DISS. ETH NO. 16527

Comparative and molecular approaches to identifying polymorphisms in genes associated with Arthrogryposis multiplex congenita (AMC) in swine

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH
for the degree of
Doctor of Sciences

presented by

SEM GENINI
Dipl. Ing. Agr. ETH

born 1st of July 1976
citizen of Cresciano, Ticino

accepted on the recommendation of
Prof. Dr. G. Stranzinger, examiner
Prof. Dr. P. Vögeli, co-examiner
Prof. Dr. C. Wenk, co-examiner
Prof. Dr. A. Archibald, external co-examiner

Zurich, 2006

Acknowledgements

I am deeply grateful to all those who contributed to this PhD thesis.

First, I would like to express my sincere gratitude to Prof. Dr. G. Stranzinger, Breeding Biology, Institute of Animal Sciences, Swiss Federal Institute of Technology Zurich (ETHZ), for giving me the opportunity to do my doctorate in his group and for making it possible for me to attend international conferences and workshops on genetics.

My special gratitude goes to my supervisor, Prof. Dr. P. Vögeli, Breeding Biology, Institute of Animal Sciences, ETHZ, for his guidance during my research and the suggestions and helpful criticism throughout my doctoral studies and the realization of my dissertation. His support and supervision were decisive for this work. I really enjoyed working with him.

I am extremely grateful to Prof. Dr. A. Archibald, Head of Genetics and Genomics, Roslin Institute (Edinburgh), Scotland (UK), for agreeing to be my co-examiner and for his invaluable help with the BAC clones. In this regard, I also would like to give my warmest thanks to Dr. S. Anderson.

I am particularly grateful to Prof. Dr. C. Wenk, Nutrition Biology, Institute of Animal Sciences, ETHZ, for agreeing to be my co-examiner.

For introducing me to the microarray technique, I am especially grateful to Mr. R. Talbot, A. Downing and the entire group at the ARK-Genomics Center at the Roslin Institute in Edinburgh.

I acknowledge Dr. S. Neuenschwander for his enthusiastic introduction to many different areas of molecular biology.

I am deeply indebted to Dr. C. Schelling for his ever-open door and his readiness to discuss all kinds of problems.

I am grateful to T.T. Nguyen and Š. Špilar for their technical assistance and their patience in the laboratory.

Many thanks to A. Frei, D. Steiger, G. Bärtschi, E. Wenk, K. Wüthrich and S. Beerli for their help in the laboratory and with administrative matters.

For the outstanding care of the pigs at the experimental farm of the ETHZ at Chamau, I would like to express my appreciation to Dr. H. Leuenberger, Dr. S. Gebert and all the workers at the facility, especially B. Stadler.

For excellent co-operation special thanks to all my colleagues and friends at the institute, who contributed in many different ways to the accomplishment of this work: B. Pineroli, Dr. B. Bucher, Dr. C. Hagger, D. Joller, Dr. H. Jörg, Prof. Dr. H.U. Bertschinger, Dr. L. Hasan, Dr. M. Schneeberger, Dr. M. Asai, Dr. M. Goe, Dr. P. Python and Dr. J. Kaiser.

This project was financed by an ETH grant (No 0-43113-03) and supported by other public and private organizations in Switzerland (e.g. SUISAG, Sempach, led by Dr. A. Hofer). The contribution from the Krämer Foundation of the Department for Agriculture and Food Sciences, which covered the costs of printing this dissertation, is gratefully acknowledged.

Per concludere voglio di cuore ringraziare le persone che piu' mi sono state vicine moralmente e fisicamente durante questo periodo a Zurigo.

Sicuramente mi mancheranno i pranzi e le discussioni "sportive" D. Christen, D. Zingg, F. Ménétrey, M. Frapolli e T. Paternoster, così come le lunghe serate in compagnia A. Bettio, A. Franchini, C. Fleischli, H. Nguyen, I. Taiana, M. Malek, M. Zürcher, P. Barlovic, P. Joset, R. Babey, S. Rezzonico e W. Bisson (rigorosamente in ordine alfabetico e sperando di non aver dimenticato nessuno).

I miei pensieri più intensi sono per la mia famiglia; Silvana, Marzio, Marica, Omi, Graziella e Danko senza di voi non sarei mai arrivato fin qui.

Last but not least, grazie Anna per l'aiuto, la pazienza e la comprensione durante questo duro periodo e grazie Fernando, Mari e Ferran per avermi sempre dato un grandissimo supporto.

TABLE OF CONTENTS

Summary

Arthrogryposis multiplex congenita (AMC), defined as permanent joint contractures present from birth, is one of the most common congenital defects in piglets and other mammals. For example, it is found in one of 3000 newborn babies. A genetic form of Arthrogryposis was recently identified in Swiss Large White (LW) pigs. At the beginning of 2004, at least 14 LW AI (artificial insemination) boars were known to be carriers of the defective allele and caused considerable economic losses to the Swiss pig industry. Piglets born with this syndrome had multiple defects of the legs and spinal column and did not survive. In order to study the disease, a highly inbred experimental herd with a total of 358 pigs was established. Of these, 84 (23.5%) were found to be affected, while the remaining 274 (76.5%) were normal. In the population under study, the disease was controlled by a single autosomal recessive allele designated as *amc*. A total of 219 pigs were used for linkage analysis, including seven full or half-sib founders (F_1), three F_0, 160 F_2 and 49 F_3 animals. A comprehensive genome scan revealed that the defective *amc* allele is located on SSC5. Significant pair-wise linkage (LOD > 6.00) was found for *AMC* and eight marker loci. The order that best fit with the data was *SW963-SW1987-SW152-AMC-(SW904, SW1094)-SWR1526-(SWR1974, SW310)*. *AMC* was mapped by both linkage and QTL analyses to the relative position 92 cM, between *SW152* and *SW904/SW1094*, which are located in bands q12-q23. To identify further genetic markers that can help to eliminate the disease and to determine the similarity between SSC5 and human chromosomes, porcine homologues of 18 human genes, which are interesting positional and functional candidates for AMC, were placed onto this swine chromosome with the INRA-Minnesota swine radiation hybrid panel (IMpRH). Fifteen genes (*CACNA1C*, *COL2A1*, *CPNE8*, *C12ORF4*, *DDX11*, *GDF11*, *HOXC8*, *KCNA1*, *MDS028*, *MGC5576*, *PHB2*, *PRICKLE1*, *SCN8A*, *TUBA8* and *USP18*) were selected from pig-human comparative database analysis. The remaining three functional candidates, *C3F*, *NR4A1* and *Q6ZUQ4* were selected from a microarray experiment comparing the genes expressed in an AMC-affected piglet with genes expressed in a normal piglet in three tissues (brain, muscle and spinal cord). Tissue samples were isolated at birth from the cerebellum, the musculus longissimus dorsi and the upper part of the spinal cord (where the AMC piglets showed scoliosis). The RNA was extracted from the tissues and reverse transcribed to cDNA. The cDNA was labeled with Cy3 and Cy5 for a dye swap experiment and hybridized on self-printed Qiagen pig 13K arrays containing probes for about 13,000 porcine transcripts. Twelve measurements were obtained from six slides and the results were analyzed with the BlueFuse and GeneSpring programs. Only the genes with a high confidence tag in each pair of dye

swap slides were considered. In the diseased piglet, 308 genes (~2.4%) were differentially expressed (fold change >2), 131 were upregulated and 177 downregulated. Only one gene (*SFRS6*) was overexpressed and none underexpressed in all three tissues of the AMC piglet. The differential expression of *C3F*, *NR4A1*, *Q6ZUQ4*, *SFRS6* and *SLC2A1* was also examined by real-time PCR with a TaqMan assay. The microarray findings were confirmed in all the genes but not in all the tissues, the exceptions being *SFRS6* in the brain and *SLC2A1* in the muscle. Five genes (*CPNE8*, *PRICKLE1*, *Q6ZUQ4*, *TUBA8* and *USP18*) mapped to the interval believed to contain the gene that causes AMC. The mapping data suggested that the chromosomal regions from *TUBA8* to *USP18* on HSA22 and from *CPNE8* to *PRICKLE1* on HSA12, spanning about 4.5 Mb and containing 16 genes, are the human counterparts of the region containing the porcine *AMC* locus. Furthermore, the findings of radiation hybrid mapping for SSC5q12-q22 revealed that, between the segments of HSA12p13 and HSA12q12, there is a small chromosomal interval of HSA22q11.2, spanning less than one million base pairs and containing human homologues of microsatellite *SW152* and genes *Q6ZUQ4*, *TUBA8* and *USP18*. Seven breakpoint blocks were identified and further comparative information was obtained about the extensive rearrangements in the order of the genes between HSA12, HSA22 and SSC5 in the proximity of the AMC region. Hence, 17 partial gene sequences of *ABCD2*, *CNTN1*, *CPNE8*, *FOXM1*, *KCNA1*, *KIF21A*, *SLC2A13*, *TUBA8* and *YAF2* were compared between normal and diseased piglets in order to provide additional markers for fine mapping, linkage mapping and association studies and to describe the AMC region in more detail. Genetic differences, strongly associated with the defect in the research population, were found in *TUBA8* and *CNTN1*. Three SNPs in *TUBA8* and one SNP in *CNTN1* co-segregated with the AMC phenotype in 230 pigs of the experimental herds without recombination events (LOD scores of 38.5 and 24.1, respectively). However, based on examinations of affected pigs on commercial farms, these mutations are not considered to cause AMC. Furthermore, a new microsatellite (*bE77*) locus was identified in the region. The allele $bE77^{306}$ co-segregated in the resource family with *amc* without recombination (LOD score = 51.2). An already assembled physical contig, containing BAC clones carrying the microsatellites of interest, combined with the information of 15 recombinant pigs in the studied family, enabled the delineation of a new order of loci on SSC5: *SW152-bE77-TUBA8-(AMC/CNTN1)-SW904/SW1094*. The microsatellites *bE77* and *SW904* were used as genetic markers to predict the susceptibility to AMC of 80 LW AI boars in Switzerland. The alleles $bE77^{306}$, $SW904^{180}$ and $SW904^{172}$, strongly associated with the disease in the experimental herd, indicated a higher risk of susceptibility to AMC for 17 boars (21.3%). This marker test

was also applied to 41 piglets from 14 commercial families, believed to be affected with AMC. The results confirmed the previous diagnosis in 34 cases (83%) and revealed incorrect paternities in three families. In conclusion, we developed a powerful and highly reliable marker test to discover *AMC* carriers and, thus, to limit the incidence and spread of the disease and to reduce economic losses in the Swiss pork industry. Furthermore, the results reported here are an important roadmap for future studies of poorly understood forms of AMC in humans and other species. They are a useful tool for collecting data for determining the order of genes for the sequencing project of the pig genome.

Zusammenfassung

Arthrogryposis multiplex congenita (AMC), definiert als dauerhafte Kontraktion der Gliedmassen bei Geburt, ist eine der häufigsten kongenitalen Defekte bei Ferkeln und anderen Säugetieren. Sie kommt in einem von 3000 Neugeborenen vor. Eine genetische Form von AMC wurde kürzlich in der Schweiz bei Edelschweinen (ES) entdeckt. Anfangs 2004 waren mindestens 14 ES KB (künstliche Besamung)-Eber Träger des defekten Allels und verursachten beträchtliche ökonomische Verluste in der schweizerischen Schweineproduktion. Ferkel mit diesem Syndrom hatten steife und meist gekrümmte Gliedmassen und eine gewölbte Wirbelsäule. Sie waren nicht lebensfähig. Um die Krankheit genetisch zu untersuchen, züchteten wir eine experimentelle Herde mit ingesamt 358 Schweinen. Von diesen waren 84 (23.5%) krank, während die restlichen 274 (76.5%) normal waren. In der Population wurde die Krankheit durch ein einzelnes autosomal rezessives Allel vererbt, das als *amc* bezeichnet wurde. Im Ganzen wurden 219 Schweine für die Kopplungsanalyse benützt, einschliesslich sieben Voll- oder Halbgeschwistergründer (F_1), drei F_0-, 160 F_2- und 49 F_3-Tiere. Mit einem umfassenden Genomscan konnte *AMC* Schweine-Chromosom 5 (SSC5) zugewiesen werden. Eine signifikante paarweise Kopplung (LOD-Score > 6.00) wurde für *AMC* und acht weitere Marker gefunden. Die wahrscheinlichste Reihenfolge der Loci war *SW963-SW1987-SW152-AMC-*(*SW904, SW1094*)*-SWR1526-*(*SWR1974*, *SW310*). *AMC* wurde mit Kopplungs- und QTL-Analysen auf die relative Position 92 cM kartiert, zwischen *SW152* und *SW904*/*SW1094*, die in den Bändern q12-q23 liegen. Um weitere genetische Marker zu identifizieren, die in einem diagnostischen Gentest zur Reduktion der Häufigkeit von AMC-Schweinen verwendet werden könnten, und um die Ähnlichkeit zwischen SSC5 und menschlichen Chromosomen festzustellen, wurden 18 Schweinehomologe zu menschlichen Genen, die sowohl als positionelle wie auch als funktionelle Kandidaten für AMC interessant sind, mit dem INRA-Minnesota Radiation Hybrid Panel (IMpRH) auf SSC5 kartiert. Fünfzehn Gene (*CACNA1C*, *COL2A1*, *CPNE8*, *C12ORF4*, *DDX11*, *GDF11*, *HOXC8*, *KCNA1*, *MDS028*, *MGC5576*, *PHB2*, *PRICKLE1*, *SCN8A*, *TUBA8* und *USP18*) stammten aus der vergleichenden Datenbank Schwein/Mensch. Die restlichen drei funktionellen Kandidaten *C3F*, *NR4A1* und *Q6ZUQ4* wurden aus dem Mikrochip-Experiment ausgewählt, welches die exprimierten Gene eines AMC-kranken Ferkels mit jenen eines normalen Ferkels in drei Geweben (Gehirn, Muskel und Rückenmark) verglich. Die Gewebeproben wurden kurz nach der Geburt vom Kleinhirn, dem musculus longissimus dorsi und dem oberen Teil des Rückenmarks (wo die AMC-Ferkel Skoliose zeigten) genommen, anschliessend die RNA extrahiert und in die cDNA revers transkribiert.

Die cDNA wurde mit Cy3 und Cy5 für ein Färbungstauschexperiment markiert und mit Oligonukleotiden auf selbst aufgespritzten Schweine Qiagen 13K Microchip-Platten hybridisiert. Der Mikrochip enthielt ungefähr 13000 Schweinegenproben. Auf den sechs Microchips wurden zwölf Messungen durchgeführt. Die Resultate wurden mit den Programmen BlueFuse und GeneSpring ausgewertet. Nur die Gene mit einem hohen Vertrauensgrad in jedem Färbungstauschexperiment wurden betrachtet. Im AMC-Ferkel wurden 308 Gene (~2.4%) unterschiedlich exprimiert (Expressionsunterschied > 2), davon waren 131 über- und 177 unterreguliert. Nur ein Gen (*SFRS6*) war in allen drei Geweben des AMC-Ferkels überexprimiert. Die unterschiedliche Expression von *C3F*, *NR4A1*, *Q6ZUQ4*, *SFRS6* und *SLC2A1* wurde auch mit Realzeit-PCR in einem TaqMan-Experiment überprüft. Die Resultate der Mikrochipanalyse wurden in allen Genen aber nicht in allen Geweben bestätigt, mit Ausnahme von *SFRS6* im Gehirn und *SLC2A1* im Muskel. Fünf Gene (*CPNE8*, *PRICKLE1*, *Q6ZUQ4*, *TUBA8* und *USP18*) wurden im Interval kartiert, das das verursachende Gen für AMC enthalten könnte. Die Kartierungsdaten zeigten, dass die chromosomalen Regionen von *TUBA8* bis *USP18* auf HSA22 und von *CPNE8* bis *PRICKLE1* auf HSA12, die ungefähr 4.5 Mb lang sind und 16 Gene enthalten, die menschlichen Gegenstücke zur AMC-Region bilden. Ausserdem ergaben die Resultate der Radiation Hybrid Panel-Kartierung für SSC5q12-q22, dass zwischen den Segmenten HSA12p13 und HSA12q12 ein kleiner chromosomaler Abschnitt von HSA22q11.2 besteht, der weniger als eine Million Basen umfasst, und der die menschlichen Gegenstücke zum Mikrosatelliten *SW152* und zu den Genen *Q6ZUQ4*, *TUBA8* und *USP18* enthält. Wir identifizierten sieben Blöcke und erhielten weitere vergleichende Informationen über die umfangreichen Anordnungen der Gene zwischen HSA12, HSA22 und SSC5 in der Nähe der AMC-Region. Weiter wurden 17 Teilgensequenzen von *ABCD2*, *CNTN1*, *CPNE8*, *FOXM1*, *KCNA1*, *KIF21A*, *SLC2A13*, *TUBA8* und *YAF2* zwischen normalen und kranken Ferkeln verglichen, um zusätzliche Marker für die Feinkartierung, Kopplungsanalyse und Assoziationsstudien zu erhalten und um die AMC-Region ausführlicher zu beschreiben. Genetische Unterschiede, die stark mit dem Defekt in der experimentellen Herde gekoppelt waren, wurden in *TUBA8* und *CNTN1* gefunden. Drei SNPs in *TUBA8* und ein SNP in *CNTN1* kosegregierten mit dem AMC-Phänotyp in 230 Schweinen der experimentellen Herde ohne Rekombination (LOD-Score von 38.5 beziehungsweise 24.1). Basierend auf Untersuchungen von betroffenen Schweinen in kommerziellen Populationen, waren diese Mutationen nicht kausal für AMC verantwortlich. Ausserdem wurde ein neuer Mikrosatellit (*bE77*) gefunden. In der Ressourcenfamilie kosegregierte das Allel $bE77^{306}$ mit *amc* zu 100% (LOD-Score = 51.2).

Ein bereits zusammengesetztes physikalisches Contig, das BAC-Klone mit den interessierenden Mikrosatelliten enthielt, kombiniert mit den Informationen von 15 rekombinanten Schweinen der untersuchten Familie, ermöglichte folgende neue Anordnung der Loci auf SSC5: *SW152-bE77-TUBA8-(AMC/CNTN1)-SW904/SW1094*. Die Mikrosatelliten *bE77* und *SW904* wurden als genetische Marker für *AMC* benützt, um mögliche Träger aus den 80 LW KB Ebern zu eruieren. Die Allele $bE77^{306}$, $SW904^{180}$ und $SW904^{172}$, die in der experimentellen Herde in einem hohen Kopplungsungleichgewicht mit *AMC* waren, kamen bei 17 Ebern (21.3%) vor. Dieser Markertest wurde bei 41 verdächtigen AMC-Ferkeln angewendet, die aus 14 Familien aus der Praxis stammten. Die Resultate bestätigten die Krankheitsdiagnose in 34 Fällen (83%) und deckten in drei Familien falsche Vaterschaften auf. Zusammenfassend gesagt, entwickelten wir einen leistungsfähigen und in hohem Ausmass zuverlässigen Markertest, um AMC-Träger zu entdecken und damit die Ausbreitung der Krankheit zu verringern und ökonomische Verluste in der schweizerischen Schweineproduktion zu vermindern. Ausserdem stellen die Resultate einen wichtigen Beitrag für zukünftige Studien von kaum erforschten Formen von AMC beim Menschen und anderen Spezies dar. Sie sind auch ein nützliches Werkzeug für die Organisation einer Datensammlung und für die Bestimmung der Anordnung der Gene im Hinblick auf das Sequenzierungsprojekt des Schweinegenoms.

Riassunto

L'Arthrogryposis multiplex congenita (AMC) rappresenta uno dei difetti congeniti più comuni nel maiale come, in generale, in tutti i mammiferi. Si manifesta fin dalla nascita con permanenti contratture delle articolazioni ed è, per esempio, riscontrata negli umani con una frequenza di un neonato affetto su 3000 nati. Una forma genetica d'artrogrifosi è stata recentemente identificata in Svizzera nella razza suina Large White (LW). Analisi condotte hanno permesso di identificare, ad inizio 2004, almeno 14 verri utilizzati in un programma d'inseminazione artificiale (IA) portatori dell'allele difettoso, con considerevoli perdite economiche per il settore suinicolo svizzero. I maialini nati con questa sindrome presentano molteplici difetti alle gambe ed alla colonna vertebrale tali da non consentirne la sopravvivenza. Allo scopo di analizzare geneticamente questa patologia è stata creata una famiglia sperimentale di 358 maiali altamente consanguinea, di cui 84 animali (23.5%) presentavano la tipica sintomatologia, mentre 274 (76.5%) non risultavano presentare alcun sintomo. Appare così evidente come, in questa popolazione, la malattia risulti controllata da un singolo allele autosomale recessivo indicato come *amc*. Un campione di 219 maiali, costituito da sette animali "half o full sib" fondatori (F_1), tre F_0, 160 F_2 e 49 F_3, è stato usato in un'analisi di "linkage". Attraverso analisi comprensiva del genoma, è stato dimostrato come l'allele difettoso *amc* sia posizionato sul cromosoma 5 di maiale (SSC5). Un significativo "pair-wise linkage" (LOD > 6.00) è stato trovato per *AMC* e otto marcatori. L'ordine che meglio si adattava con i dati era *SW963-SW1987-SW152-AMC-(SW904, SW1094)-SWR1526-(SWR1974, SW310)*. Attraverso sia analisi di "linkage" che di QTL, è stato possibile mappare *AMC* alla posizione relativa di 92 cM, tra *SW152* e *SW904/SW1094*, situati nelle bande q12-q23. Per l'identificazione d'ulteriori marcatori genetici utili ad eliminare la malattia ed a determinare la similarità di SSC5 con cromosomi umani, si è proceduto alla mappatura sul cromosoma SSC5 di 18 geni suini omologhi di geni umani, riconosciuti come interessanti candidati posizionali e funzionali per AMC, mediante l' "INRA-Minnesota swine radiation hybrid panel" (IMpRH). Quindici di questi geni (*CACNA1C*, *COL2A1*, *CPNE8*, *C12ORF4*, *DDX11*, *GDF11*, *HOXC8*, *KCNA1*, *MDS028*, *MGC5576*, *PHB2*, *PRICKLE1*, *SCN8A*, *TUBA8* e *USP18*) sono stati scelti tramite analisi comparativa delle sequenze geniche depositate in banca dati, mentre i rimanenti tre (*C3F*, *NR4A1* e *Q6ZUQ4*) da un'analisi d'espressione genica basata su tecnologia microarray tra maiali affetti e soggetti sani condotta su tre diversi tessuti (cervello, muscolo e midollo spinale). I campioni di tessuto sono stati prelevati alla nascita rispettivamente dal cervelletto, dal muscolo *longissimus dorsi* e dalla parte superiore del midollo spinale (dove i soggetti

affetti da AMC presentano scoliosi). Dopo averne estratto l'RNA si è quindi proceduto alla produzione dei rispettivi cDNA tramite trascrizione inversa. I cDNA sono stati marcati con i fluorofori Cy3 e Cy5 come indicato nel protocollo per "dye swap experiments" e quindi ibridizzati su self-printed Qiagen pig 13K arrays contenenti circa 13000 geni porcini. Un totale di dodici misure sono state ottenute da sei vetrini ed i relativi risultati analizzati con i programmi GeneSpring e BlueFuse. Solamente i geni presentanti valori affidabili per ogni coppia di vetrini sono stati considerati. Nel soggetto affetto, 308 geni (~2.4%) hanno manifestato espressione differenziale (differenza d'espressione > 2), e precisamente 131 sono risultati essere sovraregolati mentre 177 sottoregolati. In particolare il solo gene *SFRS6* è risultato essere sovraespresso nel soggetto affetto in tutti e tre i tessuti contemporaneamente, mentre nessuno sottoespresso. In seguito si è proceduto all'analisi d'espressione di *C3F*, *NR4A1*, *Q6ZUQ4*, *SFRS6* e *SLC2A1* attraverso "real-time" PCR basata su chimica TaqMan. Questa ha confermato i risultati ottenuti in precedenza tramite analisi microarray, con le sole eccezioni rappresentate da *SFRS6* nel cervello e *SLC2A1* in muscolo. Cinque geni (*CPNE8*, *PRICKLE1*, *Q6ZUQ4*, *TUBA8* e *USP18*) sono stati mappati nell'intervallo ipotizzato contenere il gene responsabile della patologia. Questi dati di mappatura suggerivano che le regioni cromosomiche comprese tra *TUBA8* e *USP18* su HSA22 e tra *CPNE8* e *PRICKLE1* su HSA12, lunghe circa 4.5 Mb e contenenti 16 geni, siano le controparti umane della regione contenente il locus *AMC* suino. Nello stesso modo il mappamento per SSC5q12-q22 ha rivelato che, fra i segmenti di HSA12p13 e di HSA12q12, è presente un piccolo intervallo cromosomico di HSA22q11.2, che misura meno di un milione di basi, e che contiene gli omologhi umani del microsatellite *SW152* e dei geni *Q6ZUQ4*, *TUBA8* e *USP18*. Sette "breakpoints blocks" sono stati identificati ed ulteriori informazioni comparative sono state ottenute sulle molteplici riorganizzazioni nell'ordine dei geni fra HSA12, HSA22 e SSC5 in prossimità della regione AMC. Partendo da ciò, 17 sequenze parziali dei geni *ABCD2*, *CNTN1*, *CPNE8*, *FOXM1*, *KCNA1*, *KIF21A*, *SLC2A13*, *TUBA8* e *YAF2* sono state comparate fra individui sani ed affetti nell'intento di ottenere sia dei marcatori supplementari da utilizzare in ulteriori studi di "linkage" e di associazione, che una descrizione e mappatura piu' fine della regione di AMC. Differenze genetiche fortemente connesse con il difetto sono state trovate nella popolazione in analisi in *TUBA8* ed in *CNTN1*. Tre SNPs in *TUBA8* ed uno SNP in *CNTN1* sono risultati co-segregare senza eventi di ricombinazione (valori LOD di 38.5 e 24.1 rispettivamente) con il fenotipo AMC in 230 maiali. Tuttavia, dopo aver esaminato capi affetti provenienti da allevamenti commerciali, queste mutazioni sono state escluse essere la causa della patologia. In seguito un nuovo microsatellite (*bE77*) è stato

identificato nella regione. L'allele *bE77*^{306} è risultato co-segregare con *amc* senza ricombinazioni (LOD = 51.2). Un "physical contig" gia' assemblato, contenente cloni BAC portanti i microsatelliti d'interesse, uniti alle informazioni di 15 maiali ricombinanti appartenenti alla famiglia studiata, ha permesso la delineazione di un nuovo ordine di loci su SSC5: *SW152-bE77-TUBA8-(AMC/CNTN1)-SW904/SW1094*. I microsatelliti *bE77* e *SW904* sono stati poi utilizzati come marcatori genetici per la predizione della suscettibilità verso questa patologia di 80 verri LW utilizzati per l'inseminazione artificiale in Svizzera. Gli alleli *bE77*^{306}, *SW904*^{180} e *SW904*^{172}, strettamente associati alla malattia, hanno indicato un elevato rischio predispositivo per AMC in 17 sugli 80 verri analizzati (21.3%). Questo test di marcatori è stato poi applicato a 41 maialini, appartenenti a 14 famiglie commerciali, ritenuti essere affetti da AMC. I risultati hanno confermato la diagnosi precedente in 34 casi (83%) mentre hanno rivelato casi di falsa paternità in tre delle 14 famiglie. In conclusione, in questo lavoro è stato sviluppato un potente ed altamente affidabile test, basato su marcatori molecolari, per l'identificazione di capi portatori di AMC, utilizzabile per limitare l'incidenza e la diffusione della malattia e per ridurre così le perdite economiche al settore suinicolo svizzero. I risultati riportati costituiscono inoltre un'importante base per futuri studi su forme poco conosciute d'AMC negli umani, come pure in altre specie. Essi rappresentano anche un utile mezzo di raccolta dati per la determinazione dell'ordine dei geni nell'ormai prossimo progetto di sequenziamento del genoma del maiale.

