it an ideal organism for studies in development and evolution, along with applications in agriculture and medicine.

Avian genomics has its origins in genetic linkage mapping (Burt and Cheng 1998), but our knowledge of the chicken genome has been transformed in recent years, mostly through the analysis of large numbers of partial cDNA sequences (Abdrakhmanov et al. 2000; Tirunagaru et al. 2000; Boardman et al. 2002) and culminating with the chicken genome sequence (Hillier et al. 2004). These were landmark events in our understanding of avian biology, developmental biology, and the evolution of vertebrates and will facilitate applications in agriculture and medicine.

Chicken research has had a significant impact on fundamental biology and the chicken has been a popular model organism for at least 100 years, for example, with the discovery of B cells and tumor viruses (Brown et al. 2003). Ready access to the chicken embryo using incubated eggs and the ease of manipulation make this system ideal for studies of vertebrate development (Stern 2004, 2005). The chicken has been used in many of the classical studies on the molecular basis of patterning in the vertebrate embryo, in particular, the limb bud. In recent times, other model organisms, such as the mouse and zebrafish, have been in greater demand because of increased genetic resources and the ability to manipulate their genomes. The chicken EST and genome programs have removed many of these limitations in the chicken. In addition, new tools such as the electroporation of chicken embryos and the use of RNAi to knock down gene expression are likely to make the chicken embryo a powerful model for the molecular study of development in vertebrates (Stern 2004, 2005).

During the past 80 years, modern selective breeding has made spectacular progress in both egg and meat production traits (Burt 2002). World egg production has increased to 795 billion/year in 2002 (Commodity Research Bureau [CRB]) and broiler meat to 6.5 million tons/year (USDA foreign agricultural service [FAS]), during this period. Associated with these successes have been a number of undesirable traits. In meat-type chickens, there has been an increase in the incidence of congenital disorders, such as ascites and lameness, reduced fertility, and reduced resistance to infectious disease. In egg-type chickens, there has

been an increase in the incidence of osteoporosis associated with increased egg production. Given the possibility that genetic progress in egg and meat production will reach its limit within the next twenty years (Burt 2002), priorities in the poultry industry will be to reduce these costs and develop new products. The consumer wants high-quality products (e.g., increased egg shell strength), which requires greater uniformity and predictability in production. With an increased requirement for food safety, there will be a need to reduce the use of chemicals and antibiotics and increase genetic resistance to pathogens. These new traits are difficult and costly to measure by conventional genetic selection, and the developments in poultry genomics in the last few years promises new solutions to these problems.

In this report, key features and limitations of the draft chicken genome sequence will be discussed. More detailed reviews have been presented elsewhere from the viewpoint of genomics (Burt 2004a; Dequeant and Pourquié 2005), developmental biology (Burt 2004b; Stern 2005), evolution (Ellegren 2005), and genomic tools (Antin and Konieczka 2005).

## Chicken genome

### Genome sequence

The current draft of the chicken genome (Hillier et al. 2004) was assembled using a whole-genome sequencing strategy, including BAC, fosmid, and plasmid paired-end reads (WASHU). This approach produced a high-quality assembly, in part because of the relatively small size of the chicken genome, one third that of a typical mammal. However, it was the low repetitive DNA content, only 11% compared with 40%–50% found in mammals, that was a key contributing factor to the quality of the final assembly. This sequence employed DNA from a single inbred female Jungle Fowl (*Gallus gallus gallus*, the ancestor of domesticated chickens; Fumihito et al. 1994) and represented a 6.6-fold coverage of the genome. Together with genetic and BAC maps, almost 100,000 contigs were assembled into a scaffold of 907 Mb, or 86% of a 1050-Mb genome. In birds, it is the female that is the heterogametic sex, with single copies of the Z and W chromosomes. Therefore, these chromosomes were poorly represented in the final assembly. In addition, unlike the rest of the genome, the

E-mail [Dave.Burt@bbsrc.ac.uk](mailto:Dave.Burt@bbsrc.ac.uk); fax +44-131-440-0434.

Article and publication are at <http://www.genome.org/cgi/doi/10.1101/gr.4141805>.

W chromosome has a high repeat content and so very little sequence was assembled. Targeted sequencing of the sex chromosomes will be necessary to complete their assemblies. For autosomes, sequence coverage was 98% based on overlaps with an independent set of BAC clones sequenced to high quality. Overlaps with cDNA clones suggested 5%–10% of genes were missing from the final assembly; gene duplications and GC-rich sequences were a particular problem. The MHC region on chromosome 16, a rich source of duplicated genes, was very poorly represented. Further work to complete the chicken genome sequence to a high quality for comparative genomics and gene discovery is required.

### Genome organization

A unique characteristic of avian genomes is the large variability in chromosome size. In addition to a pair of sex chromosomes, chickens have 38 pairs of autosomes: 5 macro-, 5 intermediate, and 28 microchromosomes. Since each chromosome arm must have at least one obligate crossover, it follows that the microchromosomes will have the highest rate of recombination. Comparison of genetic maps (Schmid et al. 2005) and genome sequences confirms this expectation, with crossover rates of 2.8 cM/Mb for macrochromosomes and 6.4 cM/Mb for microchromosomes. This is in contrast to 1–2 cM/Mb for most human chromosomes, making the chicken ideal for genetic linkage studies. High-resolution genetic maps will be necessary to define variation in recombination rates within chromosomes.

