Muscle ESTs II: Cloning, Sequencing, and Mapping the Pig Gene for the Intermediate Filament Protein Desmin

C. K. Tuggle, associate professor,
B. Smith, undergraduate student,
I. Sanchez-Serrano, technician,
C. Ernst, postdoctoral research associate,
and L. Marklund, postdoctoral research associate

Department of Animal Science
Iowa State University

ASL-R1576

Summary and Implications

The results of sequencing large numbers of random cDNAs in the human genome project has clearly shown value in gene discovery and mapping. We previously demonstrated a simple approach to rapidly identify muscle-specific pig cDNAs for sequencing by strong hybridization to muscle cDNA probes. In this report, we sequenced additional muscle cDNAs, emphasizing clones that were not strongly expressed in muscle. Fewer clones matched known muscle-specific genes in the pig or other species, indicating selection by hybridization is most accurate for strongly hybridizing clones showing weak signals with non-muscle probes. Some muscle-specific clones, however, also were identified by sequencing the weakly hybridizing class of cDNAs. One of these was M239, encoding muscle-specific desmin, which is an intermediate filament protein involved in holding muscle fibers together. As desmin may be an interesting candidate gene for muscle tenderness, the desmin gene (*DES*) was fully sequenced and mapped both physically and genetically. Pig desmin is highly conserved, being more than 97% identical at the amino acid level to human and mouse desmin. Linkage of *DES* was observed for three loci already mapped to pig chromosome 15 (SSC15), and physical mapping placed *DES* on SSC15q23-26. This mapping information will be useful in further studies on the role of *DES* in muscle biology and muscle tenderness traits.

Introduction

Although economically important traits such as growth and carcass characteristics have been studied and genetically improved, less attention has been paid to muscle- or meat-quality traits. This is partly due to the fact that comprehensive muscle and meat quality values can only be obtained after slaughter, making genetic selection less direct and greatly increases the costs of phenotyping animals. In contrast, if genetic markers for these quantitative traits (termed QTL) can be identified, then marker-assisted selection could be used to more quickly improve pork quality. For example, the pig genome map was recently used to locate RN, a meat-quality gene controlling the technological yield in Paris ham production (1). One of the first steps toward use of such markers will be the production of genetic and physical framework maps in direct analogy to the human genome effort. To most effectively use such maps, genes expressed in muscle tissue will need to be cloned, sequenced, and mapped onto the genome map of the pig. The major goal of our long-term work is to identify and map genes expressed in muscle, while the initial goal is to identify muscle-specific cDNAs. These cDNAs will then be sequenced to establish expressed sequence tags (ESTs) that are unique sequences that tag each expressed gene in the muscle for rapid genetic analysis. We have previously established a method to identify muscle-specific cDNAs through a hybridization selection process (2). In this report, we present development of ESTs for a number of new cDNAs cloned from a muscle cDNA library, as well as sequencing and mapping of a complete cDNA for the pig gene encoding the muscle-specific intermediate filament protein, desmin.

Materials and Methods

Expression analysis and sequencing of pig cDNAs

Each cDNA was amplified and tested for muscle specific expression on dot blots as previously described (2). In brief, PCR products of each clone are spotted onto a nitrocellulose filter and hybridized to either muscle cDNA probe or a non-muscle (spleen/liver/lung) cDNA probe. Each selected cDNA clone was sequenced from one or both ends using vector primers at the Iowa State University DNA Sequencing and Synthesis facility. The resulting sequences were used to search the DNA sequence databases using BLAST (3). For desmin (M239), the rest of the 2.3 kb cDNA was sequenced at the Facility and assembled into a final sequence shown in Figure 1.

Physical mapping of desmin

Primers designed to amplify the 3′ untranslated region. The forward primer was 5′-GTAGGGACCCCGCAGAGTCA-3′ and the reverse primer was 5′-ACCCCTGCCACCTGCCTGAG-3′. The PCR reaction used 1.5 mM MgCl2 standard PCR with the following program: 3 min 94°C followed by 30 cycles of 1 min 94°C, 1 min 65°C, 1 min 72°C. A final step of 10 min 72°C was used. The template used was the somatic cell hybrid panel of INRA-Toulouse. Positive results were seen with hybrids 2,4,16,23. These results were evaluated for concordance with chromosome using the method of Chevalet (4).