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS, SYMBOLS, UNITS

A:	Ampère
A, C, G, T:	adenine, cytosine, guanine, thymine
AI:	artificial insemination
AMC:	Arthrogryposis Multiplex Congenita
APS:	ammonium peroxydisulfate
At:	annealing temperature
A_x:	absorbance at wavelength of x nm (x = 260, 280)
BAC:	bacterial artificial chromosome
BES:	BAC end sequences
BSA:	bovine serum albumin
bp:	base pair
C:	concentration
°C:	degree Celsius
cDNA:	complementary DNA
CDS:	coding sequence
cM:	centimorgan
c*x*:	centi-*x* (= 10^{-2} *x*); x is a unit of the international system
C_T:	cycle threshold
DA:	Dystal Arthrogryposis
ddH_2O:	double distilled water
DEPC:	diethyl pyrocarbonate
df:	dilution factor
DGCR:	DiGeorge syndrome critical region
DNA:	deoxyribonucleicacid
dsDNA:	double stranded DNA
dNTP(s):	deoxyribonucleotide(s) = deoxynucleoside triphosphate solutions
dATP:	2'-deoxyadenosine 5'-triphosphate
dCTP:	2'-deoxycytidine 5'-triphosphate
dGTP:	2'-deoxyguanosine 5'-triphosphate
dTTP:	2'-deoxythymidine 5'-triphosphate
E. coli:	*Escherichia coli*
EDTA:	ethylenediaminetetraacetic acid
EST:	expressed sequence tag
EtBr:	ethidium bromide
*Fok*I:	Flavobacterium okeanokoites
FISH:	fluorescence in situ hybridization
g:	gram
GCG:	Genetics Computer Group
h:	hour
HCl:	hydrochloric acid
HSA:	Homo sapiens chromosome
H_2O:	water, see ultra pure H_2O
ID number:	identification number
IMpRH:	INRA and Minnesota porcine radiation hybrid panel
INRA:	Institut National de la Recherche Agronomique
IPTG:	isopropyl β-D-thiogalactopyranoside
k:	kilo-*x* (= 10^3 *x*); x is a unit of the international system
KCl:	potassium chloride
l:	liter
LB:	Luria Bertani
LOD:	logarithm of the odds
LW:	Large White
M:	molarity (= mol/liter)
MAS:	marker assisted selection
Mb:	megabase (= 10^6 base pair)
$MgCl_2$:	magnesium chloride
min:	minute
mol:	mole
MOPS:	3-(N-mopholino) propane sulfonic acid
mRNA:	messenger RNA

MW:	molecular weight
m*x*:	milli-*x* (= 10^{-3} *x*); x is a unit of the international system
µ*x*:	micro-*x* (= 10^{-6} *x*); x is a unit of the international system
n*x*:	nano-*x* (= 10^{-9} *x*); x is a unit of the international system
NaAc:	sodium/natrium acetate
(n)ACh(R):	(nicotinic) acetylcholine (receptor)
NaCl:	sodium chloride
NADH / NADPH:	nicotinamide adenine dinucleotide / NAD phosphate
NaOH:	sodium hydroxide
Na-phosphate:	sodium phosphate buffer
NCBI:	National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov)
nt:	nucleotide
OD:	optical density
ORF:	open reading frame
p:	probability
PAGE:	polyacrylamide gel electrophoresis
PB:	pancuronium bromide
PBS:	phosphate buffered saline
PCR:	polymerase chain reaction
PIC:	polymorphism information content
pmol:	picomole
ppm:	particles *per* million (= mg/kg)
QTL:	quantitative trait locus
QTN:	quantitative trait nucleotide
RFLP:	restriction fragment length polymorphism
Rn:	normalized reporter
RNA:	ribonucleic acid
RNase:	ribonuclease
RNasin:	ribonuclease inhibitor
rpm:	revolutions *per* minute
RT:	reverse transcription
RH:	radiation hybrid
SD:	standard deviation
SDS:	sodium dodecyl sulphate
s:	second (time)
SMA:	Spinal muscular atrophy
SNP:	single nucleotide polymorphism
SSC:	*Sus scrofa* chromosome
ssDNA:	single stranded DNA
STS:	sequence tagged site
SUISAG:	public company for pig breeding services in Switzerland
t:	time
TBE:	tris-borate-EDTA buffer
TE:	tris-EDTA buffer
TEMED:	N, N, N', N'-tetramethylethylenediamine
temp:	temperature
Tm:	melting temperature
*Tsp*45I:	*Thermus* species YS45
U:	unit
Ultrapure H_2O:	deionized, ultrafiltered water
UTR:	untranslated region
UV:	ultraviolet light
V:	volt
v/v:	volume *per* volume
W:	Watt
w/v:	weight *per* volume
x:	times
XCFF:	xylene cyanol FF
x g:	acceleration of gravity
X-Gal:	5-bromo-4-chloro-3-indolyl β-D-galactopyranoside
YAC:	yeast artificial chromosome

GENE SYMBOLS

ABCD2:	ATP-binding cassette, sub-family D, member 2
ASB8:	Ankyrin repeat and SOCS box protein 8
BIN2:	Bridging integrator-2
BMP11 (=*GDF11*):	Bone morphogenetic protein 11
CACNA1C:	Calcium channel, voltage-dependent L-type alpha-1C subunit
CD9:	Leukocyte antigen MIC3
COL1A2:	Alpha 2 type I collagen
COL2A1:	Alpha 1 type II collagen
CPNE8:	Copine-8
CNTN1:	Contactin 1
C1R:	Complement component 1, r subcomponent
C1S:	Complement component 1, s subcomponent
C3F:	Gene rich cluster, C3f gene
C12ORF4:	Chromosome 12 open reading frame 4
DDX11:	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 11
FOXM1:	Forkhead box protein M1
GDF11 (=*BMP11*):	Growth differentiation factor 11
HOXC8:	Homeobox gene family C8
KCNA1:	Potassium voltage-gated channel, shaker-related subfamily, member 1
KIF21A:	Kinesin family member 21A
LRRK2:	Leucine-rich repeat kinase 2
MDS028:	Uncharacterized hematopoietic stem/progenitor cells protein MDS028
MGC5576:	Hypothetical protein MGC5576
MUC19:	Mucin 19
NELL2:	Protein kinase C-binding protein *NELL2* precursor
NR4A1:	Nuclear receptor subfamily 4, group A, member 1
PAQR2:	Adiponectin receptor protein 2
PCBP2:	Poly(rC) binding protein 2
PDZRN4:	PDZ domain containing RING finger 4
PHB2:	Prohibitin-2
PPFIBP1:	Protein tyrosine phosphatase receptor type f polypeptide-interacting protein binding protein 1
PPHLN1:	Periphilin 1
PRICKLE1:	Prickle-like protein 1 (REST/NRSF-interacting LIM domain protein 1)
P11:	Placental protein 11 precursor
Q6ZUQ4:	Hypothetical protein FLJ43449
Q8IXV1:	Hypothetical protein LOC283464
SCN8A:	Sodium channel, voltage gated type VIII alpha polypeptide
SFRS6:	Splicing factor arginine/serine-rich 6
SLC2A1:	Solute carrier family 2, facilitated glucose transporter member 1
SLC2A13:	Solute carrier family 2 (facilitated glucose transporter), member 13
SLC6A12:	Sodium- and chloride-dependent betaine transporter 6, member 12
SMN:	Survival motor neuron
TUBA8:	Tubulin alpha-8 chain
UBP43 (=*USP18*):	Ubiquitin-specific protease 43-KD
USP18 (=*UBP43*):	Ubiquitin-specific protease 18
YAF2:	YY1 associated factor 2

1. Introduction

Kondrashov (1995) calculated that, in mammals, as many as 100 new neutral mutations, those which do not affect the organism's chances of survival in its natural environment because they neither enhance nor impair the organism carrying them, accumulate through the generations in the entire genome of each individual. Mutations are considered to be the driving force of evolution and keep evolution going. Those that are less favorable (*deleterious*) are removed from the gene pool by natural selection, while more favorable (that can lead to useful gene variations) and neutral mutations, the overwhelming majority, tend to accumulate. Eyre-Walker & Keightley (1999), analyzing only protein-coding regions, estimated that, in humans, gorilla and chimpanzee, there are about two to five deleterious gene mutations per individual and per generation. This is a reasonable approximation for all mammals. The focus of this dissertation is on one or more deleterious gene mutations, which have a negative effect on the health of piglets and cause Arthrogryposis Multiplex Congenita.

1.1 Arthrogryposis Multiplex Congenita (AMC)

Arthrogryposis (also known as arthrogryposis multiplex congenita (AMC), amyoplasmia congenita or multiple congenital articular rigidities) is caused by one or more unfavorable mutations and literally means "curved or hooked joints" (Greek). This term is used to describe the presence of two or more (multiplex) joint contractures (arthrogryposis) present at birth (congenita).

This disease was recently identified in the Swiss Large White (LW) breed and was previously described in Yorkshire piglets (Ely & Leipold 1979). It is one of the most common congenital defects in piglets, humans and other mammals (section 1.1.4).

It is generally accepted that AMC has many causes and can be classified into neurogenic and myogenic forms (Adams *et al.* 1962). The neurogenic forms of AMC are more common and are usually associated with spinal cord dysplasias, including various degrees of loss of ventral horn (motor) neurons and syringomyelia with fibrofatty replacement of muscles, typical of denervation atrophy (Edwards 1971). It has been proposed that prolonged prenatal immobilization will lead to fixation of the joints of the immobilized parts (Panter *et al.* 1988; Whittem 1957).

There are a number of known or suspected etiologies of AMC in piglets; the most important of these include genetic factors, nutritional deficiencies, toxic plants, pharmacological and viral agents and prolonged prenatal immobilization. For example, feeding pigs tobacco stalks (*Nicotiana tabacum*) between day 4 and 50 of pregnancy was reported to cause the condition (Crowe 1978). Between 40 and 50% of piglets in affected litters were deformed. Poison

hemlock (*Conium maculatum*) and jimsonweed (*Datura stramonium*) have also been reported to cause AMC in Large White and Landrace piglets (Dyson & Wrathall 1977; Leipold *et al.* 1973, respectively). However, Hampshire piglets born from sows administered jimsonweed in the first three months of gestation did not show symptoms of the disease (Keeler 1981).

Piglets affected with AMC may cause dystocia, as parturition has been reported to be prolonged by several hours and neonatal deaths from asphyxia and hypoglycemia were common.

Death of newborn piglets, which have AMC, is believed to result from respiratory arrest due to neurogenic malfunction of the respiratory centre, hypoplasia of the muscles involved in respiration in the thorax or a combination of both (Ely & Leipold 1979). Fetal atelectasis together with amniotic fluid in the respiratory passages is another indication of terminal respiratory failure. In the AMC form described here, it was impossible to establish unequivocally the origin of the disease.

1.1.1 Clinical and pathological findings

AMC is defined as permanent fixation or ankylosis of the joints, and can occur in the forelimbs, hind limbs and/or the vertebral column, which leads to various degrees of flexion or extension at birth and/or scoliosis (Figure 1.1, 1 and 2, respectively).

Moreover, AMC may affect ligaments, skeletal muscles, the central and peripheral nervous systems or any combination of these tissues. A shortened lower jaw (brachygnathia inferior) has also been observed in affected piglets (Figure 1.1, 3). However, cleft lip, cleft palate, hydrocephalus or other cranial anomalies have not been linked with AMC and malformation of internal organs has not been detected. All the affected piglets were stillborn (Figure 1.1, 4).

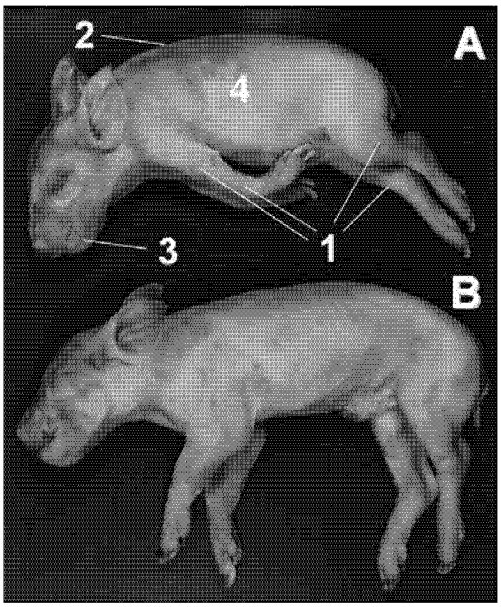

Figure 1.1: AMC in piglets.
A: piglet affected by arthrogryposis multiplex congenita (AMC)
B: normal piglet
1) The joints of the fore and hind limbs are fixed and are parallel to the thorax
2) Scoliosis
3) Brachygnathia inferior
4) Stillborn

1.1.2 Mode of inheritance

Ely & Leipold (1979) and Lomo (1985) reported purebred Yorkshire piglets with AMC and suggested that the syndrome is genetic and is controlled by a single autosomal recessive allele (*amc*).

The mutation responsible for *AMC* in the Swiss Large White population and analyzed in this study occurred in the boar 2401 JN (Hift), which was born in 1990 and then used widely for artificial insemination (AI). Its defective gene was spread through AI; however, the exact number of affected animals in the entire Swiss population is unknown, mostly due to faulty or missing records.

1.1.3 Relevance of the disease

In recent months, interest in AMC and the importance of this project has increased considerably and the number of affected LW pigs augmented dramatically. At the end of 2003, at least 14 AI boars were suspected of carrying the defect, which led to considerable economic losses for farmers and the entire pork industry in Switzerland (Hofer 2003). The present study was initiated (1) to develop a genetic test to reduce the incidence of the disease and (2) to use porcine AMC as animal model for AMC in human and other species. Recent discoveries about the human genome were made possible by applying information about

animal species (Andersson & Georges 2004; Burt 2002; Milan *et al.* 2000 A), making the project an attractive and interesting route for future studies in humans and other species.

1.1.4 AMC in humans

AMC in humans has been known for a long time and is estimated that one of every 3000 newborn babies has the disease (Anderson 1997; Staheli *et al.* 1998). There are more than 50 distinct forms of AMC, which is consistent with the variety of symptoms associated with the disease. Not all these forms have a genetic origin, and not all have been mapped to a specific human chromosome. Table 1.1 lists the 26 most common genetic forms of AMC that have been mapped to a human chromosome and shows their chromosomal locations.

Table 1.1: AMC in humans with gene and location on HSA.
The gene, in which the mutation occurred, is shown in parentheses along with its mode of inheritance (1=dominant, 2=recessive, 3=unknown)[1].

Disease name	Human chromosome	Gene symbol
Arthropathy-camptodactyly syndrome	1q24-q25	*CACP* (*PRG4*)[2]
Popliteal pterygium syndrome	1q32-q41	*PPSI*[1]
Lethal spinal muscular atrophy	5q12.2-q13.3	*SMA* (*SMN1*)[2]
Congenital contractual arachnodactyly	5q23-q31	*CCA* (*FBN2*)[1]
Diastrophic dysplasia	5q32-q33.1	*DTD*[2]
AMC neurogenic	5q35	*AMCN*[2]
Saethre-Chotzen syndrome	7p21	*SCS* (*TWIST1*)[1]
Distal arthrogryposis type 1A (DA1)	9p21-q21.2	*AMCD1*[1]
Distal arthrogryposis type 1A (DA1)	9p13.2-p13.1	*AMCD1* (*TPM2*)[1]
Lethal congenital contracture syndrome 1	9q34	*LCCS1*[2]
Cerebrooculofacioskeletal syndrome	10q11	*COFS* (*ERCC6, CSB*)[2]
Myasthenic syndrome, congenital, associated with episodic apnea	10q11.2	*CMS-EA* (*CHAT*)[2]
Saethre-Chotzen syndrome	10q26	*SCS* (*FGFR2*)[3]
Distal arthrogryposis type 2B, Freeman-Sheldon syndrome variant (DA2B)	11p15.5	*FSSV* (*TNNI2*)[1]
Distal arthrogryposis type 2B, Freeman-Sheldon syndrome variant (DA2B)	11p15.5	*FSSV* (*TNNT3*)[1]
Cerebrooculofacioskeletal syndrome	13q33	*COFS* (*ERCC5*)[3]
Spondylocostal dysostosis, Jarcho-Levin syndrome	15q26.1	*SCDO1* (*MESP2*)[2]
Arthrogryposis, renal dysfunction and cholestasis	15q26.1	*ARC* (*VPS33B*)[3]
Distal arthrogryposis type 7, trismus-pseudocamptodactyly syndrome (DA7)	17p13.1	(*MYH8*)[1]
Symphalangism proximal	17q21-q22	*SYM1*[1]
Niemann-Pick disease, type C1	18q11-q12	*NPC1*[2]
Spondylocostal dysostosis, Jarcho-Levin syndrome	19q13	*SCDO1* (*DLL3*)[2]
Cerebrooculofacioskeletal syndrome	19q13.2-q13.3	*COFS* (*ERCC2*)[3]
Lethal multiple pterygium syndrome	X linked	*LMPS*[2]
Infantile spinal muscular athrophy with arthrogryposis	Xp11.3-q11.2	*AMCX1*[2]
Arthrogryposis X-linked type V	Xq23-q27	*AMCX5*[2]

The most important category of arthrogryposis comprises ten forms of distal AMC (DA1-DA10), characterized by contractures of the fingers and toes. The most serious forms are DA6 and DA9, which to the best of my knowledge have not been mapped.
None of the AMC disorders in humans is progressive and only a few rare forms are lethal at birth.

1.1.5 AMC in other species

Arthrogryposis is known in almost every mammalian companion and farmed species, because the name of the syndrome is purely descriptive and a broad spectrum of different symptoms can be assigned to it. The most important species with naturally occurring AMC and a similar phenotype description, as it is presented in this study (section 1.1.1), are cattle, sheep, chicken, mouse, buffalo, goat, dog and horse. The cause of the disease is usually a toxic substance or a virus and only in few cases genetic. An overview of various AMC forms in these species, considered to be possible counterparts of pig AMC, is presented in the next sections.

1.1.5.1 Cattle

Inherited *AMC* has been reported for several breeds of cattle, for example in Charolais (Duchesne & Eggen 2005), Piedmont calves (Longeri *et al.* 2003), Holstein (Bahr *et al.* 2004), Brown Swiss x Braunvieh cross calves (Dirksen *et al.* 1992) and Hessian (Klein & Herzog 1990). As for pigs (section 1.1.1), AMC in cattle can also be caused by ingestion of toxic plants by pregnant cows (Panter *et al.* 1998) and often by viruses (Kitani *et al.* 2000; Tsuda *et al.* 2004).
Especially in cattle, arthrogryposis and congenital contractures have been recorded and classified as, for example, perosomus elumbis (Jones 1999), spinal muscular atrophy (*SMA*) (Pumarola *et al.* 1997) and complex vertebral malformation (*CVM*) (Agerholm *et al.* 2001).

1.1.5.2 Sheep

AMC is common in sheep. It was found in Suffolk lambs as autosomal recessively inherited (Doherty *et al.* 2000), in the uterus of ewes inoculated with the bunyaviruses La Crosse, San Angelo and Main Drain in early gestation (Edwards *et al.* 1997) and in animals grazing sorghum pastures (Bradley *et al.* 1995) and hemlock (Lopez *et al.* 1999).

1.1.5.3 Chicken

Chicken embryos are the best suited for use as models for toxicology studies to induce arthrogryposis. AMC was caused by injecting coniine, extracted from *Conium maculatum*

(Forsyth *et al.* 1996) and pancuronium bromide (PB) (Lamb *et al.* 2003) into eggs of chicken. Furthermore, Kitano *et al.* (1997) reported that lesions, occurred in chicken after inoculating the yolk sac with Aino virus, were similar to the lesions reported for calves suffering from AMC.

1.1.5.4 Mouse

Arthrogryposis is very rare in mouse. Interesting findings were reported by Jacobson *et al.* (1999) and Polizzi *et al.* (2000), who showed that injecting pregnant mice with plasma of women, whose foetuses had severe AMC, resulted in AMC in the mouse foetuses. Furthermore, contrary to the other species described above, AMC was not induced in mice and rats fed with *Conium maculatum* (Forsyth *et al.* 1996).

In conclusion, mice can suffer from muscular dysgenesis (*MDG*), transmitted as an autosomal recessive trait (Banker 1977), which has features similar to those of AMC in pigs.

1.1.5.5 Buffalo

A very similar form of AMC as found in pigs was discovered in a herd of Murrah buffalo in Brazil (Schild *et al.* 2003). The calves were stillborn, showed rigidity and fixation of the limbs, as well as scoliosis and brachygnathia. The authors suggested an autosomal recessively inherited disease and presented a picture of an affected calf, the symptoms of which were identical to the AMC piglet in Figure 1.1.

1.1.5.6 Goat, dog and horse

Reports of AMC in goat (Panter *et al.* 1990), dog (Blazej *et al.* 1998, as spinal muscolar atrophy) and horse (Mayhew 1984; Nes *et al.* 1982) are rare but confirm the presence of the disease in the most important of the mammalian species.

1.2 From the phenotype to the causative mutation

1.2.1 Establishment of an experimental family with a segregating phenotype

The factors that limit the exact mapping of a monogenic disease gene are the extent of informative meiosis, the number of available polymorphic markers and the necessity of a clear and unambiguous phenotype. The formation of an experimental herd of selected founder animals with segregating phenotypes is very expensive and is a prerequisite for the successful analysis of genomes.

In recent years, considerable advances have been made owing to the increasing number of new markers. Although microsatellites (section 1.2.4.1.1) are still the backbone of the characterization of the genome of farm animals, a growing number of SNPs (section 1.2.4.1.2) complements them.

The present study was not limited in this way. The experimental herd at Chamau comprised 358 pigs (Table 2.1), of which 14 were recombinants with flanking markers in the *AMC* region. Furthermore, the characterization of the disease was easy, clear and a high number of polymorphic markers was available.

1.2.2 Comparative candidate gene approach

The comparative approach to studying candidate genes in farm animals is an important tool for identifying genes, which have particularly well-defined phenotypes, very similar to diseases in human or other mammals.

Comparative mapping (for reviews see Burt 2002; Hayes *et al.* 2003) makes use of knowledge about chromosome segment conservation in different species to predict the position of homologous genes in species, in which the genes have not been mapped and for which information on the genome is sparse. The search for candidate genes can be based on clinical, physiological and pathological similarities. The resolution of comparative maps increased with new technologies, such as chromosome G-banding, ZOO-FISH, bidirectional chromosomal painting and mapping with radiation hybrid panels (section 1.2.6). Furthermore, the availability of complete genomic sequences of several species permits the comparison of nucleotides, which is the ultimate resolution of a map. The enormous amount of data collected from sequencing projects are available in http://www.ensemble.org and http://www.ncbi.nlm.nih.gov.

The comparative candidate gene approach was successful in various studies of pigs, for example for the identification of the L-gulono-gamma-lactone oxidase gene (*GULO*), causing vitamin C deficiency in pigs (Hasan *et al.* 2004), a marker allele cosegregating 100% with the

recessive allele involved in congenital progressive ataxia (CPA) and spastic paresis (SP) (Kratzsch *et al.* 1999), the so-called halothane gene (Vögeli *et al.* 1994) responsible for malignant hyperthermia syndrome in pigs as well as the α(1,2) fucosyltransferase 1 (*FUT1*) gene controlling resistance and susceptibility to the adhesion and colonization of *E. coli* F18 in the small intestine (Meijerink *et al.* 2000).

Another successful example of comparative mapping and candidate gene studies of mouse and pig showed that the porcine gene *KIT*, encoding the mast cell growth factor (*MGF*) receptor, is the causative gene for the dominant white color of pigs (Moller *et al.* 1996).

The comparative candidate gene strategy in this study was applied to determine whether the gene causing AMC in swine is the same as that responsible for the similar cases of AMC in humans, i.e. the neurogenic *AMCN*, which maps to HSA5q35 (Shohat *et al.* 1997) or spinal muscular atrophy (*SMA*). *SMA* is a subtype of heritable human *AMCN* and is often caused by several exon or gene duplications and deletions or point mutations in the gene survival motor neuron (*SMN*) (Lefebvre *et al.* 1995; Lorson *et al.* 1999), which maps to HSA5q13. Both these regions in humans correspond to evolutionarily conserved segments on porcine chromosome 2, segments q27-28 and q22-23, respectively (Rettenberger *et al.* 1995; http://www.toulouse.inra.fr/lgc/pig/compare/compare.htm).

The hypothesis of linkage between *AMC* and microsatellite markers on SSC2 (section 2.2.3.2.1) was investigated.

1.2.3 Functional candidate gene approach

1.2.3.1 Analysis of gene expression

1.2.3.1.1 Microarray analysis

Microarray technology is a powerful tool for analyzing gene function and has revolutionized life-science research. It was first described by Schena *et al.* (1995) and enables the role played by genes in living organisms and how they are expressed in tissues or organs to be studied. Microarrays are a necessary high throughput tool for profiting from the avalanche of new data. Furthermore, they are a fast growing technology used to investigate the expression levels of thousands of genes simultaneously. They are particularly useful in studies of gene expression, analysis of chromatin structure, genome mapping, SNP discrimination, detection of transcription factor activity as well as in many other applications. Figure 1.2 shows the most important steps of this technique. The first stage is the isolation of mRNA from the tested and the reference animal, which will be compared in the experiment (a). The mRNA is reverse transcribed to cDNA and for example that of the test animal is labeled with Cy3 (green), that of the reference animal with Cy5 (red). This process is repeated with the

reversed colors in a dye swap experiment to eliminate possible influences of the different dyes on the results (b). Equal amounts of the two differently labeled cDNAs of the two animals are mixed together (c) and hybridized to the microarray plate (f). The microarray plate consists of purified oligonucleotides or spotted cDNA probes (d) attached by a robot to several thousand positions (spots) on a glass surface (e). After washing to eliminate unbound labelled cDNA, the arrays are scanned and the fluorescence intensity is measured. Spots for genes that are equally expressed in both animals appear yellow; in overexpressed genes the color used for labeling (red or green) predominates and when expression of the gene does not occur the spot is black. The spots are shown in a diagram (left) and in a real array (right) (g). Finally, the results are evaluated using specific programs and a list of up- and downregulated genes is provided (h).

The signals resulting from this technology are proportional to the relative concentration of the labeled target, its specificity, the duration and temperature of hybridization and the amount of probe material in each spot. The poor reproducibility of microarray experiments is one of the biggest obstacles that must be overcome. For these reasons, this method should be regarded as a primary screening method. Therefore, the use of quantitative PCR for re-examining the levels of transcription for genes selected by microarray data is indispensable and routinely performed. The significant genes discovered with the microarray technology in this study were confirmed by real-time PCR with TaqMan probes (section 1.2.3.1.2).

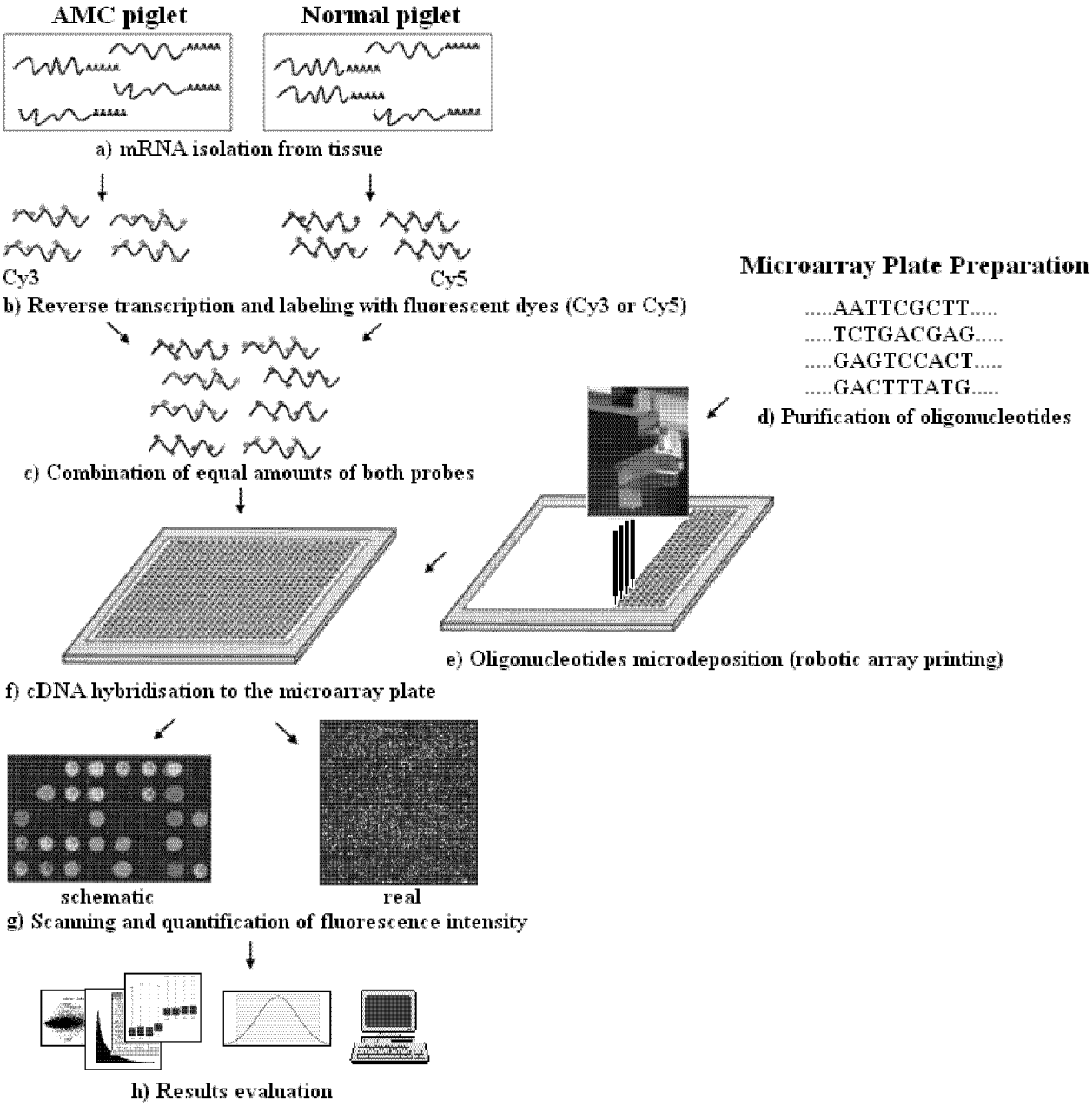

Figure 1.2: Work flow of a microarray experiment.
Letters a to h are explained in the text.

1.2.3.1.2 Real-time PCR with TaqMan probes

Real-time PCR is today the method of choice to confirm differential gene expression detected in microarray experiments. This amplification system permits the direct quantification of the initial amount of template by monitoring amplicon production during each PCR cycle, in contrast to other methods, by which the final amount produced is detected (Freeman *et al.* 1999). When the TaqMan assay is used, the detection and quantification of a fragment is achieved by visualizing a fluorescent signal emitted by the TaqMan primer. The measured fluorescence is proportional to the amount of PCR product present in the reaction. The entire assay relies on the 5'-3' exonuclease activity of the Taq DNA polymerase, which cleaves the

TaqMan probe. The probe contains a fluorescent reporter dye at the 5' end and a quencher dye at the 3' end. The quencher (TAMRA) absorbs the fluorescent energy from the reporter dye (FAM) when the probe is intact. When the probe is cleaved by the 5'-3' exonuclease of the DNA polymerase, the reporter dye emits fluorescence, which is monitored by the computer software until amplification stops (Figure 1.3).

