Poultry Genome Sequences: Progress and Outstanding Challenges

J.B. Dodgson^a M.E. Delany^b H.H. Cheng^c

^aDepartment of Microbiology and Molecular Genetics, Michigan State University, East Lansing, Mich.,

^bDepartment of Animal Science, University of California, Davis, Calif., and ^cUSDA-ARS, Avian Disease and Oncology Laboratory, East Lansing, Mich., USA

Key Words

BAC contig map · Comparative genome map · Copy number variation · Genome sequence · Genome-wide marker-assisted selection · Single nucleotide polymorphism

Abstract

The first build of the chicken genome sequence appeared in March, 2004 – the first genome sequence of any animal agriculture species. That sequence was done primarily by whole genome shotgun Sanger sequencing, along with the use of an extensive BAC contig-based physical map to assemble the sequence contigs and scaffolds and align them to the known chicken chromosomes and linkage groups. Subsequent sequencing and mapping efforts have improved upon that first build, and efforts continue in search of missing and/or unassembled sequence, primarily on the smaller microchromosomes and the sex chromosomes. In the past year, a draft turkey genome sequence of similar quality has been obtained at a much lower cost primarily due to the development of ‘next-generation’ sequencing techniques. However, assembly and alignment of the sequence contigs and scaffolds still depended on a detailed BAC contig map of the turkey genome that also utilized comparison to the existing chicken sequence. These 2 land fowl (Galliformes) genomes show a remarkable level of similarity, despite an estimated 30–40 million years of separate evolution since their last common ancestor. Among the advantages

offered by these sequences are routine re-sequencing of commercial and research lines to identify the genetic correlates of phenotypic change (for example, selective sweeps), a much improved understanding of poultry diversity and linkage disequilibrium, and access to high-density SNP typing and association analysis, detailed transcriptomic and proteomic studies, and the use of genome-wide marker-assisted selection to enhance genetic gain in commercial stocks.

Copyright © 2011 S. Karger AG, Basel

Getting the Genomes Sequenced

By the end of the 20th century, it was clear that the draft human genome sequence would soon be completed and quickly would be followed by a ‘finished’ human genome and a high-quality mouse genome sequence. The National Human Genome Research Institute (NHGRI) became interested in the genomes of other species that would provide evolutionary perspective for comparison to the human genome. Given the extensive history of the chicken as a model organism employed in many fundamental discoveries and the existence of an internationally-coordinated effort to map its genome, the chicken provided the logical ‘outgroup’ species [reviewed in Siegel et al., 2006]. A proposal to generate a draft (6× coverage) sequence of the chicken genome was submitted in 2002

Table 1. Summary of initial chicken and turkey draft genome sequences

	DNA source	Method ^a	Contigs >1 kb	Scaffolds >1 kb	Total coverage	Total aligned ^b	chrUn_random	Alignment method ^c	Cost ^d
Chicken ^e	inbred UCD001 Red Jungle Fowl	WGS Sanger	98,612	32,767	1,050 Mb	905 Mb	165 Mb ^f	BAC contig map	>10 M USD
Turkey ^g	partially inbred bird from closed commercial line ^h	WGS Roche/ 454 and Illumina	128,271	26,917	936 Mb	917 Mb	18.6 Mb	BAC contig comparative map	~0.2 M USD

^a Primary method of sequence acquisition. WGS = Whole genome shotgun. ^b Scaffolds ordered and oriented on specified chromosomes, excludes chrUn_random. ^c Primary method of placing and orienting scaffolds (supplemented with linkage map data). ^d Estimated cost of sequence acquisition (excludes physical map

construction and annotation). ^e International Chicken Genome Sequencing Consortium [2004]. ^f Size of chrUn_random in build 1 of the chicken genome; this was reduced to 63.9 Mb in build 2. ^g Dalloul et al. [2010]. ^h Chaves et al. [2009].