Identification of polymorphism at desmin and linkage mapping

Primers were designed by using the OLIGO 5 program from pig sequences and genomic organization information in other species. The forward primer was 5′-GCCCAGCTTCAGGAACAAC-3′ (putative exon 4) and

the reverse primer was 5′-ATCCAGGGCCATCTTGACAT-3′ (putative exon 6). The PCR reaction used 3 mM MgCl2, 25 ng of pig genomic DNA, and 4 pmol of primer with standard Taq polymerase enzyme. The following program was used: 95°C 5 min; then, 35 cycles of 95°C 30 sec, 59°C 45 sec, 72°C 3 min; then 72°C 5 min. A 1.0-kb fragment was observed upon PCR amplification, and the following restriction enzymes were used to look for polymorphisms: Alu, Bfa I, BstU I, Dde I, Dpn I, Dpn II, Hae III, Hinf I, Mnl I, Mse I, Msp I, Nla III, Nla IV, Rsa I, Taq I. With Nla III, a polymorphism was seen. The two allelic forms were identified as uncut (1kb) and cut (550 bp and 425 bp). The grandparental, F1, and F2 genomic DNA from the PigMap family were genotyped and used for linkage mapping with CRI-MAP (5).

Results and Discussion

We previously demonstrated that random amplification of muscle cDNAs coupled with selection by hybridization to muscle mRNA pools could efficiently identify muscle-specific cDNAs (2). One question that remained was the identity of cDNAs that hybridized poorly to both muscle and nonmuscle probes, and if such clones were useful to sequence as a class. Thus, we sequenced additional clones which were not strongly expressed in muscle and results are shown in Table 1. Several muscle-specific genes were identified, including the gene for CRC, otherwise known as the stress gene (6), a gene for muscle-specific caveolin, and the gene for the muscle-specific intermediate filament protein desmin. The majority of cDNAs sequenced, however, were not known muscle genes. Thus, to identify primarily muscle-specific clones, those clones that weakly hybridize to both probes should be avoided. One clone, M239, however, clearly was highly related to the muscle intermediate filament protein desmin, and we have further analyzed this gene. Sequencing of M239 confirmed we have cloned a full-length cDNA for pig desmin. Conceptual translation of this cDNA indicates pig desmin is 472 amino acids in length. Sequence comparisons show this cDNA is clearly the desmin isoform of the intermediate filament protein family, as comparison to another IF protein, vimentin, shows much less similarity than to other known mammalian desmin protein sequences (Table 2).

To physically map pig desmin, primers were designed in the 3′ untranslated region to amplify a 334-bp region. Sequencing confirmed the identity of the PCR product. Amplification was not detected from the rodent genomic DNA that is present in all hybrids in the French somatic cell hybrid panel (SCHP), but the expected fragment was observed with pig DNA. Thus, these primers were then used to physically map desmin using the SCHP. Statistical analysis of the PCR results mapped desmin to SSC15q23-26.

To genetically map desmin, primers were designed to amplify genomic DNA between putative exons 5 and 7. An approximately 1.0-kb fragment was observed and sequencing of the exon regions of this fragment demonstrated that the PCR product was correct. A number of restriction enzymes were surveyed to find polymorphisms between the Chinese pigs and the commercial breeds within the PiGMaP families. One enzyme, NlaIII, detected a sequence difference and the two allelic forms (allele A or B) within several of the PiGMaP families. A general survey of a number of pig breeds also showed that the desmin allele A was predominant, but that polymorphism existed in Chester White, Duroc, and Yorkshire (Table 3). DNA from PiGMaP families was amplified and NlaIII digestion was used to genotype the members of the families. Segregation was observed and linkage analysis was performed. Desmin was linked to three loci on pig chromosome 15 (SSC15): S0088, S0284 and S0015. This linkage mapping confirms the physical mapping. Desmin has been mapped in humans to chromosome 2 (HSA2), and a large region of HSA2 hybridizes to probes produced from SSC15 (7). Thus, the mapping of *DES* to SSC15 is consistent with this larger scale comparative mapping result, and further demonstrates the conservation of this genomic region between pigs and humans.