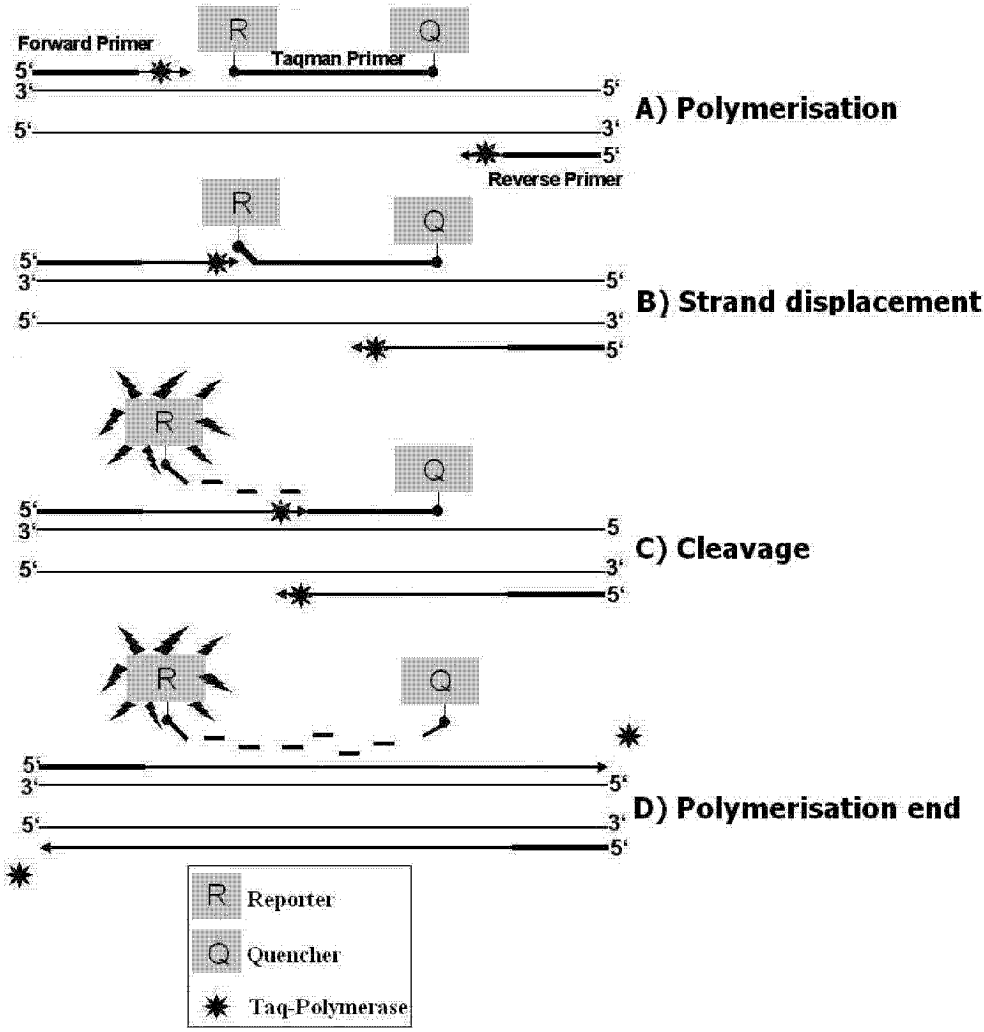

Figure 1.3: Principle of real-time PCR with a TaqMan probe.

Representation of real-time PCR with a fluoregenic probe to detect a specific product during amplification. The forward, reverse and TaqMan (containing a reporter and a quencher) primers anneal to the target area. The Taq DNA polymerase starts to add nucleotides (A: polymerisation) to the primers and extends the newly formed DNA strand (B: strand displacement). When it reaches the TaqMan primer, it cleaves the quencher from the reporter due to its 5' exonuclease activity (C: cleavage). The fluorescence emitted by the cleaved reporter dye is detected and accumulates with each cycle until polymerisation ends (D), increasing the intensity of fluorescence proportionally to the amount of amplicon production (slightly modified from TaqMan® Universal Master Mix Protocol, Applied Biosystems).

The magnitude of the fluorescent signal generated by real-time PCR is calculated according to the formula

$$\Delta Rn = (Rn^{+}) - (Rn^{-})$$

where **ΔRn** is the difference between Rn^{+} and Rn^{-}, $\mathbf{Rn^{+}}$ the value of the reaction containing all the components (the sample of interest, blue in Figure 1.4) and $\mathbf{Rn^{-}}$ the normalized fluorescent intensity of a passive reference dye (baseline value, red in Figure 1.4).

ΔRn plotted against cycle numbers produces the amplification curves and gives the threshold cycle (C_T) value (Figure 1.4), the most important parameter for quantification. C_T is defined as the cycle, at which fluorescence emission crosses the fixed threshold (horizontal line in Figure 1.4, set by user), in other words when the system begins to detect the increase in the fluorescent signal (and therefore cDNA) associated with exponential growth of the PCR product during the logarithmic phase. This phase provides the most useful information about the reaction. C_T is proportional to the amount of starting material (Heid *et al.* 1996); the greater the initial amount of cDNA, the sooner the accumulated product is detected in the PCR process and the lower the C_T value.

The baseline value represents the minimum level of expression (Figure 1.4), where PCR has a large quantity of background signals due to changes in the reaction medium and before the significant increase in the target amplicon. Therefore, the baseline value is the number of cycles used to determine the baseline fluorescence.

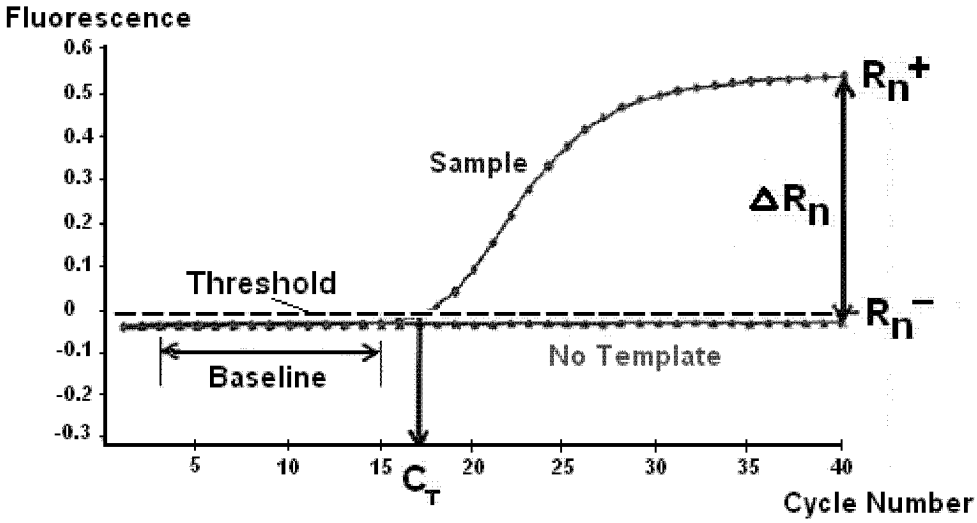

Figure 1.4: Log view of amplicon production during each cycle in a standard real-time PCR.

Rn is the normalized reporter, Rn^+ is the Rn value of a reaction containing all components, Rn^- is the Rn value of an unreacted sample, ΔRn is the magnitude of the signal generated by a given reaction and C_T (cycle threshold) is the cycle, during which a statistically significant increase of ΔRn is first detected (slightly modified from TaqMan® PCR Universal Mastermix manual).

1.2.4 Genetic mapping

The term "mapping a gene" (genetic mapping or linkage analysis) refers to the determination for a particular gene of its location on a chromosome and/or its position relative to molecular markers or to other genes. Many methods have been developed to detect genetic variants (molecular markers), which are essential in genetic mapping and association studies. A high number of markers have been detected and ordered in maps, which constitutes the most useful tool for genetic mapping (section 1.2.4.1). Genetic maps are constructed according to the number of observed recombinations between two or more loci in related individuals. The distances in the genetic maps are measured in centimorgan (named after the pioneer geneticist and winner of the Nobel Prize, Thomas Hunt Morgan). One cM represents a 1% chance that a marker at one genetic locus will be separated from a marker at another locus due to crossing over in a single generation (in humans 1 cM is more or less equivalent to one million base pairs). Hence, without recombinants, the order of the markers in a chromosomal segment can not be determined. In swine, for example, genetic mapping and mapping information were applied for marker-assisted selection (Vögeli *et al.* 1994) and for indirect selection of genotypes based on closely linked microsatellite markers (for the CPA disease in Switzerland, Kratzsch *et al.* 1999) and on closely linked RFLPs (for susceptibility to *E. coli* F4 bacteria, Jørgensen *et al.* in preparation).

1.2.4.1 Marker maps

Genetic maps are used to find genes that cause inherited disorders. They are the basis of whole-genome studies, including QTL analysis, analysis of comparative candidate genes, fine mapping and marker-assisted selection. DNA markers, the major component of genetic maps, are important because, when sufficiently close to a putative disease gene, they are inherited together as haplotypes. These haplotypes can then be observed in pedigrees and used for marker-assisted selection (section 1.2.4.4). The best marker maps are those saturated with polymorphic markers, which are evenly distributed on each chromosome.

The most important markers on current farm animal linkage maps are the microsatellites (section 1.2.4.1.1) followed by the single nucleotide polymorphisms (SNPs, section 1.2.4.1.2). The other relevant markers are restriction fragment length polymorphisms (RFLPs), randomly amplified polymorphic DNA (RAPD), single stranded conformational polymorphisms (SSCP), serum proteins and blood groups (reviewed by Vignal *et al.* 2002).

The use of SNPs as high throughput method for the characterization of genomes considerably augmented and is more and more important. As large amounts of sequence data become available, SNPs are detected more rapidly and are easily typed.

Microsatellites and SNP markers are evenly distributed across the entire genome, regardless of age, gender or environmental factors.

A continuously updated database (http://roslin.thearkdb.org) that includes the genetic maps of several species such as cat, cattle, chicken, cow, deer, horse, pig, salmon, sheep, tilapia and turkey is a good tool in genome mapping projects.

The pig map included 4081 loci (status 1. November 2005). Of these, 1588 were genes and 1673 were microsatellites.

1.2.4.1.1 Microsatellites

Microsatellites are DNA sequences of about 100 to 200 nucleotides consisting of tandemly repeated short nucleotide motifs. Due to their extremely high level of polymorphism, they have been used in comparative studies of genetic variation, parentage assessment, genetic differentiation among closely related populations, genetic linkage maps, analysis of population structure and studies of population variation in geographically isolated or endangered species.

Their use was the method of choice in genotyping farm animals (for a review see Beattie 1994) and they form the backbone of all livestock linkage maps. Thus, microsatellite resources are not the limiting factor for attempting to map trait loci; however, a higher density of marker coverage is required for fine mapping (section 1.2.5.1). The discovery of

SNPs and the construction of physical contigs (section 1.2.5.1.2) help to improve the resolution of maps. A specific database for pig microsatellites (http://www.marc.usda.gov/genome/genome.html) permits the selection of hundreds of microsatellites on the entire pig genome and form the starting point of further analysis, such as full genome scans (section 1.2.4.2).

1.2.4.1.2 Single nucleotide polymorphisms (SNPs)

SNPs are the most common types of genetic variation (over 90% of total polymorphisms) in eukaryotic genomes and in genes causing disease (Körkkö *et al.* 1998). They are allelic variants of true Mendelian traits and, therefore, represent a rich source of genetic markers. According to the definition of Brookes (1999) SNPs are single base pair positions in genomic DNA, where different sequence alternatives (alleles) exist in normal individuals in some population(s); the least frequent allele has an abundance of 1% or higher. SNPs are found at a frequency of about one in 800 to 1000 bp in human genomic DNA (Dawson 1999). They are not equally distributed over the genome because, within coding sequences, SNPs tend to occur less frequently than in non-coding regions. Using P↔Q (P_1↔Q_1), where P and Q are the allelic nucleotides of an SNP on one strand and P_1 and Q_1 their base pair on the second strand, the four SNP alternatives include one transition (C↔T (G↔A)) and three transversions (C↔A (G↔T), C↔G (G↔C) and T↔A (A↔T)). Two thirds of the SNPs involve a transition and the three transversions, which occur at similar frequencies, share the remaining one third (Brookes 1999). At present, few SNP markers are available compared to the number of microsatellites in farm animals.

Although SNPs are a better alternative than microsatellites, because they have higher incidence throughout the genome, can be localized in coding sequences, and can, therefore, be functional, they have a lower mutation rate.

SNP genotyping methods are still under development. First important breakthroughs have been realized. Hence, a panel of 41 SNPs for animal identification and/or determination of parentage in pigs has been described and made available by Rohrer *et al.* (2005).

Moreover, first high throughput projects combining microarray technology with the testing of SNPs have started.

1.2.4.2 Whole genome scan

Scanning the whole genome is the method of choice for determining a chromosomal region carrying a major gene responsible for a monogenic disease.

This method is applied if no obvious or too many candidate genes are identified in other species, leading to no or too many chromosomal regions to screen for the gene.
The method is based on the previous construction of complete marker maps (section 1.2.4.1) with a high number of evenly spaced markers, covering the entire genome and having a known chromosomal position. The ultimate goal of this method is to type a family, that segregates a particular phenotype, with a grided set of such polymorphic markers and to find one or more markers linked to the phenotype. With the entire genome scanned, each marker can be assigned a LOD score, which is a measure of the probability that the marker is linked to a gene affecting that trait. A LOD score of 2 means that the marker is 100 times more likely to be linked to a gene than not. As a rule, a LOD score of >3 is considered to be significant.
In performing a whole genome scan it is critical to overcome the fact that close relatives in highly inbred families usually share large areas of the genome. Hence, the discovery of polymorphic loci may represent a problem, leading to the lack of significant areas.

1.2.4.3 QTL analysis

Microsatellites and SNP markers are important in localizing monogenetic traits such as inherited disease genes. However, a great number of traits in livestock, especially those which are economically important, are multifactorial. Quantitative trait loci (QTL) mapping is used to identify chromosomal regions or genes involved in these traits. The term was first introduced by Geldermann (1975) and defines a polymorphic locus (site), a stretch of DNA on a chromosome containing alleles that differentially affect the expression of a continuously distributed phenotypic trait. The former is indicative of complex inheritance controlled by many genes, each usually having a small effect on the trait, which can have any value within a range. Because many QTLs may be associated with a single trait, they can be found on different chromosomes.
The presence of QTLs is inferred from genetic mapping (section 1.2.4). In recent years, following the development of high-density genetic marker maps (section 1.2.4.1) and the discovery of numerous polymorphic markers, it has become possible to dissect genetically quantitative traits. As for genetic mapping, QTL analysis always starts with finding chromosomal regions, which tend to be shared by affected relatives and to differ between affected and unaffected individuals. Thus, statistical analysis is required to demonstrate that different genes interact with each another and to determine whether they produce a significant effect on the phenotype. Moreover, the discovery of genes underlying QTL is considered to be one of the most difficult tasks in genetic research (Andersson & Georges

2004). The insulin growth factor 2 gene (*IGF2*) in swine (Van Laere *et al.* 2003), which augments the muscle growth and heart size and reduces back-fat thickness, and the bovine acylCoA:diacylglycerol acyltransferase (*DGAT1*) gene (Grisart *et al.* 2002), which influences the milk yield and composition, are successful examples in animal genetics research for the identification of QTN (quantitative trait nucleotide, i.e. the causative polymorphism in the QTL). These studies on *IGF2* and *DGAT1* reinforce the evidence that, today, successful mapping of QTLs seems to be limited to traits where a single gene or mutation explains most of the phenotypic variation and the trait can be practically considered monogenic (Andersson & Georges 2004).

See http://www.animalgenome.org/QTLdb/ for a review of QTL studies in pig and QTL loci already mapped to SSC (to date 791).

1.2.4.4 Marker-assisted selection (MAS)

MAS was developed to combine genomic and phenotypic information in order to meliorate and speed up the selection response of "traditional methods" to meet increasing demands for efficiency and quality in the marketing of livestock. The advances in animal molecular genetics described above led to the identification of multiple genes or genetic markers associated with genes that affect traits of interest in livestock, including genes for qualitative traits and QTL or genomic regions that affect the quantitative traits. The use of genetic markers provides a great opportunity for enhancing the response to selection of a linked characteristic, trait or disease-associated gene, in particular for traits that are difficult to improve by conventional selection. Such traits, for example, have low heritability or a difficult and unclear phenotype, are expensive and only appear late in life. The genotypic information provided by the markers can be achieved at early stages of development, regardless of sex and environment. Furthermore, a huge number of animals can be analyzed, many more than could be phenotyped. The importance of several tests for MAS in pig breeding programs, including QTLs affecting litter size, growth, body composition, feed intake, meat quality and disease resistance, are reported in a critical review (Dekkers 2004).

1.2.5 Physical mapping

A broad definition of physical mapping suggests that it consists of placing nucleotidic sequences (loci) with respect to a DNA matrix. Physical mapping refers to two classical methods, somatic cell hybridization (and its derivate radiation hybrid panel, section 1.2.6) and the fluorescence in situ hybridization (FISH) technique. On physical maps the location of identifiable landmarks on the DNA (e.g., microsatellites, SNPs, restriction-enzyme cutting sites and genes) do not depend on inheritance and the distances are measured in base pairs

(bp), in contrast to genetic maps. For a species, the lowest-resolution physical map of the chromosomes is the banding pattern while the highest is the complete nucleotide sequence. Physical maps are particularly important when searching for disease genes by means of positional cloning strategies and for DNA sequencing. They also lead to the identification of conserved synteny among diverse vertebrate genomes.

1.2.5.1 Fine mapping

Recombinant mapping has shown little success once the critical region has been reduced to one or two Mb by linkage analysis or FISH.

This bottleneck is caused by the improbability that recombinants will be observed in extant family material (Boehnke 1994). Therefore, fine mapping aims at refining the critical location of a disease gene within the bounds provided by linkage analysis or FISH and using techniques with a significantly higher resolution.

Various methodologies and approaches, such as the construction of BAC libraries and contigs, allow higher resolution of the genome with sequencing being the ultimate resolution.

1.2.5.1.1 BAC libraries

Many easy to handle vector systems have been developed and used to construct large-insert chromosomal DNA libraries required for both genomic sequence determination and functional studies.

Bacterial artificial chromosome (BAC) libraries contain the entire genome of an organism segmented in thousands of DNA fragments, up to 200 kb long and referred to as clones.

BACs have emerged as the vector systems of choice because they have the advantage of being more stable and less susceptible to artifacts compared to cosmids and yeast artificial chromosomes (YACs).

Another advantage of these types of vectors is that they can be propagated *in vitro* in appropriate host cells, such as *E. coli*, and yield a large amount of the starting clone material. Furthermore, they are easily accessible to other research groups. While uncharacterized BAC libraries are useful, more extensive characterization of the clones in a library and organization of the library clones into a physical continuous map (contig) significantly widen the range of uses to which the BAC clones can be put.

A coordinated international consortium was created to develop a porcine BAC map; the Wellcome Trust Sanger Institute, the USDA ARS, Roslin Institute, INRA and the University of Illinois are current participants.

1.2.5.1.2 Construction of physical contigs

Correctly assembled and ordered BAC clones form contiguous DNA stretches or blocks (contigs). Thus, contig maps consist of overlapping clones, clones with a common part, representing and covering unbroken chromosomal segments. Contig maps are important for finding genes localized in a small area (<2 Mb) and enable the researcher to examine a complete and often large segment of the genome. They are also a necessary template for fine mapping a region of interest and, in the near future, for sequencing the entire swine genome.

The techniques used to construct contigs are high throughput fingerprinting and BAC end sequencing (BES). Marra *et al.* (1997) first described restriction enzyme fingerprinting in humans to construct physical maps. In brief, fingerprints are generated by digesting clones with *Hin*dIII. Then electrophoresis of the restriction digests on agarose gels is performed and the data collected by a fluorimager. The relative mobility of the fragments is used to identify other clones, which share a large proportion of bands with the same relative mobility. In this way overlap of the clones is inferred and a contig constructed, where the relative positions of the clones reflect the extent to which they overlap. The analysis of the abundant data, and especially the assembly of the clones, takes place using appropriate software such as finger printed contigs (FPC) (Soderlund *et al.* 2000).

On the other hand, in BES the ends of BAC clones are sequenced and used to generate STS in order to identify new overlapping BAC clones basing on shared fragments in the entire library. The order of the clones with shared segments is also done with specific software, of which FPC is the most important. In pig, physical maps of BAC clones representing 15.3x coverage across the 2.7 Gb genome were constructed at the Sanger Institute using both of the above-mentioned methods. Currently, 267,884 fingerprints are assembled into <200 contigs and over 400,000 BES were generated in a coordinated international collaboration of four laboratories (Sanger Institute, Roslin Institute, INRA and University of Illinois) providing four libraries. Current information on the fingerprinting of BAC clones and end-sequencing in pig is available on the website of the Sanger Institute http://www.sanger.ac.uk/Projects/S_scrofa/.

1.2.5.1.3 Sequencing genomic DNA and ESTs

The purpose of sequencing is to determine the order of the nucleotides of a DNA fragment.

DNA sequencing of candidate gene regions is still the method for identifying mutations between different phenotypes and of analyzing expressed sequence tags (ESTs). There are two basic methods of sequencing: the Maxam-Gilbert and the Sanger methods, which differ primarily in the way the nested DNA fragments are produced. Maxam-Gilbert sequencing

(chemical degradation method) relies on chemicals to cleave DNA at specific bases, resulting in fragments of different lengths. Sanger sequencing (chain termination method) is the more commonly used and is fully automated. It involves the use of enzymes to synthesize DNA chains of varying length in four different reactions, stopping DNA replication at positions occupied by one of the four bases and then determining the resulting fragment lengths.

ESTs are single pass nucleotide sequences generated from the ends of (randomly) selected clones from cDNA libraries. Today, they are primarily used for identifying genes in large sequencing projects and for constructing physical gene-based maps. While sequencing genomic DNA is the only definitive way to access all the genes of an organism, this process is both expensive and slow. By accumulating EST sequences, it is possible to identify important genes in livestock by leaping across taxonomic boundaries from genes identified in "model organisms" (e.g. humans, mice and rats).

EST studies of a diverse collection of organisms and tissues are being carried out. As of November 2005, there were over 31 million entries in the dbEST database (http://www.ncbi.nlm.nih.gov/dbEST/), the vast majority from genes of humans, mice and *Xenopus tropicalis*. However, there is also a large number of entries for farm animals and plants.

The current state and the further development of pig ESTs (at present more than half a million entries) can be followed by visiting the site http://pigest.genome.iastate.edu/.

1.2.6 Radiation hybrid mapping

Radiation hybrid mapping has proven to be the key to constructing whole genome maps. This technique has increased the resolution of genetic maps and made the mapping of new marker loci much faster and easier than FISH.

Radiation hybrid panels are an improved derivate of somatic cell panels used previously. In a panel of radiation hybrids, cell lines are used, which retain a different portion of pig chromosomal material (donor genome) on the background of a rodent genome (recipient). After cell fusion (hybridization), the chromosomes of the donor are preferentially lost from the hybrid cells until only a few fragments remain. Those individual hybrid cells are propagated and maintained as cell lines and contain specific pig chromosomes and chromosome fragments. They are scored for the presence or absence of specific PCR fragments of interest (e.g. microsatellites, STSs or ESTs).

The frequency with which these fragments co-segregate corresponds to their proximity to one another on the genome. Statistical analysis of the similarity of the retention patterns (section 2.2.5.4) allows definition of marker order and intermarker distances. The most important

radiation hybrid panel available for pig, developed in a joint effort of the INRA and the University of Minnesota (http://www.toulouse.inra.fr/lgc/pig/RH/IMpRH.htm) was used in this work (section 2.2.3.8).

1.2.7 Candidate genes for AMC

After excluding HSA5 (section 3.2.1) interesting positional and physiological human candidate genes for AMC were chosen based on their chromosomal location on human chromosomes 12 or 22, expected to be the counterpart of SSC5, where pig AMC was mapped. An overview follows and includes a description of the genes known to affect developmental processes or to have mutant phenotypes with similar abnormalities as pig AMC.

1.2.7.1 Voltage-gated ion channels

Piglets born to dams that were fed toxic plants during pregnancy, such as *Nicotiana tabacum* (burley tobacco, Crowe 1978), C*onium maculatum* (poison hemlock, Dyson & Wrathall 1977; Lopez *et al.* 1999) or *Datura stramonium* (jimsonweed, Leipold *et al.* 1973), exhibited the same symptoms as the diseased piglets in this study. Thus, candidate genes were hypothesized as being correlated with the functionality of some ion channels or receptors present in brain, muscle or spinal cord. The mechanisms underlying the toxicity of these compounds have not been fully delineated. It appears that the toxic compounds, piperidine alkaloids in most cases, such as coniine or coniceine (Lopez *et al.* 1999), block ion channels either directly or through their actions on a receptor. This in turn inhibits fetal movement and leads to the typical symptoms (Panter *et al.* 1988).

Maternal autoantibodies, specific for fetal antigens, were also thought to be another cause of AMC; such antibodies could inhibit ion channel function in the fetal acetylcholine receptor (AChR), as reported for swine by Vincent *et al.* (1995) and for humans by Matthews *et al.* (2002). Furthermore, as proposed by Lopez *et al.* (1999), toxic compounds may affect the regulation of amniotic fluid levels, thereby leading to malformation.

Many recent studies indicate that voltage-gated ion channels frequently contribute to inherited disorders of movement and to seizure. Important reviews on the recent findings of disease caused by voltage-gated channel mutations were published (Ashcroft 2000; Lehmann-Horn & Jurkat-Rott 1999). Dozens of such disorders with striking clinical similarities have been described and led to the coining of the terms “ion channel diseases” or “ion channelopathies”, a group of hereditary disorders associated with ion channel mutations in both humans and animals. Voltage-gated ion channels are both highly selective and remarkably efficient and are responsible for the generation of conducted electrical signals in

excitable cell membranes. They also are the basis of neuronal communication. According to Celesia (2000), some of the important directly gated channels are voltage-gated for cations (sodium, calcium and potassium) and anions (chloride) and are ligand-gated (ACh, glutamate, GABA and glycine).

A review of the function and the channellopathies of the voltage-gated genes *CACNA1C*, *KCNA1* and *SCN8A* is reported here.

1.2.7.1.1 *CACNA1C*

The *CACNA1C* gene mediates the influx of calcium ions into excitable cells upon membrane polarization. It is involved in a wide range of calcium-dependent processes, including muscle contraction, hormone and/or neurotransmitter release, gene expression, cell motility, cell division and cell death. Furthermore, calcium channels with the alpha-1C subunit are crucial for excitation-contraction coupling in the heart.

The most important disease caused by a mutation in the *CACNA1C* gene is the Timothy syndrome, characterized by multiorgan malfunction, including lethal arrhythmias, webbing of fingers and toes, congenital heart disease, immune deficiency, intermittent hypoglycemia, cognitive abnormalities and autism (Splawski *et al.* 2004).

1.2.7.1.2 *KCNA1*

The *KCNA1* gene mediates the voltage-dependent potassium ion permeability of excitable membranes. These channels are important for a variety of cell functions, including modulation of action potentials, determination of resting membrane potential and development of memory and learning. In addition to their role in regulating myocyte excitability, cardiac K+ channels control heart rate and coronary vascular tone and are implicated in the development of arrhythmias (Curran *et al.* 1992). Defects in *KCNA1* are the cause of the autosomal dominant disorder episodic ataxia with myokymia (EAM) described by Browne *et al.* (1994).

1.2.7.1.3 *SCN8A*

The *SCN8A* gene, which is abundantly expressed in neurons of the central and peripheral nervous systems, encodes a voltage-gated neuronal sodium channel localized in axonic, dendritic, presynaptic and postsynaptic membranes (Plummer *et al.* 1998). In mouse, a defective *SCN8A* gene causes neurological disorders including muscular atrophy, ataxia, tremor, progressive paralysis and dystonia (Kearney *et al.* 2002).

1.2.7.2 *CNTN1*

CNTN1 is not a direct voltage-gated channel gene but, as reported by Kazarinova-Noyes *et al.* (2001), it influences the functional expression and distribution of Na^+ channels in neurons. This gene is a glycosylphosphatidylinositol (GPI)-anchored neuronal cell membrane adhesion molecule of the immunoglobulin superfamily. It is of critical importance in the formation and growth of axon connections and mediates cell surface interactions in the nervous system of developing embryos during the development of the nervous system. Furthermore, it participates in oligodendrocyte generation by acting as a ligand and promoting *NOTCH1*, which was shown to control oligodendrocyte differentiation in the spinal cord (Genoud *et al.* 2002). The results of Berglund *et al.* (1999) also demonstrated that contactin controls axonal and dendritic interactions of cerebellar interneurons and contributes to cerebellar microorganization. The contactin -/- mice mutants in their study displayed a severe ataxic phenotype consistent with defects in the cerebellum and survived only until postnatal day 18. Poor control of voluntary movements, posture and balance was accompanied by failure to gain weight and progressive weakening.

1.2.7.3 *COL2A1*

The gene *COL2A1* specifically encodes for cartilage collagen, contrarily to *COL1A2*, which is found in skin, tendon and bone. It is essential for the normal embryonic development of the skeleton, for linear growth and for the ability of the cartilage to resist compressive forces.
Defects in this gene are well described and are very frequent (>10 records). The most important syndromes are osteoarthritis, achondrogenesis-hypochondrogenesis type II (*ACG2*), a disease characterized by the absence of ossification in the vertebral column, sacrum and pubic bones and multiple epiphyseal dysplasia (*EDM*), a generalized skeletal dysplasia associated with significant morbidity and joint pain, joint deformity and short stature. Gaiser *et al.* (2002) and Garofalo *et al.* (1991) also found that transgenic mice carrying a partially deleted human *COL2A1* displayed severe chondrodysplasia with short limbs and trunk, craniofacial deformities, dwarfism, scoliosis, cleft palate and delayed mineralization of the bone. The affected pups died at birth of acute respiratory distress caused by the inability to inflate the lungs.