Many sequence characteristics, such as %GC content, CpG island density, and gene density, show clear relationships with chromosome size and therefore recombination rate (Table 1). However, we must be cautious about making any conclusions on cause and effect with these correlations (Fazzari and Grealley 2004). The density of genes is highest on the microchromosomes, confirming earlier conclusions based on mapping genes (Smith et al. 2000) and CpG islands (McQueen et al. 1996). The estimated number of CpG islands based on bioinformatics approaches depends on the definition in use. In this case (Hillier et al. 2004), ~70,000 CpG islands were predicted in the chicken, with 38% of these located in regions of conserved synteny with mammalian genomes. Since 48% are associated with a gene, CpG island density mimics gene density and is highest on microchromosomes. Conversely, sizes of introns and intergenic regions and density of repetitive elements correlate negatively with gene den-

sity and are reduced on microchromosomes. If we assume that genomes balance selective constraints favoring DNA loss over those that favor expansion and that selection will be most efficient in regions of high recombination where linkage of alleles are more readily broken (Hill and Robertson 1966), then the correlation of the densities of genes, CpG islands, repeats, etc. with chromosome size (and therefore recombination rate) is to be expected.

Comparison of orthologous chicken and turkey sequences revealed that different chromosome size classes are subject to different evolutionary forces (Axelsson et al. 2005). Microchromosomes show 18% higher sequence divergence in introns and a 26% higher rate of synonymous substitution in coding sequences than macrochromosomes, indicating that the smaller chromosomes are more susceptible to germline mutations. A possible cause for the differences in mutation rate is “biased-gene-conversion” (Meunier and Duret 2004), a recombination-induced mutation mechanism.

Ever since the first gene maps were created (Haldane 1927), comparative maps have been used to examine the evolution of the vertebrate genome. Comparisons between the early gene maps of human and chicken (Burt et al. 1999) suggested extensive conservation of synteny, possibly more than found between mouse and human. The comparison of chicken with mammalian and fish genomes has confirmed and extended this view (Bourque et al. 2005). The estimated number of interchromosomal rearrangements between the mammalian ancestor and chicken, during an estimated period of 500 million years (Myr), is almost the same as the number found in the mouse lineage, over the course of ~87 Myr.

### Genes and proteins

A major benefit of the chicken genome sequence has been the set of gene predictions. The most conservative evidence-based approach of Ensembl generated 17,709 predictions (Table 2). The comparative *ab initio* methods, TWINSCAN (Korf et al. 2001) and SGP-2 (Syntenic Gene Prediction-2) (Parra et al. 2003), predict larger gene sets but likely include false positives. In total, there may be 20,000–23,000 genes; suggesting we still have more to learn about gene prediction (Eyras et al. 2005). When used to identify novel genes missed in the current human gene set (Ensembl 22,287 genes), only an additional 37 were predicted (Castelo et al. 2005), which suggests we have identified most of

the “conserved” genes found in birds and mammals. Only 75 processed (or retrotransposed) pseudogenes were found in the chicken genome (Hillier et al. 2004), compared with 15,000 in mammals. The reason for this low number may be the sequence specificity of reverse transcription by avian LINES (long interspersed elements). Mammalian LINES are more promiscuous and able to retrotranspose most mRNAs. It was hoped that the lack of pseudogenes in the chicken would help to identify functional noncoding RNA genes in mammalian genomes via conservation of chromosomal gene location. (Because of their noncoding charac-

**Table 1.** General characteristics of macro- and microchromosomes

Feature	Macrochromosomes	Microchromosomes	Reference
Cytogenetic band type	G-band	R-band	Ponce de Leon et al. 1992
Gene density (per Mb)	9.0 to 15.4	13.8 to 41.2	Hillier et al. 2004
Intron length (bp)	4066 to 5742	1867 to 4128	Hillier et al. 2004
Exon length (bp)	164 to 171	157 to 172	Hillier et al. 2004
Intergenic gap length (kb)	18 to 31	8 to 24	Hillier et al. 2004
Intron length/exon length	24.6 to 34.6	11.7 to 23.2	Hillier et al. 2004
%G + C content	38.4 to 40.1	40.9 to 50.1	Hillier et al. 2004
CpG island density (per Mb)	29 to 49	73 to 266	Hillier et al. 2004
LINES (%)	6.0 to 11.9	2.5 to 10.0	Hillier et al. 2004
SINEs (%)	~ none	~ none	Hillier et al. 2004
Synonymous rate	0.090 to 0.125	0.111 to 0.156	Axelsson et al. 2005
Nonsynonymous rate	0.011 to 0.021	0.007 to 0.016	Axelsson et al. 2005
Ka/Ks ratio	0.128 to 0.360	0.066 to 0.177	Axelsson et al. 2005
GC3%	49 to 53	56 to 65	Hillier et al. 2004
DNA replication	late	early	Ponce de Leon et al. 1992
Recombination rate (cM/Mb)	2.5 to 3.2	2.5 to 17.1	Hillier et al. 2004

