

An Update on Sequencing the Swine Genome: Opportunities and Challenges

Max F. Rothschild
C.F. Curtiss Distinguished Professor
M.E. Ensminger Chair in International Agriculture
Center of Integrated Animal Genomics
Iowa State University
mfrothsc@iastate.edu

Introduction

Imagine for the moment, using not just 2 or 3 or 5 or 10 genes to select for traits but 100s or 1000s of genes to improve pig production and create specialized pork products. To do this, sequencing of the pig genome is required, and this will revolutionize pork production. For the past year serious efforts to sequence the pig genome have been underway. These efforts, the progress that has been made, and opportunities and challenges for further sequencing are discussed in this brief report.

What is sequencing?

Sequencing is the unraveling of the DNA to understand the genetic code. It is equivalent to breaking down books into individual sentences and even specific letters in these sentences and words. The letters in the genetic code (A, T, G, and C) are combined into “words” and these words are the genes that control traits or contribute to phenotypes of the animal like rate of growth, level of fat, reproductive performance and disease susceptibility. Knowing the genetic code requires that we apply modern molecular biology or laboratory methods to break up the code into smaller pieces and then “read” the code. Funding to sequence the pig genome is an international effort provided by the USDA, National Pork Board, Iowa Pork Producers Association, University of Illinois, Iowa State University, North Carolina Pork Council, North Carolina State University, the Wellcome Trust Sanger Institute, UK and a number of research institutions from around the world including those from China, Denmark, France, Japan, Korea, Scotland and England.

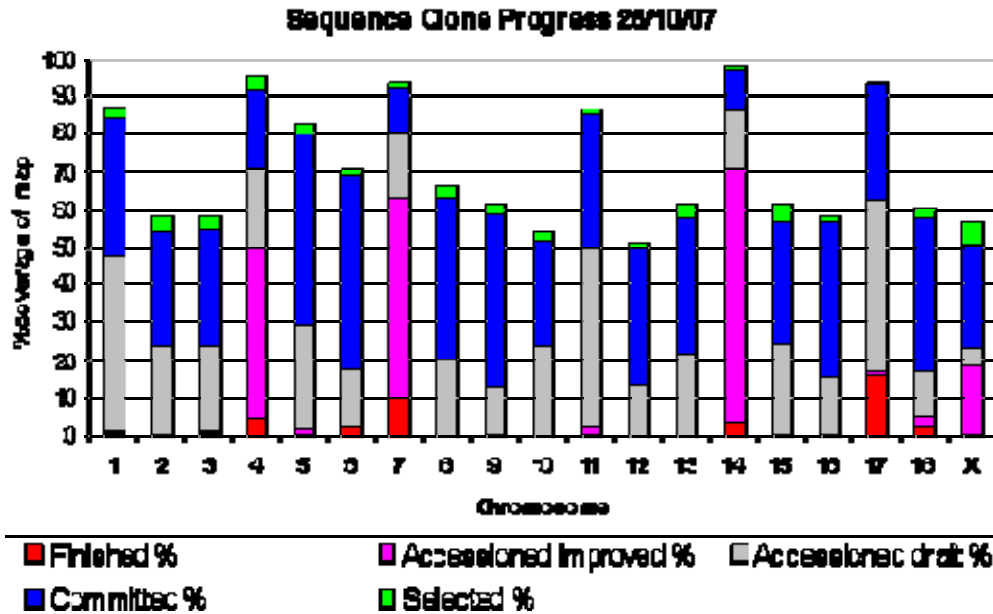
A complete description of sequencing was given in last year’s NSIF report (Rothschild, 2006).

Progress of the sequencing efforts

The pig genome sequencing began in part when a Danish-Chinese project was initiated several years ago. This project produced a 0.6 X coverage. To have excellent sequence, a 6X copy of sequence is needed. The new effort initiated recently by US, UK and other country partners has as its goal a 3X -4X coverage with additional sequencing coverage being obtained from foreign lab contributions. Updates can be seen at <http://www.animalgenome.org/pigs/genomesequence/>. The International Genome Consortium Sequencing Newsletter (<http://www.animalgenome.org/pigs/newsletter/index.html>) contains updates also. A comparison

with a similar graph from last year would suggest progress is being made but at a slower rate than expected. Now a final rough draft of the sequence is expected by 2009 instead of 2008.

Figure 2. Update of sequencing efforts at a glance (see <http://www.animalgenome.org/pigs/newsletter/index.html>)



Lessons learned from initial sequencing

Because of the nature of the sequencing and assembly of the genome the quality of sequence depends greatly on the amount of coverage which is typically measured in number of copies (X). Initial sequencing was expected to be at the level of 3X-4X. Funding to make an improved draft (4X-5X) is now being sought. The improvement from 3X to 5X coverage is considerable. For example, this level of sequencing will allow for more complete understanding of individual genes and their attributes. It will allow for a more correct order of the genes and will reduce the gaps (places without sequence) by over 50%. Such improvements will be extremely important in the discovery of genetic differences called single nucleotide polymorphisms (SNPs) and the mapping of them correctly. A goal for the sequencing project is to find out how different animals are and how this translates to the genetic causes of trait differences. While a 3X-4X coverage will be quite helpful, extensive work on chromosome 17 by Rothschild and colleagues (Hart et al., 2007) has shown that 5X or even a 6X coverage will yield better prediction of gene order, and a greater likelihood of discovery of important SNPs.