1.2.7.4 *GDF11*

The *GDF11* gene plays an important role in mesodermal and neural tissues in establishing the pattern of the axial skeleton and in regulating skeletal muscle mass (McPherron *et al.* 1999). It is also referred to as *BMP11* and acts mainly during early embryogenesis in the developing nervous system, in the dorsal root ganglia and in dorsal lateral regions of the spinal cord.

McPherron *et al.* (1999) gave evidence of a malfunction of *GDF11*, which leads to mutant mice with posterior displacement of the hind limbs.

1.2.7.5 *HOXC8*

The *HOXC8* gene is a specific transcription factor and is part of a developmental regulatory system that provides cells with specific positional identities on the anterior-posterior axis. Similar to *GDF11*, it is involved in anterior and posterior patterning and innervation of the limbs. Yueh *et al.* (1998) demonstrated that overexpression of *HOXC8* in transgenic mice causes defective cartilage, whose severity depends on transgene dosage, and suggested that *HOXC8* regulates skeletal development.

This homeobox gene family has been suggested (Jones 1999) as candidate gene for perosomus elumbis, a set of congenital abnormalities that includes deformation of the lumbosacral spinal cord and vertebrae and arthrogryposis of the hind limbs.

1.2.7.6 *TUBA8*

The gene *TUBA8* encodes for tubulin, the main constituent of microtubules that have many distinct functions and are involved in mitosis, cell movement, intracellular movement and other biologic processes (Stanchi *et al.* 2000). To the best of my knowledge, no human diseases are correlated with mutations in this gene.

1.2.7.7 *USP18*

The principal function of the nuclear protein encoded by the gene *USP18*, also referred to as *UBP43*, is to cleave ubiquitin adducts from a broad range of protein substrates. It is necessary to maintain a critical cellular balance in both healthy and stressed organisms. *USP18* maps to HSA22q11.2. This region (an area of about 2 Mb) is known as the "DiGeorge syndrome critical region" and is consistently deleted in this disease and related disorders.

The human syndrome is marked by thymic aplasia or hypoplasia, parathyroid hypoplasia and congenital cardiac abnormalities (Schwer *et al.* 2000). In mice, the symptoms are very similar, homozygous null mutants die prematurely and show cellular necrosis, breakdown of the blood-brain barrier, hydrocephaly with enlarged ventricles and severe neurological abnormalities including convulsions, tremor, ataxia, rolling and circling.

1.3 Hereditary diseases mapped in swine

Using the various approaches and mapping methods described in section 1.2, several pig genes were successfully mapped to SSC (http://omia.angis.org.au/). Of the total 214 known pig disorders, 35 implicated a single locus and only in 11 genes had the causative mutation been discovered. Table 1.2 gives a detailed list of selected traits with gene, chromosomal location, description of the causative mutation and reference.

Table 1.2: Examples of monogenetic traits mapped in swine.

Trait	Gene	SSC	Mutation	Reference
Meat quality (acidic meat, Rendement Napole (RN) mutation)	*PRKAG3*	15	A→G transition	Milan *et al.* (2000 A)
Oedema disease, post weaning diarrhea	*FUT1*	6	G→A transition	Meijerink *et al.* (2000)
Vitamin C deficiency	*GULO*	14	Exon 8 deletion	Hasan *et al.* (2000)
Malignant hyperthermia	*RYR1*	6	C→T transition	Vögeli *et al.* (1994)
Coat color, agouti	*AGRP*	6	Intronic substitution	Kim & Rothschild (2001)
Coat color, dominant white	*KIT*	8	Duplication G→A transition in intron 17	Moller *et al.* (1996) Marklund *et al.* (1998)
Coat colour, extension	*MC1R*	6	Missense mutations	Kijas *et al.* (1998)
Hypercholesterolaemia	*LDLR*	2	Amino acid substitution	Hasler-Rapacz *et al.* (1998) Grunwald *et al.* (1999)
Blood group system A0	*A*	1	Deletion in the 0 gene	Yamamoto & Yamamoto (2001)
Membranoproliferative glomerulonephritis type II	*Factor H*	?	C→G (position 1590) and T→G (position 3610)	Hegasy *et al.* (2002)
Pseudo-vitamin D deficiency rickets (*PDDR*)	*P450C1*	?	Deletion	Chavez *et al.* (2003)

1.4 Objectives of the study

The final goal of this study was to find the causative mutation for pig AMC in the Swiss LW population.

In the first part of the study the principal aims were:

1) to confirm the previously described autosomal recessive inheritance of *AMC* in Yorkshire pigs.
2) to map the AMC phenotype to the porcine genome by assessing linkage of the gene encoding *AMC* and microsatellites evenly distributed over the genome.

To accomplish those goals a whole genome scan followed by a SSC5 scan and a QTL analysis with microsatellite markers were performed.

In the second part of the project, techniques were applied with the following objectives:

3) to characterize the human chromosomes corresponding to the pig *AMC* region (flanked by microsatellites *SW152* and *SW904*) by mapping, with the INRA-Minnesota swine radiation hybrid panel, pig ESTs and functional candidate genes.
4) to find additional functional candidate genes for *AMC* with the microarray technology.
5) to identify new markers (SNPs and microsatellites) to further delineate the *AMC* region and to develop a powerful diagnostic genetic test, applicable to commercial herds, to distinguish healthy pigs from carriers.
6) to scan porcine candidate genes for sequence differences between *AMC/AMC* (normal) and *amc/amc* (diseased pigs).

As well as these very specific purposes, the aim of this project was to contribute to the analysis of the porcine genome by improving the resolution of its genomic map and by elucidating new chromosomal rearrangements and new evolutionary breakpoints of the human-pig comparative map.

2. Materials and Methods

2.1 Materials

2.1.1 Animals

All the pigs used in this study are Swiss Large White.

2.1.1.1 Pigs at the experimental farm

AMC was first observed in two litters (Table 2.1, matings 1 and 2). These litters were derived from two sisters (359 HNM and 360 HNM) mated with the same sire 1077 GT (F_0 animals). Six offspring of these two matings (one sire, 68 HNM, and five dams, 71 HNM, 72 HNM, 75 HNM, 78 HNM and 82 HNM, F_1 founder generation) were born. These pigs were kept for breeding at the experimental farm of the Swiss Federal Institute of Technology Zurich (ETHZ) in Chamau where they formed the basis of the highly inbred AMC herd. The sire 68 HNM was mated with ten dams, three of which were his sisters (Table 2.1, matings 3, 4 and 5), two his half-sisters (matings 6 and 7) and five his daughters (matings 8, 9, 10, 11 and 12), born of previous matings.

After the death of sire 68 HNM, another sire (1481 CH), also a descendant of the founder boar 2401 JN and already kept at the research station, was found to be carrier of the disease. 1481 CH was mated with six dams, one sister (72 HNM, mating 14) and five daughters of sire 68 HNM (matings 15, 16, 17, 18 and 19).

A total of 358 piglets were analyzed (until 20.11.2005). The 84 diseased piglets (23.5%) were easily identified because they exhibited all the typical symptoms described in the introduction and shown in Figure 1.1. Therefore, they had the genotype *amc/amc*. It was phenotypically impossible to distinguish between healthy piglets (*AMC/AMC*) and carriers (*AMC/amc*) in the remaining 274 normal piglets (76.5%).

Table 2.1: The AMC-resource family.
Mating pairs, identification number, *AMC* genotype and number of affected and normal offspring with their generation (*DNA not analysed) (Status 20.11.2005).

			Offspring number			
Mating	Father	Mother	Affected (*amc/amc*)	Normal (*AMC/AMC, AMC/amc*)	Total	Generation
1	1077 GT (*AMC/amc*)	360 HNM (*AMC/amc*)	2*	4	6	F_1
2	1077 GT (*AMC/amc*)	359 HNM (*AMC/amc*)	1*	3	4	F_1
3	68 HNM (*AMC/amc*)	71 HNM (*AMC/amc*)	3	18	21	F_2
4	68 HNM (*AMC/amc*)	72 HNM (*AMC/amc*)	14	54	68	F_2
5	68 HNM (*AMC/amc*)	75 HNM (*AMC/amc*)	9	17	26	F_2
6	68 HNM (*AMC/amc*)	78 HNM (*AMC/amc*)	7	39	46	F_2
7	68 HNM (*AMC/amc*)	82 HNM (*AMC/amc*)	2	6	8	F_2
8	68 HNM (*AMC/amc*)	1154 CH (*AMC/AMC*)	0	9	9	F_3
9	68 HNM (*AMC/amc*)	1731 CH (*AMC/amc*)	4	15	19	F_3
10	68 HNM (*AMC/amc*)	1733 CH (*AMC/AMC*)	0	11	11	F_3
11	68 HNM (*AMC/amc*)	1734 CH (*AMC/amc*)	1	8	9	F_3
12	68 HNM (*AMC/amc*)	2215 CH (*AMC/amc*)	3	16	19	F_3
13	68 HNM (*AMC/amc*)	2884 CH (*AMC/amc*)	5	3	8	F_4
14	1481 CH (*AMC/amc*)	72 HNM (*AMC/amc*)	7	6	13	F_4
15	1481 CH (*AMC/amc*)	2215 CH (*AMC/amc*)	8	22	30	F_4
16	1481 CH (*AMC/amc*)	2884 CH (*AMC/amc*)	5	11	16	F_4
17	1481 CH (*AMC/amc*)	3290 CH (*AMC/amc*)	8	19	27	F_4
18	1481 CH (*AMC/amc*)	3647 CH (*AMC/amc*)	0	8	8	F_4
19	1481 CH (*AMC/amc*)	3669 CH (*AMC/amc*)	5	5	10	F_4
TOTAL			**84**	**274**	**358**	

2.1.1.2 Pigs from commercial farms

Forty-one piglets from 14 commercial farms and suspected of having AMC (Table 2.2) were tested with the markers *bE77* and *SW904*, flanking the *AMC* region. The parent's identification numbers and the number of affected piglets per litter are given.

Table 2.2: Piglets from commercial herds suspected of having AMC.
(*DNA not available).

Mating	Father	Mother	Number of piglets suspected of having AMC (*amc/amc*)
1	4218 ELE	1595 ICT*	1
2	4591 KE	2184 KE*	6
3	4631 PU	1972 VFT*	3
4	5587 HZO*	488 FYG*	2
5	5527 WO*	245 NIE*	5
6	3133 X1*	8975 BKE*	3
7	9695 KPN	2477 CHR*	1
8	3367 HT	1724 HT*	5
9	9695 KPN	3390 IAG*	2
10	7667 EBA	1530 JBS*	4
11	965 RBA	6408 FJ1	1
12	3099 DUC	868 MS1*	2
13	8389 PU	292 MS1*	4
14	9975 HZP	8497 CHR*	2
	TOTAL		**41**

Table 2.3 shows the identification numbers of 103 boars tested for susceptibility to AMC with *bE77* and *SW904*.

Boars 1 to 80 are presently used in Switzerland for artificial insemination and 81 to 103 are strongly suspected of being carriers of *AMC*.

Table 2.3: AI boars tested with the AMC flanking markers bE77 and SW904.

Boar number	Boar ID	Boar number	Boar ID	Boar number	Boar ID
1	2 MW1	36	571 UZ1	71	4351 XE1
2	636 TK	37	630 KV	72	8119 HZP
3	705 CRM	38	838 UZ1	73	8389 PU
4	986 TK	39	1060 W2	74	8493 EBA
5	1218 WOZ	40	1087 WOZ	75	8919 HZP
6	1475 HZR	41	1148 WOZ	76	9589 KE
7	1558 HZR	42	1218 UZ1	77	9591 KE
8	1787 WOZ	43	1219 UZ1	78	9841 W1
9	1896 RBA	44	1331 HB1	79	9859 KE
10	4840 ICT	45	1332 HB1	80	9975 HZP
11	5554 RJ	46	1402 WOZ	81	8216 RZ
12	6450 HSB	47	1418 HZR	82	9542 TW
13	7222 HSB	48	1522 HZR	83	7364 KPN
14	7290 HSB	49	1767 JRB	84	1371 LP
15	7796 EBA	50	1775 JRB	85	1491 LP
16	7899 EBA	51	1867 WOZ	86	6714 RZ
17	7920 HP	52	1869 WOZ	87	4218 ELE
18	7924 EBA	53	1870 WOZ	88	2399 GT
19	7990 HP	54	2605 JRB	89	4122 KE
20	8508 ZS	55	2607 JRB	90	1997 XE
21	9467 HZP	56	2960 XE1	91	7703 EBZ
22	9477 MWM	57	3091 XE1	92	287 XE
23	9857 MWM	58	3156 KPU	93	9200 NFE
24	9882 MWM	59	3176 KPU	94	8612 RZ
25	9889 RZ	60	3200 XE1	95	4591 KE
26	257 HZR	61	3329 XE1	96	4631 PU
27	311 UZ1	62	3331 XE1	97	3367 HT
28	424 UZ1	63	3335 KPU	98	9231 KPN
29	444 UZ1	64	3505 KPU	99	9695 KPN
30	471 WOZ	65	3548 KPU	100	7667 EBA
31	501 KV	66	3594 KPU	101	965 RBA
32	503 KV	67	3680 XE1	102	3099 DUC
33	533 UZ1	68	3719 XE1	103	5232 KE
34	545 UZ1	69	3821 XE1		
35	555 ELW	70	3822 XE1		

2.2 Methods

2.2.1 Isolation and extraction of nucleic acids

2.2.1.1 Genomic DNA

Genomic DNA was extracted from EDTA-anticoagulated whole blood of piglets, approximately two months of age, according to the method of Vögeli *et al.* (1994). In brief, 0.6 ml frozen whole blood was mixed with 0.5 ml lysis buffer (A). After 15 min, the solution was centrifuged for 25 s at 13000 x g and the supernatant discarded. The pellet was resuspended with 0.5 ml lysis buffer (A); this step was repeated three more times with 1 ml lysis buffer (A). The pellet was then resuspended in 0.2 ml turbo PCR buffer and 20 µl proteinase K (20 mg/ml). After incubation at 54°C for 2 h and at 95°C for 10 min the DNA was ready to use. According to Bolt (1992), the DNA concentration obtained with this quick method was about 40 ng/µl. Genomic DNA was obtained from the tail of diseased piglets and from piglets that died of normal causes, according to the protocol of Laird *et al.* (1991). In short, tissue samples were transferred to 500 µl lysis buffer (B) containing 2.5 µl proteinase K (20 mg/ml) and incubated overnight on a rotating shaker at 56°C. After complete lysis, the tubes were spun out for 10 min (14000 x g) and the supernatant transferred to labeled tubes containing 0.5 ml isopropanol. The precipitate was extracted from the solution with a tip and washed with 70% EtOH. The resulting DNA was resuspended in 20 to 500 µl TE buffer, according to the pellet size, and stored at -20°C. The DNA concentration was also about 40 ng/µl.

2.2.1.2 Tissue RNA

Total tissue RNA was isolated at birth from the cerebellum, the musculus longissimus dorsi and the upper part of the spinal cord (where the AMC piglets showed scoliosis) using the standard protocol of animal tissue extraction of the RNeasy Maxikit (Qiagen, Basel, CH). One to two g of tissue were ground in liquid nitrogen with a mortar and pestle, transferred to a 50 ml-falcon tube and homogenized in 15 ml buffer RTL with 0.143 M ®-mercaptoethanol for 60 s. Muscle RNA was additionally incubated in 15 ml ddH_2O and 150 µl proteinase K at 55°C for 20 min. The homogenized lysate was centrifuged at room temperature (3000 x g) for 10 min and the supernatant transferred to a 50-ml falcon tube. One volume of 70% ethanol was added to the tube, mixed by shaking the tube and the entire solution was transferred to an RNeasy maxi spin column. The column was centrifuged for 5 min (3000 x g). The flow-through was discarded; 15 ml buffer RW1 was added to the column and centrifuged again for 5 min. After discarding the flow-through, 10 ml of buffer RPE were added to the column.

After 2 min centrifugation, the previous step was repeated and the column centrifuged for 10 min. The RNeasy column was transferred to a new 50-ml falcon tube and 0.8 ml of RNase-free water added to the column filter. The column was subsequently centrifuged for 3 min and the RNA, dispersed in the flow-through, was stored at -80°C.

2.2.2 Quantification of nucleic acids

2.2.2.1 Spectrophotometer

The concentration and purity of the genomic DNA and the tissue RNA was measured with a spectrophotometer (Lambda Bio UV/VIS, Perkin Elmer, Rotkreuz, CH) at wavelengths of 260 and 280 nm according to the following formula

$$c\ (\mu g/\mu l) = (OD_{260} \times df \times ec)/1000$$

where $\mathbf{OD_{260}}$ is the measured optical density at 260 nm, **df** the dilution factor and **ec** the extinction coefficient. For double-stranded DNA, the absorbance of 1 OD at 260 nm (known as ec) is 50 µg/ml, for RNA 40 µg/ml.

The purity was determined from the ratio A_{260}/A_{280}; a ratio of 1.8 is considered to be pure DNA, while a ratio of 2 is pure RNA.

2.2.2.2 Ethidium bromide plates

The concentration of DNA for sequencing was estimated on EtBr plates. DNA (0.5 µl) was compared to the same amount of λ-phage DNA standards with dilutions of 5, 10, 25, 50, 75 and 100 ng/µl.

Samples and standards were dried on the plates for 15 min at room temperature, visualized in an UV light box (365 nm) and photographed with Photo Finish® ZSoft 3.0 (WordStar Atlanta Technology Center Inc., Atlanta, USA).

2.2.3 DNA techniques

2.2.3.1 Primer design

Primers should generally be 20 to 30 nucleotides in length (Newton *et al.* 1994), have a GC content of 55 to 60% and a similar melting point. They are designed to avoid self- and forward/reverse complementarities and to prevent the development of hairpin loops (Rozen & Skaletsky 2000).

The Primer3 website (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) is a helpful software to facilitate primer design and was used in the present study to develop the primers for PCR, sequencing and radiation hybrid mapping. They were designed from published

porcine EST gene sequences or according to published human genomic sequences for cross-species PCR. The complete list of primer sequences used in this study is given in Table 3.1.

2.2.3.2 Microsatellite markers

2.2.3.2.1 SSC2 screening

Five microsatellites (*SWR2516*, *SW1650*, *SW1026*, *S0370* and *SWR345*) on SSC2 were evaluated first. SSC2 is the corresponding pig chromosome of HSA5, where the human forms of AMC (*AMCN* and *SMA*) most similar to pig AMC are located (section 1.2.2).

2.2.3.2.2 Full genome screening

Forty-eight evenly spaced fluorescently labeled (TET, FAM or HEX) microsatellite markers were used to perform a full genome scan of 130 animals (from matings 1-7, Table 2.1). Depending on the length of the chromosomes, two to five informative microsatellites for each chromosome, distributed over the entire swine genome at intervals of about 40 cM were selected (Rohrer *et al.* 1996; http://www.marc.usda.gov/genome/genome.html).

2.2.3.2.3 SSC5 screening

After the initial full genome linkage analysis with five microsatellite markers spread over the entire SSC5 (*ACR*, *SW1482*, *SW2*, *SW1987* and *SW995*), 12 additional polymorphic markers (*SW2425*, *SW1071*, *S0005*, *TNFR*, *SW963*, *SW152*, *SW904*, *SW1094*, *SWR1526*, *SWR1974*, *SW310* and *SW1200*) and 89 additional pigs (from matings 4 to 11, Table 2.1) were added for further genotyping. This enabled a multipoint map, which covers the *AMC* region, to be generated. The details of each microsatellite are available at http://www.marc.usda.gov/genome/genome.html.

Furthermore, a total of 230 pigs (Table 2.1) were tested with the microsatellite *UMNp1275* (Fahrenkrug *et al.* 2005) and all the experimental pigs with *bE77*.

Each animal was manually haplotyped and each genotype was checked for plausible parentage and double recombinants. In the case of disputable or unclear genotypes, animals were re-evaluated and, if necessary, re-genotyped. This procedure permitted the precise location of *AMC* with the aid of recombinants (section 3.7).

2.2.3.3 Polymerase chain reaction (PCR)

PCR amplification was carried out in a total volume of 25 µl containing 100 ng of porcine genomic DNA, 1x PCR buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl and 1.5 mM $MgCl_2$), 200 µM of each deoxynucleotide, 0.4 µM of each primer (if microsatellites fluorescently labeled forward and normal reverse) and 1.25 u of Taq DNA-Polymerase (Sigma-Aldrich Chemie, Buchs, CH). The PCR profile of a PCR Express Machine (Hybaid, Münchenstein,

CH) included an initial 5 min denaturation at 95°C followed by 35 cycles of 95°C for 30 s, 54 to 68°C annealing temperature for 30 s and 72°C for 30 s. The final elongation was at 72°C for 7 min. Table 3.1 gives an overview with the product sizes and the annealing temperatures.

2.2.3.4 Identification and control of PCR products

The quality of the PCR amplification was controlled on agarose gels stained with ethidium bromide to permit the visualization of the DNA bands. The length, purity and approximate quantity of the PCR fragments were analysed by separation on a 1.5% (*w/v*) agarose gel in 0.5x TBE. The solutions were heated in a microwave oven to complete melting of the agarose and then cooled to 55 to 60°C. After polymerization of the gel (30 min) and the addition of EtBr (final concentration 0.1 µg/ml), 14 µl PCR product and 2.6 µl triple or double loading dye were mixed and loaded into each slot adjacent to a size marker (12 µl H_2O, 1 µl triple loading dye and 1 µl ladder).

The power was constant at 100 V (Desagatronic 3x 500/1000, Desaga, Heidelberg, D) for about 1 h. DNA was viewed under UV light and photographed with Photo Finish® ZSoft 3.0.

2.2.3.5 Genescan analysis

With the genescan analysis, the length of the microsatellites was precisely determined. Electrophoresis was carried out in 50% denaturing urea-4.5% polyacrylamide/bis-acrylamide (29:1) gels according to the manufacturer's instructions.

The gel polymerization (1 h) was initiated by adding 350 µl APS (10% stock) and 15 µl TEMED. The run was performed in 1x TBE buffer for 4 h.

After adding 0.5 µl genescan-350 TAMRA size standard to 0.4 PCR product mixed with 2.5 µl formamide and 0.5 µl commercial loading buffer (50 mg/ml Blue dextran and 25 mM EDTA pH 8), the samples (denaturated at 95°C for 5 min) were analyzed on a 377 ABI Prism sequencer (Applied Biosystems, Rotkreuz, CH). The results were evaluated with the ABI 672 Genescan program and genotype software (version 2.1, Applied Biosystems).

2.2.3.6 Identification of BAC clones and extraction of BAC DNA

The experiments described in this section were performed in collaboration with the group of Genetics and Genomics, Roslin Institute, Edinburgh, UK, led by Prof. Dr. A. Archibald and his co-worker Dr. S. Anderson.

The Roslin PigE BAC library (Anderson *et al.* 2000) was screened. It provides about a five-fold coverage of the pig genome and consists of approximately 100,000 independent BAC clones with an average size of 150,000 bp, stored as individual clones in 96-well plates and 384-well microplates.

Positive BAC clones were identified using PCR primers from markers *SW1987*, *SW152*, *S0018*, *SW904* and *SW1094* surrounding the *AMC* region. *E. coli* bacteria carrying the BACs containing the microsatellites of interest were struck out on 10 cm 1.5% LB-agar medium petri dishes containing 25 µg/ml chloramphenicol, 50 µg/ml X-Gal and 50 µg/ml IPTG and grown overnight at 37°C. A single colony was picked at random from the plate and grown at 37°C in 3 ml LB medium with 10 µg/ml chloramphenicol under constant shaking (300 RPM) for 8 h. After that, the starter culture was diluted 1/500 into chloramphenicol selective LB medium and grown at 37°C under constant shaking (300 RPM). After 14 h, the culture reached a cell density of approximately 3 x 10^9 cells per ml, the amount necessary to extract enough DNA for direct sequencing. The DNA was prepared using the Nucleobond AX100 kit (Macherey-Nagel, Düren, D) according to the manufacturer's recommendations. The cell culture was centrifuged for 15 min (3500 RPM) at 4°C and, after discarding the supernatant, the BAC was treated with the buffers S1, S2 and S3. SDS was eliminated using filters instead of centrifugation and the subsequent elution step was performed with buffer N5, preheated to 65°C. After isopropanol precipitation and washing with ethanol, the DNA pellet was resuspended in 500 µl TE; a yield of about 20 µg DNA was obtained. The isolated DNA was checked for the presence of the selected microsatellite and sequenced as described in section 2.2.3.7.2.

2.2.3.7 Sequencing

2.2.3.7.1 DNA

Partial gene sequences of the genes *ABCD2*, *CACNA1C*, *COL2A1*, *CNTN1*, *CPNE8*, *C3F*, *C12ORF4*, *DDX11*, *FOXM1*, *GDF11*, *HOXC8*, *KCNA1*, *KIF21A*, *MDS028*, *MGC5576*, *NR4A1*, *PHB2*, *PRICKLE1*, *Q6ZUQ4*, *SCN8A*, *SLC2A13*, *TUBA8*, *USP18* and *YAF2* were amplified and sequenced with the primers listed in Table 3.1. The sequences of one diseased and one genetically normal animal were compared. Furthermore, the partial gene sequences of the genes *TUBA8* and *CNTN1* of 230 pigs belonging to the AMC research herd (Table 2.1) were analyzed as well.

An overview of the degree of sequence identity between the entire pig sequences, reported with their accession numbers and the putative human homologues, the primer sequences, the product sizes and the annealing temperatures is given in Table 3.1.

DNA was sequenced on an ABI377 sequencer, based on the method of Sanger *et al.* (1977). PCR fragments were purified with spin columns (Millipore, Volketswil, CH) in order to remove unbound primers and PCR buffer. After purification, sequencing of PCR was performed with an amount of template DNA (in ng) that corresponded to 30% of the length

(in bp) of the fragment being sequenced. DNA quantification was done with EtBr plates as described in section 2.2.2.2. The calculated amount of DNA was combined with 1.6 pmole primer, 2 µl Big Dye™ terminator v3.0 sequencing ready reaction, 2 µl Half Big Dye and ddH_2O to give a total volume of 10 µl. The profile consisted of 5 min at 95°C followed by 35 cycles of 30 s at 95°C, 15 s at 50°C and 4 min at 60°C.

Before loading the samples onto the gel, the DNA was precipitated by adding 10 µl 3 M sodium acetate (pH 5.2), 80 µl ddH_2O and 200 µl absolute ethanol followed by vortexing. The samples were stored at –80°C for 30 min and centrifuged at 21000 x g for 30 min at 4°C. The pellets were washed twice by adding 200 µl 70% ethanol and were centrifuged for 5 min at 21000 x g in a cold room (4°C).

The samples were resuspended in 1.5 µl of formamide/loading buffer (*v/v*) 5:1 and denatured at 95°C. Electrophoresis (12 h) was carried out in 4.5% polyacrylamide gels as illustrated for gene scan (section 2.2.3.5).

The raw data were collected using the ABI Prism™ 377 sequence analysis software and the sequences were then determined and analysed with the program Chromas, version 2.0 (available at http://www.technelysium.com.au/chromas_lite.html) and the GCG software package version 10.1 (Genetics Computer Group, Madison, USA). The identity of the porcine PCR products was verified by sequence analysis and BLAST searches of the GenBank non-redundant (nr) sequence database (http://www.ncbi.nlm.nih.gov/BLAST/Blast.cgi). Newly discovered pig sequences (Appendix, Table A-1) were submitted to the GenBank database and made available to the public.

2.2.3.7.2 BAC clones

BAC clones were directly sequenced as described for DNA in section 2.2.3.7.1, with a few differences. The PCR reaction for sequencing comprised 2.5 ng BAC DNA, 4 µl Big Dye™ terminator v3.0 sequencing ready reaction, 4 µl Half Big Dye and 30 pmol primer (T7, Sp6, *SW152* forward, *SW152* reverse, *SW904* forward or *SW904* reverse). The reaction was cycled 99 times instead of 35. The generated sequences were also BLASTed against the non-redundant nucleotide database at NCBI website.

2.2.3.8 Radiation hybrid mapping

The INRA-University of Minnesota whole genome porcine radiation hybrid (RH) panel IMpRH (Yerle *et al.* 1998) was used to map the pig-specific gene sequences. The 118 hybrid line clones of the RH panel were subjected to PCR amplification as described in section

2.2.3.3. Table 3.1 gives the annealing temperatures for PCR. Furthermore, pig DNA was the positive control and hamster genomic DNA and no DNA were the negative controls. The PCR products were examined by electrophoresis on Etbr-stained 1.5% agarose gels (section 2.2.3.4) and tested for the presence of the porcine sequences.

2.2.3.8.1 Enzymatic digestion of *SCN8A* and *KCNA1*

In order to distinguish the pig from the hamster DNA bands, an enzymatic digestion of the 118 probes of the genes *KCNA1* and *SCN8A* was necessary. The PCR product (20 µl) of the *KCNA1* probe was mixed with 2 µl H_20, 2.5 µl NEBuffer 4 and digested with 0.5 µl of the restriction enzymes *Fok*I (New England Biolabs, Allschwil, CH) for 3 h at 37°C. The PCR product (20 µl) of the *SCN8A* probe was mixed with 1 µl water, 1 µl BSA, 2.5 µl NEBuffer 1 and 0.5 µl enzyme *Tsp*45I (New England Biolabs) for 3 h at 65°C.

