THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Genetics. Chicken genome--science nuggets to come soon

OPEN ACCESS

GENETICS

Chicken Genome—Science Nuggets to Come Soon

Dave Burt and Olivier Pourquie

Many will be pleased to hear that the assembly of the draft sequence of the chicken genome is expected in March 2004. The chicken community has long suffered from its division between the agricultural world, interested in improving existing breeds by genetics in their immune function and associated pathologies, and the academic world, for which the chick embryo has constituted an important model for decades (1). In recent discussions aimed at providing a broad range of genetic resources, these worlds are coming together.

The chicken genome project had its origins in a number of genome mapping projects started 10 years ago, in which the nature of quantitative traits, such as growth, body weight, carcass composition, egg production, fatness, ascites, feather pecking, stress, etc., was of primary interest. The development of e-tools and resources in the chicken has facilitated large-scale studies to map monogenic (2–4); and more than 200 quantitative trait loci (QTLs) for a wide range of traits (5). These studies have demonstrated that the chicken is a powerful tool for studying the molecular basis of quantitative genetic variation.

To date, the species most closely related to mammals whose genomes have been sequenced are puffer fish with a common ancestor 400 million years ago. Because birds and mammals shared a common ancestor 300 million years ago, the chicken genome, as the first avian genome to be sequenced, will bridge a large evolutionary gap.

The chicken genome project will include a mixture of targeted and whole-genome sequencing (WGS). The high repeat content (~50%) in mammalian genomes has been a problem for the WGS strategy, however in the chicken the repeat content is only 15% (6). More than 75% of WGS sequences from chicken bacterial artificial chromosome (BAC) clones were contained in a single contig compared with only 25% in the mouse.

D. Burt is at the Roslin Institute (Edinburgh), Midlothian EH25 9PS, UK. E-mail: dave.burt@bbsrc.ac.uk O. Pourquie is at the Stowers Institute for Medical Research, Kansas City, MO 64110, USA. E-mail: OLP@Stowers-Institute.org

Now, a physical map based on 20-fold genome coverage of overlapping BAC clones is being assembled by investigators at Washington University, Texas A&M, Michigan State University, and Wageningen University. The physical map will be integrated with ARKdb (7–9), and other chicken mapping resources (10).

The Ensembl project (11), has expressed interest in coordinating the annotation of the chicken genome, provided dedicated funds to finance appropriate personnel are secured. Both automatic and community annotation will be necessary. At present, the sequences of chicken proteins available in the databases include two complementary sets of 1064 chicken proteins from the 122,564 proteins of the SwissProt database and 2047 chicken sequences of the 830,525 proteins of the TrEMBL databases. Development of new gene ontology terms will be required.

During the last 12 months there has been dramatic progress in the development of chicken EST resources with the chicken as one of the top five in the current release of dbEST (12). In total, there are now more than 600,000 chicken ESTs. Future plans are to sequence 10,000 full-length cDNA sequences within the next 12 months from these various collections (13). Thus far, there are already 3000 full-length chick cDNAs in GenBank.

Chicken complementary DNA microarrays have already been produced for studies on metabolism and immune responses (14). International collaboration between the UK and USA form the basis of the current generic cDNA microarrays being fabricated on two 14K arrays. These arrays will be made available to the chick community either from ARK-genomics (UK) or the Fred Hutchinson Cancer Research Center (USA). An initiative to develop a database of gene expression patterns during embryonic development has recently been launched and will include whole-mount chicken in situ hybridization images (15). It is also planned to link this effort to an adaptation to the chick embryo of the mouse developmental atlas (16). Plans are also under way to create RNA interference vectors to perform a systematic analysis of gene function during chick development (17, 18). A single-nucleotide polymorphism (SNP) map of the chicken would be a crucial tool in the search for QTLs. The Beijing Genome Institute, which recently completed the rice genome sequence, intends to generate one-fold genome coverage of both White Leghorn (L) and Broiler (B) chicken lines. When compared with the Jungle Fowl (JF) sequenced by Washington University, these sequences will generate a rich resource of SNPs.

More than 40 scientists representing the different areas of biology using the chicken as a working model, as well as specialists of genome science involved in the chick genome project, recently held a workshop at the Wellcome Trust Genome Campus in Cambridge (UK). During the workshop, an International Chicken Genome Consortium was established, that will provide coordination and leadership for research that will benefit from the chicken genome sequence. The consortium has been organized around a steering group and a number of technical groups, each responsible for delivering the research (sequencing, gene expression, chicken biology, proteomics). A Web site, ChickNET (19), is planned to serve as a gateway for investigators with an interest in the chicken genome, developmental biology, genetics, biodiversity, immunology, and model organism databases for other species (Flybase, Zfin, etc). A mailing list and general information are available (20, 21). A genome workshop is planned for mid-2004, when the chicken genome data will be ready to interpret.

References and Notes

1. See www.genome.gov/Pages/Research/Sequencing/ SeqProposals/Chicken_Genome.pdf
2. E. A. Pitel, *Genet. Sel. Evol.* (*Versailles*) **32**, 73 (2000).
3. M. Schmid *et al.*, *Cytogenet. Cell Genet.* **90**, 169 (2000).
4. D. W. Burt *et al.*, *Mol. Vis.* **9**, 164 (2003).
5. D. J. de Koning, *J. Anim. Sci.* **81**, 1158 (2003).
6. J. McPherson, W. Warren, personal communication.
7. Roslin mode, see www.thearkdb.org/browser? species=chicken
8. See www.genome.iastate.edu/chickmap/dbase.html
9. J. Hu *et al.*, *Nucleic Acids Res.* **29**, 106 (2001).
10. M. Morisson *et al.*, *Genet. Sel Evol.* (*Versailles*) **34**, 521 (2002).
11. M. Clamp *et al.*, *Nucleic Acids Res.* **31**, 38 (2003).
12. See www.ncbi.nlm.nih.gov/dbEST
13. S. Hubbard, J.-M. Buerstedde, personal communication.
14. P. E. Neiman *et al.*, *Oncogene* **22**, 1073 (2003).
15. GEISHA, http://geisha.biosci.arizona.edu
16. EMAP, http://genex.hgu.mrc.ac.uk/
17. W. R. Brown, S. J. Hubbard, C. Tickle, S. A. Wilson, *Nature Rev. Genet.* **4**, 87 (2003).
18. S. Wilson, personal communication.
19. See www.chickengenome.org
20. See chickengenome@lists.bbsrc.ac.uk
21. See www.lists.bbsrc.ac.uk/mailman/listinfo/chicken-genome