Pig Genome Update

No. 102 May 1, 2010

Sequencing pig genome assembly and the publications that are likely to result will continue to be of primary importance. Pig Genome Update 101 (March 1) was devoted to primarily explaining the process and much of the opportunities. Please see:
http://www.animalgenome.org/pigs/newsletter/No.101.html .

Updates on the process are continuing. The pig genome is being sequenced and characterized under the auspices of the Swine Genome Sequencing Consortium (SGSC). In accordance with the Bermuda and Fort Lauderdale agreements and the more recent Toronto Statement the data have been released into pubic sequence repositories in a timely manner and in advance of publication. An assembly (Sscrofa9) covering about 90% of the pig genome has been annotated in Ensembl automated pipelines and was released with Ensembl 56 in September 2009. Another announcement of SScrofa9b is expected soon. The publication of the pig genome sequence will be based on the next assembly (Sscrofa10) which will incorporate additional data providing greater coverage and greater contiguity.

The SGSC has submitted a marker paper for publication in which the Consortium sets outs its plans for the analysis and publication of a draft pig genome sequence. These plans were presented to participants in the Pig Genome III conference held at the Wellcome Trust Sanger Institute, 2-4 November 2009 and a series of analysis working groups were established.

The SGSC would welcome the input of members of the community in the analysis and annotation of the draft sequence. **"Annotation"** comprises manual correction and improvement of the gene models for individual genes on a gene-by-gene basis. We are using the Wellcome Trust Sanger Institute's Otterlace tools for these manual annotation tasks – training can be provided as necessary. Individuals who make significant contributions to the manual annotation effort will be recognized through authorships of the relevant paper(s). **"Analysis"** comprises global analyses of particular aspects of genomic structure (e.g. repetitive sequences, segmental duplications,..) or of gene families (e.g. the immunoglobulin superfamily) or of genes with effects on specific traits (e.g. immunity, reproduction, olfaction). The outputs from "analysis" groups will be a few sentences or paragraphs in the main sequence paper plus possibly a companion paper that describes and discusses the specific analyses in greater detail. Authorship would be recognized in accordance with scientific convention. The paper describing the draft cattle genome sequence (Science 324, 522) provides an indication of "analysis" groupings – the text of the paper illustrates the outputs from such analyses. Members of the community who wish to contribute to the analysis and annotation of the draft pig genome sequence should contact the relevant analysis group leader or the coordinators, for example, if they wish to establish additional analysis groupings (Larry Schook [schook@illinois.edu] or Alan Archibald [alan.archibald@roslin.ed.ac.uk]).

Swine Genome Sequencing Consortium genome sequence analysis groups

Analysis group	Lead contact
Assembly	Alan Archibald (alan.archibald@roslin.ed.ac.uk)
Structural variation, segmental duplication, copy number variation	Christian Bendixen (christian.bendixen@agrsci.dk)
Repetitive DNA, transposable elements	Geoff Faulkner (geoff.faulkner@roslin.ed.ac.uk)
Speciation, wild and related suids and selection	Lawrence Schook (schook@uiuc.edu)
Evolution	Leif Andersson (Leif.andersson@imbim.uu.se)
Comparative genomics	Martien Groenen (Martien.groenen@wur.nl)
Imprinting	Ole Madsen (Ole.madsen@wur.nl)
SNP	Martien Groenen (martien.groenen@wur.nl)
ncRNA	Jan Gorodkin (gorodkin@genome.ku.dk)
Gene builds	Steve Searle (Searle@sanger.ac.uk)
Protein interactions	Soren Brunak (brunak@cbs.dtu.dk)
Immune genes	Chris Tuggle (cktuggle@iastate.edu)
Reproduction	Max Rothschild (mfrothsc@iastate.edu)
Obesity	Max Rothschild (mfrothsc@iastate.edu)
Olfaction, neuropeptide and prohormone	Sandra Rodriguez-Zas (rodrgzzs@illinois.edu)
Manual annotation	Jim Reecy (jreecy@iastate.edu)
Biomedical Models	Lawrence Schook (schook@illinois.edu)

Companion Manuscripts. The Steering Committee encourages topic groups as well as others to develop companion manuscripts. The Steering Committee would request that groups forward their intent to develop a companion manuscript (including a tentative title) to their attention (either Larry Schook or Alan Archibald) in order to coordinate data exchange and to ensure no overlaps or competing efforts. Also, developing a list of potential companion manuscripts will permit the Steering Committee to negotiate with journals for a dedicated issue focused on the pig genome project. Also, please forward your ideas as to potential journal(s) for the Steering Committee to contact. Submission of companion manuscripts has a target date of September 1, 2010.