2.2.4 RNA techniques

2.2.4.1 Gel electrophoresis

The quality and purity of the RNA was examined by gel electrophoresis. A total of 2.5 µg RNA was mixed with 1.5 µl RNA loading buffer (1x MOPS, 16.6% formaldehyde and 50% formamide). The samples were heated to 60°C for 15 min to denature secondary structures and, then, cooled immediately on ice. Two µl of RNA loading dye (1 mM EDTA, 50% glycerol and 0.04% bromophenol blue) were added and the samples run for 4 h at 60 V on a 1.25% denaturing agarose gel containing 0.66 M formaldehyde and 1x MOPS. The gel was stained in a methylene blue solution for 15 min at room temperature and destained with ddH_2O until two bands, representing the 18S and 28S ribosomal RNA, were visible, the latter being about twice as intense as the former in intact RNA.

2.2.4.2 Reverse transcription (RT)

Reverse transcription (RT) of the RNA for first-strand cDNA synthesis was carried out according to a standard protocol (Promega, Madison, USA). In brief, 20 µl reaction mixture contained 1 µg RNA, 5 mM $MgCl_2$, 1x reverse transcription buffer (10 mM Tris–HCl, pH 9, 50 mM KCl and 0.1% Triton X-100), 1 mM of each dNTP, 1u/µl recombinant RNasin ribonuclease inhibitor, 15 u/µg AMV reverse transcriptase and 0.5 µg oligo $(dT)_{15}$ primer per µg RNA. The samples were incubated at 42°C for 15 min and then at 95°C for 5 min; thereafter they were cooled for 5 min on ice.

2.2.4.3 Microarray analysis

The microarray analysis was performed at the ARK-Genomics center, Roslin Institute, Edinburgh, UK, under the supervision of Mr. R. Talbot and Miss A. Downing.

The transcriptional profiles of three different tissues (brain, muscle and spinal cord) of two newborn Swiss Large White piglets, one with AMC, the other normal, were compared on self-printed Operon pig 13K oligo arrays. The array consisted of 10,665 70mer gene sequences (Pig Genome Oligo Set Version 1.0) representing 10,665 Sus scrofa gene sequences with a hit to a known human, mouse or pig gene transcript and some sequences containing a 3' expressed sequence tag (EST). The array also contained 2,632 70mer probes (Pig Genome Oligo Extension Set Version 1.0) representing 2,632 Sus scrofa genes with at least one 3' EST. All the probes had a Tm of 78°C (±5°C) and a 5' amino linker. Furthermore, negative and positive control probes, as well as "landing lights" (β-actin) for easy alignment of the grids during analysis were added. The gene list corresponding to the printed oligonucleotides is available on the Operon website http://omad.operon.com/pig/query.php.

RNA samples were extracted as described in section 2.2.1.2 and loaded into an RNA 6000 LabChip. The Chip was transferred to an Agilent 2100 Bioanalyser, which monitored the fluorescence between 670 and 700 nm, enabling to determine the purity and quantity of the RNA. Labeling the total RNA with Cy3 or Cy5 was carried out with the Fairplay kit (Stratagene, La Jolla, USA) before hybridizing to the arrays on a Genomic Solutions GeneTAC Hybridization Station (Genomics Solutions, Cambridgeshire, UK). The protocols for labeling and hybridization are available from http://www.ark-genomics.org/protocols/. Each slide had AMC tissue and normal tissue; each pair was measured twice with the dyes reversed in the second measurement. The slides were scanned at an appropriate level (less than 0.3% of features with maximum signal) with a Scan Array 5000 XL (GSI Lumonics, Watertown, USA). Data were extracted from the images using the image analysis program BlueFuse (BlueGnome, Cambridge, UK). The quantified images with the aligned features were checked for anomalies (e.g. dust particles); such features were not analyzed. Each array was printed with duplicate features for each oligonucleotide printed. The proprietary fusion function of the software was used to merge the two features for each oligonucleotide. The fused data were analyzed further. BlueFuse relies on a Bayesian approach to analyze the image, which gives a weight to each component of a measured feature. These weights are used to determine an overall "confidence" in the value of each measurement, which is used to grade the data (A "good" to E "very poor"). Because each pair of samples was analyzed on

two slides, the quality of a gene value was taken as a combination of both confidence values. Only those genes with one grade of A or B and the other of C or higher were compared with the gene lists derived from the GeneSpring analysis.

The extracted, quantified data were imported into the GeneSpring 5 (Silicon Genetics, Redwood City, USA) program for normalization and determination of the differentially expressed genes. Gene lists were generated based on a change in expression (by more than 2) between the normal and the diseased pigs.

A shorter list of genes for further evaluation was obtained applying the confidence grade of each measurement to the lists of genes generated for each tissue from GeneSpring.

2.2.4.4 Real-time PCR

Real-time PCR with a TaqMan assay was used to amplify the RT products of the genes *C3F*, *NR4A1*, *Q6ZUQ4*, *SFRS6* and *SLC2A1*, selected from the microarray experiment. The primers were designed using the Primer Express™ Software v1.5 (PE) (Applied Biosystems). The TaqMan probe was selected in an area between the F and R primers. It was designed to be longer than the primers and with a Tm of 10°C higher than the primers. The probe was labeled with FAM at the 5' end (reporter dye) and with TAMRA at the 3' end (quencher dye). The primer sequences are listed in Table 3.1.

The reaction mixture (total of 100 µl) included 10 µl first-strand cDNA (section 2.2.4.2), 200 µM dNTPs, 2 mM $MgCl_2$, 1x reverse transcription buffer (10 mM Tris–HCl, pH 9, 50 mM KCl and 0.1% Triton X-100), 50 pmol of each of the three forward, reverse and TaqMan primers and 2.5 u of Taq DNA Polymerase (according to the protocol of Promega, Madison, USA).

The real-time PCR reaction was performed in a thermal cycler ABI Prism 7700 Sequence Detector (Applied Biosystems) with MicroAmp optical tubes (Applied Biosystems). The samples were cycled at 95°C for 10 min followed by 40 cycles at 95°C for 15 s and at 60°C for 1 min. The default setting for the baseline fluorescence calculation was baseline cycles 3 to 15. The threshold default was 10 standard deviations above the mean fluorescence generated during the baseline cycles and the C_T value was determined by the point at which the fluorescence crossed the threshold.

The following formula was used to calculate the fold changes in the initial amount of cDNA of the five genes in the normal piglet compared to the piglet with AMC:

$$I = 2^{|\Delta C_T|}$$

where **I** is the fold increase and $|\mathbf{\Delta C_T}|$ the absolute value of the difference C_T (normal)-C_T (AMC-piglet).

If $|\Delta C_T|>1$, a difference in gene expression in the starting cDNA by more than two was found.

2.2.5 Statistical analysis

2.2.5.1 Chi square test

The hypotheses of a recessive and sex-independent mode of inheritance of AMC were tested with a chi square test. This test is a non-parametric test of statistical significance and roughly estimates the degree of confidence in accepting or rejecting a hypothesis. Typically, the hypothesis tested with chi square is whether or not two samples (in our case observed and expected number of piglets with AMC) are sufficiently different in one characteristic (3:1 ratio of normal to diseased piglets for a recessive mode of inheritance; 1:1 ratio of male to female piglets with AMC for a sex-independent mode of inheritance) that it is possible to generalize that the tested samples are also different in that particular characteristic in a normal population under normal conditions. The program used for the test was downloaded from http://www.georgetown.edu/faculty/ballc/webtools/web_chi.html.

2.2.5.2 Genetic linkage

Marker linkage maps were computed with Crimap software, version 2.4 (Green *et al.* 1990), by using the flips and all options to determine the best ordering of the markers and the fixed option to obtain the map distances.

The correlation between the markers and the recombination frequency gave the distances in cM (two-point analysis) as well as the most probable location and order of the markers on SSC5.

2.2.5.3 QTL mapping

The map resulted from the linkage analysis (section 2.2.5.2) was used to analyze the single-trait QTL map of chromosome 5 by using the QTL-Express program (Seaton *et al.* 2002; http://qtl.cap.ed.ac.uk/), based on the least-squares regression interval mapping program developed by Haley *et al.* (1994). Marker information was used to calculate the probabilities that an F_2 offspring inherited no, one or two alleles from each founder for a putative QTL

allele at 1 cM intervals in SSC5. The analysis included information from flanking markers in addition to information from the marker itself to determine the origin of marker alleles in the F_2 progeny, according to the method of Haley *et al.* (1994). The 'quantitative trait' refers to the presence or absence of the disease (all or none trait). This binary trait was regressed on probabilities of allele transmission for additive and dominance components of the statistical model.

Detection of QTL was based on F statistics that were computed from the sums of squares explained by the additive and dominance coefficients for the QTL. Empirical significance thresholds of the F statistics were derived at the chromosomal level on a single trait basis by the permutation test using the QTL-Express program. A total of 10,000 random permutations of the data were analyzed. Only individuals from F_2 population (Table 2.1) were considered.

2.2.5.4 Chromosomal assignment

The results of the RH panel were statistically evaluated with the IMpRH mapping tool (Milan *et al.* 2000 B), available from the IMpRH Web Server (http://imprh.toulouse.inra.fr/).

The closest linked marker of each gene was identified using the statistical two-point analysis.

2.2.5.5 Allele frequencies

The allele frequencies and the observed heterozygosity of the microsatellites *bE77* and *SW904* in 80 pigs (1-80 in Table 2.3), representing the Swiss LW population, were calculated using the Microsatellite Toolkit version 3.1 (available from http://oscar.gen.tcd.ie/_sdepark/ms-toolkit/).

3. Results

3.1 AMC inheritance

We analyzed a total of 358 piglets (20.11.2005). Of these, 338 (Table 2.1, matings 1-7, 9 and 11-19) had heterozygous parents. The 84 (23.5%) diseased descendants exhibited severe arthrogryposis and were of both sexes (52 males and 32 females). The X^2 test, which was calculated from the segregation data, showed that the observed ratios of the *AMC* to *amc* alleles did not deviate significantly from the 3:1 ratio of normal to diseased piglets, as would be expected when both parents are heterozygous (X^2 = 0.002; 0.95 < P < 0.97; 1df). Furthermore, the observed ratios of male to female diseased piglets did not deviate significantly from the 1:1 ratio that would be expected to result from a sex chromosome-independent disease (X^2 = 2.4; 0.1 < P < 0.2; 1df).

3.2 Linkage mapping

3.2.1 SSC2

No correlation was found between the markers *SWR2516*, *SW1650*, *SW1026*, *S0370* and *SWR345* and *amc* (LOD scores of 0.18, 0.0, 0.0, 0.0 and 0.04, respectively).

3.2.2 Full genome

The preliminary genome scan of 130 pigs revealed a correlation between the *amc* allele and the allele *165* (165 = size in bp) of microsatellite *SW1987*, located on pig chromosome 5. The LOD scores between *AMC* and 43 microsatellite markers, located on the entire swine genome, with the exception of SSC5, ranged from 0.00 to 0.35 and, therefore, did not even come close to the significant value of 3.

3.2.3 SSC5

Porcine *AMC* mapped to SSC5; there was significant pair-wise linkage to eight markers. For each of these markers, the two-point recombination frequencies and LOD scores greater than 6 are as follows: *SW963* (0.12, 8.3), *SW1987* (0.07, 10.5), *SW152* (0.01, 14.9), *SW904* (0.03, 16.8), *SW1094* (0.03, 14.9), *SWR1526* (0.03, 13.3), *SWR1974* (0.03, 15.3) and *SW310* (0.03, 13.3). Figure 3.1 shows the sex-averaged linkage map with the most likely order of the microsatellites and *AMC* on SSC5, calculated with the "fixed" option. The distance between *SW152* and *AMC* is 2.5 cM and the distance between *SW904*/*SW1094* and *AMC* is 2.4 cM. No recombination events between *SW904* and *SW1094* were found. The microsatellites *SW1468* and *S0018*, also located in this region, were not informative in the experimental family.

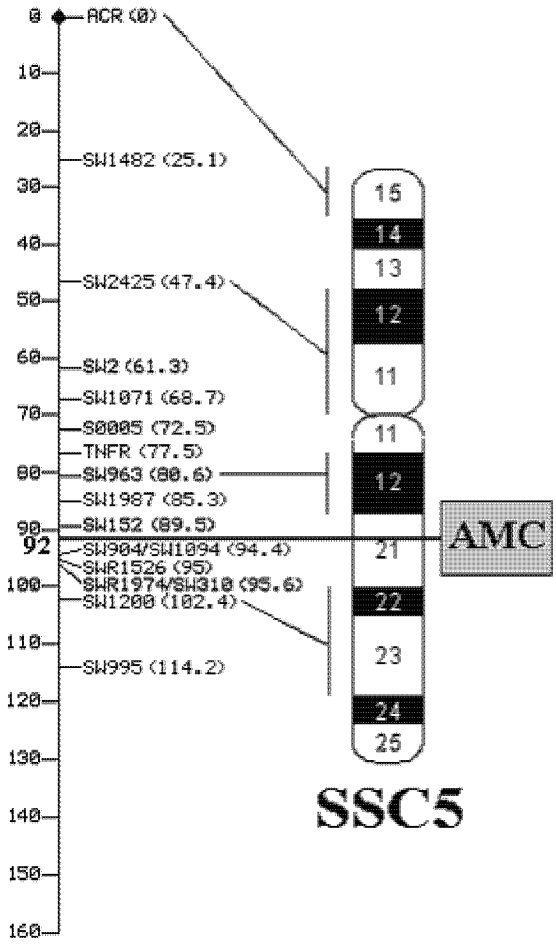

Figure 3.1: Genetic and physical maps of markers and AMC on SSC5.
Sex-averaged genetic linkage maps with genetic positions and distances in Kosambi cM of microsatellite markers and *AMC* on swine chromosome 5 (SSC5).

3.3 QTL mapping

Significant QTL with a maximal F value of 24.3 was detected on SSC5 in the region of *AMC* (position 92 cM). Figure 3.2 shows the QTL graph for significance at the 1 and 5% chromosome-wise levels, representing plots of the F statistic across the chromosome. Thresholds for chromosome-wise significance at the 1 and 5% levels, as determined by the permutation test, were 5.1 and 7.1, respectively.

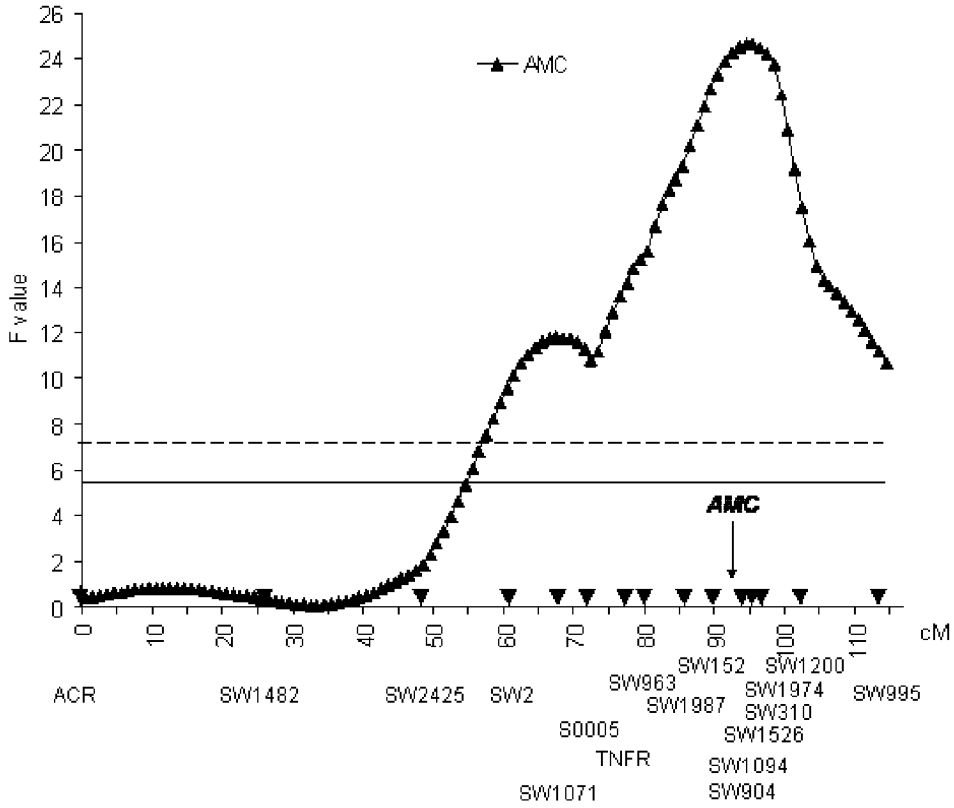

Figure 3.2: F ratio curves for evidence of a QTL on SSC5.
The x axis indicates the relative position on the linkage map. The y axis represents the F-ratio. Arrows on the x axis indicate the position of a marker.
(-----): 1% chromosome-wise significance
(___): 5% chromosome-wise significance

3.4 Sequencing

3.4.1 BAC clones

Using the primers T7 and Sp6 the end sequencing results of the clones PigE-258K14 and PigE-214F17, containing *SW152* and *SW904* respectively, confirmed the recently published sequences at the Sanger Institute. The sequences (about 350 bp in length), obtained with the microsatellite primers *SW152* forward and reverse and *SW904* forward and reverse, could not be assigned with statistical significance to a putative human sequence using the program blast.

3.4.2 DNA

Table 3.1 lists 35 partial gene sequences of 24 different genes and gives the PCR conditions. Underlined primers were used to map the genes with the RH panel.

Nine porcine homologues to human *ABCD2*, *CNTN1*, *KCNA1*, *KIF21A* and *SCN8A*, sequenced from genomic DNA by cross-species PCR, were submitted to GenBank (Appendix, Table A-1) and received accession numbers with the initials DQ.

The sequence identities between the pig EST gene sequences (GenBank accession numbers in Table 3.1) and the putative human sequences ranged from 83% of *USP18* to 98% of *ABCD2* and *SCN8A*.

SNPs were found in the partial gene sequences of *CNTN1* (DQ180325) and *TUBA8* (BW967428) (further discussed in section 3.4.3).

The remaining 33 gene sequences of a diseased and a normal piglet were identical.

Table 3.1: Porcine sequences of positional candidate genes for AMC.

*: TC accession number (http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?species=pig)
**: TaqMan primer sequence
***: UniProt name (http://www.pir.uniprot.org)
Underlined: primers used for radiation hybrid mapping

Gene Symbol	Porcine GenBank accession number	Sequence similarity with human (%)	Primer sequence (5'-3')	PCR product size (bp)	Annealing temperature
ABCD2	DQ338453	91	F-atcattcccggtagatgcag R-gcactcgatctggtccattt	275	65°C
	DQ338454	95	F-atgattgccatccctgctac R-gccacagattgggagaacat	212	62°C
	DQ338455	98	F-ttgatggtgaaggaggttgg R-gcactgagtcttctcccaaaa	159	60°C
CACNA1C	CR910212	87	F-acatcaacaatgccaacaacac R-ctgcttacctcttggagctttg	156	65°C
COL2A1	TC218959*	94	F-ttttctgggtggcctccatg R-tgggaaacctcgttcaccctg	185	58°C
CNTN1	DQ180325	87	F-caacaaaaccatatcctgctga R-ccccagaggtgctaatctca	324	60°C
	DQ180326	89	F-ggctgcacacagagttcaag R-tgctttcttggtgaaggtttc	154	62°C
	CK453087	90	F-gcctttggcatccttgtcta R-cttttggatgaaggggttcc	218	62°C
CPNE8	CA779293	84	F-gcaagggaagggaaagattc R-tcagaacccaaactgcatga	169	60°C
	CA779293	91	F-aatattgccctcgactgctg R-gaatctttcccttcccttgc	224	60°C
C3F	TC103650*	91	F-aacccccgcttcactggt R-agtaacgagccacccaggc **T-ccatcgcctctttcaacatcaacacca	67	60°C
C12ORF4	BV102651	87	F-ctttcctttgtttcccccac R-ttaactgcatggctatgggc	242	58°C
DDX11	AW315438	88	F-agctgacgtcttcagatggg R-gggtcaaacccagagttgtc	285	58°C
FOXM1	TC49301	89	F-tttgagtaggacagcaggca R-agggcacagatgcaaacagt	235	60°C
GDF11	AY722392	87	F-aaaggtagaccaggaatgcaaa R-atctctcctcaactcaccttgc	247	57°C
HOXC8	AC145073	86	F-agaatctgttgccaaaaggaag R-atttcgaagtaaaaggggaagc	229	57°C
KCNA1	DQ180323	96	F-ttctccttcgagcttgtggt R-tcggggatactggaaaagtg	388	56°C
	DQ180323	97	F-gaggcggaagaagctgagt R-gtagaggaactgcggcgac	305	56°C
	DQ355786	92	F-tcgagacgcagctcaagac R-ctcggggtactcgaagagc	341	56°C

Table 3.1: Porcine sequences of positional candidate genes for AMC. (continuation)

*: TC accession number (http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?species=pig)
**: TaqMan primer sequence
***: UniProt name (http://www.pir.uniprot.org)
Underlined: primers used for radiation hybrid mapping

Gene Symbol	Porcine GenBank accession number	Sequence similarity with human (%)	Primer sequence (5'-3')	PCR product size (bp)	Annealing temperature
KIF21A	DQ180324	89	F-cttcgtcggcaagtaagacc R-aaacactcggccagtcaatc	389	60°C
	BQ600558	88	F-catgtgcagtgggaaaagc R-ccctggctaagctttcatgt	251	60°C
	BQ600558	92	F-tgggaaacatttttgatgtcc R-cttttcccactgcacatgat	217	60°C
MDS028	TC63101*	97	F-aggttgtctcatcctaccaagtgt R-aacctataccacaactcacagcaa	361	68°C
MGC5576	TC202281*	85	F-tggcagtacctgtcagaagaga R-ctccgagtgcctactatttgct	222	58°C
NR4A1	TC120775*	96	F-gggccagctgttcccag R-tgagtttatgtacacgtggaaggg **T-ctcctcctgtttgctccctcttctcgc	70	60°C
PHB2	G72834	92	F-tttccctcgttttgctctgt R-tgacccttcacccttgaatc	241	58°C
PRICKLE1	BV102540	85	F-tgaactcttccatgctgcac R-cccctctcgtcaaagtggta	256	62°C
*Q6ZUQ4****	TC118731*	91	F-gaagggacgtggccatatga R-aggtgtccaggcgctcct **T-aggcttgcgtggcccgatgtag	62	60°C
SCN8A	DQ180322	98	F-ccaggccctgatagtttcaa R-actcttccctgcttccaggt	168	58°C
SLC2A13	BF194506	93	F-ggctgtgtacatgtggtgct R-taaaagcagggcagtgaggt	196	68°C
TUBA8	BW967428	91	F-ggaatgcatatccgtccac R-ggctccagatcaaccatgac	210	65°C
	BP444494	90	F-acggcaagaagtccaagcta R-caccaggttggtctggaact	300	60°C
USP18	TC222862*	83	F-gggaaactgcatatcttctggt R-ctggagagtcggtgaccaatac	553	58°C
YAF2	AJ667918	95	F-aggttactgttggcgacctg R-tctgcatggtacgacagagg	241	65°C
	AJ667918	89	F-ggtcatcttcacccagagga R-ggcatctttaactgcggaga	265	65°C

3.4.3 Detection of SNP markers

3.4.3.1 *TUBA8* gene

In the partial gene sequence of *TUBA8* (size 210 bp, exon 2), three SNPs in complete linkage disequilibrium with *AMC* (*T→C* transitions) were discovered at relative positions 124 bp, 130 bp and 163 bp of the PCR product (Figure 3.3). Therefore, a total of 230 pigs, including 14 important recombinants (Table 3.3), were sequenced. Diseased piglets had only the *T* allele at the three mutation sites, carriers were heterozygous (*T/C)* at all three sites and homozygous normal pigs had only the *C* allele at the three loci. No recombinant pig between *TUBA8* and *AMC* was found in our experimental family. However, one diseased piglet, from a commercial farm, showed recombination (Table 3.2). The LOD score between *TUBA8* and *AMC* was 38.5. Three pig EST sequences (accession numbers BP444494, CJ005941 and CJ006528) showed the *C* at the three SNP loci, as in normal pigs. Compared with the corresponding human sequence, normal piglets were identical only for the first site at relative position 124 bp. For the other 2 SNPs at positions 130 bp and 163 bp, humans and diseased piglets had the same genotype (*T/T*).

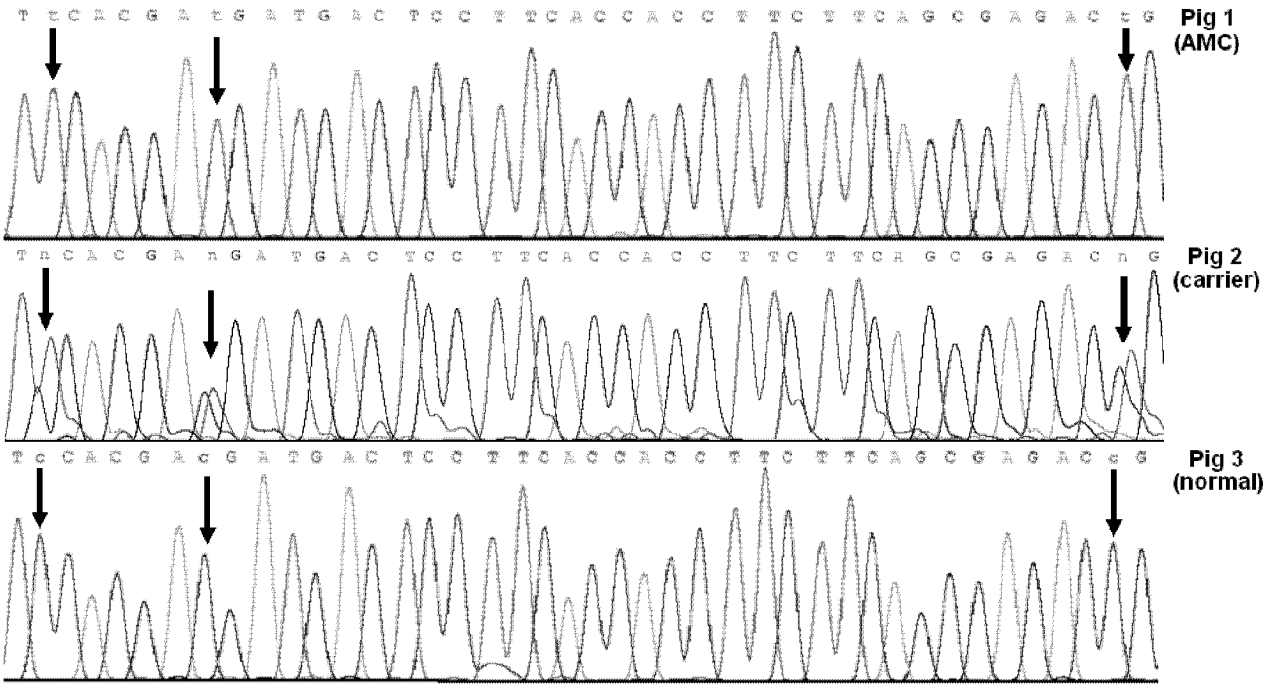

Figure 3.3: Visualization of SNPs in TUBA8.
Nucleotide sequences in *TUBA8* surrounding the three SNPs. The arrows indicate the locations of the three SNPs in pig 1 (genotype *T/T*), pig 2 (genotype *C/T*) and pig 3 (genotype *C/C*).

3.4.3.2 *CNTN1* gene

A polymorphism in complete linkage disequilibrium with *AMC* (*G→A* transition) was found in the partial gene sequence of *CNTN1* (size 324 bp, intron 7, DQ180325) at relative position 165 bp (Figure 3.4).

The major limitation of this SNP for linkage analysis in our family was that boar 1481 CH (matings 14-19, Table 2.1) had only the *A* allele and was, therefore, not informative. A total of 230 pigs, including 15 recombinants (Tables 3.2 and 3.3), were sequenced. Diseased piglets had the *A* allele at the mutation site, carriers were heterozygous (*A/G*) and homozygous normal pigs had only the *G* allele. The LOD score between *CNTN1* and *AMC* for all the 358 pigs of the experimental herd was 24.1 and the recombination fraction was 0.

The corresponding base in humans could not be defined and a comparison with the human sequence was impossible, because the SNP occurred in an intron.

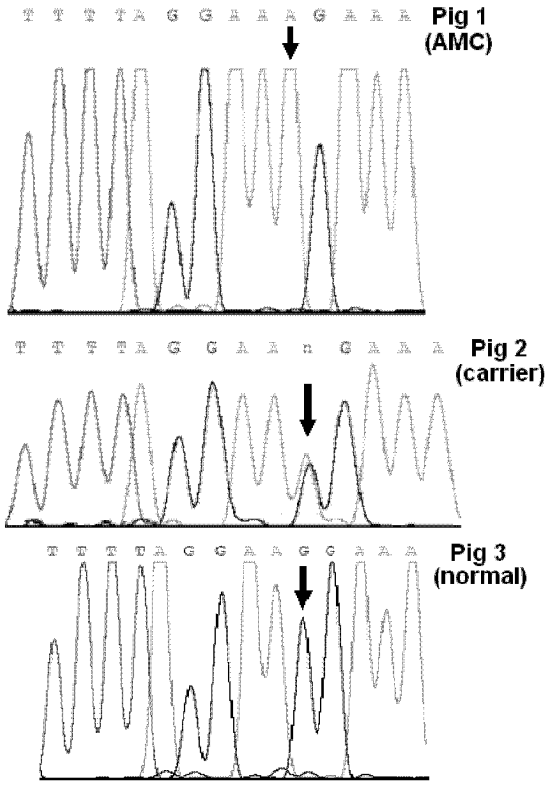

Figure 3.4: Visualization of a SNP in CNTN1.
Nucleotide sequences in *CNTN1* surrounding the SNP. The arrows indicate the location of the SNP in pig 1 (genotype *A/A*), pig 2 (genotype *A/G*) and pig 3 (genotype *G/G*).