(http://www.genome.gov/Pages/Research/Sequencing/SeqProposals/Chicken_Genome.pdf) and subsequently given 'high priority' by NHGRI. When NHGRI and the Washington University Genome Sequencing Center (WUGSC) decided to sequence the chicken genome, the Beijing Genomics Institute (BGI) chose to complement that framework sequence with 'sample' (one fourth of the genome each) sequencing of broiler, layer (White Leghorn) and Silkie chicken DNAs in order to provide a dense single nucleotide polymorphism (SNP) map [International Chicken Polymorphism Map Consortium, 2004].

The 'chicken' genome that was sequenced is actually that of a single female bird of the UCD001 inbred red jungle fowl line [Abplanalp, 1992], red jungle fowl being the primary 'wild-type' version of the domestic chicken. DNA from this bird previously had been used to make several bacterial artificial chromosome (BAC) libraries [Lee et al., 2003]. These BACs allowed for the construction of a physical 'BAC contig' map by BAC fingerprinting and hybridization techniques. Both fingerprinting and sequencing benefit from using DNA from a single inbred individual because heterozygous polymorphisms that could be confused with sequencing errors or fingerprint differences are minimal. The sequence acquisition approach used was primarily 'whole genome shotgun (WGS)' sequencing of small insert clone libraries, in some cases using 'mate pairs', sequences of both ends of a DNA fragment whose spacing in the genome can be approximated based on the selected size of library inserts. The WGS approach yields many thousands of sequence contigs (overlapping sequence reads) that are assembled into larger scaffolds by linking and spacing contigs with mate

pair reads (table 1). However, the sequence remained in tens of thousands of scaffolds that could only be assembled into a higher-order alignment with chicken chromosomes using BAC end sequences, the BAC contig map [Wallis et al., 2004] and other mapping data. The initial sequence assembly was made public on March 1, 2004, and its initial analysis was published that December [International Chicken Genome Sequencing Consortium, 2004], along with the results of the BGI SNP sample sequencing [International Chicken Polymorphism Map Consortium, 2004]. A second build based on directed sequencing of problem areas plus an assembly using additional linkage and radiation hybrid map data appeared in May 2006. A third build based only on 12× Roche/454 sequences (see below) has been done by the WUGSC, and this is currently being melded with the earlier Sanger sequence data into a new assembly in collaboration with the University of Maryland Center for Bioinformatics and Computational Biology.

It is interesting to contrast the chicken genome sequence effort to the recent sequencing of the turkey genome [Dalloul et al., 2010]. The former required over USD 10,000,000 of NIH support and was done by one of the major NHGRI sequencing centers, whereas the latter was done by the Virginia Bioinformatics Institute at the Virginia Tech and the USDA-ARS Beltsville Agricultural Research Center with a consortium of seed funding totaling about USD 200,000. The final products are roughly similar in terms of coverage and contig and scaffold length (table 1). The difference in cost relates to the use of 'next-generation' sequencing (NGS) technology employing both the Roche/454 and Illumina (formerly Solexa) platforms. Once again, a necessary adjunct to assembling

turkey sequence scaffolds and aligning them with chromosomes was a high-resolution BAC contig map (not included in the cost figures cited in table 1). That map also was based on a comparative alignment between turkey sequences and probes and the existing chicken sequence [Zhang et al., manuscript in preparation].

Some Highlights from the Genomes

The initial conclusions based on the chicken sequence are described in detail by the International Chicken Genome Sequencing Consortium [2004] and in a series of companion studies [Abril et al., 2005; Axelsson et al., 2005; Bourque et al., 2005; Hubbard et al., 2005; Kellner et al., 2005; Ovcharenko et al., 2005a, b; Paulsen et al., 2005; Wicker et al., 2005; Yokomine et al., 2005]. The chicken genome fulfilled its promise in terms of evolutionary comparisons to mammalian genomes. The distance (~300 M years to the last common ancestor) between birds and mammals is such that essentially no random background homology in their sequences exists, so any regions of similarity found between them can be assumed to be under evolutionary ('purifying') selection, which strongly suggests that they are functionally significant. This made it relatively straightforward to identify the approximately 70 Mb (~6% of the chicken genome, ~2% of mammalian genomes) that are likely to be functional in both birds and mammals. Surprisingly, over half of the regions under selective pressure are not within or near gene-coding sequences. Indeed, they are often within so-called 'gene deserts' [Ovcharenko et al., 2005b]. The reasons for this remain unclear, although non-coding RNAs and regulatory regions that exert their effect on distant genes are likely to provide at least some of the explanation. Comparison between chicken and mammalian genomes also confirmed the earlier estimates [Burt et al., 1999; Suchy et al., 2001] of relatively long blocks of common gene order between the two. This implies that the nearly 3-fold difference in their genome sizes is mostly due to the expansion and contraction of interspersed repetitive sequence families (transposable elements) that rarely disrupt gene order in the process. In particular, avian genomes generally lack short interspersed nuclear element (SINE) repeats, indicating that this class of transposable element has been inactive for at least 50 M years in birds. Retrotranscribed pseudogenes are also rare in avian genomes, apparently due to the high specificity of the reverse transcriptase encoded by the primary LINE

