THE PIG GENE MAPPING PROJECT - PiGMaP

C.S. Haley[1], A. Archibald[1],L. Andersson[2], A.A. Bosma[3], W. Davies[4], M. Fredholm[5], H. Geldermann[6], M. Groenen[7],I. Gustavsson[2], L. Ollivier[8], E. M. Tucker[9], A. Van de Weghe[10]

[1]AFRC Institute of Animal Physiology and Genetics Research, Roslin, Midlothian, EH25 9PS, UK
[2]Dept. of Animal Breeding and Genet., Swedish University of Agric. Sciences, S-750 07 Uppsala 7, Sweden
[3]Faculty of Vet. Sciences, State University, Yatelaan 1, P.O. Box 80 157, 3508 TD Utrecht, Netherlands
[4]Norwegian College of Veterinary Medicine, P.O. Box 8146, Dept N-0033, Oslo 1, Norway
[5]Royal Vet. and Agricultural University, Bülowsvej 13, DK-1870 Frederiksberg C, Copenhagen, Denmark
[6]Inst. Tierzucht Vererbungsforschung, Tierärztlichen Hochschule,Bünteweg 17P, 3000 Hannover 71,FRG
[7]Wageningen Agricultural University, POB 338, 6700 AH Wageningen, Netherlands
[8]INRA Station de Génétique Quantitative et Appliquée, Domaine de Vilvert, 78350 Jouy-en-Josas,France
[9]AFRC Institute of Animal Physiology and Genetics Research Babraham, Cambridge CB2 4AT, UK
[10]State University of Ghent - Faculty of Veterinary Medicine, Heidestraat 19 B9220 Merelbeke, Belgium.

SUMMARY

This paper describes a recently initiated European project for the production of a gene map of the pig, with polymorphic molecular genetic markers evenly spaced through the genome and landmark loci physically assigned to chromosomes. Potential benefits include:
- Understanding the organisation and action of genes controlling valuable quantitative traits.
- The ability to improve breeding stock using marker assisted selection.
- The possibility of isolating for study or manipulation important genes by 'reverse genetics'.
- Studying the genomic organisation and evolutionary relationships of pigs, man and other mammals.
- The development of porcine models of human disease.

The pig has several advantages over other farm animal species for genetic mapping including a well defined karyotype, large full-sib families, short generation interval, the availability of diverse genetic stocks and an advanced breeding industry capable of exploiting developments. European commercial breeds, Chinese Meishan pigs and the European Wild Boar will be used to create reference families. The genetic distance between these stocks will facilitate the isolation of informative genetic markers and allow genes controlling the physical differences in economically important traits to be studied. The comparative conservation of DNA sequences and linkage relationships between mammalian species allows the porcine genetic map to be built on a skeleton of molecular markers which have already been mapped in man or in the mouse. This can be fleshed out with molecular markers based on minisatellite or microsatellite loci (variable number tandem repeat loci or VNTRs) which are very polymorphic and thus highly informative for linkage studies both between and within breeds. The polymerase chain reaction (PCR) for typing microsatellite loci provides the prospect of rapid, automated genotyping of animals for marker loci in the future.

A sample of cloned marker sequences will be physically assigned to chromosomes using hybrid cell lines and regionally located by *in situ* hybridization to provide at least one proximal and distal landmark locus on each chromosome arm. The porcine karyotype is well suited to chromosome sorting using fluorescence activated cell sorting (FACS) and a porcine flow karotype will be developed allowing rapid assignment of cloned sequences to chromosomes and the production of chromosome specific libraries.

In the final stage of the production of the molecular map, genes controlling physical variation (quantitative trait loci or QTLs) will be mapped and studied. Located genes or chromosomal segments of major economic value can be selected between or within breeds using marker assisted selection.

Work on the production of the gene map has begun. Heterologous and homologous probes are being used to examine the genetic distance between the breeds involved and are being located by *in situ* hybridization. The production of reference families is under way. Porcine locus specific VNTR probes are being isolated. Hybrid somatic cell lines are available and a porcine flow karyotype is being developed. It is envisaged that a low resolution genetic and physical map could be achieved in 3-4 years, with preliminary mapping of QTLs of large effect possible shortly afterwards.

