Porcine genomics delivers new tools and results: This little piggy did more than just go to market

MAX F. ROTHSCHILD

1 Department of Animal Science, Center for Integrated Animal Genomics, Iowa State University, Ames, Iowa 50011, USA

Summary

The past decade has yielded new tools for pig geneticists and breeders thanks to the considerable developments resulting from efforts to map the pig genome. The pig genetic linkage map now has nearly 5000 loci including several hundred genes, microsatellites and amplified fragment length polymorphisms (AFLP) markers. Using tools that include somatic cell hybrid panels and radiation hybrid panels, the physical genetic map is also growing rapidly and has over 4000 genes and markers. Scientists using both exotic and commercial breeds for quantitative trait loci (QTL) scans and candidate gene analyses have identified a number of important chromosomal regions and individual genes associated with growth rate, leanness, feed intake, meat quality, litter size and disease resistance. Using marker-assisted selection (MAS) the commercial pig industry is actively incorporating these gene markers and traditional performance information to improve traits of economic importance in pig production. Researchers now have novel tools including pig gene arrays and advanced bioinformatics that are being exploited to find new candidate genes and to advance the understanding of gene function in the pig. Sequencing of the pig genome has been initiated and further sequencing is now being considered. Advances in pig genomics and directions for future research and the implications to both the pig industry and human health are reviewed.

1. Introduction

The pig was likely one of the first animals domesticated over 7000 years ago and pork is now the major red meat consumed (43%) worldwide (Rothschild & Ruvinsky, 1998). As one of the world’s most important livestock it has a vast geographic distribution and is represented by nearly 500 breeds worldwide. More recently, the pig’s role has expanded beyond just being a food source to potentially serving as an important model system for human health and representing a significant future source of organs for transplantation.

Coordinated efforts to understand the pig genome began in earnest in the early 1990s with the development of the international PiGMaP gene mapping project (Archibald et al., 1995) and efforts by the USDA and US agricultural universities (Rothschild et al., 1994; Rohrer et al., 1996). Both the PiGMaP and the USDA Cooperative State Research Education and Extension (CSREES) programs were designed to provide a structure that included collaboration and cooperation. In the US, a Pig Genome Coordinator position was developed. The Pig Genome Coordinator facilitated collaboration between scientists from state and private universities and federal labs that operated cooperatively in a Swine Genome Technical Committee, which has met yearly since 1994. The US Pig Genome Coordinator activities, in concert with activities of the scientists from USDA-ARS and international gene mapping projects, such as PiGMap and others, have allowed the status of the pig gene map to evolve more quickly and developments in functional genomics to advance rapidly in the last several years.

2. Pig Gene Mapping

Two significant linkage maps were published by the mid 1990s (Archibald et al., 1995; Rohrer et al., 1996). However, linkage progress has recently slowed but new gene markers consisting of microsatellites, amplified fragment length polymorphisms (AFLPs) and single nucleotide polymorphisms (SNPs) continue to be identified and mapped, and some integration of the maps continues to take place. The largest single published map contains about 1200 markers (Rohrer et al.,...
1996), but recently a large number of additional markers have been added to the on line version of the map (http://www.genome.iastate.edu/maps/marcmap.html). In total there are nearly 1000 genes and 1800 markers in the public database (www.thearkdb.org/browsing?species=pig) developed by the Roslin Institute. There is a developing AFLP map with about 2300 AFLPs that is likely to be added to the PiGMaP linkage map some time in the future. Together, these markers now total over 5000 and composite maps can be roughly combined using markers found on more than one individual map.

With the development and use of chromosome painting (Goureau et al., 1996), a somatic cell hybrid map (Yerle et al., 1996) and a 7000 rad radiation hybrid (RH) panel (ImpRH) (Yerle et al., 1998; Hawkins et al., 1999) integration of the linkage, cytogenetic and physical maps is well underway (Lahbib-Mansais et al., 2003). This RH map is growing rapidly and now contains nearly 4000 markers, including microsatellites, and over 2000 new expressed sequence tags (ESTs), of which many are human orthologs and enable comparative mapping (Rink et al., 2002; Tuggle et al., 2003; Cerera et al., 2003). These valuable resources continue to be employed and there is further development and use of an advanced 12000 rad RH map (Yerle et al., 2002). In particular, this will build a rapidly developing comparative map which will accelerate the identification of the genes explaining variation in traits of interest, either those identified by QTL studies or through direct approaches such as candidate gene association analyses.