(long interspersed nuclear element) repeat element in birds, CR1 (chicken repeat 1).

Comparisons between the turkey and chicken genomes illuminate a smaller window of separate evolution (30–40 million years), but they confirm many of the above conclusions [Dalloul et al., 2010]. First, the overall structure of the 2 genomes is remarkably similar, much more than is seen in comparing rat to mouse or human to gibbon, even though these 2 pairs have been separated from their last common ancestors by only about half the time as have chicken and turkey. To date, only 2 possible translocations between the chicken and turkey genomes have been found, and these are small enough that they could have been repeated sequences in the ancestral genome of which turkey retained 1 copy and chicken another, or they could be translocations due to the action of transposable elements. In total, there are about 30 large-scale (>100 kb) chromosomal rearrangements that distinguish the chicken and turkey genomes, most of which are inversions, with the most common site of inversions being towards the p end of orthologous chromosomes. In addition to confirming the high stability of avian genomes, the turkey sequence and its companion physical/comparative map demonstrate a couple of other evolutionary trends in avian chromosome evolution. First, turkey chromosomes tend to be acrocentric with small, if any, p arms in comparison to those of chicken. As one example, figure 1 shows that the p arm of GGA11 is inverted in the orthologous MGA13 chromosome. MGA13 forms a single contig in the comparative turkey-chicken BAC contig map with no internal centromere. Other than MGA1 and probably MGA5 and 25, there are no turkey autosomes for which there is clear evidence of a p arm that comprises a significant portion of the chromosome. Second, several of the rearrangements that distinguish the 2 genomes are likely due to unequal recombination between members of repeated gene families. This observation is in line with another conclusion from the initial chicken/mammalian genome comparison: that expansion and contraction of multigene families seems to be a major factor in genome evolution [International Chicken Genome Sequencing Consortium, 2004].

Avian Genome Sequences as Tools for the Future

Having their genomes sequenced makes chickens and turkeys 'members of the club' of species in which molecular genetic research can be conducted using the full power of modern genomics. The sequences directly pro-

Fig. 1. Inversion of the GGA11 p arm when compared to the orthologous MGA13 chromosome. CHORI-260 BAC 91J22 (GGA11 end coordinates: 183,843 and 3,729,465) is labeled in green (FITC), while 109I12 (GGA11 end coordinates: 75,337 and 270,206) is labeled with red (Enzo Red). Both were used for FISH analysis of chicken and turkey pachytene chromosomes. The chicken chromosome spread (left frame) affirms that 91J22 contains sequences that are on both the distal GGA11p and the proximal end of GGA11q as indicated by the BES alignments noted above (the

centromere on GGA11 is located in the sequence at coordinates 2,688,052–3,188,051). BAC 109I12, as expected, labels a single site on distal GGA11p. In the turkey spread (right frame), 91J22 labels a single site that co-localizes with 109I12 internal to MGA13. The inversion joining sequences homologous to distal GGA11p to those on proximal GGA11q is confirmed in the BAC contig comparative map by fingerprint and probe hybridization analysis (not shown), in which the entire length of MGA13 forms a single BAC contig, indicative of no internal centromere.

vide lists of genes, non-coding RNAs and repetitive elements and their individual sequences and arrangements with respect to one another. This gene map is an essential tool for the identification of positional candidate genes that may encode interesting mutations or causal polymorphisms that result in quantitative trait loci (QTL, alleles that contribute to quantitative phenotypic variation). For example, the chicken sequence has already contributed to understanding the molecular basis for numerous single gene mutation phenotypes [e.g. Kerje et al., 2004; Gunnarsson et al., 2007; Elferink et al., 2008; Eriksson et al., 2008; Dorshorst and Ashwell, 2009; Wright et al., 2009; Dorshorst et al., 2010]. However, the genome sequences play an even more critical role in the generation of high-density linkage and association maps and as facilitators of transcriptomic and proteomic analyses.

