

Sequencing the Swine Genome: Progress and Prospects

Max F. Rothschild
C.F. Curtiss Distinguished Professor
Department of Animal Science
Center of Integrated Animal Genomics
Iowa State University
mfrothsc@iastate.edu

Introduction

The recent near completion of the human genome sequence in the beginning of 2001 has catapulted our understanding of our genetic complexity as human beings. Furthermore, mining this wealth of information will allow biologists to understand human diversity including traits like height and weight or eye and hair color and even more complex traits like susceptibility to various diseases. This means that in the next 10-20 years a whole new form of medicine, called genomic medicine, will make it possible to develop individualized diagnoses, treatments and cures for each person based on their individual and unique genotype. That will revolutionize medicine.

Around the world scientists are spending billions of dollars to learn more about the human genome and these results may be used to better understand pig health, reproduction, growth, and behavior by comparing the pig genome sequence to the human genome sequence. To date, we now have the ability to select animals on the basis of individual gene tests for improved reproductive performance, growth rate, leanness and meat quality. Already this has meant millions of dollars of improvement in several of these traits. But imagine for the moment, using not just 5 or 10 genes to select for a trait 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.

What is sequencing?

Sequencing is the unraveling of the DNA to understand the genetic code (Figure 1). It is equivalent to breaking down books into individual sentences and even specific letters in these sentences and words (Figure 2). The letters in the genetic code (A, T, G, 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 the U.K.

Figure 1. Unraveling of chromosomal information to the individual genes. Figure adapted from DOE human genome figure.

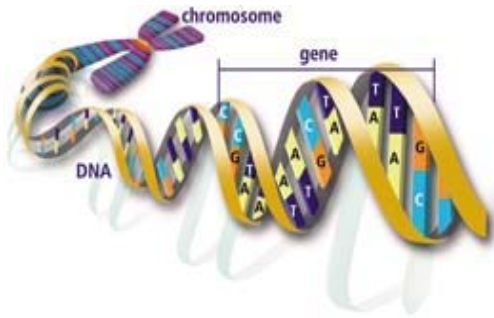
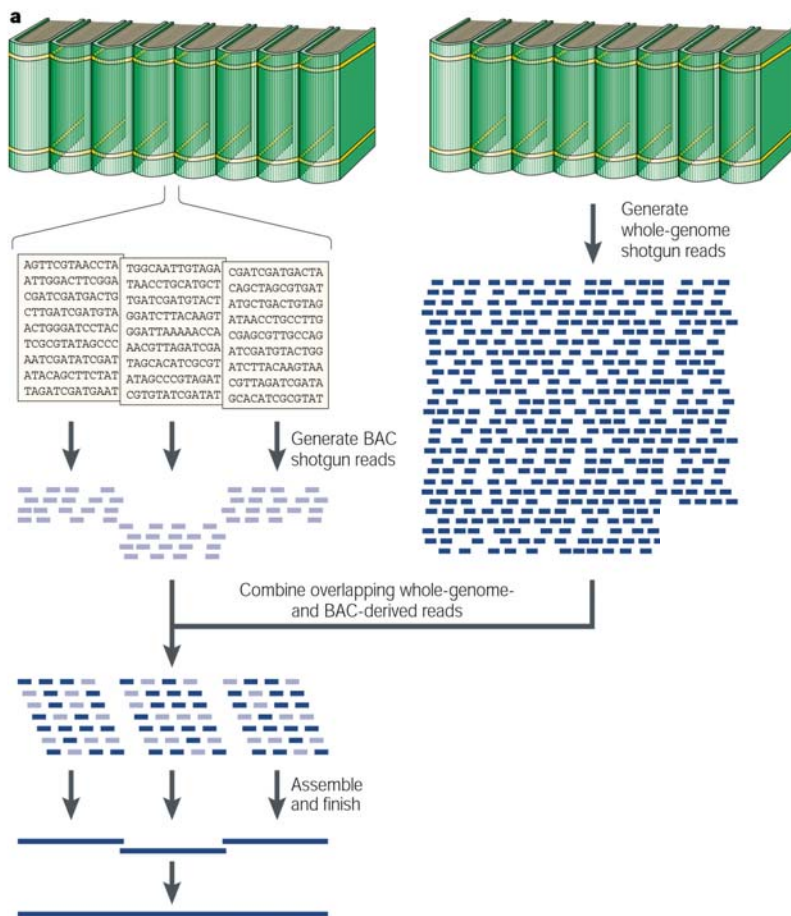