3. Database Activities

Informatics and databases provide the tools needed for future discoveries. Significant pig bioinformatics efforts have been initiated by the Roslin Institute, Scotland (www.thearkdb.org) and to a lesser extent in the US (www.genome.iastate.edu) to support the pig genome efforts and display the gene maps (Archibald and Law, 2003). PiGBASE, which can be reached through these sites, has several features, including pig gene mapping references with over 1100 citations in the database and gene maps with about 2800 loci. Last year there were over 2 million hits at these pig genome sites. Additional web sites exist for the cytogenetic map of the pig (http://www.toulouse.inra.fr/lgc/pig/cyto/cyto.htm) and the RH panel map (http://www.toulouse.inra.fr/lgc/pig/hc/Menuchr.htm). A comparative map is also on the web (http://www.toulouse.inra.fr/lgc/pig/compare/compare.htm). In addition, a new EST database (http://pigest.genome.iastate.edu) has been developed and should become a similarly useful resource. It is now accessible and contains over 98 988 pig EST entries and further development will continue and it will be updated in the near future.

<table>
<thead>
<tr>
<th>Number of</th>
<th>Number discovered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth QTL</td>
<td>15</td>
</tr>
<tr>
<td>Fat QTL</td>
<td>10</td>
</tr>
<tr>
<td>Meat Quality</td>
<td>10</td>
</tr>
<tr>
<td>Reproduction QTL</td>
<td>3</td>
</tr>
<tr>
<td>Health related QTL</td>
<td>3</td>
</tr>
</tbody>
</table>

*Best estimated from literature survey – see Bidanel & Rothschild (2002).*

Other useful gene tools are available from the US pig genome web site (http://www.genome.iastate.edu).

4. QTL and Candidate Genes

Pork is the major red meat source worldwide and contributes to 43% of the world’s red meat consumed. Therefore, its production requires efficient fast growth, reduced feed intake, high carcass merit and meat quality. In addition, there must be high levels of reproductive success among breeding animals and disease resistance and increased survivability in young pigs. Genetic variation exists for some of these traits, as illustrated by results for stress response (Desautes et al., 2002). Increasingly, other requirements are being demanded by the consumers of pork. These include removal of antibiotic growth promoters from the feed, and altering of facilities to promote better animal welfare. Finally, concerns regarding the environmental impact of swine waste have prompted producers to consider ways to reduce feed wastage and improve feed efficiency.

Using the initial linkage maps, researchers began QTL experiments to determine the underlying regions associated with traits of interest. To aid in this effort the US Pig Genome Coordinator supplied greater than 300 pairs of microsatellite primers to over 40 labs worldwide. Researchers using these pig gene mapping resources and commercial and exotic pig breeds, have identified quantitative trait loci (QTL) affecting many of these traits (Table 1). A significant number of QTL have been reported on nearly all chromosomes for growth, carcass and meat quality traits and several chromosomes for reproduction (Bidanel & Rothschild, 2002). The discovery of QTL affecting immune health (stress measures, response to vaccination, challenge to limited diseases).
response traits and disease resistance has been limited due to the difficult experimental requirements for obtaining phenotypes and hence reports of QTL are far less numerous. This is an area where gene expression approaches may be particularly valuable to further unraveling of genetic causes of disease in pigs.

Following discoveries of imprinted genes in other species, researchers have expanded their projects to find imprinted and parent of origin effects (De Koning et al., 2000). In particular, one such region on chromosome 2 has been intensively investigated (Georges et al., 2003) and IGF2 implicated in causing a major effect in muscle mass. Georges and colleagues cleverly employed a haplotype sharing strategy analysis combined with marker assisted segregation analysis to position the QTL within a 500 kb region. The causal quantitative trait nucleotide (QTN) was revealed after investigating over 180 SNPs. This work clearly points to the need for careful analysis of all gene regions and the proper animals and phenotypic information. Further evidence for imprinted regions and genes are likely to be found now that these approaches have been developed.