3.5 Identification of new microsatellites

The microsatellite *bE77* (forward sequence tgctatacagcaaattgacccact and reverse sequence gcagacacagctcagatcca) located on the BAC clone CH242-77C1 was in complete linkage disequilibrium with *AMC* in the resource population. $bE77^{306}$ co-segregated with the *amc* allele and $bE77^{281}$ was 100% linked with the normal allele.

The LOD score between *bE77* and *AMC* for the 358 pigs was 51.2 and the recombination fraction was 0. One recombinant piglet was found between the marker and *AMC* in the diseased pigs originating from a commercial herd (Table 3.2).

Six new microsatellites in the *AMC* region, located on BAC clones CH242-47G11, CH242-257E16, CH242-285P2, CH242-20I13, PigE-229I17 and PigE-194N21, were identified. As the microsatellite *UMNp1136* (Fahrenkrug *et al.* 2005), they were not informative in the founder pigs of the experimental herd and, therefore, were not investigated further.

3.6 Recombinant pigs

3.6.1 Position of *AMC* according to recombinants

Only one recombinant offspring (number 10023, Table 3.2) was found between *TUBA8* and *AMC*. The diseased piglet from a commercial herd had the genotype $TUBA8^{T/C}$. The recombination site indicates that the *AMC* locus is located towards the telomere from the SNP in exon 2 of *TUBA8*.

Table 3.2: Recombinant pig in regard to the SNPs in TUBA8.

Genotypes and haplotypes of a diseased recombinant offspring, with its parents and mating number, tested with the microsatellites *SW152*, *bE77* and *SW904* and the informative SNPs in *TUBA8*. The first allele shown for the offspring is the father's allele (underlined) and the second is that of the mother (in box).

↑: site of recombination
Bold: recombined chromosomal segment
Genotype *amc*/*amc*: diseased piglet
Genotypes *AMC*/*amc* or *amc*/*AMC*: normal piglets
?: uncertain allele, because DNA was not available

Mating (see Table 2.2)	Animal	*SW152*	*bE77*	*TUBA8*	*AMC*	*SW904*
8	Father 3367 HT	167/177	281/306	*C*/*T*	*AMC*/*amc*	172/180
8	Mother 1724 HT	177/?	306/?	*T*/?	*amc*/*AMC*	172/?
8	10023 (AMC)	**167**/177	**281**/306	**↑*C***/*T*	*amc*/*amc*	180/172

Table 3.3 gives the results of 14 manually haplotyped recombinant pigs, which permit to locate *AMC* towards the centromere from the microsatellite *SW904*.

Table 3.3: Recombinant pigs in regard to microsatellite SW904.
Genotypes and haplotypes of 14 recombinant offspring (6 normal and 8 with AMC) with their parents and mating numbers and tested with the microsatellites *bE77* and *SW904* flanking the *AMC* region. The first allele shown for each offspring is the father's allele (underlined) and the second is that of the mother (in box).

F_2/F_3: offspring generation
↑: site of recombination
Bold: recombined allele
Genotype *amc/amc*: diseased piglets
Genotypes *AMC/amc* or *amc/AMC*: normal piglets

Mating (see Table 2.1)	Animal	*bE77*	*AMC*	*SW904*
1	Father 1077 GT	281/306	*AMC/amc*	182/172
1	Mother 360 HNM	281/306	*AMC/amc*	172/180
1	F_1: 71 HNM (normal)	306/281	*amc/AMC*	↑**182**/172
2	Father 1077 GT	281/306	*AMC/amc*	182/172
2	Mother 359 HNM	281/306	*AMC/amc*	172/180
2	F_1: 79 HNM (normal)	306/281	*amc/AMC*	172/↑**180**
3	Father 68 HNM	281/306	*AMC/amc*	182/180
3	Mother 71 HNM	306/281	*amc/AMC*	182/172
3	F_2: 18 (normal)	306/281	*amc/AMC*	180/↑**182**
3	F_2: 1735 CH (normal)	306/281	*amc/AMC*	180/↑**182**
4	Father 68 HNM	281/306	*AMC/amc*	182/180
4	Mother 72 HNM	281/306	*AMC/amc*	182/180
4	F_2: 1897 CH (normal)	281/306	*AMC/amc*	↑**180**/180
4	Father 68 HNM	281/306	*AMC/amc*	182/180
4	Mother 72 HNM	281/306	*AMC/amc*	182/180
4	F_2: 74 (AMC)	306/306	*amc/amc*	180/↑**182**
9	Father 68 HNM	281/306	*AMC/amc*	182/180
9	Mother 1731 CH	306/281	*amc/AMC*	180/172
9	F_3: 80 (AMC)	306/306	*amc/amc*	↑**182**/180
13	Father 68 HNM	281/306	*AMC/amc*	182/180
13	Mother 2884 CH	306/281	*amc/AMC*	180/182
13	F_3: 104 (AMC)	306/306	*amc/amc*	↑**182**/180
14	Father 1481 CH	306/281	*amc/AMC*	180/174
14	Mother 72 HNM	281/306	*AMC/amc*	182/180
14	F_3: 129 (AMC)	306/306	*amc/amc*	↑**174**/180
16	Father 1481 CH	306/281	*amc/AMC*	180/174
16	Mother 2884 CH	306/281	*amc/AMC*	180/182
16	F_3: 124 (AMC)	306/306	*amc/amc*	↑**174**/180
16	Father 1481 CH	306/281	*amc/AMC*	180/174
16	Mother 2884 CH	306/281	*amc/AMC*	180/182
16	F_3: 125 (AMC)	306/306	*amc/amc*	180/↑**182**
16	F_3: 147 (AMC)	306/306	*amc/amc*	180/↑**182**
18	Father 1481 CH	306/281	*amc/AMC*	180/174
18	Mother 3647 CH	281/306	*AMC/amc*	182/180
18	F_3: 4975 CH (normal)	306/281	*amc/AMC*	180/↑**180**
19	Father 1481 CH	306/281	*amc/AMC*	180/174
19	Mother 3669 CH	281/306	*AMC/amc*	182/180
19	F_3: 169 (AMC)	306/306	*amc/amc*	180/↑**182**

The recombinant events in Tables 3.2 and 3.3 confirm the results of the linkage analysis (Figure 3.1) and suggest that the most probable locus order on the q-arm of SSC5 is centromere-*SW152-bE77-TUBA8* (exon 2)-*AMC/CNTN1* (intron 7)-*SW904/SW1094*-telomere (Figure 3.5).

The location of *AMC* with respect to *CNTN1* and of *SW904* with respect to *SW1094* could not be determined due to missing recombinants.

3.6.2 Position of *UMNp1275* according to recombinants

The microsatellite *UMNp1275*, previously assigned to SSC5 but not mapped with respect to the other microsatellites of the AMC region, was mapped between the microsatellites *SW1987* and *SW152* based on the two recombinants in Table 3.4. The normal offspring 3662 CH indicates that *UMNp1275* is located towards the telomere from *SW1987* and the diseased offspring 28 indicates that *UMNp1275* is located towards the centromere from *SW152*.

Table 3.4: Recombinant offspring indicating the position of UMNp1275.

Genotypes and haplotypes of two recombinant offspring (one normal and one with AMC), with their parents and mating number, tested with five polymorphic microsatellites. The first allele shown for each offspring is the father's allele (underlined) and the second is that of the mother (in box).

F_2/F_3: offspring generation
↑: site of recombination
Bold: recombined chromosomal piece

Mating (see Table 2.1)	Animal	*SW963*	*SW1987*	*UMNp1275*	*SW152*	*bE77*
13	Father 68 HNM	164/168	167/165	202/198	167/177	281/306
13	Mother 2884 CH	168/164	165/167	198/202	177/167	306/281
13	F_2: 3662 CH (normal)	168/164	165/167	**↑202**/202	**167**/167	**281**/281
5	Father 68 HNM	164/168	167/165	202/198	167/177	281/306
5	Mother 75 HNM	164/168	167/165	202/198	167/177	281/306
5	F_3: 28 (AMC)	168/**164**	165/**167**	198/**↑202**	177/177	306/306

3.7 BAC clones

The microsatellite *SW1987* was located in BAC clone PigE-58F22, *SW152* in PigE-258K14, *S0018* in PigE-190I5 and *SW904* in PigE-214F17. *SW1094* was not found in any of the clones of the library.

All the clones are physically ordered (http://www.sanger.ac.uk/Projects/S_scrofa/) in contig ctg5002, which is continuously updated. To date (20.11.2005), this contig contains 3941 ordered BAC clones.

Figure 3.5 illustrates and graphically presents the information provided by the Sanger BAC contig combined with those of the recombinant pigs (section 3.6.1) in the AMC region. Four breakpoints (A to D differenciated by color) were identified based on the BAC order in

ctg5002 and their corresponding human location. The short A segment (green), containing *CACNA1C*, is located in pig between the segment B (brown), containing *SW1987* and spanning the remaining p-arm of HSA12 and the short segment D (red) of HSA22 with *SW152* and *bE77*. Adjacent to D is segment C (blue), containing the gene *CNTN1* and the microsatellites *SW904* and *S0018*.

To summarize, the order of the segments in pig is B, A, D and C. Furthermore, segment B appears to be inverted in pig with respect to human.

The corresponding human regions of SSC5q2.1, where *AMC* is mapped, are on HSA22q11 (D segment, red) and on HSA12q12 (C segment, blue), between the gene *TUBA8* and the microsatellite *SW904*. This region spans from about 16,968,113 to 16,983,493 Mb on HSA22 and from about 37,212,050 to 41,596,622 Mb on HSA12 to give a total of about 4,5 Mb.

The BES of clone CH242-77C1, containing the microsatellite *bE77*, has not yet been assigned to a corresponding human chromosomal region.

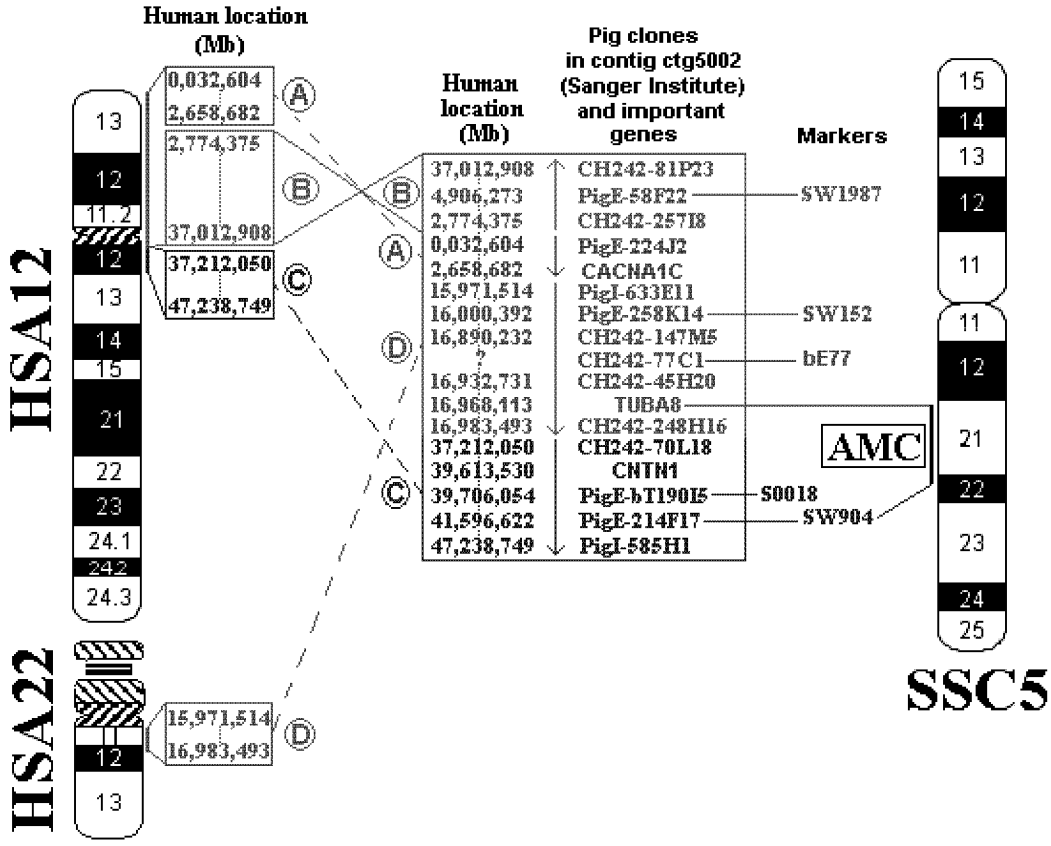

Figure 3.5: Chromosomal homologies between SSC5 and HSA12/HSA22 with location of AMC in BAC contig ctg5002.

The homologies in Mb between SSC5 and HSA12/HSA22 in the *AMC* region, based on Sanger BAC contig ctg5002 information, are shown. Four breakpoint blocks human-pig (A-D) have different colors, as well as the BAC clones carrying important microsatellites. The location of *AMC* is inferred from the information of recombinants between *TUBA8* and *SW904* on SSC5q2.1. The clone CH242-77C1 (?) has not yet been assigned to a corresponding human region.

3.8 Microarray analysis

A total of 308 genes (~2.4%) were differentially expressed under the strict conditions described in section 2.2.4.3; 131 were upregulated and 177 downregulated in the piglet with AMC. Only the gene *SFRS6* was upregulated and no genes were downregulated in all three tissues. The greatest number of differentially expressed genes was found in muscle (196) followed by brain (92) and spinal cord (35). Only 14 genes were differentially expressed at the same time in more than one tissue; the majority (11) were downregulated in the diseased piglet. A detailed representation of the number of genes differentially expressed in the three tissues is shown in a Venn diagram (Figure 3.6). The corresponding human sequences of the differentially expressed genes were evenly distributed on each human autosome and no particular chromosomal region had a significantly higher number of genes. None of the differentially expressed genes were located on the sex chromosome. Only the five differentially expressed genes, *C1S*, *C3F*, *PCBP2*, *NR4A1* and *Q6ZUQ4*, were located on HSA12 or HSA22 and were expected to map to the AMC region in pig.

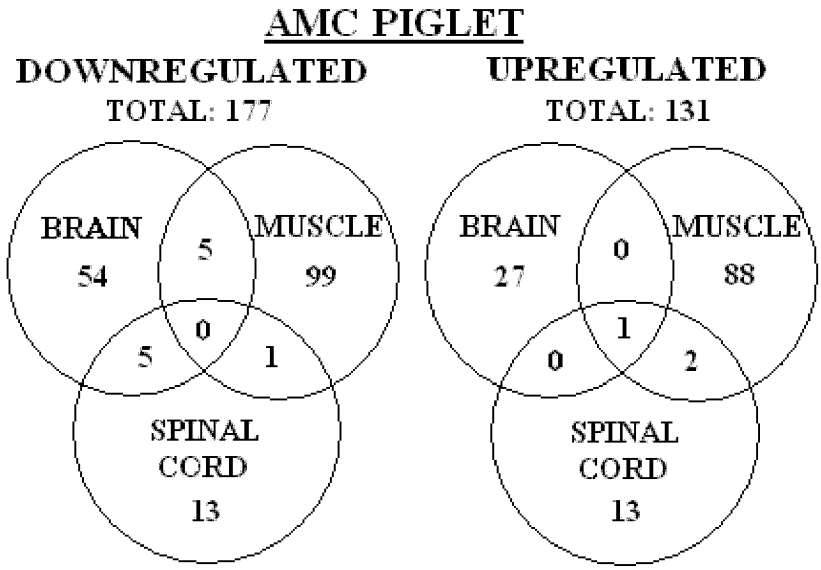

Figure 3.6: Number of genes down- and upregulated in a microarray experiment in brain, muscle and spinal cord of a diseased compared to a normal piglet.

3.9 Real-time PCR

As determined by microarray analysis, the abundance of differential gene transcripts in samples of the diseased and the normal piglet was tested by real-time PCR. As reported in Figure 3.6 for the diseased piglet, the genes *SFRS6* located on HSA20q13.11 and upregulated in all three tissues, *SLC2A1* located on HSA1p34.2 and downregulated in brain and muscle, *Q6ZUQ4* located on HSA22q11 and *C3F* located on HSA12p13, both more abundant in muscle and *NR4A1* located on HSA12q13 and less abundant in brain were analyzed.

Although neither *SFRS6* nor *SLC2A1* are on HSA12 or HSA22, they were analyzed because *SFRS6* was the only gene differentially expressed in all three tissues and *SLC2A1* belongs to the same gene family as *SLC2A13*, an important positional and functional candidate gene for *AMC*.

For all the cases, except *SFRS6* in brain and *SLC2A1* in muscle, the differential gene expression observed with the microarray analysis was validated. Figure 3.7 gives the graphic representation of the real-time PCR curves for the five genes.

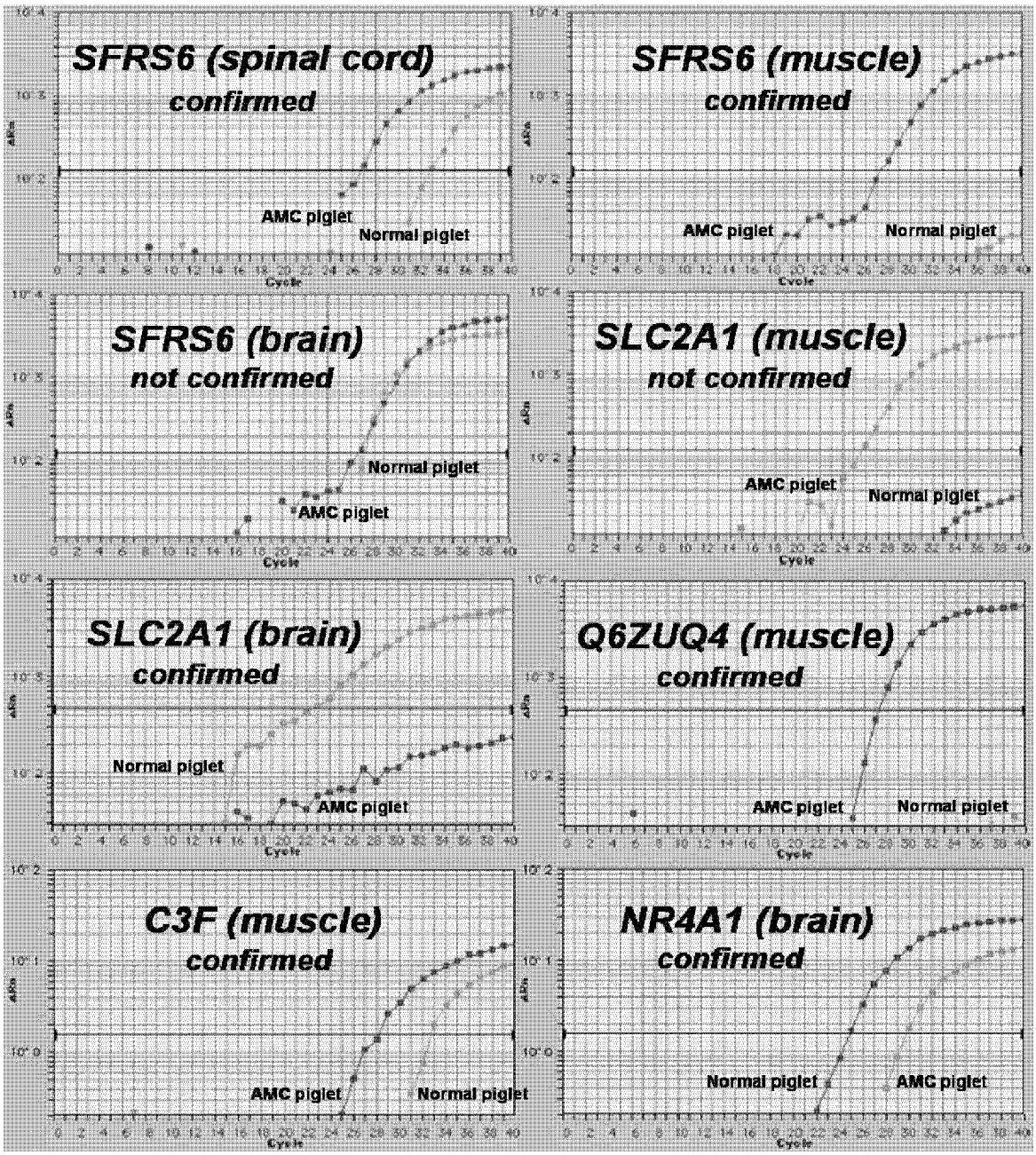

Figure 3.7: Log view of the amplicon production during each real-time PCR cycle of a AMC diseased compared to a normal piglet for SFRS6 in spinal cord, muscle and brain, for SLC2A1 in muscle and brain, for Q6ZUQ4 and C3F in muscle and NR4A1 in brain.

Confirmed: accordance between the results of the microarray and the real-time PCR analysis
Not confirmed: the results of the microarray and the real-time PCR analysis were different

The C_T values of the real-time curves of *SFRS6* in the spinal cord of the diseased and the normal piglet were 26.67 and 32.82 with a threshold value of 124.13.

The I value was $2^{(32.82-26.67)} = 71.01$, meaning that the amount of *SFRS6* cDNA in the spinal cord of the disease piglet was about 71 times higher than that of the normal piglet. An amplification of *SFRS6* was observed in the muscle of the AMC piglet (C_T=27.39) but there was no amplification in the muscle of the normal piglet.

In the brain of both piglets a similar amplification of *SFRS6* was detected (C_{TAMC}=26.67, $C_{Tnormal}$=27.18, I value = 1.42). Thus, the amount of cDNA in the brain of both piglets was statistically the same; a diverse gene expression could not be confirmed.

In contrast to the expectations following the microarray analysis, muscle *SLC2A1* amplified in the AMC piglet (C_T=25.71, threshold = 124.13) but not in the normal piglet.

The same gene amplified in the brain of the normal piglet with a C_T value of 22.67 and a threshold value of 471. In the AMC piglet the amount of cDNA was so low that the amplification did not reach the threshold value and therefore no C_T value could be evaluated.

The amplification of *Q6ZUQ4* in the muscle of the AMC piglet could be significantly detected after the 27th cycle of PCR (C_T=27.25, threshold value = 471), while no amplification of the same gene was detected in the normal piglet. The C_T value for *C3F* was 28.15 for the diseased piglet and 32.64 for the normal piglet (threshold value = 1.58). The I value was 22.47. Therefore, the amount of *C3F* cDNA in the muscle of the diseased piglet was 22 times higher than in the muscle of the normal one. Furthermore, the C_T was 29.8 for the diseased piglet and 24.91 for the normal piglet in the real-time PCR for *NR4A1* (threshold value = 1.58) with an I value of 29.65. Therefore, expression of *NR4A1* in the brain of the diseased piglet was about 30 times weaker than in the normal piglet.

3.10 Radiation hybrid mapping

The RH mapping results with assignments, the nearest linked marker, the LOD scores, the flanking markers and the location of the corresponding human gene (with accession number) for 18 gene sequences are presented in Table 3.5.

Table 3.5: Chromosomal assignments of 18 porcine candidate genes for AMC using the swine RH panel.

Gene symbol	Retention fraction (%)	Closest marker	LOD score	Flanking markers	Location (bp) in human http://www.ensembl.org (human accession number)
CACNA1C	33	*SW152*	13.5	*SW1987-SW152*	HSA12: 2094650-2672283 (NM_000719)
MDS028	47	*SW1468*	6.5	*SW1987-SW152*	HSA12: 2792124-2804367 (NM_018463)
C12ORF4	40	*SW1468*	18.4	*SW1468-SW1987*	HSA12: 4467172-4517898 (NM_020374)
KCNA1	33	*SW1987*	25.3	*SW1987-SW152*	HSA12: 4890806-4892293 (NM_000217)
PHB2	38	*SW963*	15.5	*S0005-SW963*	HSA12: 6944778-6950152 (BC014766)
C3F	41	*SW963*	14.9	*S0005-SW963*	HSA12: 6955608-6996103 (NM_005768)
DDX11	37	*SW152*	10.4	*SW1987-SW152*	HSA12: 31118077-31148999 (BC050069)
CPNE8	40	*SW152*	11.1	*SW152-S0018*	HSA12: 37332259-37587499 (BC046366)
PRICKLE1	48	*SW904*	19.1	*S0018-SW904*	HSA12: 41139341-41269745 (NM_153026)
MGC5576	44	*SW1200*	9.3	*SW1200-SW2034*	HSA12: 46644258-46648927 (NM_024056)
COL2A1	38	*SW1200*	11.3	*SW986-SW1200*	HSA12: 46653017-46684528 (NM_001844)
SCN8A	38	*SW439*	6.9	*SW70-IFNG*	HSA12: 50364210-50488564 (NM_014191)
NR4A1	38	*SW439*	7.1	*SW70-IFNG*	HSA12: 50717766-50739552 (NM_173157)
HOXC8	52	*SW70*	6.4	*SW436-SW70*	HSA12: 52689156-52692810 (NM_022658)
GDF11	39	*DK*	15.1	*DK-SW332*	HSA12: 54423333-54430332 (NM_005811)
Q6ZUQ4	35	*SW152*	14	*SW152-S0018*	HSA22: 16644986-16646112 (AK125438)
TUBA8	35	*SW152*	14.7	*SW152-S0018*	HSA22: 16968113-16989052 (BC074827)
USP18	34	*SW152*	11	*SW152-S0018*	HSA22: 17007311-17034714 (NM_017414)

The highest LOD score was found for *KCNA1* with *SW1987* (25.3) and the lowest for *HOXC8* with *SW70* (6.4). The retention frequency in the 118 clones ranged from 33% (*CACNA1C* and *KCNA1*) to 52% (*HOXC8*). The genes *CPNE8* and *PRICKLE1*, located on HSA12q12, and *Q6ZUQ4*, *TUBA8* and *USP18*, located on HSA22q11, were mapped in the AMC region between the microsatellites *SW152* and *SW904* (Figure 3.8). HSA12q12-13 represents a majority of the *AMC* interval, while HSA22q11 is homologous to the region directly adjacent to *SW152*.

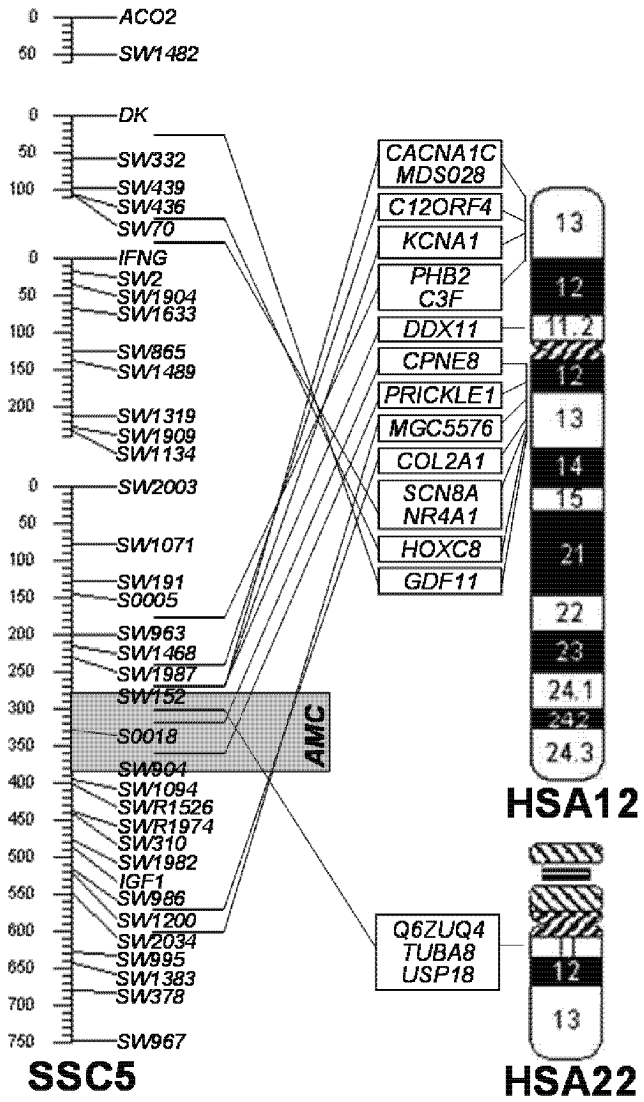

Figure 3.8: Chromosomal assignments of 18 genes on SSC5 and homologies with HSA12 and HSA22.

Cytogenetic RH map of SSC5 (Hawken *et al.* 1999, left) with the genes mapped using the radiation hybrid panel, the location of *AMC* and the comparative human maps with the location of the genes on HSA12 and HSA22.

Five major breakpoint blocks (A, D, E, F and G) and two minor blocks (B and C) contained genes mapped in this study (Figure 3.9, normal face type). Seven genes mapped with the RH panel by others (accession number and closest marker in parentheses) *BIN2* (BI336267, *SW439*), *C1R* (BI336339, *SW963*), *PAQR2* (BI336162, *SW152*), *PPFIBP1* (BI336612, *SW1134*) and *SLC6A12* (BI132403, *SW152*) from http://www.ag.unr.edu/ab/standard.htm and *CD9* (BI401319, *SW1468*) and *NELL2* (BI404672, *SW310*) from Tuggle *et al.* (2003) were added (Figure 3.9, bolt face type) in order to confirm assignments and create consistent blocks.