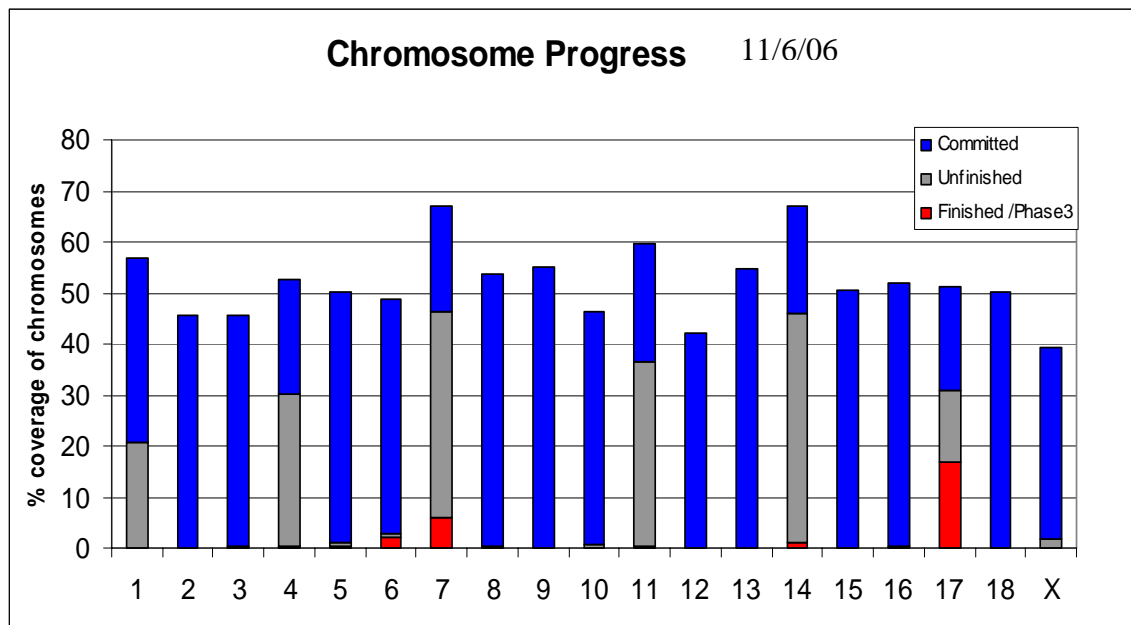
Figure 2. Comparison of ordered sequencing (left side) and shotgun sequencing (right side). This figure is from DOE human genome project. Sequencing from both methods leads to final sequence.



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. Already this new effort is progressing nicely. Updates can be seen daily at <http://www.animalgenome.org/pigs/genomesequence/>. These updates are provided as part of the USDA Bioinformatic Coordinator's team effort. Other information about the sequencing can be seen at that page and web pages at the Sanger Institute and the University of Illinois (see <http://www.piggenome.org/sequence.php>). Additional details about the sequencing efforts can be read from the Pig Genome Update also at <http://www.animalgenome.org/pigs/newsletter/index.html> or at the International Genome Consortium Sequencing Newsletter (<http://www.animalgenome.org/pigs/newsletter/index.html>).

Figure 3. Update of sequencing efforts at a glance see <http://www.animalgenome.org/pigs/newsletter/index.html> (example November 11, 2006)



Chromosome 17: an example in progress

So how in fact does the sequencing really help? At present we have good but not complete maps of the pig genome. The sequencing will provide not only the “ultimate genetic map” but will allow us to have the tools to hunt down mutations of interest in our own specialized herds and families. This genome sequence of the pig serves as a template to look into the sequence differences in pigs of interest for traits that are economically important.