**Table 2.** Frequency and class of gene/protein predictions (Ensembl, June 2004)

Gene family	Number	Function
tRNA	280	Transfer RNA, adaptor in translation
5S rRNA	12	Ribosomal RNA, component of ribosome
5.8S RNA <sup>a</sup>	~145	Ribosomal RNA, component of ribosome
18S rRNA <sup>a</sup>	~145	Ribosomal RNA, component of ribosome
28S rRNA <sup>a</sup>	~145	Ribosomal RNA, component of ribosome
snRNP U1	18	Major spliceosome
snRNP U2	6	Major spliceosome
snRNP U4	4	Major spliceosome
snRNP U5	9	Major and minor spliceosomes
snRNP U6	15	Major spliceosome
snRNP U4atac	1	Minor spliceosome
snRNP U6atac	4	Minor spliceosome
snRNP U11	1	Minor spliceosome
snRNP U12	1	Minor spliceosome
miRNA	121	Translation repression
snoRNA	83	Small nucleolar RNA, takes part in processing of rRNA
RNaseP	1	Ribozyme, processes tRNA
snRNP U7	1	3'-end processing of replication-dependent histone pre-mRNAs
SRP	3	RNA component of signal recognition particle
7SK	4	Binds P-TEFb, which activates transcription by phosphorylating C-terminal domain of RNA Pol II. This process is negatively regulated by the 7SK RNP.
Y RNA	2	Component of the Ro RNP, in association with Ro60 and La, function of Ro RNP not known.
Telomerase RNA	1	Provides the template for telomeric DNA addition
BIC	1	Cooperates with c-myc in B lymphomagenesis and erythroleukemogenesis
Total RNA genes	571 (1441 incl. rRNA genes)	
Total pseudogenes	75	
Total protein coding genes	17,709	

<sup>a</sup>Ribosomal DNA 40-kb repeat (18S, 5.8S, 28S); Muscarella et al. (1985).

ter, it is difficult to distinguish functional RNA genes from the large excess of RNA pseudogenes in mammals by ab initio methods.) In chicken, 571 RNA genes in 20 distinct families were predicted and only the miRNA and snoRNA families (that usually lie within introns of coding genes) show conserved synteny to the extent that protein coding genes do. That the other noncoding RNA families did not suggest that they may transpose throughout the genome in ways that differ from coding genes.

Comparisons between mammals and birds can also start to address questions about gene gains/losses (Hillier et al. 2004). Comparisons between human, chicken, and *Fugu* suggest a core set of almost one third of all genes (7606) is conserved in all vertebrates. These comparisons also suggest that the rates of gene loss were higher in the avian lineage and fewer gene duplications were found in birds. Careful comparisons detected some genes lost from the chicken lineage, including vomeronasal receptors, caseins, and some genes of the immune system. Similarly, birds have more keratins specific to feathers and mammals have lost the avidin egg proteins. The discovery that all enzymes in the urea cycle were present but apparently not used for this function in birds was perplexing.

### New tools for genome analysis

Important by-products of any genome project are the resources (cDNA and BAC clones, genetic markers, etc.) and information it provides for future research (Antin and Konieczka 2005). Together with chromosome paints, BAC clones (BPRC) have been used to define cytogenetically all chicken chromosomes (Masabanda et al. 2004). Because of the nearly identical sizes of the microchromosomes in mitotic chromosome spreads, this was not previously feasible. A BAC map with 20-fold redundancy or 91%

coverage of the chicken genome has been assembled into 260 contigs (Wallis et al. 2004; ChickFPC). BAC contig maps are under construction for other birds; including turkey, California condor, and zebra finch (Edwards et al. 2005). These clones can be used to target specific genomic regions and to create whole-genome BAC arrays for comparative surveys of avian genomes. These arrays may be able to classify many avian species into unique clades, a notoriously difficult task (Edwards et al. 2005). From the very start, ESTs and cDNA clones have been important (Boardman et al. 2002; ChickEST), in particular for the prediction of chicken genes. ESTs have been used to create cDNA microarrays (Burnside et al. 2005) and design DNA chips (Affymetrix) for high-throughput gene expression assays. A total of 4532 full-length cDNA clones (Caldwell et al. 2004; Hubbard et al. 2005), representing ~25% of known gene predictions in chicken, can now be used in evolutionary and functional studies (available from ARK-Genomics). RNAi and transgenic technologies are now available in the chicken, which when combined with the accessible chicken embryo, makes this a powerful system for functional studies in vivo (Brown et al. 2003; Nakamura et al. 2004; Sang 2004; Stern 2004). The application of these tools and access to the biological information they generate is a huge and complex task. There are a number of databases distributed throughout the world (Table 3), including genome browsers (Ensembl, NCBI, and UCSC), genetic maps (ARKdb and ChickACE), gene expression (GEISHA), and others, but there is a need to integrate these views into a single Model Organism Database (GMOD).

