

SGSC Conference Call
March 16, 2009
Draft Minutes

- I. Present:** A. Archibald, J. Beever, P. Burfening, C. Churcher, R. Clark, M. Groenen, D. Hamernik, V. Hansford, L. Matthews, D. Milan, L. Schook, and S. Sims.
Regrets: K. Eversole, W. Liu, and C. Rogel-Gaillard

II. Action Items:

- 1) Archibald, Groenen, and Beever will develop a strategy for coordinating RNASeq and full-length cDNA sequencing efforts.
- 2) Archibald will talk to staff at Ensembl regarding their plans for when and how to submit the swine genome sequence to Genbank.
- 3) Groenen will make a list of samples from individual animals from zoos and try to get additional samples from zoos.
- 4) If anyone has samples to contribute to a high resolution genetic map, please contact Jon Beever.

- III. Sequence Update:** Clark distributed a PowerPoint file with 2 slides showing the overall progress on the swine genome sequencing project prior to the call. A total of 16,465 clones have been selected for sequencing and sent to the pipeline. This covers about 96% of the physical map. There are 2,624 Mb of total sequence (68.42 Mb of finished quality) from 15,196 clones. About 9,411 clones have been sequenced to the "improved/finished" stage.

A new assembly with sequence from all chromosomes is expected by the end of April, 2009. The new assembly will appear in Ensembl. When the sequencing team sees the new assembly, they will identify gaps to target for further sequencing.

Schook said that he submitted a Letter of Intent (on behalf of the SGSC) to submit an application to the FY2009 USDA-CSREES-AFRI Animal Genome competitive grants program. The Letter of intent describes a project to obtain funding for additional improvement of the swine genome sequence. About 1/3 of the swine genome sequence needs to be improved. Burfening indicated that Schook will be invited to submit a full application for this project. He also said that based on the number of Letters of Intent that were received and the limited funding that is available for these types of projects, competition will be extremely keen.

Schook thanked the entire Sanger team for their continued efforts to generate a high-quality draft genome sequence for swine.

- IV. Sequence of SSY:** Churcher said that they were collecting blood from a Duroc boar to construct another library for sequencing SSCY. They are also trying to sequence a few clones from other boars.
- V. Annotation Update:** Archibald is working with staff at Sanger and Ensembl to set up an annotation pipeline.

Beever sequenced about 65 million tags from a brain library with Solexa technology. He has 3' tags from 2 runs and is now using the RNASeq approach. He also has some 454 sequence and additional Solexa sequence. All of this sequence data could be added to the new assembly that will be released by the end of April, 2009. Beever has also been working on sequencing full-length cDNAs for more than a year. He has been able to sequence to the 5' end but it has been difficult to get sequence through the 5' UTR. Archibald, Groenen, and Beever will develop a strategy for coordinating RNASeq and full-length cDNA sequencing efforts.

VI. SNP Chip Update: Groenen said that they now have data from about 1,200 animals in the HapMap project. He also has prepared a first draft of the paper that will describe sequencing for SNP discovery as well as design and validation of the SNP chip. There is some concern that there is a very limited amount of biology in this paper. The group prefers to save the biology for another paper that will have a higher impact (to be published later in 2009). In the meantime, the group feels that it is important to publish the first paper on the SNP chip because many people are using the SNP chip and they will want to cite this paper. Potential journals to publish the first paper include: *PLOS-1*; *BMC Genetics*; or *BMC Genomics*.

Groenen said that they only have samples from one individual pig for some groups (e.g., the sample from the San Diego zoo). There is room for additional samples from other individuals. He will make a list of the samples that they have and try to get other samples from zoos.

Beever said that he recently received genotypes on 672 animals. These samples were genotyped at GeneSeek, which provided quick service that was of very high quality. The animals are from a three generation, F2 pedigree. He is completing a linkage analysis now. He would like to get another ~1,300 samples to develop a high resolution genetic map. If anyone has samples to contribute, please contact Jon Beever.

VII. Publication Strategy: No updates available at this time.

VIII. Farewell from Deb: Hamernik said that she will be leaving USDA at the end of March, 2009 to move to Nebraska where she has accepted the position of Associate Dean, Agricultural Research Division and Associate Director, Nebraska Agricultural Experiment Station. She congratulated the group on their ability to generate a high quality working draft sequence of the swine genome. She also thanked the group for their willingness to work together as a team. The group thanked Hamernik for organizing the monthly SGSC teleconferences and for supporting the swine genome sequencing project. Pete Burfening (USDA-CSREES-AFRI) recently joined the SGSC and will be responsible for setting up the monthly teleconferences in the future.

IX. Other: Milan distributed a PowerPoint file describing his efforts with RH mapping of SSC1 using a 12,000 RAD panel. They found overall good agreement with ~720 SNP and the assembly. However, four markers moved from the middle of SSC1 to the top of SSC1. These

four markers correspond to a single BAC and suggest that this BAC may not be correctly positioned on the BAC map. Clark said that the sequencing/mapping team would look into this further.

Matthews said that Sanger received an inquiry from Melissa Landrum (NCBI) asking how and when the swine genome sequence would be submitted to Genbank. Archibald said that the group had not yet discussed this issue. He suggested that this topic be added to the agenda for the April SGSC teleconference. In the meantime, Archibald will discuss this topic with Carol Scott at Ensembl.

- X. Next Conference Call:** The next teleconference will be held on Monday, April 20 at 8:00 a.m. (Eastern Time, US). Members of the SGSC Technical Committee will be invited to join this teleconference.