As mentioned earlier, partial sequences of other chickens were compared to the reference red jungle fowl sequence to generate about 3 million SNPs [International Chicken Polymorphism Map Consortium, 2004]. Initial-

ly, a set of about 3,000 highly polymorphic, evenly spaced SNPs were selected and genotyped on 2,580 birds that spanned most of worldwide chicken diversity [Muir et al., 2008]. Among other things, the results showed that layer stocks are significantly less diverse than meat-type commercial chickens, and that current commercial breeds have lost about half of the alleles present in the ancestral chickens likely to have contributed to their formation. Recent high-density SNP-typing efforts have increased linkage map resolution to 13,000 markers [Groenen et al., 2009] and currently a 60K SNP chip is in use, with prospects for a 500K or larger SNP chip in the not-too-distant future. These SNP maps allow fine-scale linkage and/or association mapping of at least the primary components of multigenic QTL. Previously, the density of useful polymorphic markers was the primary limitation on resolving QTL to workable genomic intervals in which a candidate causal polymorphism might be sought. With dense SNP maps, the level of linkage disequilibrium (LD) in the population segregating for the QTL becomes the limiting

factor for causal QTL allele identification via association mapping. The caveat, however, is that with large numbers of markers (and tests), one requires substantial populations of carefully phenotyped individuals to localize QTL alleles with statistical significance, especially those alleles with only moderate phenotypic effect.

Perhaps the most significant impact of the high-density SNP maps is that they may allow breeders to skip over the QTL-encoding gene ascertainment step altogether and go directly to measuring breeding values using genome-wide SNP profiles (genome-wide marker-assisted selection; GWMAS) [Meuwissen et al., 2001]. The impact of GWMAS is already being tested in commercial chickens and promises to significantly speed up breeding progress, especially in layers. There are still questions to be answered in terms of how much effort and cost must be committed to ascertaining and regularly updating the genome-wide molecular breeding value models used in GWMAS, but genotyping costs continue to drop such that accurate phenotype analysis becomes the primary economic burden for GWMAS in chickens.

In addition to SNP analysis, attention is starting to be paid to copy number variations (CNV). Since CNV regions are problematic for most mapping approaches and for sequence assembly, they were missed or at least ignored early on. However, there are now clear indications that CNVs are involved in some dramatic phenotypic differences (e.g. late feathering [Elferink et al., 2008]; pea-comb [Wright et al., 2009]) and are deserving of more careful study.

Another technology that becomes available with access to a genome sequence is transcriptomics, the detailed measurement of relative or absolute changes in mRNA expression levels in any chicken tissue or cell line responding to any environmental effect (disease challenge, feed composition, age, etc.) of interest. For example, the chicken genome sequence led to the development of Affymetrix and Agilent oligonucleotide arrays that are now widely used to measure chicken mRNA levels [e.g. Heidari et al., 2008; Li et al., 2008]. Based on the recent completion of the turkey genome sequence, similar resources for turkeys should soon appear. An alternative to microarrays is the use of NGS to enumerate cDNA (or mRNA) copy numbers in any sample by sequencing it to great depth ('mRNA-seq' or 'digital gene expression' [reviewed in Wold and Myers, 2008]; both are advanced versions of SAGE, serial analysis of gene expression [Velculescu et al., 1995]). This approach also aids in the transcriptional annotation of the genome (locating transcribed exons, alternative splicing, start and stop