As an alternative to the extensive QTL scans, candidate genes analyses using biological or mutational candidates from other species (Rothschild & Soller, 1997) have been employed (Bidanel & Rothschild, 2002). A number of significant associations (reviewed in Bidanel & Rothschild, 2002) have been demonstrated for candidate genes (see Table 2) for a variety of traits. For litter size, four gene associations have been identified (ESR, PRLR, RBP4, FSHB) with effects ranging from 0.25 to over 1 pig per allele per gene copy, depending on breed background. For growth and backfat, a number of genes have been investigated but one, MC4R, has clearly been associated with reduced feed intake, faster growth and less backfat (Kim et al., 2000). Well known meat quality genes (HAL, RN) have been reported and gene tests to remove their negative effects have been employed by seedstock producers and breeding companies. For disease resistance, several candidate genes or gene regions (K88, FUT1, SLA, NRAMP) have been identified and FUT1 is currently being used commercially to reduce post weaning diarrhea. However, a gene polymorphism has recently been identified as associated with resistance to K88 E. coli (Jørgensen et al., 2003). In addition, coat color genes (KIT, MC1R) have been used extensively in an effort to produce white pigs that are preferred by commercial meat packers. Many of these discoveries are being used by the commercial pig industry in combination with traditional performance information in marker-assisted selection programs to improve pig production (Table 2). The effectiveness of this approach has varied but large pig breeding companies are expanding strategies to use newly discovered genes and markers. Furthermore, as QTL regions have become clearly identified, positional candidate gene analyses are being employed to elucidate other known QTL. In particular, Rothschild and colleagues have used the positional candidate approach to find QTN in PRKAG3 associated with pH, drip loss and meat color (Ciobanu et al., 2001) and QTN in CAST associated with sensory tenderness (Ciobanu et al., 2004). Continuing efforts to exploit QTL maps and crossing experiments, combined with the improving comparative map, will allow for additional positional candidates to be identified and the causative QTN to be discovered.

Table 2. Candidate genes and gene tests identified and used in the industry

<table>
<thead>
<tr>
<th>Candidate genes</th>
<th>Traits</th>
<th>Industry use</th>
</tr>
</thead>
<tbody>
<tr>
<td>HAL</td>
<td>meat quality/stress</td>
<td>yes – extensive</td>
</tr>
<tr>
<td>KIT</td>
<td>white color</td>
<td>yes – exclusive</td>
</tr>
<tr>
<td>MC1R</td>
<td>red/black color</td>
<td>yes – exclusive</td>
</tr>
<tr>
<td>MC4R</td>
<td>growth and fatness</td>
<td>yes – exclusive</td>
</tr>
<tr>
<td>RN, PRKAG3</td>
<td>meat quality</td>
<td>yes – exclusive/exclusive</td>
</tr>
<tr>
<td>AFABP, HFABP</td>
<td>intramuscular fat</td>
<td>unknown</td>
</tr>
<tr>
<td>CAST</td>
<td>tenderness</td>
<td>yes – exclusive</td>
</tr>
<tr>
<td>IGF2</td>
<td>carcass composition</td>
<td>yes – exclusive</td>
</tr>
<tr>
<td>ESR, PRLR, RBP4</td>
<td>litter size</td>
<td>yes – exclusive</td>
</tr>
<tr>
<td>FSHB</td>
<td>reproduction</td>
<td>unknown</td>
</tr>
<tr>
<td>NRAMP, SLA</td>
<td>disease susceptibility</td>
<td>unknown</td>
</tr>
<tr>
<td>FUT1, NRAMP, SLA</td>
<td>disease susceptibility</td>
<td>yes – exclusive</td>
</tr>
<tr>
<td>Trade secret tests</td>
<td>several traits</td>
<td>yes – exclusive</td>
</tr>
</tbody>
</table>

5. Sequencing Efforts

Analysis of the porcine genome has provided strong evidence that it is of similar chromosomal organization (2n=38, including meta- and acrocentric chromosomes), size (3 x 10^9 bp), and complexity to the human genome. Recent efforts by many researchers have generated ESTs from cDNA clones randomly picked from libraries from many tissues. These ESTs...
projects have varied in size and in the tissues used (i.e. Winterø et al., 1996; Davoli et al., 2002; Rink et al., 2002a; Yao et al., 2002; Nobis & Cousens, 2003). The largest of these types of projects published to date was sponsored by the USDA and reported the sequencing and initial analysis of 66 245 ESTs (Fahrenkrug et al., 2002). In addition, 21 499 sequences from reproductive tissue were produced by a consortium of several research groups (Tuggle et al., 2003). Genbank now contains approximately 120 000 sequences, and the October 2002 TIGR release has combined these to form nearly 17 350 clusters and 31 000 singletons. Several additional deposits of 5000 to 20 000 EST sequences are expected soon. Most importantly however, a major Sino-Danish effort to sequence the pig genome (http://www.piggenome.dk/) has resulted in approximately 800 000 EST sequences that are expected to be deposited in the database in the next several months and also a 1X sequence coverage of the porcine genome is likely to be publicly released soon (Fredholm, personal communication). Sequencing of a large number of these ESTs will continue to help assist comparative mapping efforts, candidate gene discovery and expression analysis.