### Applications of the chicken genome sequence

Birds and mammals shared a common ancestor ~310 million years ago (Mya) (Hedges 2002). Sequence comparisons between

**Table 3. Online resources for avian genomics**

Web site	Description
<a href="http://bacpac.chori.org/">http://bacpac.chori.org/</a>	BPRC: BAC resources center.
<a href="http://chick.umist.ac.uk/">http://chick.umist.ac.uk/</a>	ChickEST: BBSRC chicken EST database.
<a href="http://chicken.genomics.org.cn/index.jsp">http://chicken.genomics.org.cn/index.jsp</a>	ChickVD: chicken variation database.
<a href="http://geisha.biosci.arizona.edu/">http://geisha.biosci.arizona.edu/</a>	GEISHA: <i>Gallus gallus</i> EST and in situ hybridization analysis.
<a href="http://genex.hgu.mrc.ac.uk/intro.html">http://genex.hgu.mrc.ac.uk/intro.html</a>	EMAP: Edinburgh mouse atlas project.
<a href="http://genome.ucsc.edu/cgi-bin/hgGateway">http://genome.ucsc.edu/cgi-bin/hgGateway</a>	UCSC genome browser.
<a href="http://genomeold.wustl.edu/">http://genomeold.wustl.edu/</a>	WASHU: Washington University genome sequencing center.
<a href="http://www.affymetrix.com/">http://www.affymetrix.com/</a>	Affymetrix.
<a href="http://www.animalsciences.nl/ChickFPC/">http://www.animalsciences.nl/ChickFPC/</a>	ChickFPC: Chicken FPC BAC map.
<a href="http://www.ark-genomics.org/">http://www.ark-genomics.org/</a>	ARK-Genomics: Center for functional genomics in farm animals.
<a href="http://www.chicken-genome.org/">http://www.chicken-genome.org/</a>	AvianNET: the avian genome information network.
<a href="http://www.crbtrader.com/fund/articles/eggs.asp">http://www.crbtrader.com/fund/articles/eggs.asp</a>	CRB: commodity research bureau.
<a href="http://www.ensembl.org/Gallus_gallus/index.html">http://www.ensembl.org/Gallus_gallus/index.html</a>	Ensembl genome browser.
<a href="http://www.fas.usda.gov/currwmt.asp">http://www.fas.usda.gov/currwmt.asp</a>	FAS: USDA foreign agricultural service.
<a href="http://www.gmod.org/">http://www.gmod.org/</a>	GMOD: Generic Model Organism Database.
<a href="http://www.nature.com/nature/focus/avianflu/index.html">http://www.nature.com/nature/focus/avianflu/index.html</a>	Avian flu: Web Focus.
<a href="http://www.ncbi.nlm.nih.gov/genome/guide/chicken/">http://www.ncbi.nlm.nih.gov/genome/guide/chicken/</a>	NCBI genome browser.
<a href="http://www.thearkdb.org/browser?species=chicken">http://www.thearkdb.org/browser?species=chicken</a>	ARKdb: genetic mapping databases.
<a href="http://acedb.asg.wur.nl/">http://acedb.asg.wur.nl/</a>	ChickACE: Wageningen Animal Sciences Group ACEbrowser.

these groups are characterized with a high signal-to-noise ratio for the detection of functional elements. Taken together with the ready access to chicken embryos and as a major food source, chicken genomics is likely to have major applications and benefits in comparative genomics, evolutionary biology and systematics, models of development and human disease, and agriculture.

#### Comparative genomics

A major reason for sequencing the chicken genome was to increase our understanding of the human genome through comparative genomics, for example, to define regions under selection such as coding and regulatory elements (Hillier et al. 2004). Comparisons with known functional sequences suggested that 75% of coding regions and 30%–40% of regulatory elements are conserved. Only 2.5% of the chicken sequence could be aligned with that of the human (44% coding, 25% intronic, and 31% intergenic) and, given that 5% of the mammalian genome is under selection, almost all of this is likely to be of functional significance.

Comparative genomics has identified ~400 ultra-conserved regions (UCR) greater than 200 bp sharing at least 95% sequence identity between human and chicken (Sandelin et al. 2004). Surprisingly, highly conserved, noncoding regions like the UCR often exist far from any predicted gene within so-called “gene deserts” that are apparently free of any known protein-coding genes and are often clustered (Ovcharenko et al. 2005). Genes with a role in transcriptional regulation and development flank many of these UCR and gene deserts. These regions are often far from genes and represent distant regulatory signals.

Parent-specific gene expression by genomic imprinting is only found in mammals and not birds or lower vertebrates. Therefore, comparison of imprinted genes in mammals with orthologs in the chicken may uncover features about the origins of imprinting. Comparative mapping suggests these genes cluster on macrochromosomes in regions that preferentially undergo asynchronous DNA replication (Dunzinger et al. 2005). Analysis of the chicken region orthologous to the imprinted mammalian *ASCL2-H19* region (Yokomine et al. 2005) revealed extensive conservation of gene organization, except *H19*, a critical noncoding imprinted gene. This gene and its regulatory elements were absent from the chicken genome. These studies suggest that

imprinted genes were clustered before the evolution of imprinting, an event that occurred after the divergence of birds and mammals ~310 Mya. Subsequently, imprinting control elements, such as the *H19* gene region, must have evolved by duplication and/or transposition into these gene clusters.

A long-standing question in genome evolution has been the question of genome size. The chicken genome is 35% the size of the human and 45% of mouse. In part, this can be explained in terms of the low frequency of repeats, pseudogenes, segmental duplication, and gene duplications (Hillier et al. 2004). However, these factors only account for 20%–25% of the variation in genome size, so other factors are at work, possibly a dearth of ancient repeats (that are no longer detectably repetitive) or reduction in cell size and energy conservation (Hughes and Piontkivska 2005).