INTRODUCTION

A genetic map would be made up of linked polymorphic marker loci approximately evenly distributed through the genome and is made possible by the availability of large numbers of genetic markers based on DNA polymorphisms (e.g. RFLPs and VNTRs). The map is constructed by estimating recombination frequencies from individuals in three generation pedigrees genotyped for each of the marker loci. In addition to DNA polymorphisms, all other polymorphic loci (e.g. blood groups or proteins) could be mapped, many for the first time. As the map is being constructed, or at a later date, individuals are measured for a variety of

performance traits (e.g. growth rate, litter size, disease resistance, etc.) which enables identification of QTLs linked to the marker loci. Physical mapping, using hybrid cell lines and *in situ* hybridisation, is used to place detected linkage groups on particular chromosomes or chromosomal regions.

Our understanding of the genetic architecture of variation for economically important traits is very poor. Present techniques are not capable of showing whether appropriate models should be based on a few genes or a few tens of genes. Using the genetic map and statistical techniques currently being developed (e.g. Lander and Botstein, 1989) it will be possible to detect and map QTLs of moderate or large effect. This will show for the first time whether there are genes of this size contributing to genetic variation in economically important traits and will also tell us the genomic organisation of the loci involved. At first the techniques will be applied to the study of the large genetic differences which exist between breeds. As the genetic map and techniques become more refined, it will also be possible to study the causes of genetic variation within breeds or lines and show whether this is controlled by the same loci as between breed variation.

The identification and measurement of linkage relationships between QTLs and marker loci allows the use of marker assisted selection (e.g Soller and Beckman, 1987). Animals are selected on their genotypes at marker loci as well as on their phenotypes. Marker loci could be used to delimit those portions of the genome that are required to be transferred from one breed to another. For example, a RFLP based genetic map has been produced for tomatoes and has been used to map QTLs of commercial importance with a view to producing a commercial variety by combining the best alleles from two species (Paterson *et al.*,.1989) The potential value of this technique is large, for example, if the genetic map enables genes controlling half the additional prolificacy of the Meishan to be transferred into European pigs, this would reduce pig production costs by over £30 x 10^6 per annum in the U.K. and by over 500 x 10^6 ECU per annum in the EC. As techniques improve, marker assisted selection could be applied within breeds, potentially leading to significant improvements in selection responses.

In the longer term, it may be possible to isolate and clone QTLs known only by their map position. Such 'reverse genetics' has been successfully applied to clone disease loci in man (e.g. cystic fibrosis, Riordan *et al.*, 1989). Once cloned, the structure, expression and function of the loci can be studied. This technique will reveal some of the estimated greater than 95% of genes which have yet to be located and provides potential material for other developments, such as transgenic programmes. Additionally, the alignment of the porcine and human maps will allow the development of models of human disease which can be experimentally manipulated, as is already happening with malignant hyperthermia (Halothane sensitivity/PSE in pigs).

FOUNDATIONS OF A PORCINE GENE MAP

A preliminary RFLP map of man has been produced (Donis-Keller *et al.*, 1987). This provides a foundation upon which to build genetic maps of other mammalian species as human DNA probes are likely to detect homologous DNA sequences in other mammals, but the proportion that detect RFLPs has yet to be determined. Conservation of linkage relationships between man and other mammals (Nadeau, 1989) enables human DNA probes to be selected which are distributed evenly across the genome. However, mapping other mammalian genomes will be a major project requiring collaboration between laboratories in several countries.

The pig has several advantages for gene mapping such as a short generation interval and large full-sib family size. Unlike cattle, sheep and poultry, the pig has a well defined and easily determined karyotype with high resolution banding, making physical mapping more tractable, especially by *in situ* hybridization (Yerle *et al* , 1986; Chowdray *et al.*, 1989). Finally, the forward looking pig breeding organisations in Europe mean that advanced techniques, such as marker assisted selection, are likely to be exploited in future.