Acknowledgments

This study was supported in part by the National Pork Producers Council, and the Iowa Agricultural and Home Economics Experiment Station.

References

1. Milan, D., P. Le Roy, N. Woloszyn, J.C. Caritez, J.M. Elsen, P. Sellier & J. Gellin. 1995. The RN locus for meat quality maps to pig chromosome 15. Genet. Sel. Evol. 27: 195–199.
2. Tuggle, C. K. and C. B. Schmitz. 1994. Cloning and characterization of pig muscle cDNAs by an expressed sequence tag approach. Anim. Biotech. 5: 1–13.
3. Altschul S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. (1990) Basic local alignment search tool. Journal of Molecular Biology 215: 403–410.
4. Chevalet, C., J. Gouzy, and M. Sancristobal-Gaudy. (1997). Regional assignment of genetic markers using a somatic cell hybrid panel: a WWW interactive program available for the pig genome. Comput. Appl. Biosci. 13: 69–73.
5. Green, P., K. Falls, and S. Crooks. (1990). Documentation for CRI-MAP, version 2.4. Washington University School of Medicine, St. Louis, MO.
6. Fujii J., K. Otsu, F. Zorzato, S. de Leon, V.K. Khanna, J.E. Weiler, P.J. O'Brien, and D.H. MacLennan. (1991). Identification of a mutation in porcine ryanodine receptor associated with malignant hyperthermia. Science 253: 448–451.
7. Goureau A., Yerle M., Schmitz A, J. Riquet, D. Milan, P. Pinton, G. Frelat, and J. Gellin. (1996). Human and porcine correspondence of chromosome segments using bidirectional chromosome painting. Genomics 36: 252–262.

Table 1. Comparison of new pig expressed sequence tag (EST) sequences to Genbank database.

EST	Clone Length (kb)	Genbank Accession Number	EST Length (bp)	Match Accession No. and Description[a]	Match Score
M10	0.9	AA063647	443	(n) X62880 Muscle CRC (pi)	1100
M165	1.0	AA063655	403	(n) U31968 (p) U31968 muscle-specific Caveolin (n) Z18951 caveolin	1313 493 685
M185	1.8	AA063657	394	(n) X91349 (p) P45974 De-Ubiquitinase (h)	367 180
M186	1.8	AA063658	443	(n) M24088 Transaminase A (pi)	1153
M198	1.3	AA063660	441	no matches	
M218	1.7	AA063661	402	no matches	
M221	1.8	AA063662	472	(n) Z48042 (p) S52289 GPI-anchored p137 protein (h)	1104
M223	1.3	AA063663	456	no matches	
M239	2.3	AA063665	330	(n) X94252 muscle desmin-like 3'UT (pi)	1121
M248	1.6	AA063667	348	(n) U07802 Tis 11d gene (h)	385
M256	0.7	AA063668	420	no matches	
M327	1.1	AA063669	352	no matches	

h: human; pi: pig; m: mouse; r: rat; rb: rabbit; y: yeast; ha: hamster.

[a]Accession number of highest scoring, significant database matches is given.
(n) nucleotide level match, (p) protein level match. Score shown was calculated by the blast program (3), cutoff was 300 for blastn (n scores) and 100 for blastx (p scores).

Table 2. Pig desmin sequence comparison with intermediate filament protein sequences from other organisms.

Protein (species)	% Similarity	% Identity
Desmin (human)	98.9	98.5
Desmin (mouse)	98.1	97.2
Desmin (frog)	87.3	83.2
Desmin (zebrafish)	79.3	71.7
Vimentin (human)	72.9	63.6
Vimentin (mouse)	71.3	62.1

Table 3. Desmin genotype population study.

Breed (No.)	% of animals with indicated genotype		
	AA	AB	BB
Meishan (12)	50	33	17
Chester White (5)	80	20	0
Duroc (10)	50	50	0
Hampshire (5)	100	0	0
Landrace (11)	100	0	0
Yorkshire (9)	89	11	0
Large White(6)	100	0	0