#### Developmental biology

Applications in developmental biology are likely to be another major beneficiary of the genome sequence (Burt 2004b; Stern 2005). The chicken has always been a favorite among developmental biologists (Brown et al. 2003; Stern 2005) because of easy access to the chick embryo and ease of manipulation. These features, when combined with the new tools of genomics, are ideal for testing gene function and predicted regulatory sequences in vivo. For example, studies on the conservation of the avian *SOX2* genes have identified neural specific enhancers, confirmed in vivo by electroporation of chick embryo neural tubes (Uchikawa et al. 2004).

In the mouse and other model systems, whole-mount in situ hybridization screens have been useful in identifying patterns of expression that may suggest developmental functions of novel genes (EMAP). A similar effort has started in the chicken using the large collection of sequenced chicken ESTs (Boardman et al. 2002; ARK-Genomics; ChickEST). Data can be accessed at GEISHA and standard three-dimensional embryo reconstructions are under development (EMAP).

#### Genetic variation and complex trait analysis

In parallel with the chicken genome sequencing project, a consortium (Wong et al. 2004; Wang et al. 2005b; ChickVD) gener-

ated 2.8 million SNPs from a comparison of the Red Jungle Fowl reference sequence and partial genome scans of Silkie, Broiler, and Layer lines. Nucleotide diversity ( $5 \times 10^{-3}$  per nucleotide) was six times the rate found in humans (Ellegren 2005). Resequencing confirmed 94% of the total and 83% of the nonsynonymous SNPs. An initial surprise was that ~70% of SNPs were common to all breeds, suggesting an origin prior to domestication 5,000–10,000 years ago. Another possibility is that their ancestry has been lost because of extensive cross breeding between Asian and western poultry populations. The next steps are to verify a larger sample of SNPs and create high-resolution genetic and linkage disequilibrium maps of chicken populations. These assays will be used to map and identify genes controlling traits of economic and biological interest at quantitative trait loci (QTL). Currently, more than 600 QTL have been mapped using microsatellites (Andersson and Georges 2004; Hocking 2005; Wang et al. 2005b). The availability of a standard set of 10,000 or more SNPs combined with the ease of building structured large resource populations hold much promise towards the identification of genes controlling these traits.

#### *Animal health and the avian immune system*

One area that has benefited most from genomic approaches has been the characterization of the genes and proteins in the avian immune system. The MHC was the first major chicken genome sequence to be assembled (Kaufman et al. 1999) and was a surprise, being relatively compact and simpler than those of mammals. Since then, there has been slow progress in the isolation of avian cytokines and other signaling molecules. The main problem has been their high rate of evolution, limiting their detection using homology to mammalian sequences (Staeheli et al. 2001). Even now, one must be careful in concluding that avian homologs to mammalian immune genes do not exist, as several examples known from ESTs or directed sequencing were not found in the genome assembly. This started to change when analysis of large EST data sets identified 185 immune-related sequences (Lynn et al. 2003; Smith et al. 2004). This compared with the 80 genes identified by Tirunagaru et al. (2000) and the 28 genes listed in the review by Staeheli et al. (2001). Sequences included interleukins, transcription factors, chemokines, differentiation antigens, receptors, genes involved in the Toll pathway, and MHC-associated genes. The discovery of *IL4* and other cytokines involved in the Th2 response (Smith et al. 2004) was a surprise, since it had previously been speculated that the chicken does not elicit a typical Th2 response (Staeheli et al. 2001). The receptors for *IL10* and *IL13* were also identified, indicating that the chicken probably also contained these genes, which are typical Tr1 and Th2 cytokines. This was confirmed by sequencing specific BAC clones identified assuming conservation of synteny between chicken and mammalian genomes (Avery et al. 2004; Rothwell et al. 2004).

A comprehensive analysis of the chicken genome sequence has identified many cytokines, chemokines, and their receptors (Hillier et al. 2004; Kaiser et al. 2004, 2005; Wang et al. 2005a). Even genes once thought to be mammalian-specific, including *IL3*, *IL7*, *IL9*, *IL26*, *CSMF*, *LIF*, and Cathelicidin, were found (Hillier et al. 2004). These are proteins that evolve rapidly and require more effort to detect. A number of orthologs to human chemokines are absent from the chicken genome, including *CCL2*, *7*, *8*, *11*, *15*, *18*, *23*, *24*, and *26*; *CXCL1-7*, *9*, *10*, and *11*, possibly products of independent gene duplications in mam-

mals. Similarly, missing chemokine receptors included *CCR1*, *CCR3*, *CCR10*, *CXCR3*, and *CXCR6*. The lack of functional eosinophils correlates with the absence of the eotaxin genes (*CCL22*, *CCL24*, *CCL26*) and their receptor (*CCR3*). Chickens lack lymph nodes and also the genes for the lymphotoxins (LT- $\alpha$  and - $\beta$ ) and their receptors. *TNF* is also absent, but its receptor, *TNFRSF1A* (*ENSGALG00000014890*) is present, suggesting that further sequencing will reveal this gene in the chicken. Similar analyses have been performed on the leukocyte receptor complex (Nikolaidis et al. 2005) that regulates the activity of T- and B-lymphocytes and NK cells. A model of evolution by repeated birth and death of these Ig-like receptors' genes was proposed.