sites, etc.). Such an approach has already been used to characterize the chicken microRNA set [Burnside et al., 2008] and is currently being expanded to avian mRNA analyses. Although the microarray approach is still the most economical for large sample sizes, the RNA-seq techniques have the advantage that no a priori assumptions are made as to what genes or exons are likely to be transcribed. Thus, new, unannotated genes of interest can be discovered. Similarly, the genome sequence provides the required substrate for proteomic analysis of chicken tissues by high-throughput mass spectrometry, and the results of such experiments feed back to aid in further annotation of the sequence itself [Liu et al., 2006; Buza et al., 2007]. Basically, NGS provides an open-ended and powerful method to assay the sequence composition of any nucleic acid sample, RNA or DNA, whether that composition has been influenced by experimental conditions (bird treatments), tissue of origin or by fractionation methods, such as chromatin immunoprecipitation, DNase or chemical treatment of nuclei, DNA methylation, etc.

NGS techniques also open the way for economical 're-sequencing' of a variety of commercial and experimental chicken lines [Rubin et al., 2009; Eriksson et al., 2010; H. Cheng, personal communication]. This approach should be particularly powerful in identifying segments of the genome that have been fixed by selective sweeps for traits of commercial advantage or experimental design. With new 'third-generation' single-DNA polymerase sequencing techniques now coming on-line [Eid et al., 2009; Lipson et al., 2009], the cost of sequencing a vertebrate genome is rapidly approaching the USD 1,000 NHGRI-set goal, opening the way for every chicken of special commercial or research interest to have its genome sequenced. Note, however, that this is really only feasible when a high-quality, aligned reference sequence is already in hand, and that, just as with the turkey genome sequence, the cost and effort burdens shift mainly to data analysis rather than sequence acquisition. In addition, we can expect a host of other avian genome sequences to appear soon, providing context in which to place the evolution of the chicken and the turkey. A 6× assembly of the zebra finch genome has been completed by the WUGSC, and the BGI has completed the duck genome sequence [Warren et al., 2010]. The Genome 10K Community of Scientists [2009] includes over 5,000 bird species in their list of 16,203 vertebrates proposed for possible genome sequencing.

Some Outstanding Challenges

The first challenge for the chicken genome community is to actually ‘finish’ the sequence. While most of the chicken genome sequence appears to be of surprisingly high accuracy, about 5–10% of it is either missing altogether or exists as small, un-aligned segments in the randomly ordered collection referred to as chrUn_random (63.9 Mb in chicken build 2). Other than the small gaps between sequence contigs scattered throughout the genome, most of this missing sequence originates from small ‘microchromosomes’ and the W sex chromosome. Chicken chromosomes 29–31 and 33–38 have no assembled sequence, and chromosomes 16, 32, W and perhaps a couple of others are clearly missing much of their lengths. There are 2 likely explanations for this problem, both of which probably contribute. First, there is evidence that some of this sequence is never obtained, possibly because it is deleterious in *E. coli* and cannot be cloned, has a high GC content such that it is difficult to Sanger sequence, and/or is poorly recovered in DNA preparations. Second, since the density of BACs and BAC end sequences, as well as SNP markers, for these regions is low (in part due to poor clonability), and they may also contain unusually high numbers of simple repetitive sequences, it is often impossible to align these sequence contigs and scaffolds, leading to their placement on chrUn_random. As mentioned previously, a 12× Roche/454 chicken sequence was done by WUGSC for their third build. Since NGS approaches require no cloning step, it was hoped that this would obviate the problem. However, preliminary analysis of both chicken build 3 and of the turkey genome, sequenced totally with NGS technology, suggests that the smaller microchromosome sequences are still mostly among the missing. While their apparent size is not a problem, a similar situation exists with the Z and especially the W sex chromosome assemblies. Since a female bird was used for all sequencing and mapping libraries, both sex chromosomes are represented at only half the coverage of the autosomes, and the W chromosome in particular is known to be very repeat-rich. As a result, the Z chromosome assembly is less complete than those of most autosomes and the W chromosome assembly is nearly non-existent. Using libraries of selected Z chromosome DNA, the Page lab at the Whitehead Institute, MIT has completed the chicken Z sequence [Bellott et al., 2010], so the promise exists that the missing segments on microchromosomes and the W can be greatly improved. Most likely, both additional mapping and targeted NGS will be required.