At present the gene map of pig is rudimentary when compared to the approximately 1600 and 1450+ genes which have been mapped in Man and mice respectively (HGM10, 1989): only 38 genes are assigned to 17 linkage or syntenic groups, with only 27 tentative chromosomal assignments (O'Brien *et al.*, 1988; Thomsen, 1989), of which half a dozen are by *in situ* hibridization. An ideal tool for the production of a genetic map is a cross between two diverse strains which differ in allele frequencies at individual marker loci. This increases the number of animals which provide information on linkage in the F2 (at the extreme, when the two strains are fixed for different alleles at a marker locus, all F1 matings are informative). Additionally, a large phenotypic strain difference means that QTLs segregating in the F2 cross between strains can be mapped.

Three diverse pig stocks are the Chinese Meishan, European commercial lines and the European Wild Boar. Insufficient information is available to compare the 3 stocks directly, but the genetic distance between the Meishan and European breeds has been estimated to be on par with that between two mouse species (Oishi *et al.*, 1989)(*Mus musculus* and *M. spretus* ; a cross between these has been shown to be a powerful tool for gene mapping; Avner *et al.*,1988). At the phenotypic level, the Meishan differs from European breeds for many traits (40% greater litter size, with advantages in ovulation rate and prenatal survival, puberty at less than half the age, resistance to stress and K88 *E. coli*, docile temperament, slow growth rate and twice the subcutaneous fat thickness, e.g. Bidanel *et al.*, 1989), and the European Wild Boar is different again for traits including litter size and seasonality (Spitz, 1986). Therefore, crosses between pairs of these breeds will

provide a good foundation for the production of a marker map and the mapping of QTLs. Ultimately, there is the possibility of studying the function of individual QTLs or of using marker assisted selection to cross potentially valuable Meishan or Wild Boar QTLs into European breeds (or vice versa).

PROGRAMME

The work programme is divided into two linked areas, genetic and physical mapping and an outline and timetable is shown in Figure 1. For genetic mapping the first priority is the accumulation of cloned DNA sequences (probes) detecting RFLPs between the foundation breeds. These probes would be a mixture of porcine coding sequences and heterologous sequences, the former would be isolated using standard techniques (Schumm *et al.*, 1988; Bowden *et al.*, 1989) and the latter would be largely from man and which had already been placed on the human map. The markers detected would largely be di-allelic, but their human map position and assumed human/pig synteny conservation would enable them to be selected so as to be approximately evenly spread through the pig genome. These markers would be used to assess the degree of genetic divergence between the foundation breeds (European commercial stocks, European Wild Boar and Chinese Meishan). This would determine the most useful crosses and the degree of complementarity between the various crosses

Ultimately, the genetic map should be composed of highly polymorphic loci in order to make it informative for work within breeds as well as between breeds. VNTRs , either based on minisatellite or microsatellite sequences, provide such loci. Minisatellite sequences would be used to screen pig genomic libraries to provide locus specific probes for the detection of RFLPs on Southern blots (Jeffreys *et al.*, 1985; Nakamura *et al.*, 1987). VNTR loci based on simple oligonucleotide sequences (microsatellite sequences) are also highly polymorphic and may have a more random distribution than minisatellite sequences, are amenable to PCR typing and are found in many species (Litt and Luty, 1989; Weber and May, 1989; Tautz, 1989). Porcine genomic libraries would be screened with simple repeat sequences, positive clones sequenced and used to design unique oligonucleotides for flanking regions. These sequences provide PCR primers enabling animals to be genotyped by PCR followed by polyacrylamide gel electrophoresis.

Di-allelic and VNTR markers would be assessed on F2 reference families (50-100 F2 individuals per group) to confirm monogenic segregation of the markers and to detect close linkage or allelism between markers. Polymorphic markers would be exchanged between collaborating groups and linkage information accumulated over all groups of reference families. This would provide a powerful test for linkage and allow comparison of distances in different types of crosses (e.g. Meishan x Large White versus Wild Boar x Large White) to provide further information on comparative genomic organisation.

Physical mapping will be performed using two techniques. *In situ* hybridization of cloned genes on metaphase or prometaphase chromosomes allows precise localization of cloned genes within chromosomal regions. The rationalisation of the panels of hybrid cell lines which contain different combinations of porcine chromosomes makes it possible to rapidly assign cloned sequences to chromosomes. The availability of porcine cell lines and animals carrying chromosomal translocations will be particularly useful.