#### Conclusions

When the first issue of *Genome Research* appeared 10 years ago, avian genomics was still in a mapping phase (Burt and Cheng 1998). The idea of sequencing the chicken genome was only a dim possibility and comparative maps were hailed as an alternative mapping resource. As the first livestock species to be fully sequenced, the chicken genome sequence is a landmark in both avian biology and agriculture. The avian community was small but has grown rapidly in the last two years thanks to the EST and genome sequencing programs. The challenge now is to keep the momentum going and to exploit these resources. The creation of AvianNET, an organization to encourage the exchange of tools and resources in avian biology, is a start but only a beginning. The chicken genome was determined to inform us about the nature and function of the human genome. It has also informed us about the nature of birds and other vertebrates. With 9600 extant avian species, there is still a lot to learn. Birds, in particular, poultry and ducks are a source of many infectious diseases (Avian Flu: Web Focus 2005) and genomics is going to tell us a lot about host responses to these pathogens. There is therefore a need to sequence and characterize other avian genomes. This time these sequences will be used to inform us about responses to pathogens that infect both humans and birds.

#### Acknowledgments

I would like to thank many colleagues and collaborators for their continued support and enthusiasm on issues related to avian genomics and acknowledge financial support from the Biotechnology and Biological Science Research Council (UK). In addition, I would like to thank the useful comments and suggestions from the anonymous reviewers.

#### References

- Abdrakhmanov, I., Lodygin, D., Geroth, P., Arakawa, H., Law, A., Plachy, J., Korn, B., and Buerstedde, J.M. 2000. A large database of chicken bursal ESTs as a resource for the analysis of vertebrate gene function. *Genome Res.* **10**: 2062–2069.
- Andersson, L. and Georges, M. 2004. Domestic-animal genomics: Deciphering the genetics of complex traits. *Nat. Rev. Genet.* **5**: 202–212.
- Antin, P.B. and Konieczka, J.H. 2005. Genomic resources for chicken. *Dev. Dyn.* **232**: 877–882.
- Avery, S., Rothwell, L., Degen, W.D., Schijns, V.E., Young, J., Kaufman, J., and Kaiser, P. 2004. Characterization of the first non-mammalian T2 cytokine gene cluster: The cluster contains functional single-copy genes for IL-3, IL-4, IL-13, and GM-CSF, a gene for IL-5 that appears to be a pseudogene, and a gene encoding another cytokine-like transcript, KK34. *J. Interferon Cytokine Res.* **24**: 600–610.