The newly formed seven breakpoint blocks (Figure 3.9) comprised the genes *SLC6A12*, *PAQR2*, *CACNA1C*, *MDS028*, *C12ORF4*, *CD9*, *PHB2*, *C3F*, *C1R* and *PPFIBP1* on

HSA12p13-p11 (block A), *KCNA1* on HSA12p13 (block B), *DDX11* on HSA12p11 (block C), *CPNE8*, *PRICKLE1* and *NELL2* on HSA12q12 (block D), *MGC5576* and *COL2A1* on HSA12q12 (block E), *BIN2*, *SCN8A*, *NR4A1*, *HOXC8* and *GDF11* on HSA12q13 (block F) and *Q6ZUQ4*, *TUB8A* and *USP18* on HSA22q11 (block G). The most likely block order on SSC5, based on the LOD scores and the flanking microsatellites for the genes mapped, was F, A, G, D, E with the insertion of the blocks B and C into the A block. Block B was mapped between the genes *C12ORF4* and *MDS028* and block C between *MDS028* and *CACNA1C*. Of the seven identified blocks, the gene orders in blocks A, F and E seem to be the reverse of that of humans.

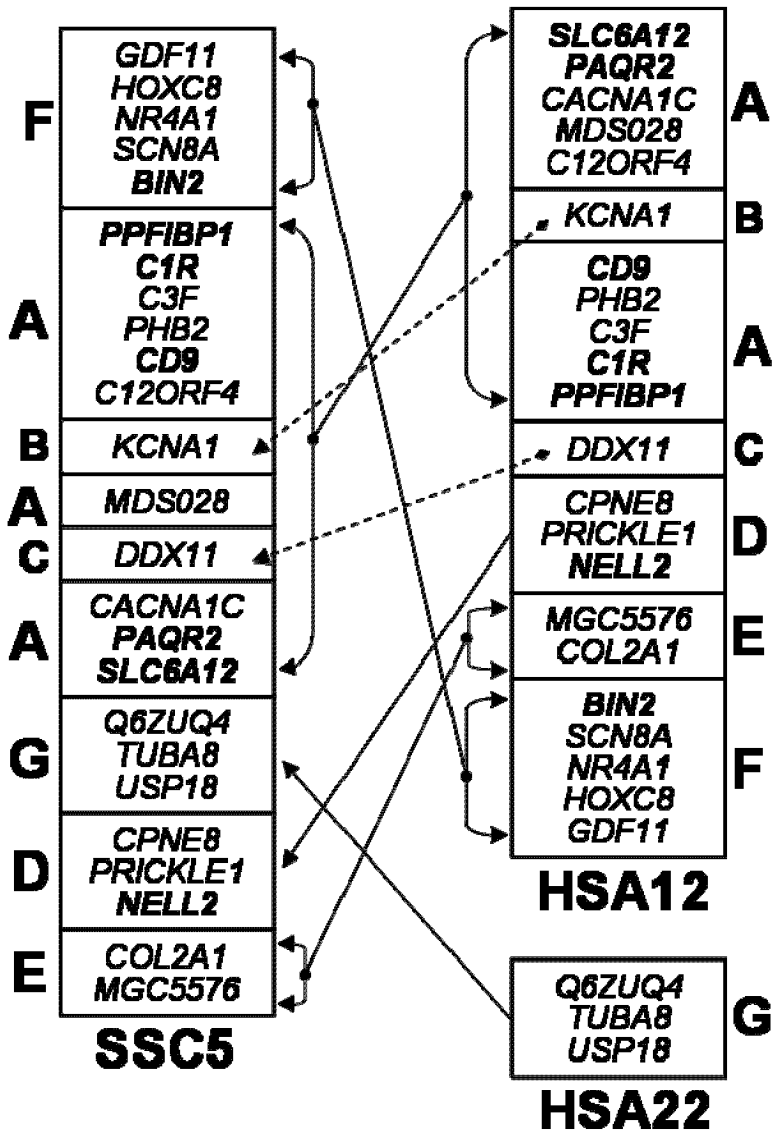

Figure 3.9: Most likely order of 25 genes on HSA12, HSA22 and SSC5.
Genes mapped in this study and genes reported in two previous studies (bold) *BIN2*, *C1R*, *PAQR2*, *PPFIBP1* and *SLC6A12* (http://www.ag.unr.edu/ab/standard.htm); *CD9* and *NELL2* (Tuggle *et al.* 2003).

3.11 Genetic marker test for *AMC*

A genetic AMC marker test comprising the microsatellites *bE77* and *SW904*, which flank the *AMC* region, was carried out.

The genotypes of 80 AI boars (1-80) and 23 sires (81-103), suspected of being carriers of AMC, are presented in Table 3.6. The classification of susceptibility groups (see last column) is inferred from haplotype results between *AMC* and *bE77* and *AMC* and *SW904*, for the experimental herd. The pigs were classified as follows: 1. high probability of not carrying *AMC* (without $bE77^{306}$ and $SW904^{180}$); 2. no decision possible (with $bE77^{306}$ but without $SW904^{180}$ or $SW904^{172}$); 3. no decision possible (without $bE77^{306}$ but with $SW904^{180}$ or $SW904^{172}$ or $SW904^{176}$); 4. carrier allele possible (with $bE77^{306}$ and with $SW904^{180}$ or $SW904^{172}$ or $SW904^{176}$).

Of the 80 AI boars, 50 (62.5%) belonged to the first, 9 (11.3%) to the second, 4 (5%) to the third and 17 (21.2%) to the fourth grade. The allele frequencies of the markers were 81.9% for $bE77^{281}$, 18.1% for $bE77^{306}$, 0.6% for $SW904^{166}$, 15.6% for $SW904^{172}$, 7.5% for $SW904^{176}$, 5% for $SW904^{180}$, 70.6% for $SW904^{182}$ and 0.6% for $SW904^{184}$. The observed heterozygosity of *bE77* was 0.28 and of *SW904* 0.45.

Of the 23 boars suspected of carrying *AMC*, 13 (56%) were assigned to grade 4 (possible carrier) and 10 (44%) to grade 1 (low risk of being carrier). None of these pigs were assigned to the grades 2 or 3.

Table 3.6: Genotypes of bE77 and SW904 for 103 AI-boars as indicators of susceptibility to AMC.

The allele lengths in the microsatellites *bE77* and *SW904* are given in bp. Boars 1-80 are presently used in Switzerland for AI; 81-103 were suspected of being carriers of *AMC*. Grades 1-4 explained in the text.

Boar	*bE77*	*SW904*	Grade	Boar	*bE77*	*SW904*	Grade	Boar	*bE77*	*SW904*	Grade
1	281/281	182/182	1	**36**	281/281	182/182	1	**71**	281/281	176/182	1
2	281/306	182/182	2	**37**	281/281	172/182	1	**72**	281/306	182/182	2
3	281/281	182/182	1	**38**	281/281	180/182	3	**73**	281/281	182/182	1
4	281/306	172/182	4	**39**	281/281	182/182	1	**74**	281/306	172/182	4
5	281/306	182/182	2	**40**	281/281	172/182	1	**75**	281/281	182/182	1
6	281/281	176/182	1	**41**	281/281	182/182	1	**76**	281/306	172/182	4
7	281/281	182/182	1	**42**	281/306	172/182	4	**77**	281/306	182/182	2
8	281/281	172/182	1	**43**	281/306	172/182	4	**78**	281/281	182/184	1
9	281/281	176/182	1	**44**	281/281	172/182	1	**79**	281/281	182/182	1
10	281306	166/172	4	**45**	281/281	182/182	1	**80**	281/281	176/182	1
11	281/281	182/182	1	**46**	281/281	172/180	3	**81**	281/306	180/180	4
12	281/281	172/176	1	**47**	281/281	182/182	1	**82**	281/281	172/182	1
13	281/281	172/182	1	**48**	281/281	182/182	1	**83**	281/306	176/180	4
14	281/306	182/182	2	**49**	281/281	172/176	1	**84**	281/306	180/180	4
15	281/306	180/180	4	**50**	281/281	172/182	1	**85**	281/306	172/182	4
16	281/281	182/182	1	**51**	281/306	182/182	2	**86**	281/306	180/180	4
17	281/306	176/182	4	**52**	306/306	172/182	4	**87**	281/281	176/182	1
18	281/281	182/182	1	**53**	281/306	172/182	4	**88**	281/281	182/182	1
19	281/306	180/180	4	**54**	281/281	182/182	1	**89**	281/281	176/182	1
20	281/281	176/182	1	**55**	281/281	182/182	1	**90**	281/281	182/182	1
21	281/281	176/182	1	**56**	281/281	172/182	1	**91**	281/306	180/182	4
22	281/281	172/180	3	**57**	281/306	182/182	2	**92**	281/281	182/182	1
23	306/306	182/182	2	**58**	281/306	176/182	4	**93**	281/306	172/182	4
24	281/281	182/182	1	**59**	281/281	182/182	1	**94**	281/281	182/182	1
25	281/281	180/182	3	**60**	281/281	182/182	1	**95**	281/306	180/182	4
26	281/281	182/182	1	**61**	281/281	182/182	1	**96**	281/306	172/182	4
27	281/281	182/182	1	**62**	281/281	182/182	1	**97**	281/306	172/180	4
28	281/281	176/182	1	**63**	281/281	172/182	1	**98**	281/306	180/182	4
29	281/281	182/182	1	**64**	306/306	176/182	4	**99**	306/306	180/182	4
30	281/281	182/182	1	**65**	281/281	182/182	1	**100**	281/306	180/182	4
31	281/281	182/182	1	**66**	281/281	182/182	1	**101**	281/281	172/182	1
32	281/281	182/182	1	**67**	281/306	172/182	4	**102**	281/281	182/182	1
33	281/306	182/182	2	**68**	281/306	172/182	4	**103**	281/281	182/182	1
34	281/281	182/182	1	**69**	281/306	172/172	4				
35	281/281	182/182	1	**70**	281/306	172/182	4				

Table 3.7 gives the results of the marker test for *AMC* for 41 pigs from commercial herds and diagnosed by farmers to be diseased with *AMC* and for their parents (when DNA was available). The grades 1 to 4 in the last column represent the same grades as reported in Table 3.6, grade 5 indicates a high probability for *AMC* (with $bE77^{306/306}$, except for 10023, a recombinant pig (Table 3.2), and with $SW904^{180/180}$ or $SW904^{180/182}$ or $SW904^{172/180}$ or $SW904^{172/176}$) and grade 6 includes with high confidence normal piglets (without $bE77^{306/306}$ and without $SW904^{180/180}$ or $SW904^{172/180}$ or $SW904^{172/176}$). Two offspring (10007 and 10008) could not be assigned a grade (?).

The markers indicated that 34 piglets (82.9%) were diseased (grade 5); while for five of the piglets (12.2%) *bE77* and *SW904* did not confirm the AMC phenotype (grade 6).

For three matings (12, 13 and 14, 21.4%), parentage was incorrect. Piglet 10033 may have been an offspring of mating 12, in contrast to piglet 10034.

In three cases (matings 4, 5 and 6) the parent's DNA was not available.

Of 11 offspring, five of the fathers had a grade of 1 and six a grade of 4. The only tested dam (6408 FJ1) had a grade of 1.

Table 3.7: Genotypes of bE77 and SW904 for 41 piglets from commercial farms as indicators of susceptibility to AMC.

The allele lengths in the microsatellites *bE77* and *SW904* are given in bp. The 41 piglets were diagnosed as having AMC by farmers. The genotypes of the parents (if DNA was available) with the ID of the boars in parentheses (see Table 3.6), are also shown. Grades 1 to 6 explained in the text. ?: not classifiable.

Mating	Member	Animal ID (number in Table 3.6)	*bE77*	*SW904*	Grade
1	Father	4218 ELE (87)	281/281	176/182	1
1	Offspring	10000	281/281	182/182	6
2	Father	4591 KE (95)	281/306	180/182	4
2	Offspring	10001	306/306	180/180	5
2	Offspring	10002	306/306	180/180	5
2	Offspring	10003	306/306	180/180	5
2	Offspring	10004	306/306	180/180	5
2	Offspring	10005	306/306	180/180	5
2	Offspring	10006	281/306	180/182	6
3	Father	4631 PU (96)	281/306	172/182	4
3	Offspring	10007	281/306	172/182	?
3	Offspring	10008	281/306	172/176	?
3	Offspring	10009	281/281	182/182	6
4	Offspring	10010	306/306	172/180	5
4	Offspring	10011	306/306	172/180	5
5	Offspring	10012	306/306	180/180	5
5	Offspring	10013	306/306	180/182	5
5	Offspring	10014	306/306	180/182	5
5	Offspring	10015	306/306	180/180	5
5	Offspring	10016	306/306	180/180	5
6	Offspring	10017	306/306	172/176	5
6	Offspring	10018	306/306	172/176	5
6	Offspring	10019	306/306	172/176	5
7	Father	9695 KPN (100)	306/306	180/182	4
7	Offspring	10020	306/306	180/180	5
8	Father	3367 HT (97)	281/306	172/180	4
8	Offspring	10022	306/306	172/180	5
8	Offspring	10023	281/306	172/180	5
8	Offspring	10024	306/306	172/180	5
8	Offspring	10025	306/306	172/180	5
8	Offspring	10026	306/306	172/180	5
9	Father	9695 KPN (100)	306/306	180/182	4
9	Offspring	10027	306/306	172/180	5
9	Offspring	10028	306/306	172/180	5
10	Father	7667 EBA (101)	281/306	180/182	4
10	Offspring	10029	306/306	180/180	5
10	Offspring	10030	306/306	180/180	5
10	Offspring	10031	306/306	180/180	5
10	Offspring	10041	306/306	180/180	5
11	Father	965 RBA (102)	281/281	172/182	1
11	Mother	6408 FJ1	281/281	172/182	1
11	Offspring	10032	281/281	172/182	6
12	Father	3099 DUC (104)	281/281	182/182	1
12	Offspring	10033	281/306	172/182	6
12	Offspring	10034	306/306	180/180	5
13	Father	8389 PU (73)	281/281	182/182	1
13	Offspring	10035	306/306	180/180	5
13	Offspring	10036	306/306	180/180	5
13	Offspring	10037	306/306	180/180	5
13	Offspring	10038	306/306	180/180	5
14	Father	9975 HZP (80)	281/281	176/182	1
14	Offspring	10039	306/306	180/180	5
14	Offspring	10040	306/306	180/180	5

4. Discussion

4.1 Inheritance of *AMC*

The findings of the breeding experiments provide information, which suggests that the *AMC* gene, responsible for Arthrogryposis in pigs, is inherited in an autosomal recessive manner. Both dominant and sex-linked inheritance can be excluded, because the Mendelian rules for recessive inheritance were confirmed and both male and female piglets were affected, even if males to a greater extent.

4.2 Linkage and QTL analyses

This study is the first to investigate AMC in swine on a genetic level by specifically utilizing linkage analysis and gene mapping. We analyzed the most similar forms of AMC found in humans, the neurogenic *AMC* (*AMCN*, Shohat *et al.* 1997) and the spinal muscular atrophy (*SMA,* Lefebvre *et al.* 1995; Lorson *et al.* 1999), which were also shown to be important candidate genes for cattle AMC (Longeri *et al.* 2003) and sheep AMC (Murphy *et al.* 2004).

Five microsatellite markers (*SWR2516*, *SW1650, SW1026*, *S0370* and *SWR345*) were evaluated on SSC2, which corresponds to HSA5, but we did not find a correlation between these markers and *amc*.

Linkage analysis showed that the mutated gene was located on SSC5, most likely between *SW152* and *SW904*/*SW1094* at position 92 cM.

QTL interval mapping (Haley *et al.* 1994) was applied to the data to determine whether a chromosome scan gives the same results as typical linkage analysis when the 'trait' is strictly inherited as monogenic. A statistically significant QTL was mapped on SSC5 (position 92 cM) in the region of *AMC*. The QTL results confirmed and reinforced the previous linkage results.

The calculated order of the markers compares well to previous linkage maps of swine, despite the fact that the estimated genetic distances and recombination rates were not fully consistent. This can most easily be explained by the examination of different breeds and genetically diverse family material and the high level of inbreeding within the AMC family. In all cases but one, the order of the markers in the map was the same as that in the MARC map (http://www.genome.iastate.edu/maps/marcmap.html). The exception was that, in our map, *SW1094* is located at the same position as *SW904* on SSC5 (Figure 3.1), whereas in the MARC map *SW1094* is located closer to the telomere. However, our results are consistent with the findings of Cirera *et al.* (2003) and Hawken *et al.* (1999).

The microsatellites *SW152* and *SW904/SW1094* are located between the physically mapped microsatellites *SW963* and *SW1200* (http://www.genome.iastate.edu/maps/marcmap.html) and, thus, between bands q12-q23. This region is likely to correspond and shares homology with an evolutionarily conserved region on human chromosome 12, bands p13-p12, as shown by chromosome painting between human and pig (Goureau *et al.* 2001; http://www.toulouse.inra.fr/lgc/pig/compare/SSC.htm). Furthermore, some ESTs on HSA12 and HSA22 were mapped to SSC5 bands q12-q23 (Cirera *et al.* 2003; Meyers *et al.* 2005; http://www.ag.unr.edu/beattie/research/first_generation.htm) and the position of *SW152* was assigned to HSA22 (Robic *et al.* 2003). Consequently, the porcine *AMC* gene has its human counterpart on HSA12p13-p12 and HSA22q11 and human genes located there represent positional candidate genes for *AMC*.

Based on comparative mapping, the genes reported to cause the human forms of Arthrogryposis that are not located on chromosome 12 or 22 are not considered to be the most likely candidates for *AMC* in swine. However, it would be wrong to exclude potentially appropriate genes based solely on comparative mapping, because minor chromosomal rearrangements may occur.

4.3 Microarray and real-time analyses

The small amount of good quality RNA available did not permit further repetitions, and this limits the interpretation of the gene expression experiments. With only one sample it is impossible to conclude that the differences seen are differences that would always occur between the normal and diseased state or if they merely represent the differences between the two samples under analysis. Using a statistical power analysis, based on a moderately variable sample, it has been shown that at least six biological replicates are required for a two-fold difference to be statistically significant. It is also unknown in which tissue the deficient gene for AMC is expressed or if the gene is expressed at birth; therefore, we investigated three different tissues (brain, muscle and spinal cord). Only 14 genes were differentially expressed at the same time in more than one tissue, showing high heterogeneity and a low correlation in the expression pattern of the three tissues, which may be due to the above-mentioned difficulties. The sparse location of the differentially expressed genes on all of the 18 chromosomes, except the sex chromosome, did not indicate significantly important regions, which might be involved in the pathway of the AMC disease. The high number of genes differentially expressed in muscle compared to the other tissues may be caused by the difficulty in extracting comparative muscle RNA due to the small amount of muscle tissue in newborns.

Despite these problems, the above factors played a secondary role in this preliminary microarray study. The main purpose of the microarray analysis was to find functional candidate genes for AMC to map with the radiation hybrid panel, in order to validate their significance also as positional candidates. This process led to the detection of significant new human chromosomal regions to examine for new candidate genes, as demonstrated by the results of gene *Q6ZUQ4* on HSA22.

The TaqMan real-time PCR results confirmed the microarray findings for the three genes *C3F*, *NR4A1* and *Q6ZUQ4*, which were subsequently selected for RH mapping.

The gene *SFRS6*, a splicing factor located on HSA20, was confirmed by real-time to be upregulated in the spinal cord and muscle, but not in the brain of the affected animal. Even though *SFRS6* is not located on HSA12 or HSA22, it is an interesting gene for further investigation, because it is involved in many different pathways. Its upregulation in the diseased piglet may indicate the absence of another component of the pathway, which may be of interest for the AMC pathology.

SLC2A1, a sugar transporter located on HSA1, is upregulated in the brain but not in the muscle of the normal piglet. Even though the microarray analysis was only partially confirmed, this gene furnishes important indications of the involvement of the sugar family in the AMC disease. Hence, gene *SLC2A13*, a member of this family and located in the AMC region, represents a suitable candidate for further studies.

Thus, the real-time PCR analyses of selected genes provide evidence that the microarray results are reliable.

Furthermore, the genes *PCBP2* and *C1S*, located on HSA12, were differentially expressed on the microarray. *PCBP2* was mapped by FISH (Goureau *et al.* 2001) and *C1S* with the RH panel (http://www.ag.unr.edu/ab/standard.htm, EST-AR100E12) to chromosomal regions on SSC5 outside the AMC interval. Therefore, we did not check these results by real-time PCR and we did not map the genes with the RH panel.

4.4 Gene mapping and analysis of BAC clones

The results of the radiation hybrid panel indicate that the porcine genes *CPNE8* and *PRICKLE1*, located in human on HSA12q12 and *Q6ZUQ4*, *TUBA8* and *USP18*, located in human on HSA22q11, map in the expected AMC interval. *CPNE8*, *Q6ZUQ4*, *TUBA8* and *USP18* were mapped towards the centromere (near *SW152*) and *PRICKLE1* towards the telomere of SSC5 (near *SW904*). Hence, of the eight interesting physiological candidate genes for AMC, *CACNA1C*, *COL2A1*, *GDF11*, *HOXC8*, *KCNA1*, *SCN8A*, *TUBA8* and *USP18* (section 1.2.7), only the latter two mapped to the AMC interval. Unfortunately, microsatellite

S0018, located between *SW152* and *SW904* (Rohrer *et al.* 1996), was not polymorphic in the experimental family and, thus, this microsatellite did not deliver more information about the AMC region.

Likewise, the microsatellite *UMNp1275*, reported to be linked with *SW152* (Fahrenkrug *et al.* 2005), was situated outside the *AMC* region and was not evaluated further.

Based on the RH findings, genes mapped outside the interval *SW152* to *SW904* are not considered to be the most likely candidates for swine *AMC*. *SCN8A* has been mapped to HSA12 segment q13 (Plummer *et al.* 1998) and in chicken to the linkage group E22C19W28 (Gao *et al.* 2005). Therefore, comparison of these three genomes shows an evolutionary conserved segment between SSC5p11-12, chicken linkage group *E22C19W28* and human chromosome 12q13. Although *SCN8A* is a plausible candidate gene for the fray mutation in chicken and a mutation in human *SCN8A* causes ataxia and paralysis, this gene does not cause AMC in pig because it is located on the other chromosomal arm of SSC5.

However, based solely on the results of mapping, it would be wrong to exclude potentially appropriate genes, which map very closely to *SW152* (such *CACNA1C*), because small mapping errors or resolution of the RH panel may affect the results. For these reasons, the position of *CACNA1C* was confirmed by determining its location in the BAC clones ordered in contig ctg5002 (http://www.sanger.ac.uk/Projects/S_scrofa/).

The assignments of the 18 porcine genes to SSC5 were in line with those published recently (Cirera *et al.* 2003; Farber *et al.* 2003; Lahbib-Mansais *et al.* 2003; Meyers *et al.* 2005; Tuggle *et al.* 2003), and our mapping results are in agreement with the comparative map between pig and human. Furthermore, our comparative findings combined with seven genes mapped at http://www.ag.unr.edu/ab/standard.htm and by Tuggle *et al.* (2003), enabled us to predict a total of six new evolutionary conserved blocks between HSA12 and SSC5 and one between HSA22 and SSC5. Our RH panel results suggest that the most likely order of the genes in pig in block A (from *SLC6A12* to *PPFIBP1*), block E (*MGC5576* and *COL2A1*) and block F (*GDF11* to *BIN2*) is the reverse of that in humans. The two blocks containing a single gene, B (*KCNA1*) and C (*DDX11*), were mapped in the A block. The position of *KCNA1* in that block was clearly determined based on these results. In contrast, it was impossible to predict with certainty the location of *DDX11*, because, to the best of my knowledge, no other human genes close to this one have been mapped in pig. Furthermore, the RH findings revealed that one interval, spanning less than one million base pairs between *Q6ZUQ4* and *USP18* on HSA22q11 and containing the human counterpart of microsatellite *SW152*, is located between HSA12p13 and HSA12q12 segments. The human–pig comparative maps

resulted from the chromosomal assignments of the RH panel (Figure 3.8) and from the end-sequenced BAC clones in the Sanger contig 5002 (Figure 3.5) were very similar and compared very well. In summary, the similar results of both approaches provide a reliable basis for further studies.

In general, the differences in the gene order between HSA12, HSA22 and SSC5 showed broad genomic remodelling, confirming and extending previous observations that, despite extensive conservation of synteny between HSA12 and SSC5, considerable small intra- and interchromosomal rearrangements have occurred during mammalian evolution.

4.5 Sequencing and *AMC* position

A SNP in complete linkage disequilibrium between normal and diseased piglets in our research family was discovered in intron 7 of gene *CNTN1*. Unfortunately, due to the lack of recombinant pigs and to its low informative content (one of the two founder boars was not polymorphic), this mutation did not provide new information about the location of *AMC* in the interval and we did not develop a genetic test based on this marker.

Three SNPs were also discovered in the gene sequence of porcine *TUBA8*. The mutations are silent and probably do not cause AMC. A comparison of the sequence of the normal pigs and published sequences of pig ESTs confirmed the similarity. A difference was found when normal pig sequences were compared to human sequences at the second and third SNPs. For the latter two SNPs the human sequence was the same as for the diseased pigs. This clearly implies that they do not cause AMC. In our experimental family it was possible to use these SNPs as genetic markers to determine and distinguish heterozygous carriers from genetically healthy swine without recombinants and, therefore, errors. However, pigs, which do not belong to our experimental herd, must be analyzed to test the validity and linkage disequilibrium in commercial populations of swine. First experiments were conducted and the results confirmed the proximity of *AMC* to the *TUBA8* gene and revealed that an affected pig was recombinant. This recombinant was decisive for excluding the *AMC* region from the chromosomal interval *SW152* to *TUBA8* on HSA22.

The five genes mapped with the RH panel in the *AMC* interval, as well as the recombinant piglets, represent an important starting point for searching additional candidate genes and polymorphisms related to *AMC*.

Thus, the porcine chromosomal intervals corresponding to *TUBA8-USP18* on HSA22 (from 16,968,113 to 17,034,714 bp) and *CPNE8-PRICKLE1* on HSA12 (from 37,332,259 to 41,269,745 bp) should be analyzed carefully for genetic diversity, which may be associated with the disease, both as an additional marker for fine mapping and association studies and as

the causative mutation. No further genes are known in the chromosomal region on HSA22 between *TUBA8* and *USP18*. However, this region, which is known as DiGeorge syndrome critical region (DGCR), is particularly interesting because it contains one of the most important microdeletions (about 2 Mb) of the human genome with an incidence of 1 in every 4000 live births (Lindsay 2001). In addition, Castro-Gago *et al.* (2005) recently described for the first time an association between Arthrogryposis and this microdeletion. Hence, both these genes and the surrounding region will not be excluded from the analysis and will be investigated further.

In the interval between *CPNE8* and *PRICKLE1* on HSA12, 12 human genes are known (from centromere to telomere): *KIF21A*, *ABCD2*, the hypothetical protein *FLJ40126*, *SLC2A13*, *LRRK2*, *MUC19*, *CNTN1*, *PDZRN4*, *Q8IXV1*, *YAF2*, *MADP-1* protein and *PPHLN1*.

Human proximal 12q deletion is a very rare chromosomal abnormality. Heterozygous deletions of genes from *CPNE8* to *P11* in one patient (Gallego *et al.* 2000) and from *PDZRN4* to *ASB8* in another patient (Tonoki *et al.* 1998) have been reported. Miyake *et al.* (2004) compared both patients and neither showed typical symptoms of AMC, with the exception of a mild lower limb hypertonia in patient 1 and a mild scoliosis, brachygnathia and moderate hypotonia in patient 2.

As well as the analysis of the 18 gene sequences mapped with the RH panel, further sequencing of 14 partial gene regions in *TUBA8*, *CPNE8*, *KIF21A*, *ABCD2*, *SLC2A13*, *CNTN1* and *YAF2*, the physiologically most interesting genes in regard to AMC, enabled the identification of new pig orthologues of human genes, identical in normal and diseased pigs, and enabled us to exclude a major chromosomal deletion. Based on the number of recombinants found, the *AMC* locus seems to be much closer to *TUBA8* than to *SW904*, providing evidence that the genes *USP18*, *CPNE8*, *ABCD2*, *SLC2A13* (supported also by the microarray results) and *CNTN1* are the "hottest" *AMC* candidates.

4.6 Marker test for *AMC*

The microsatellites *bE77* and *SW904* are excellent markers for a diagnostic test to reduce the incidence of AMC in the Swiss LW population.

The combination of two markers was necessary to confer a sufficient degree of reliability, because some recombinant pigs were found in preliminary analyses of randomly chosen families.

The testing of *bE77* was preferred to the sequencing of the SNPs in *TUBA8*, even though *TUBA8* is closer to the *AMC* locus. The reason was that, with the available facilities, microsatellite analysis was easier and cheaper to perform. Because of the very close proximity

of *bE77* to *TUBA8*, it was felt that this choice would not influence the results or significantly increase the detection of wrong positive or negative animals.

The test was applied to commercial populations and the same haplotype associated with *AMC* in our experimental herd was detected in 82.9% of the diseased piglets and in 21.2% of the AI boars, confirming previous assumptions (Hofer 2003), that about 20% AI LW boars in Switzerland are carriers of the defected allele.

SW904 showed high heterozygosity, while *bE77* had only two alleles. The high variability of *SW904* made it difficult to interpret the results, because the same genotypes were present in diseased and normal piglets (e.g. $SW904^{180/182}$). On the other hand, greater linkage disequilibrium between the alleles of *AMC* and *bE77* was found. Consequently, the genotypes of *bE77* should be the parameter of choice for predicting susceptibility to AMC.

Surprisingly, a significantly high number of incorrect pedigrees (21.4%) were detected in the families of the commercial herds, where diseased piglets were recorded, compared to a recent study (Keller 2005). In effect, the total number of false parentages in Swiss elite breeding herds reported by Keller (2005) in 2003, 2004 and 2005 was estimated to be only 5.7%.