The porcine karyotype is particularly suited to FACS chromosome sorting (Grunwald *et al.*, 1989) and a dual laser FACS sorted flow karyotype will be developed within the project. This will enable the distribution of filters carrying FACS sorted chromosomes to participants to allow rapid assignment of cloned sequences to chromosomes. Future attempts to isolate coding sequences associated with QTLs will be greatly aided by chromosome specific libraries, and the production of these for selected chromosomes will also be undertaken.

The final part of the project will be the initiation of the experiments required to map QTLs. This involves the development of the computer technology, the coordination of the traits to be measured at participating centres and the initiation of the crosses. It is anticipated that the genotyping and performance testing of the animals in this part of the project and the mapping of QTLs will be undertaken in the years 1993 to 1995.

CONCLUSION

The production of a pig gene map would lead to many scientific and applied benefits. To achieve this aim and to initiate mapping of QTLs requires a major effort. This project brings together research workers with complementary expertise in several laboratories and European countries. This will provide the momentum required to make the aims realistic with minimisation of unnecessary duplication and coordination of effort.

REFERENCES

AVNER, P., AMAR, L., DANDOLO, L. and GUÉNET, J.L. 1988. *Trends in Genet.* **4**: 18-23.
BIDANEL, J., CARITEZ, J.C. and LEGAULT, C., 1989. *Journées Rech. Porcine en France* **21**: 345-360.
BOWDEN, D.W., MÜLLER-KAHLE, H., GRAVIUS, T., et al. 1989. *Am. J. Hum. Genet.* **44**: 671-678.

CHOWDRAY, B., HARBITZ, I., MAKINEN,A., DAVIES, W. and GUSTAVSSON,I. 1989.*Hereditas* **111**: 73-78.
DONIS-KELLER, H. et al., 1987. *Cell* **51**: 319-337.
GRUNWALD, D., FRELAT, G. and VAIMAN, M.1989. In *Flow Cytometry* (Ed. A. Yen) CRC Press: 131-142.
HGM10, 1989 *Cytogenet. Cell Genet.* **51**: 503-532.
JEFFREYS, A.J., WILSON, V. and THEIN, S.L., 1985. *Nature* **314**: 67-73.
LANDER, E.S. and BOTSTEIN, D. 1989. *Genetics* **121**: 185-199.
LITT, M. and LUTY, J.A. 1989. *Am J. Hum. Genet.* **44**: 397-401.
NADEAU, J.H. 1989. *Trends in Genet.* **5**: 82-86.
NAKAMURA, Y., LEPPART, M., O'CONNELL, P., WOLFF, R., et al. 1987. *Science* **235**: 1616-1622.
O'BRIEN S.J., SEUANEZ, H. and WOMACK, J.E., 1988. *Ann. Rev. Genet.* **22**: 323-351.
OISHI, T., TANAKA, K., OTANI, T. and TAMADA, S., 1989. *Bull. Nat. Inst. Anim. Ind. (Jpn)* **48**: 1-10.
PATERSON, A.H., LANDER, E.S., HEWITT, J.D., PETERSON, S, et al. 1989. *Nature* **335**: 721-726.
RIORDAN, J.R., ROMMENS, J.M., KEREM, B., ALON, N., ROZMAHEL, R. et al. *Science* **245**: 1066-1073.
SCHUMM, J.W., KNOWLTON, R.G., BRAMAN, J.C., et al.1988. *Am. J. Hum. Genet.* **42**: 143-159.
SOLLER, M. and BECKMAN, J.S. 1987. *Proc. 2nd Int. Conf. Quant. Gen.*, 161-188.
SPITZ, F. 1986. *Pig News and Information* 7: 171-175.
TAUTZ, D. 1989. *Nucleic Acids Res.* **17**: 6463- 6471.
THOMSEN, P.D. 1989. *J. Reprod. Fert.* Suppl **38**: 135-144.
WEBER, J.L. and MAY, P.E.1989. *Am. J. Hum. Genet.* **44**: 388-396.
YERLE, M., GELLIN, J., ECHARD,G., LEFEVRE,F. and GILLOIS,M.1986.*Cytogenet. Cell Genet.* **42**: 129-32.

Figure 1. The Pig Gene Mapping Project.