- Axelsson, E., Webster, M.T., Smith, N.G.C., Burt, D.W., and Ellegren, H. 2005. Comparison of the chicken and turkey genomes reveals a higher rate of nucleotide divergence on microchromosomes than macrochromosomes. *Genome Res.* **15**: 120–125.
- Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R., Wilson, S.A., and Hubbard, S.J. 2002. A comprehensive collection of chicken cDNAs. *Curr. Biol.* **12**: 1965–1969.
- Bourque, G., Zdobnov, E.M., Bork, P., Pevzner, P.A., and Tesler, G. 2005. Comparative architectures of mammalian and chicken genomes reveal highly variable rates of genomic rearrangements across different lineages. *Genome Res.* **15**: 98–110.
- Brown, W.R., Hubbard, S.J., Tickle, C., and Wilson, S.A. 2003. The chicken as a model for large-scale analysis of vertebrate gene function. *Nat. Rev. Genet.* **4**: 87–98.
- Burnside, J., Neiman, P., Tang, J., Basom, R., Talbot, R., Aronszajn, M., Burt, D.W., and Delrow, J. 2005. Development of a cDNA array for chicken gene expression analysis. *BMC Genomics* **6**: 13.
- Burt, D.W. 2002. Applications of biotechnology in the poultry industry. *Worlds Poult. Sci. J.* **58**: 5–13.
- . 2004a. Chicken genomics charts a path to the genome sequence. *Brief. Funct. Genomic Proteomic* **3**: 60–67.
- . 2004b. The chicken genome and the developmental biologist. *Mech. Dev.* **121**: 1129–1135.
- Burt, D.W. and Cheng, H.H. 1998. Chicken gene maps. *ILAR J.* **39**: 229–236.
- Burt, D.W., Bruley, C.K., Dunn, I., Jones, C.T., Ramage, A., Law, A.S., Morrice, D.R., Paton, I.R., Smith, J., Windsor, D., et al. 1999. Dynamics of chromosome evolution: Clues from comparative gene mapping in birds and mammals. *Nature* **402**: 411–413.
- Caldwell, R.B., Kierzek, A.M., Arakawa, H., Bezzubov, Y., Zaim, J., Fiedler, P., Kutter, S., Blagodatski, A., Kostovska, D., Koter, M., et al. 2004. Full-length cDNAs from chicken bursal lymphocytes to facilitate gene function analysis. *Genome Biol.* **6**: R6.
- Castelo, R., Reymond, A., Wyss, C., Camara, F., Parra, G., Antonarakis, S.E., Guigó, R., and Eyra, E. 2005. Comparative gene finding in chicken indicates that we are closing in on the set of multi-exonic widely expressed human genes. *Nucleic Acids Res.* **33**: 1935–1939.
- Dequeant, M.L. and Pourquie, O. 2005. Chicken genome: New tools and concepts. *Dev. Dyn.* **232**: 883–886.
- Dunzinger, U., Nanda, I., Schmid, M., Haaf, T., and Zechner, U. 2005. Chicken orthologues of mammalian imprinted genes are clustered on macrochromosomes and replicate asynchronously. *Trends Genet.* **21**: 488–492.
- Edwards, S.V., Jennings, B.W., and Shedlock, A.M. 2005. Phylogenetics of modern birds in the era of genomics. *Proc. Royal Soc. B.* **272**: 979–992.
- Ellegren, H. 2005. The avian genome uncovered. *Trends Ecol. Evol.* **20**: 180–186.
- Eyra, E., Reymond, A., Castelo, R., Bye, J.M., Camara, F., Flicke, P., Huckle, E.J., Parra, G., Shteynberg, D.D., Wyss, C., et al. 2005. Gene finding in the chicken genome. *BMC Bioinformatics* **6**: 131.
- Fazzari, M.J. and Grealley, J.M. 2004. Epigenomics: Beyond CpG islands. *Nat. Rev. Genet.* **5**: 446–455.
- Fumihito, A., Miyake, T., Sumi, S., Takada, M., Ohno, S., and Kondo, N. 1994. One subspecies of the red junglefowl (*Gallus gallus gallus*) suffices as the matriarchic ancestor of all domestic breeds. *Proc. Natl. Acad. Sci.* **91**: 12505–12509.
- Haldane, J.B.S. 1927. The comparative genetics of color in rodents and carnivora. *Biol. Rev. Camb. Philos. Soc.* **2**: 199–212.
- Hedges, S.B. 2002. The origin and evolution of model organisms. *Nat. Rev. Genet.* **3**: 838–849.
- Hill, W.G. and Robertson, A. 1966. The effect of linkage on limits to artificial selection. *Genet. Res.* **8**: 269–294.
- Hillier, L.W., Miller, W., Birney, E., Warren, W., Hardison, R.C., Ponting, C.P., Bork, P., Burt, D.W., Groenen, M.A., Delany, M.E., et al. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* **432**: 695–716.
- Hocking, P. 2005. Review on QTL mapping results in chickens. *Worlds Poult. Sci. J.* **61**: 215–226.
- Hubbard, S.J., Grafham, D.V., Beattie, K.J., Overton, I.M., McLaren, S.R., Croning, M.D., Boardman, P.E., Bonfield, J.K., Burnside, J., Davies, R.M., et al. 2005. Transcriptome analysis for the chicken based on 19,626 finished cDNA sequences and 485,337 expressed sequence tags. *Genome Res.* **15**: 174–183.
- Hughes, A.L. and Piontkivska, H. 2005. DNA repeat arrays in chicken and human genomes and the adaptive evolution of avian genome size. *BMC Evol. Biol.* **5**: 12.
- Kaiser, P., Rothwell, L., Avery, S., and Balu, S. 2004. Evolution of the interleukins. *Dev. Comp. Immunol.* **28**: 375–394.
- Kaiser, P., Poh, T.Y., Rothwell, L., Avery, S., Balu, S., Pathania, U.S., Hughes, S., Goodchild, M., Morrell, S., Watson, M., et al. 2005. A genomic analysis of chicken cytokines and chemokines. *J. Interferon Cytokine Res.* **25**: 467–484.
- Kaufman, J., Milne, S., Gobel, T.W., Walker, B.A., Jacob, J.P., Auffray, C., Zoorob, R., and Beck, S. 1999. The chicken B locus is a minimal essential major histocompatibility complex. *Nature* **401**: 923–925.
- Korf, I., Flicke, P., Duan, D., and Brent, M.R. 2001. Integrating genomic homology into gene structure prediction. *Bioinformatics* **17 Suppl 1**: 140–148.
- Lynn, D.J., Lloyd, A.T., and O'Farrelly, C. 2003. In silico identification of components of the Toll-like receptor (TLR) signaling pathway in clustered chicken expressed sequence tags (ESTs). *Vet. Immunol. Immunopathol.* **93**: 177–184.
- Masabanda, J.S., Burt, D.W., O'Brien, P.C., Vignal, A., Fillon, V., Walsh, P.S., Cox, H., Tempest, H.G., Smith, J., Habermann, F., et al. 2004. Molecular cytogenetic definition of the chicken genome: The first complete avian karyotype. *Genetics* **166**: 1367–1373.
- McQueen, H.A., Fantes, J., Cross, S.H., Clark, V.H., and Archibald, A.L. 1996. CpG islands of chicken are concentrated on microchromosomes. *Nat. Genet.* **12**: 321–324.
- Meunier, J. and Duret, L. 2004. Recombination drives the evolution of GC-content in the human genome. *Mol. Biol. Evol.* **21**: 984–990.
- Muscarella, D.E., Vogt, V.M., and Bloom, S.E. 1985. The ribosomal RNA gene cluster in aneuploid chickens: Evidence for increased gene dosage and regulation of gene expression. *J. Cell Biol.* **101**: 1749–1756.
- Nakamura, H., Katahira, T., Sato, T., Watanabe, Y., and Funahashi, J.-I. 2004. Gain- and loss-of-function in chick embryos by electroporation. *Mech. Dev.* **121**: 1137–1143.
- Nikolaidis, N., Makalowska, I., Chalkia, D., Makalowski, W., Klein, J., and Nei, M. 2005. Origin and evolution of the chicken leukocyte receptor complex. *Proc. Natl. Acad. Sci.* **102**: 4057–4062.
- Ovcharenko, I., Loots, G.G., Nobrega, M.A., Hardison, R.C., Miller, W., and Stubbs, L. 2005. Evolution and functional classification of vertebrate gene deserts. *Genome Res.* **15**: 137–145.
- Parra, G., Agarwal, P., Abril, J.F., Wiehe, T., Fickett, J.W., and Guigó, R. 2003. Comparative gene prediction in human and mouse. *Genome Res.* **13**: 108–117.
- Ponce de Leon, F.A., Li, Y., and Weng, Z. 1992. Early and late replicative chromosomal banding patterns of *Gallus domesticus*. *J. Hered.* **83**: 36–42.
- Rothwell, L., Young, J.R., Zoorob, R., Whittaker, C.A., Hesketh, P., Archer, A., Smith, A.L., and Kaiser, P. 2004. Cloning and characterization of chicken IL-10 and its role in the immune response to *Eimeria maxima*. *J. Immunol.* **173**: 2675–2682.
- Sandelin, A., Bailey, P., Bruce, S., Engström, P.G., Klos, J.M., Wasserman, W.W., Ericson, J., and Lenhard, B. 2004. Arrays of ultraconserved non-coding regions span the loci of key developmental genes in vertebrate genomes. *BMC Genomics* **5**: 99.
- Sang, H. 2004. Prospects for transgenesis in the chick. *Mech. Dev.* **121**: 1179–1186.
- Schmid, M., Nanda, I., Hoehn, H., Scharlt, M., Haaf, T., Buerstedde, J.M., Arakawa, H., Caldwell, R.B., Weigend, S., Burt, D.W., et al. 2005. Second report on chicken genes and chromosomes 2005. *Cytogenet. Genome Res.* **109**: 415–479.
- Smith, J., Bruley, C.K., Paton, I.R., Dunn, I., Jone, C.T., Windsor, D., Morrice, D.R., Law, A.S., Masabanda, J., Sazanov, A., et al. 2000. Differences in gene density on the Chicken macrochromosomes and microchromosomes: A tool for gene discovery in vertebrate genomes. *Anim. Genet.* **31**: 96–103.
- Smith, J., Speed, D., Law, A.S., Glass, E.J., and Burt, D.W. 2004. In-silico identification of chicken immune-related genes. *Immunogenetics* **56**: 122–133.
- Staehele, P., Puehler, F., Schneider, K., Göbel, T.W., and Kaspers, B. 2001. Cytokines of birds: Conserved functions—a largely different look. *J. Interferon Cytokine Res.* **21**: 993–1010.
- Stern, C.D. 2004. The chick embryo—Past, present and future as a model system in developmental biology. *Mech. Dev.* **121**: 1011–1013.
- . 2005. The chick: A great model system becomes even greater. *Dev. Cell* **8**: 9–17.
- Tirunagaru, V.G., Sofer, L., Cui, J., and Burnside, J. 2000. An expressed sequence tag database of T-cell enriched activated chicken splenocytes: Sequence analysis of 5251 clones. *Genomics* **66**: 144–151.
- Uchikawa, M., Takemoto, T., Kamachi, Y., and Kondoh, H. 2004. Efficient identification of regulatory sequences in the chicken genome by a powerful combination of embryo electroporation and genome comparison. *Mech. Dev.* **121**: 1145–1158.
- Wallis, J.W., Aerts, J., Groenen, M.A., Crooijmans, R.P., Layman, D., Graves, T.A., Scheer, D.E., Kremitzki, C., Fedele, M.J., Mudd, N.K., et