A second major challenge is annotation of the chicken and turkey genomes. The turkey sequence provides a good example of this problem, since with its NGS approach, most of the cost and effort devoted to this genome is going into annotation and analysis rather than sequencing itself. Most annotated genes and non-coding RNAs in avian genomes are identified and named by their homology to mammalian genes (usually human and mouse genes where there are large dedicated annotation projects and a much larger information base). This works well for most genes, but often fails for those known to evolve more rapidly (such as those involved in pathogen recognition and host defense). The orthology approach works even less well for non-coding RNAs and DNA regulatory regions. Furthermore, genes responsible for traits of greatest agricultural interest (e.g. egg laying) are often among those most likely to be specific to the avian lineage and therefore refractile to annotation by orthology. It will remain a major challenge to identify the genes and regulatory networks that control these traits and phenotypes. The annotation challenge is only a part of the larger problem of analyzing and mining the exponentially growing data sets being generated by NGS and other genomic tools. This is especially problematic for the relatively small and underfunded groups of animal agriculture geneticists. Much of the progress outlined above has relied on the value of the chicken as a model organism for biomedicine, but this support mechanism often does not work for those aspects of avian biology most relevant to agriculture.

As noted above, the rapid progress of various genomic technologies in many cases has made collecting adequate numbers of accurate phenotypes the rate-limiting step in genetic analysis. Although still a challenge, the chicken is probably the best agricultural animal in which to collect phenotypes. The chicken has clear advantages over other domestic animals in terms of cost per bird and relatively short generation times. The standing world population of chickens is approximately 11 billion birds [Dohner, 2001], so it is clear that large populations can be maintained and studied. Chicken breeding companies are highly integrated and use sophisticated methods to collect multifactorial phenotypic data. New imaging techniques are being developed and employed to collect phenotypes such as carcass traits on live birds that previously were only available postmortem. Among the most costly and challenging phenotypes to study are those related to susceptibility or resistance to infectious diseases.

Finding and assembling the missing sequence, annotating the genomes, and collecting adequate numbers of

phenotypes are basically technical problems that eventually will be worked out. The over-arching problem, however, is to translate genome sequences, RNA and protein abundances and other outputs of genome technology into a true functional understanding of avian biology. In other words, how do we translate genomic data into systems biology? Here we are handicapped in part by the lack of inexpensive transgenic approaches and the cost of generating and maintaining large collections of mutant lines. However, the prospect of new approaches using RNA interference in live birds [Chen et al., 2009] and Zn-finger nuclease-targeted gene knockouts [Geurts et al., 2009] may provide new methods for specific gene modification and verifying the role of candidate QTL-encoding genes. An even greater handicap is the difficulty of melding the current data onslaught into cogent understanding or at least predictive utility. Unfortunately, there is no selective advantage to organisms having evolved in a simple fash-

ion (indeed, it could be argued that simple biological mechanisms would be disfavored by evolution as being inadequately homeostatic). Therefore, we need to construct, understand and communicate complex biological models to explain our species of interest. This is probably the greatest challenge that will be keeping poultry geneticists busy well through the 21st century. In facing this formidable challenge, we can be optimistic that new, currently unforeseeable tools, approaches and discoveries will continue to arise throughout biological research that can be applied to advance poultry science as has genomics technology over the last 20 years.

Acknowledgement

Research support was provided by the USDA National Institute for Food and Agriculture (AFRI 2008-35205-18720 and Multi-State Research Committee NRSP-8).