An example of this effort is that of our lab on chromosome 17. A genome scan, performed using an F2 population derived from a Berkshire x Yorkshire cross, identified several meat quality QTL on pig chromosome 17 (SSC17) (Malek et al., 2001). These QTL included three meat color traits (color, 48 hour loin Hunter L value and 48 hour loin Minolta L value) and two lactate related traits (average lactate and average glycolytic potential). The identification of the mutations responsible for these QTL has proved to be a very challenging task, but it is necessary to allow the utilization of these QTL in modern pig breeding schemes. The objective of this study, part of a PhD project of A. M. Ramos, was to first more clearly identify the chromosomal region(s) more likely to contain the causative mutation(s) responsible for the observed meat quality QTL and secondly to find the mutations themselves (Ramos et al., 2007).

A total of 19 genes and 6 anonymous markers were first mapped to SSC17 after the initial scan. The developing linkage map then contained 33 genetic markers and was 119.3 cM long. The SSC17 marker density was substantially increased, especially under the QTL peaks, where the distance between markers was always less than 3 cM. From this information it was clear that progress could only then be made using the full pig genome sequence. Funding provided in part by Iowa State University and the Iowa Pork Producers Association was then used to obtain the finished sequence directly in the region under the QTL.

Previously, a physical map of overlapping pieces (called BAC ends) of the swine genome was developed by members of the Swine Genome Sequencing Consortium. A BAC tiling path (large overlapping pieces) spanning the SSC17 QTL region was selected and sequenced at the Sanger Institute. These clone sequences were aligned in order to generate a consensus sequence. The final alignment contained 65 clones and was approximately 7.8 Mbp long (paper in preparation). Initial annotation done on the pig sequence confirmed the extensive conservation between SSC17 and HSA20 (all human genes were found in the pig sequence and all genes maintained their human order in the pig sequence). Detailed manual annotation to Human Genome Havana standards is nearly completed at the Sanger Institute. The sequence was then used to develop Berkshire and Yorkshire specific sequence and a total of over 25 SNPs were found, mostly in the coding regions of the many genes underlying the QTL graph. To date at least three and possibly four of these SNPs are in close association with the traits of interest. It is clear that this is a region, specific to Berkshire containing many favorable alleles. Further work is continuing and is likely to produce results in the near future. Without the sequencing, real efforts to find these genes would be futile.

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. 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). 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 outperform 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 nicely. Results of these efforts are already being used to help select markers for improved growth and meat quality. Given the funding available, about \$15 million presently, it is likely we will have a draft sequence of the pig genome by late 2007 or early 2008. Will companies and seedstock breeders be ready to take advantage of these discoveries? Producers must ask the difficult questions. Are they ready to use the new genetics and genomics information? Are they positioned to first understand the information and second to use it effectively? Are there genetic systems in which they can use this information more effectively to improve pig production? Do they have niche markets they wish to fill or new products to produce? 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. Only then will producers, companies and geneticists help members of the pig industry really bring home the bacon.

Acknowledgements:

Collaborative research efforts of Drs. G. Plastow, A. Mileham and members of the Rothschild lab and those of Dr. A. M. Marcos 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 and that of the chromosome 17 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 the U.K. Funding for individual research has been provided in part by Sygen and PIC USA and by Hatch, Iowa Agricultural Experiment Station and State of Iowa funds.

References

Hu, J. Sean Humphray, Carol Scott, Jane Rogers, Antonio Marcos Ramos, James M. Reecy and Max F. Rothschild. 2006. Regional Genome Sequence Assembly for a Targeted Segment on Porcine Chromosome 17. Proc. Plant and Animal Genome XIV, San Diego, CA, Jan. 14-18.

Malek, M., J. C. M. Dekkers, H. K. Lee, T. J. Baas, K. Prusa, E. Huff-Lonergan, and M. F. Rothschild. 2001. A molecular genome scan analysis to identify chromosomal regions influencing economic traits in the pig. II. Meat and muscle composition. Mammal. Genome 12:630-636.

Ramos, A.M., Z.L. Hu, S J. Humphray, J. Rogers, J. Reecy and M. F. Rothschild. 2007. Combining fine mapping and genome sequencing for the molecular dissection of pig ssc17 meat quality QTL. Genetics (submitted)