Challenges and opportunities which revealed themselves in 2007

This year has proved to be one of both pluses and minuses. New funding, necessary for the advance sequencing has been limited in 2007. This threatens the final level of sequencing possible. While the International Sequencing Committee has remained active, two key members, Drs. Jane Rogers and Sean Humphray, have recently announced they are leaving the Sanger center in the UK. The Sanger Center has announced that this will not alter the rate of progress but it is important to note that these two individuals were important leaders in the swine sequencing effort.

On the positive side a recent meeting in October in Paris discussed the development of a consortium to have a SNP genotyping chip. This consortium, consisting of members from CSREES and ARS USDA, Iowa State University and the University of Illinois, are working together to develop a 50K SNP chip. Such a tool would be extremely useful for genetic trait discovery. The hope is to develop the chip in early 2008 and have it ready for mid to late 2008.

A second and even bigger set of developments surrounds the new sequencing technologies that are rapidly coming on line. These machines sequence by new methods and are best used to “resequence” organisms and do new SNP discovery. New sequencing is also possible. The two most popular machines use technology called the “454 method” or the “Solexa” sequencing approaches. Both machines cost about \$.5 million but offer the ability to do sequencing at a fraction of the previous cost and at nearly unbelievable speeds. Once the first draft sequence for the pig is established, resequencing entire individuals may cost only \$100,000 and would be extremely useful to find SNPs associated with special traits that one or only a few animals have. The new sequencing technologies have the opportunity to rapidly speed discovery and may represent the first new paradigm switch in sequencing and biology in some years.

The benefits pork producers might see

Sequencing the swine genome is an investment in basic research with both long- and short-term goals. The potential usefulness of genes in selection for improved pig performance will be determined more quickly if the pig genome sequence is available. Discovery and elimination of undesirable forms or alleles of these genes will be accelerated. Early past examples include the removal of mutant or negative alleles of the stress gene (HAL) and Rendement Napole (RN) gene. In the last 10 years several genes have been identified which improve performance and leanness (IGF2, MC4R), meat quality (CAST, PRKAG3) and reproduction (ESR, PRLR). Just recently many of these genes have been released to ALL pork producers for their use in genetic testing and improvement.

Sequencing of the pig genome offers the ability to multiply these discoveries into the 1000s and speed the rate of these discoveries. Greater federal funding for pig genomic research can be leveraged to provide more rapid application in these areas. For the average pork producer the many benefits include improved growth and litter size performance due to identification of genes affecting these traits. The genome sequence is a powerful tool, which will enable discoveries for improving traits of interest for producers regardless of their operational size, but those producers and companies associated with more advanced research groups or breeding companies may have

the opportunity to leap-frog with new genomic strategies. For these better positioned producers and early adopters more advanced opportunities are likely to include in the next 5-20 years the ability to produce pigs with improved immune response abilities (vaccine-ready pigs), growth-primed sire lines and development of increased niche and branded products representing unique or special attributes that one producer or one company wishes to use to increase market share and profits. It is likely that producers will have the ability to select certain genetic lines in the future that will require specialized feeds but that will out-perform existing lines by 20% to 40%.

Given that our competitors in the chicken and beef industries already have the chicken and cattle genomes sequenced we must move forward if we are to be competitive. The pig genome sequence will be essential for identifying specific genes and improving those traits that are difficult to measure, occur late in life or are evaluated on animals after harvest such as disease resistance, sow longevity and meat quality. Already we know of specific genes associated with sturdier sows and improved meat quality. Insights may be gained into how genes work together. This will allow better genetic planning to allow pig breeders and producers to select animals possessing certain sets of genes that interact in a favorable manner for a particular production system or niche market. This approach, termed “genomic selection,” will mean not just selecting on improved traits but selection on 1000s of genes.

Conclusions

Sequencing efforts have started and are moving along at a reasonable speed, though perhaps not at the speed desired. Results of these efforts are already being used to help select markers for improved growth and meat quality. It is likely we will have a draft sequence of the pig genome by late 2009. Team work and partnerships with the right seedstock breeders or breeding companies and university research faculty are likely to be keys in transforming this public information from a useful resource to a real payoff.

Acknowledgements:

Comments and suggestions provided by Dr. S. Newman are appreciated. A number of researchers world wide have participated in the sequencing efforts and in particular Dr. Larry Schook, UI and Jane Rodgers and Sean Humphray, Sanger Center are noted. The author wishes to thank financial support received from the USDA NRSP8 which supports the National Pig Genome Coordination project. Support for the pig genome sequencing efforts comes from the USDA, National Pork Board, Iowa Pork Producers Association, University of Illinois, Iowa State University, North Carolina Pork Council, North Carolina State University, the Wellcome Trust Sanger Institute, UK and a number of research institutions from around the world including those from China, Denmark, France, Japan, Korea, Scotland and England. Funding for individual research has been provided in part by Hatch, Iowa Agricultural and Home Economics Experiment Station and State of Iowa funds.

References

Hart, E.A., M. Caccamo, J.L. Harrow, S. Humphray, J. G.R. Gilbert, S. Trevanion, T. Hubbard, J. Rogers and M.F. Rothschild. 2007. Lessons learned from the initial sequencing of the pig genome: comparative analysis of an 8 MB region of pig chromosome 17. *Genome Biology* 8: R168.

Rothschild, M.F. 2006 Sequencing the Pig Genome. Progress and Prospects. NSIF proceedings. <http://www.nsif.com/Conferences/2006/pdf/SequencingSwineGenome.pdf>