- al. 2004. A physical map of the chicken genome. *Nature* **432**: 761–764.
- Wang, J., Adelson, D.L., Yilmaz, A., Sze, S.H., Jin, Y., and Zhu, J.J. 2005a. Genomic organization, annotation, and ligand-receptor inferences of chicken chemokines and chemokine receptor genes based on comparative genomics. *BMC Genomics* **6**: 45.
- Wang, J., He, X., Dai, M., Ruan, J., Chen, J., Zhang, Y., Hu, Y., Ye, C., Li, S., Cong, L., et al. 2005b. ChickVD: A sequence variation database in the chicken genome. *Nucleic Acids Res.* **33**: D438–D441.
- Wong, G.K., Liu, B., Wang, J., Zhang, Y., Yang, X., Zhang, Z., Meng, Q., Zhou, J., Li, D., Zhang, J., et al. 2004. A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. *Nature* **432**: 717–722.
- Yokomine, T., Shirohzu, H., Purbowasito, W., Toyoda, A., Iwama, H., Ikeo, K., Hori, T., Mizuno, S., Tsudzuki, M., Matsuda, Y., et al. 2005. Structural and functional analysis of a 0.5-Mb chicken region orthologous to the imprinted mammalian *Ascl2/Mash2-Igf2-H19* region. *Genome Res.* **15**: 154–165.