The announced completion of the human genome sequencing efforts has led to the intent by NIH to sequence other species, and a number of species are being considered for future sequencing. A swine genome community effort produced a ‘White Paper’ (Rohrer et al., 2003) for consideration by NHGRI that outlined the role the pig plays agriculturally and as a model for human biology. The White Paper received very solid backing from colleagues from several countries and from industry. It received a ‘high priority status’ from NHGRI but must await sufficient funding support to proceed with the project. Recent international efforts are underway to raise the estimated 50 million dollars required for this project.

6. Functional Genomic Analysis

To fully understand the physiological complexity of the pig transcriptome, researchers are initiating expression and other functional gene analyses. A limited number of genes and techniques, such as Northern analysis and differential display PCR (Wilson et al., 2002; Gladney et al., 2003), were first explored by researchers. More recently, other approaches have included quantitative real time PCR to determine mRNA levels for immune response and disease infection levels (Dawson et al., 2003; Okomo-Adhiambo et al., 2003). These techniques, while quite useful, have proved to be limited in the numbers of genes that can be considered. Other approaches have included use of limited numbers of cDNA on macroarrays (Zhao et al., 2003). Given the initial lack of development of large scale cDNA arrays for the pig, human arrays were first tested and used (Moody et al., 2003; Gladney et al., 2004). These initial experiments with such materials have proved quite valuable as reproducibility was generally high. The outstanding progress to produce large numbers of pig ESTs has now allowed for large scale expression analysis using porcine materials only. In one of the first studies, Pomp and colleagues (Caetano et al., 2002; Caetano et al., 2003) have used cDNAs derived from ovary and follicular RNA from animals from either an index line selected for higher litter size or a control line and co-hybridized them with 4600 follicle derived probes to study gene expression patterns related to reproductive efficiency. Other projects exist, including two large scale efforts in Europe. The first European Community supported project is called PathoCHIP (http://www.pathochipproject.com) and uses spotted cDNA arrays for disease organism and immune response genes, while the second, called QualityPorkGENES (www.qualityporkgenes.com) looks at the co-expression of genes related to meat quality. Cooperative efforts by the US Pig Genome Coordinator and US and international researchers have now been directed to developing a first stage cDNA or oligo spotted array for the pig genome community. The 13 000 element oligo array is now commercially available, and experiments are underway to use it to answer a variety of different experimental questions. This array and others to be developed later will advance functional analysis in the pig.

7. Pig in Biomedical Research and Transplantation

There has been biomedical interest in the pig as a model for human biology for several years (Tubleson & Schook, 1996). This research has covered nutrition, digestive physiology, kidney and heart function, diabetes and obesity. One such recent example is the discovery of a mutant MC4R gene in pigs, variants of which also cause obesity in humans (Kim et al., 2002).

Shortages of tissues and organs worldwide have increased interest in xenotransplantation. Because of its size and physiology, the pig is seen as a preferred donor. Recent concerns about retroviruses and the difficulty in creating transgenic pigs that meet all the needed requirements for safe transplantation have slowed progress in this area and several major companies have scaled back their active research.

8. Conclusions

Advances to date have been important and discoveries have moved from the lab to the farm in just a short number of years. Large-scale gene and trait identification and mapping continue to make progress and it is likely that a number of gene tests to improve pork production will be developed yearly for use in the pig industry. Commercial efforts to fully realize their
value are progressing. Sequencing and expression analysis has been initiated and offer new avenues to understand the biological complexity of the pig. The pig genomics community does not need to rely solely on developments from other organisms such as the human, the mouse or rat anymore. The pig genome continues to offer new insights for both agricultural purposes and for its importance to human biomedical concerns. Further discoveries and understanding of the genome complexity remain a significant but accomplishable challenge.

This is a journal paper of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, Project No. 3600, and was supported in part by Hatch Act and State of Iowa funds and funding from NRSP-8 and the USDA/CSREES Pig Genome Coordination program. The efforts of all scientists involved in the pig gene mapping and genomics programs worldwide are gratefully acknowledged including efforts by Alan Archibald, Andy Law, and colleagues, Roslin Institute, and Zhiliang Hu to support database activities. Comments and information supplied by Graham Plastow and Jack Dekkers are appreciated.

References