References

- Abplanalp H: Inbred lines as genetic resources of chickens. *Poult Sci Rev* 4:29–39 (1992).
- Abril JF, Castelo R, Guigo R: Comparison of splice sites in mammals and chicken. *Genome Res* 15:111–119 (2005).
- Axelsson E, Webster MT, Smith NG, Burt DW, Ellegren H: Comparison of the chicken and turkey genomes reveals a higher rate of nucleotide divergence on microchromosomes than macrochromosomes. *Genome Res* 15:120–125 (2005).
- Bellott DW, Skeletsky H, Pyntikova T, Mardis ER, Graves T, et al: Convergent evolution of chicken Z and human X chromosomes by expansion and gene acquisition. *Nature* 466:612–616 (2010).
- Bourque G, Zdobnov EM, Bork P, Pevzner PA, Tesler G: Comparative architectures of mammalian and chicken genomes reveal highly variable rates of genomic rearrangements across different lineages. *Genome Res* 15:98–110 (2005).
- Burnside J, Ouyang M, Anderson A, Bernberg E, Lu C, et al: Deep sequencing of chicken microRNAs. *BMC Genomics* 9:185 (2008).
- Burt DW, Bruley C, Dunn IC, Jones CT, Ramage A, et al: The dynamics of chromosome evolution in birds and mammals. *Nature* 402:411–413 (1999).
- Buza TJ, McCarthy FM, Burgess SC: Experimental confirmation and functional annotation of predicted proteins in the chicken genome. *BMC Genomics* 8:425 (2007).
- Chaves LD, Harry DE, Reed KM: Genome-wide genetic diversity of 'Nici', the DNA source for the CHORI-260 turkey BAC library and candidate for whole genome sequencing. *Anim Genet* 40:348–352 (2009).
- Chen M, Payne WS, Dunn JR, Chang S, Hunt H, et al: Retroviral delivery of RNA interference against Marek's Disease virus in vivo. *Poult Sci* 88:1373–1380 (2009).
- Dalloul RA, Long JA, Zimin AV, Aslam L, Beal K, et al: Multi-platform next-generation sequencing of the domestic turkey (*Meleagris gallopavo*): genome assembly and analysis. *PLoS Biol* 8:e1000475 (2010).
- Dohner JV: The Encyclopedia of Endangered Livestock and Poultry Breeds (Yale University Press, New Haven 2001).
- Dorshorst B, Okimoto R, Ashwell C: Genomic regions associated with dermal hyperpigmentation, polydactyly and other morphological traits in the Silkie chicken. *J Hered* 101:339–350 (2010).
- Dorshorst BJ, Ashwell CM: Genetic mapping of the sex-linked barring gene in the chicken. *Poult Sci* 88:1811–1817 (2009).
- Eid J, Fehr A, Gray J, Luong K, Lyle J, et al: Real-time DNA sequencing from single polymerase molecules. *Science* 323:133–138 (2009).
- Elferink MG, Vallée AA, Jungerius AP, Crooijmans RP, Groenen MA: Partial duplication of the *PRLR* and *SPEF2* genes at the late feathering locus in chicken. *BMC Genomics* 9:391 (2008).
- Eriksson J, Larson G, Gunnarsson U, Bed'hom B, Tixier-Boichard M, et al: Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. *PLoS Genet* 4:e1000010 (2008).
- Eriksson J, Rubin CJ, Zody MC, Meadows JRS, Sherwood E, et al: Identification of loci under selection during chicken domestication through whole genome sequencing. Proceedings, Plant and Animal Genomes XVIII, abstract (2010).
- Genome 10K Community of Scientists: Genome 10K: a proposal to obtain whole-genome sequence for 10,000 vertebrate species. *J Hered* 100:659–674 (2009).
- Geurts AM, Cost GJ, Freyvert Y, Zeitler B, Miller JC, et al: Knockout rats via embryo microinjection of zinc-finger nucleases. *Science* 325:433 (2009).
- Groenen MA, Wahlberg P, Foglio M, Cheng HH, Megens HJ, et al: A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. *Genome Res* 19:510–519 (2009).
- Gunnarsson U, Hellström AR, Tixier-Boichard M, Minvielle F, Bed'hom B, et al: Mutations in *SLC45A2* cause plumage color variation in chicken and Japanese quail. *Genetics* 175:867–877 (2007).
- Heidari M, Huebner M, Kireev D, Silva RF: Transcriptional profiling of Marek's disease virus genes during cytolytic and latent infection. *Virus Genes* 36:383–392 (2008).
- Hubbard SJ, Grafham DV, Beattie KJ, Overton IM, McLaren SR, et al: Transcriptome analysis for the chicken based on 19,626 finished cDNA sequences and 485,337 expressed sequence tags. *Genome Res* 15:174–183 (2005).
- International Chicken Genome Sequencing Consortium: Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432:695–716 (2004).