SGSC Genome Manuscript Steering Committee
Alan Archibald, Roslin Institute, UK (alan.archibald@roslin.ed.ac.uk)
Mario Caccamo, The Genome Analysis Center (TGAC), UK (mario.caccamo@bbsrc.ac.uk)
Carol Churcher, Wellcome Trust Sanger Institute, UK (carol@sanger.ac.uk)
Merete Fredholm, U Copenhagen, Denmark (mf@life.ku.dk)
Martien Groenen, Wageningen U, Netherlands (Martien.Groenen@wur.nl)
Kyung-Tae Lee, Korean National Institute of Animal Science (ktlee2@gmail.com)
Denis Milan, INRA, France (denis.milan@toulouse.inra.fr)
Jane Rogers, The Genome Analysis Center (TGAC), UK (jane.rogers@bbsrc.ac.uk)
Max Rothschild, Iowa State U, USA (mfrothsc@iastate.edu)
Lawrence Schook, U of Illinois, USA (Chair) (schook@illinois.edu)
Jun Wang, BGI, China (wangjun30@gmail.com)
Hirohide Uenishi, (huenishi@affrc.go.jp)

What's new on NAGRP animal genome web site. (1)Release 11 of the Animal QTLdb is complete. As of April 23, 2010, a total of 1643 new QTL have been added to the Animal QTLdb since last release before the New Year. The new data included 1532 cattle QTL and 111 pig QTL. We encourage all users to add your own new QTL data through the Curator web form: found at the

URL: http://www.animalgenome.org/QTLdb/app.html. (2) In addition to the previously released QTLdb results in GFF3 format, the same data is now available in SAM data format http://www.animalgenome.org/cgi-bin/QTLdb/SS/index. The SAM data format is to assist users using SAM Tools. For details using SAM tools, see: http://samtools.sourceforge.net/. (3) A major improvement to the CateGOrizer online tool is made to improve its performance. The improvement included faster algorithms and more efficient ways to handle large data sets and can be seen at URL: http://www.animalgenome.org/bioinfo/tools/catego/ . (4) Animal QTLdb frequently asked questions has been updated with new items. The most recent additions are to address possible confusions that some downloaded QTL data may "appear" like "duplicates". The FAQ explained why and provided hints as how to correctly examine the downloaded data. (5) A generic online tool has been in development (http://www.animalgenome.org/cgi-bin/util/2gff) to help users to convert their mapping data file into GFF3 format, which is useful to upload to GBrowse for map alignments of the genome features. Please let us know if you have data that are not on our input file format list. Check out the "what's new" page (http://www.animalgenome.org/news/whatsnew.html) for more up-to-date information. (Kindly provided by the NAGRP Bioinfo Team who may be contacted at bioinfo-team@animalgenome.org).

Upcoming meetings (see: http://www.animalgenome.org/pigs/community/meetings.html**)**

Animal Genomics for Animal Health International Symposium, 31 May - 2 June 2010, at the Maison de la Chimie, Paris France, See https://colloque.inra.fr/agah2010/ for more info.

International Society of Animal Genetics conference will take place in Edinburgh (UK), July 26-30, 2010. For details see http://www.isag.org.uk/society/conferences.asp

The 9th World Congress on Genetics Applied to Livestock Production (WCGALP), Aug. 1-6, 2010, Leipzig, Germany. For more details visit http://www.wcgalp2010.org/.

The 9th International Veterinary Immunology Symposium, August 16-20, 2010, Tokyo, Japan For more details visit http://9th-ivis.jtbcom.co.jp.

Items for *Pig Genome Update 103* can be sent to me by no later than June 15 please.

Max F. Rothschild
U.S. Pig Genome Coordinator
2255 Kildee Hall
Department of Animal Science
Iowa State University
Ames, Iowa 50011
Phone: 515-294-6202
Fax: 515-294-2401
mfrothsc@iastate.edu

Supported by Multi-State Research Funds to the National Research Service Program: NRSP-8.
National Animal Genome Research Program, Muquarrab Qureshi, NAGRP Director, NIFA