- International Chicken Polymorphism Map Consortium: a genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. *Nature* 432:717–722 (2004).
- Kellner WA, Sullivan RT, Carlson BH, Thomas JW: Uprobe: a genome-wide universal probe resource for comparative physical mapping in vertebrates. *Genome Res* 15:166–173 (2005).
- Kerje S, Sharma P, Gunnarsson U, Kim H, Bagchi S, et al: The *Dominant white*, *Dun* and *Smoky* color variants in chicken are associated with insertion/deletion polymorphisms in the *PMEL17* gene. *Genetics* 168:1507–1518 (2004).
- Lee MK, Ren CW, Yan B, Cox B, Zhang H-B, et al: Construction and characterization of three BAC libraries for analysis of the chicken genome. *Anim Genet* 34:151–152 (2003).
- Li X, Chiang HI, Zhu J, Dowd SE, Zhou H: Characterization of a newly developed chicken 44K Agilent microarray. *BMC Genomics* 9: 60 (2008).
- Lipson D, Raz T, Kieu A, Jones DR, Giladi E, et al: Quantification of the yeast transcriptome by single-molecule sequencing. *Nat Biotechnol* 27:652–658 (2009).
- Liu HC, Soderblom EJ, Goshe MB: A mass spectrometry-based proteomic approach to study Marek's Disease Virus gene expression. *J Virol Methods* 135:66–75 (2006).
- Meuwissen T, Hayes B, Goddard M: Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819–1829 (2001).
- Muir WM, Wong GK, Zhang Y, Wang J, Groenen MA, et al: Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. *Proc Natl Acad Sci USA* 105:17312–17317 (2008).
- Ovcharenko I, Loots GG, Giardine BM, Hou M, Ma J, et al: Mulan: multiple-sequence local alignment and visualization for studying function and evolution. *Genome Res* 15:184–194 (2005a).
- Ovcharenko I, Loots GG, Nobrega MA, Hardison RC, Miller W, Stubbs L: Evolution and functional classification of vertebrate gene deserts. *Genome Res* 15:137–145 (2005b).
- Paulsen M, Khare T, Burgard C, Tierling S, Walter J: Evolution of the Beckwith-Wiedemann syndrome region in vertebrates. *Genome Res* 15:146–153 (2005).
- Rubin CJ, Eriksson J, Ingman M, Jiang L, Meadows J, et al: Whole genome resequencing – a major leap forward in chicken genomics. *Proceedings, Plant and Animal Genomes XVII*, abstract (2009).
- Siegel PB, Dodgson JB, Andersson L: Progress from chicken genetics to the chicken genome. *Poult Sci* 85:2050–2060 (2006).
- Suchyta SP, Cheng HH, Burnside J, Dodgson JB: Comparative mapping of chicken anchor loci orthologous to genes on human chromosomes 1, 4 and 9. *Anim Genet* 32:12–18 (2001).
- Velculescu VE, Zhang L, Vogelstein B, Kinzler KW: Serial analysis of gene expression. *Science* 270:484–487 (1995).
- Wallis JW, Aerts J, Groenen M, Crooijmans R, Layman D, et al: A physical map of the chicken genome. *Nature* 432:761–764 (2004).
- Warren WC, Clayton DF, Ellegren H, Arnold AP, Hillier LW, et al: The genome of a songbird. *Nature* 464:757–762 (2010).
- Wicker T, Robertson JS, Schulze SR, Feltus FA, Magrini V, et al: The repetitive landscape of the chicken genome. *Genome Res* 15:126–136 (2005).
- Wold B, Myers RM: Sequence census methods for functional genomics. *Nat Methods* 5:19–21 (2008).
- Wright D, Boije H, Meadows JR, Bed'hom B, Gourichon D, et al: Copy number variation in intron 1 of *SOX5* causes the Pea-comb phenotype in chickens. *PLoS Genet* 5:e1000512 (2009).
- Yokomine T, Shirohzu H, Purbowasito W, Toyoda A, Iwama H, et al: 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 (2005).