It is obvious that incorrect phenotypes and pedigrees combined with unreported cases of AMC represent a crucial problem in finding the common ancestors of the disease, in quantifying the exact number of AI AMC carriers and in trying to perform linkage disequilibrium studies in other commercial swine populations. Accurate and precise pedigree records and complete family material would improve the efficiency and reliability of the test, the differentiation between similar haplotypes and, finally, help eliminate the disease.

5. Conclusions and perspectives

The present study aimed at identifying the causative mutation for AMC in swine according to a precise diagram (Neuenschwander 2001), which permits the detection of an undesired allele in a family with the help of genetic methods (see introduction). First, the phenotypic appearance and the unambiguous identification of a phenotype with an inherited component for AMC in pig were assessed. The segregation pattern of the disease was studied in a newly formed experimental family, where informative parents were mated. This was tedious because the majority of AMC carriers were not as fit and finding fertile boars was very difficult.

After discarding the most plausible human, cattle and sheep counterparts as candidate genes for pig *AMC*, a full genome scan revealed that the locus of the disease is on SSC5.

Saturating the chromosomal region of interest with informative markers was done at the same time as the search for positional and functional candidate genes using comparative gene maps and microarray technology.

Several human genes in the homologous regions on HSA12 and HSA22, with altered gene products, similar to the observed symptoms in pig, were selected and mapped to SSC5.

Sequencing of genomic DNA of candidate genes led to new intragenous markers (microsatellites and SNPs), closely linked with the phenotype and the discovery of recombinants, which is important for the precise localization of the disease. Although the causative mutation for AMC was not found, a marker test, independent of age, sex and environment, was developed and used to type pigs on a large scale. However, this marker is not fully linked with the defective allele; it is a predictor for selection in the Swiss LW population and may make it possible to reduce the incidence of carrier pigs and pigs with AMC as well as the resulting economic losses.

2001 was marked by outstanding and very important achievements in genome research, especially the complete draft sequence of the human genome (Genome International Sequencing Consortium 2001). In the following four years, several genomes, such as chimpanzee, mouse, rat, dog, cow and chicken, were entirely sequenced and a large body of data became available.

A consortium (Schook *et al.* 2005) was recently established with the aim of finding financial support to sequence the entire pig genome, which, of the most important farmed animal species, is still missing. The sequencing project has begun and a draft sequence should be completed by the end of 2007. In terms of chromosomal homologies, the very important contribution made by this study expands and refines the comparative map between human and swine on SSC5 and defines new evolutionary breakpoints between both species.

Up to now, much effort was spent in locating genes in pig and developing primers for cross-species PCR. The search for causative mutations is complicated and takes a great deal of time. After the clarification of the complete sequence, emphasis will shift to the characterization of gene expression and protein function. The recent development of improved tools for gene expression studies (RNA), such as tissue, developmental stage and chromosome-specific microarrays, the discovery of new important regulatory sequences, such as siRNA and miRNA, and the analysis of expressed sequences at the protein level (2D-gel electrophoresis, mass spectrometry and amino acid analysis) will dramatically accelerate the elucidation of the genetic background of simple hereditary disorders and complex traits.

The most important focus of this study in the future should be the screening of mutations in candidate genes, which are located in the *AMC* region, the extraction of RNA samples for q-PCR in different tissues and at different developmental stages of the embryo and the examination of knockout animals lacking the most promising genes.

Delineating the genetic factors responsible for AMC will enable the identification of key proteins and enzymes leading to the disease; their successful isolation may be the basis for designing new supplemental substances, which may prevent diseases such as AMC. The swine represents a particularly attractive animal model for studying the common genetic forms of human AMC due to the phenotypic similarity of the disease in humans and in pigs. Moreover, this study may represent an important roadmap for future research in humans and in other species and may trigger the identification and analysis of AMC forms, which are still not understood.

6. References

Adams R.D., Denuy-Brown D. & Pearson C.M. (1962) Diseases of Muscle: A Study in Pathology, 2nd ed. New York: Harper & Row.

Agerholm J.S., Bendixen C., Andersen O. & Arnbjerg J. (2001) Complex vertebral malformation in Holstein calves. *Journal of Veterinary Diagnostic Investigation* **13**, 283-9.

Anderson S.I., Lopez-Corrales N.L., Gorick B. & Archibald A.L. (2000) A large-fragment porcine genomic library resource in a BAC vector. *Mammalian Genome* **11**, 811-4.

Anderson T. (1997) Earliest evidence for arthrogryposis multiplex congenita or Larsen syndrome? *American Journal of Medical Genetics* **71**, 127-9.

Andersson L. & Georges M. (2004) Domestic animal genomics: deciphering the genetics of complex traits. *Nature Reviews Genetics* **5**, 202–12.

Ashcroft F.M. (2000) Ion channels and disease. Academic Press, San Diego, CA, USA.

Bahr C., Kuiper H., Peters M., Scholz H. & Distl O. (2004) Arthrogryposis associated with facial scoliosis and torticollis in German Holstein calves. *Deutsche Tierarztliche Wochenschrift* **111**, 403-7.

Banker B.Q. (1977) Muscular dysgenesis in the mouse (mdg/mdg). I. Ultrastructural study of skeletal and cardiac muscle. *Journal of Neuropathology and Experimental Neurology* **36**, 100-27.

Beattie C.W. (1994) Livestock genome maps. *Trends in Genetics* **10**, 334-8.

Berglund E.O., Murai K.K., Fredette B., Sekerkova G., Marturano B., Weber L., Mugnaini E. & Ranscht B. (1999) Ataxia and abnormal cerebellar microorganization in mice with ablated contactin gene expression. *Neuron* **24**, 739-50.

Blazej R.G., Mellersh C.S., Cork L.C. & Ostander E.A. (1998) Hereditary canine spinal muscular atrophy is phenotypically similar but molecularly distinct from human spinal muscular atrophy. *Journal of Heredity* **89**, 531–7.

Boehnke M. (1994) Limits of resolution of genetic linkage studies: implications for the positional cloning of human genetic diseases. *American Journal of Human Genetics* **55**, 379–90.

Bolt R. (1992) Diagnose der Halothanempfindlichkeit des Schweines durch die Identifizierung von Polymorphismen im Gen eines Kalziumfreisetzungskanals mit molekulargenetischen Methoden. Dissertation Nr. 9846, ETH Zurich.

Bradley G.A., Metcalf H.C., Reggiardo C., Noon T.H., Bicknell E.J., Lozano-Alarcon F., Reed R.E. & Riggs M.W. (1995) Neuroaxonal degeneration in sheep grazing Sorghum pastures. *Journal of Veterinary Diagnostic Investigation* **7**, 229-36.

Brookes A.J. (1999) The essence of SNPs. *Gene* **234**, 177-86.

Browne D.L., Gancher S.T., Nutt J.G., Brunt E.R., Smith E.A., Kramer P. & Litt M. (1994) Episodic ataxia/myokymia syndrome is associated with point mutations in the human potassium channel gene, KCNA1. *Nature Genetics* **8**, 136-40.

Burt D.W. (2002) Comparative mapping in farm animals. *Briefings in Functional Genomics and Proteomics* **1**, 159-68.

Castro-Gago M., Iglesias-Meleiro J.M., Blanco-Barca M.O., Grande-Seijo M., Barros-Angueira F. & Eiris-Punal J. (2005) Neurogenic arthrogryposis multiplex congenita and velopharyngeal incompetence associated with chromosome 22q11.2 deletion. *Journal of Child Neurology* **20**, 76-8.

Celesia G.G. (2001) Disorders of membrane channels or channelopathies. *Clinical Neurophysiology* **112**, 2-18.

Chavez L.S., Serda R., Choe S., Davidi L., Harmeyer J. & Omdahl J.L. (2003) Molecular basis for pseudo vitamin D-deficiency rickets in the Hannover pig. *The Journal of nutritional biochemistry* **14**, 378-85.

Cirera S., Jorgensen C.B., Sawera M., Raudsepp T., Chowdhary B.P. & Fredholm M. (2003) Comparative mapping in the pig: localization of 214 expressed sequence tags. *Mammalian Genome* **14**, 405-26.

Crowe M.W. (1978) Tobacco: a cause of congenital arthrogryposis. In Effects of Poisonous Plants on Livestock. New York: Academic Press, p. 419.

Curran M.E., Landes G.M. & Keating M.T. (1992) Molecular cloning, characterization, and genomic localization of a human potassium channel gene. *Genomics* **12**, 729-37.

Dawson E. (1999) New collaborations make pharmacogenomics a SNP. *Molecular Medicine Today* **5**, 280.

Dekkers J.C.M. (2004) Commercial application of marker- and gene-assisted selection in livestock: Strategies and lessons. *Journal of Animal Science* **82**, 313-28.

Dirksen G., Doll K., Hafner A., Hermanns W. & Dahme E. (1992) Spinal muscle atrophy in Brown Swiss x Braunvieh cross calves. *Deutsche tierarztliche Wochenschrift* **99**, 168-75.

Doherty M.L., Kelly E.P., Healy A.M., Callanan J.J., Crosby T.F., Skelly C. & Boland M.P. (2000) Congenital arthrogryposis: an inherited limb deformity in pedigree Suffolk lambs. *Veterinary Record* **24**, 748–53.

Duchesne A. & Eggen A. (2005) Radiation hybrid mapping of genes and newly identified microsatellites in candidate regions for bovine arthrogryposis-palatoschisis and progressive ataxia based on comparative data from man, mouse and rat. *Journal of Animal Breeding and Genetics* **122**, 28-35.

Dyson D.A. & Wrathall A.E. (1977) Congenital deformities in pigs possibly associated with exposure to hemlock (Conium maculatum). *Veterinary Record* **100**, 241-2.

Edwards J.F., Karabatsos N., Collisson E.W. & de la Concha Bermejillo A. (1997) Ovine fetal malformations induced by in utero inoculation with Main Drain, San Angelo, and LaCrosse viruses. *The American Journal of tropical Medicine and Hygiene* **56**, 171-6.

Edwards M.J. (1971) The experimental production of arthrogryposis multiplex congenita in guinea-pigs by maternal hyperthermia during gestation. *Journal of Pathology* **104**, 221-9.

Ely R. & Leipold H.W. (1979) Arthrogryposis in Yorkshire piglets. *Zeitschrift für Tierzüchtung und Züchtungsbiologie* **96**, 241-52.

Eyre-Walker A. & Keightley P.D. (1999) High genomic deleterious mutation rates in hominids. *Nature* **397**, 344–7.

Fahrenkrug S., Wagner M., Morrison L. & Alexander L.J. (2005) Map assignments of 373 previously unreported porcine microsatellites. *Animal Genetics* **36**, 76-86.

Farber C.R., Raney N.E., Rilington V.D., Venta P.J. & Ernst C.W. (2003) Comparative mapping of genes flanking the human chromosome 12 evolutionary breakpoint in the pig. *Cytogenetic and Genome Research* **102**, 139-44.

Forsyth C.S., Speth R.C., Wecker L., Galey F.D. & Frank A.A. (1996) Comparison of nicotinic receptor binding and biotransformation of coniine in the rat and chick. *Toxicology Letters* **89**, 175-83.

Freeman W.M., Walker S.J. & Vrana K. (1999) Quantitative RT-PCR: pitfalls and potentials. *BioTechniques* **26**, 112-25.

Gaiser K.G., Maddox B.K., Bann J.G., Boswell B.A., Keene D.R., Garofalo S. & Horton W.A. (2002) Y-position collagen II mutation disrupts cartilage formation and skeletal development in a transgenic mouse model of spondyloepiphyseal dysplasia. *Journal of Bone and Mineral Research* **17**, 39-47.

Gallego M., Barreiro C., Perez M., Arroyo H. & Menehem J. (2000) A new case of interstitial 12q deletion. *International Pediatrics* **15**, 37–40.

Gao Y., Hu X.X., Deng X.M., Feng J.D. & Li N. (2005) Linkage mapping of the SCN8A gene to chicken linkage group E22C19W28. *Animal Genetics* **36**, 284-84.

Garofalo S., Vuorio E., Metsaranta M., Rosati R., Toman D., Vaughan J., Lozano G., Mayne R., Ellard J., Horton W. & de Crombrugghe B. (1991) Reduced amounts of cartilage collagen fibrils and growth plate anomalies in transgenic mice harboring a glycine-to-cysteine mutation in the mouse type II procollagen alpha-1-chain gene. *Proceedings of the National Academy of Sciences* **88**, 9648-52.

Geldermann H. (1975) Investigations on inheritance of quantitative characters in animals by gene markers. I. Methods. *Theoretical and Applied Genetics* **46**, 319–30.

Genome Internation Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921.

Genoud S., Lappe-Siefke C., Goebbels S., Radtke F., Aguet M., Scherer S.S., Suter U., Nave K.A. & Mantei N. (2002) Notch1 control of oligodendrocyte differentiation in the spinal cord. *Journal of Cell Biology* **158**, 709–18.

Goureau A., Garrigues A., Tosser-Klopp G., Lahbib-Mansais Y., Chardon P. & Yerle M. (2001) Conserved synteny and gene order difference between human chromosome 12 and pig chromosome 5. *Cytogenetics and Cell Genetics* **94**, 49-54.

Green P., Falls K. & Crooks S. (1990) Documentation for CRIMAP, version 2.4., Washington University, School of Medicine, St. Louis, MO.

Grisart B., Coppieters W., Farnir F., Karim L., Ford C., Berzi P., Cambisano N., Mni M., Reid S., Simon P., Spelman R., Georges M. & Snell R. (2002) Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. *Genome Research* **12**, 222–31.

Grunwald K.A.A., Schueler K., Uelmen P.J., Lipton B.A., Kaiser M., Buhman K. & Attie A.D. (1999) Identification of a novel Arg -> Cys mutation in the LDL receptor that contributes to spontaneous hypercholesterolemia in pigs. *Journal of Lipid Research* **40**, 475-85.

Haley C.S., Knott S.A. & Elsen J.M. (1994) Mapping quantitative trait loci in crosses between outbred lines using least squares. *Genetics* **136**, 1195-207.

Hasan L., Vögeli P., Stoll P., Kramer S.S., Stranzinger G. & Neuenschwander S. (2004) Intragenic deletion in the gene encoding L-gulonolactone oxidase causes vitamin C deficiency in pigs. *Mammalian Genome* **15**, 323-33.

Hasler-Rapacz J., Ellegren H., Fridolfsson A.K., Kirkpatrick B., Kirk S., Andersson L. & Rapacz J. (1998) Identification of a mutation in the low density lipoprotein receptor gene associated with recessive familial hypercholesterolaemia in swine. *American Journal of Medical Genetics* **76**, 379-86.

Hawken R.J., Murtaugh J., Flickinger G.H., Yerle M., Robic A., Milan D., Gellin J., Beattie C.W., Schook L.B. & Alexander L.J. (1999) A first-generation porcine whole-genome radiation hybrid map. *Mammalian Genome* **10**, 824-30.

Hayes H., Elduque C., Gautier M., Schibler L., Cribiu E. & Eggen A. (2003) Mapping of 195 genes in cattle and updated comparative map with man, mouse, rat and pig. *Cytogenetic and Genome Research* **102**, 16–24.

Hegasy G.A., Manuelian T., Hogasen K., Jansen J.H. & Zipfel P.F. (2002) The molecular basis for hereditary porcine membranoproliferative glomerulonephritis type II – Point mutations in the factor H coding sequence block protein secretion. *American Journal of Pathology* **161**, 2027-34.

Heid C.A., Stevens J., Livak K.J. & Williams P.M. (1996) Real time quantitative PCR. *Genome Research* **6**, 986-94.

Hofer A. (2003) Arthrogrypose und Bekämpfungsstrategie. *Suisseporcs* **3**, 6-8.

Jacobson L., Polizzi A., Morriss-Kay G. & Vincent A. (1999) Plasma from human mothers of fetuses with severe arthrogryposis multiplex congenita causes deformities in mice. *Journal of Clinical Investigation* **103**, 1031-8.

Jones J.C. (1999) Perosomus Elumbis (Vertebral Agenesis and Arthrogryposis) in a Stillborn Holstein Calf. *Veterinary Pathology* **36**, 64–70.

Kazarinova-Noyes K., Malhotra J.D., McEwen D.P., Mattei L.N., Berglund E.O., Ranscht B., Levinson S.R., Schachner M., Shrager P., Isom L.L. & Xiao Z.C. (2001) Contactin Associates with Na+ Channels and Increases Their Functional Expression. *The Journal of Neuroscience* **21**, 7517-25.

Kearney J.A., Buchner D.A., De Haan G., Adamska M., Levin S.I., Furay A.R., Albin R.L., Jones J.M., Montal M., Stevens M.J., Sprunger L.K. & Meisler M.H. (2002) Molecular and pathological effects of a modifier gene on deficiency of the sodium channel Scn8a (Na(v)1.6). *Human Molecular Genetics* **11**, 2765-75.

Keller S. (2005) Nachkommenschaftskontrolle mittels Mikrosatelliten beim Schwein. Maturitätsarbeit, ETHZ-Kantonsschule Pfäffikon, Switzerland.

Keeler R.F. (1981) Absence of arthrogryposis in newborn Hampshire pigs from sows ingesting toxic levels of jimsonweed during gestation. *Veterinary and Human Toxicology* **23**, 413-5.

Kijas J.M.H., Wales R., Tornsten A., Chardon P., Moller M. & Andersson L. (1998) Melanocortin receptor 1 (MC1R) mutations and coat color in pigs. *Genetics* **150**, 1177-85.

Kim K.S. & Rothschild M.F. (2001) Mapping of the porcine agouti-related protein (AGRP) gene to chromosome 6. *Animal Genetics* **32**, 325.

Kitani H., Yamakawa M. & Ikeda H. (2000) Preferential infection of neuronal and astroglia cells by Akabane virus in primary cultures of fetal bovine brain. *Veterinary Microbiology* **73**, 269-79.

Kitano Y., Yasuda N., Shimizu T., Ohzono H. & Iwamoto T. (1997) Teratogenicity of Aino virus in the chick embryo. *Research in Veterinary Science* **62**, 195-8.

Klein H. & Herzog A. (1990) Arthromyodysplasia congenita hereditaria in calves of the Hessian cattle population. *Deutsche Tierarztliche Wochenschrift* **97**, 117-22.

Kondrashov A.S. (1995) Contamination of the genome by very slightly deleterious mutations: why have we not died 100 times over? Journal of Theoretical Biology 175, 583–94.

Körkkö J., Annunen S., Pihlajamaa T., Prockop D.J. & Ala-Kokko L. (1998) Conformation sensitive gel electrophoresis for simple and accurate detection of mutations: comparison with denaturing gradient gel electrophoresis and nucleotide sequencing. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 1681-5.

Kratzsch A., Stricker C., Gmür C., Rieder S., Jörg H., Ossent P., Bürgi E., Zimmermann W., Stranzinger G. & Vögeli P. (1999) Congenital progressive ataxia and spastic paresis, a hereditary disease in swine, maps to chromosome 3 by linkage analysis. *Mammalian Genome* **10**, 1036-38.

Lahbib-Mansais Y., Tosser-Klopp G., Leroux S., Cabau C., Karsenty E., Milan D., Barillot E., Yerle M., Hatey F. & Gellin J. (2003) Contribution to high-resolution mapping with 101

type I markers and progress in comparative map between humans and pigs. *Mammalian Genome* **14**, 275-88.

Laird P.W., Zijderveld A., Linders K., Rudnicki M.A., Jaenisch R. & Berns A. (1991) Simplified mammalian DNA isolation procedure. *Nucleic Acids Research* **19**, 4293.

Lamb K.J., Lewthwaite J.C., Lin J.P., Simon D., Kavanagh E., Wheeler-Jones C.P. & Pitsillides A.A. (2003) Diverse range of fixed positional deformities and bone growth restraint provoked by flaccid paralysis in embryonic chicks. *International Journal of Experimental Pathology* **84**, 191-9.

Lefebvre S., Burglen L., Reboullet S., Clermont O., Burlet P., Viollet L., Benichou B., Cruaud C., Millasseau P., Zeviani M., LePaslier D., Frezal F., Cohen D., Weissenbach J., Munnich A. & Melki J. (1995) Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* **80**, 155–65.

Lehmann-Horn F. & Jurkat-Rott K. (1999) Voltage-gated ion channels and hereditary disease. *Physiological Reviews* **79**, 1317–72.

Leipold H.W., Oehme F.W. & Cook J.E. (1973) Congenital arthrogryposis associated with ingestion of jimsonweed by pregnant sows. *Journal of the American Veterinary Medical Association* **162**, 1059-60.

Lindsay E.A. (2001) Chromosomal microdeletions: dissecting *Del22Q11* syndrome. *Nature Reviews Genetics* **2**, 858-68.

Lomo O.M. (1985) Arthrogryposis and associated defects in pigs: indication of simple recessive inheritance. *Acta Veterinaria Scandinavica* **26**, 419-22.

Longeri M., Perrone T., Bongioni G., Bona M., Zanotti M. & Galli A. (2003) Survival motor neuron (SMN) polymorphism in relation to congenital arthrogryposis in two Piedmont calves (piemontese). *Genetics Selection Evolution* **35**, 167-75.

Lopez T.A., Cid M.S. & Bianchini M.L. (1999) Biochemistry of hemlock (Conium maculatum L.) alkaloids and their acute and chronic toxicity in livestock. A review. *Toxicon* **37**, 841-65.

Lorson C.L., Hahnen E., Androphy E.J. & Wirth B. (1999) A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proceedings of the National Academy of Sciences* **96**, 6307–11.

Marklund S., Kijas J., Rodriguez-Martinez H., Rönnstrand L., Funa K., Moller M., Lange D., Edfors-Lilja I. & Andersson L. (1998) Molecular basis for the dominant white phenotype in the domestic pig. *Genome Research* **8**, 826-33.

Marra M.A., Kucaba T.A., Dietrich N.L., Green E.D., Brownstein B., Wilson R.K., McDonald K.M., Hillier L.W., McPherson J.D. & Waterston R.H. (1997) High-throughput fingerprint analysis of large-insert clones. *Genome Research* **7**, 1072–84.

Matthews I., Sims G., Ledwidge S., Stott D., Beeson D., Willcox N. & Vincent A. (2002) Antibodies to acetylcholine receptor in parous women with myasthenia: evidence for immunization by fetal antigen. *Laboratory Investigation* **82**, 1407-17.

Mayhew I.G. (1984) Neuromuscular arthrogryposis multiplex congenita in a thoroughbred foal. *Veterinary Pathology* **21**, 187-92.

McPherron A.C., Lawler A.M. & Lee S.J. (1999) Regulation of anterior/posterior patterning of the axial skeleton by growth/differentiation factor 11. *Nature Genetics* **22**, 260-4.

Meijerink E., Neuenschwander S., Fries R., Dinter A., Bertschinger H.U., Stranzinger G. & Vögeli P. (2000) A DNA polymorphism influencing α(1,2) fucosyltransferase activity of the pig FUT1 enzyme determines susceptibility of small intestinal epithelium to *Escherichia coli* F18 adhesion. *Immunogenetics* **52**, 129-36.

Meyers S.N., Rogatcheva M.B., Larkin D.M., Yerle M., Milan D., Hawken R.J., Schook L.B. & Beever J.E. (2005) Piggy-BACing the human genome II. A high-resolution, physically anchored, comparative map of the porcine autosomes. *Genomics* **86**, 739-52.

Milan D., Jeon J.T., Looft C., Amarger V., Robic A., Thelander M., Rogel-Gaillard C., Paul S., Iannuccelli N., Rask L., Ronne H., Lundstrom K., Reinsch N., Gellin J., Kalm E., Le Roy P., Chardon P. & Andersson L. (2000 A) A mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle. *Science* **288**, 1248-51.

Milan D., Hawken R., Cabau C., Leroux S., Genet C., Lahbib Y., Tosser G., Robic A., Hatey F., Alexander L., Beattie C., Schook L., Yerle M. & Gellin J. (2000 B) IMpRH server: an RH mapping server available on the Web. *Bioinformatics* **6**, 558-9.

Miyake N., Tonoki H., Gallego M., Harada N., Shimokawa O., Yoshiura K., Ohta T., Kishino T., Niikawa N. & Matsumoto N. (2004) Phenotype–genotype correlation in two patients with 12q proximal deletion. *Journal of Human Genetics* **49**, 282-4.

Moller M.J., Chaudhary R., Hellmen E., Hoyheim B., Chowdhary B. & Andersson L. (1996) Pigs with the dominant white coat color phenotype carry a duplication of the kit gene encoding the mast/stem cell growth factor receptor. *Mammalian Genome* **7**, 822-30.

Murphy A.M., MacHugh D.E., Boland M.P. & Doherty M.L. (2004) Linkage mapping and identification of the causative gene for congenital ovine arthrogryposis. 29th International Conference on Animal Genetics (Poster #D052), Tokyo, Japan.

Nes N., Lomo O.M. & Bejerkas I. (1982) Hereditary lethal arthrogryposis (muscle contracture) in horse. *Nordisk Veterinaer Medicin* **34**, 425–30.

Neuenschwander S. (2001) Structural and functional genomics in farm animals: a laboratory view point. Habilitationsschrift, ETH Zurich, Switzerland.

Newton C.R. & Graham A. (1994) PCR. Spektrum Akademischer Verlag, Heidelberg, D.

Panter K.E., Bunch T.D., Keeler R.F. & Sisson D.V. (1988) Radio ultrasound observations of the fetotoxic effects in sheep from ingestion of Conium maculatum (poison-hemlock). Journal of toxicology. *Clinical toxicology* **26**, 175-87.

Panter K.E., Bunch T.D., Keeler R.F., Sisson D.V. & Callan R.J. (1990) Multiple congenital contractures (MCC) and cleft palate induced in goats by ingestion of piperidine alkaloid-containing plants: reduction in fetal movement as the probable cause. *Journal of toxicology. Clinical toxicology* **28**, 69-83.

Plummer N.W., Galt J., Jones J.M., Burgess D.L., Sprunger L.K., Kohrman D.C. & Meisler M.H. (1998) Exon organization, coding sequence, physical mapping, and polymorphic intragenic markers for the human neuronal sodium channel gene SCN8A. *Genomics* **54**, 287-96.

Polizzi A., Huson S.M. & Vincent A. (2000) Teratogen update: maternal myasthenia gravis as a cause of congenital arthrogryposis. *Teratology* **62**, 332-41.

Pumarola M., Anor S., Majo N., Borras D. & Ferrer I. (1997) Spinal Muscular Atrophy in Holstein-Friesian calves. *Acta neuropathologica* **93**, 178–83.

Rettenberger G., Klett C., Zechner U., Kunz J., Vogel W. & Hameister H. (1995) Visualisation of the conservation of synteny between humans and pigs by heterologous chromosomal painting. *Genomics* **26**, 372-8.

Robic A., Faraut T., Iannuccelli N., Lahbib-Mansais Y., Cantegrel V., Alexander L. & Milan D. (2003) A new contribution to the integration of human and porcine genome maps: 623 new points of homology. *Cytogenetic and Genome Research* **102**, 100-8.

Rohrer G.A., Alexander L.J., Hu Z., Smith T.P., Keele J.W. & Beattie C. (1996) A comprehensive map of the porcine genome. *Genome Research* **6**, 371-91.

Rohrer G.A., Freking B.A. & Nonneman D. (2005) Single nucleotide polymorphisms for pig identification and parentage exclusion. Plant & Animal Genomes XIII Conference (Poster #203), San Diego, CA, USA.

Rozen S. & Skaletsky H.J. (2000) Primer3 on the WWW for general users and for biologist programmers. In Krawetz S. & Misener S., ed. Bioinformatics Methods and Protocols: Methods in Molecular Biology. Totowa N.J., Humana Press, 365-86.

Sanger F., Nicklen S. & Coulson A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463-7.

Schena M., Shalon D., Davis R.W. & Brown P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467-70.

Schild A.L., Soares M.P., Damé M.C., Portianski E.L. & Riet-Correa F. (2003) Arthrogryposis in Murrah buffaloes in southern Brazil. *Pesquisa Veterinária Brasileira* **23**, 13-6.

Schook L.B., Beever J.E., Rogers J., Humphray S., Archibald A., Chardon P., Milan D., Rohrer G. & Eversole K. (2005) Swine Genome Sequencing Consortium (SGSC): a strategic roadmap for sequencing the pig genome. *Comparative and Functional Genomics* **6**, 251-5.

Schwer H., Liu L.Q., Zhou L., Little M.T., Pan Z., Hetherington C.J., Zhang D.E. (2000) Cloning and characterization of a novel human ubiquitin-specific protease, a homologue of murine UBP43 (Usp18). *Genomics* **65**, 44-52.

Seaton G., Haley C.S., Knott S.A., Kearsey M. & Visscher P.M. (2002) QTL express: mapping quantitative trait loci in simple and complex pedigrees. *Bioinformatics* **18**, 339–40.

Shohat M., Lotan R., Magal N., Shohat T., Fischel-Ghodsian N., Rotter J.I. & Jaber L. (1997) A gene for arthrogryposis multiplex congenita neuropathic type is linked to D55394 on chromosome 5qter. *American Journal of Human Genetics* **61**, 1139-43.

Soderlund C., Humphray S., Dunham A. & French L. (2000) Contigs built with fingerprints, markers and FPC V4.7. *Genome Research* **10**, 1772–87.

Splawski I., Timothy K.W., Sharpe L.M., Decher N., Kumar P., Bloise R., Napolitano C., Schwartz P.J., Joseph R.M., Condouris K., Tager-Flusberg H., Priori S.G., Sanguinetti M.C. & Keating M.T. (2004) Ca (V) 1.2 calcium channel dysfunction causes a multisystem disorder including arrhythmia and autism. *Cell* **119**, 19-31.

Staheli L.T., Hall J.G., Jaffe K.M. & Paholke D.O. (1998) Arthrogryposis: a text atlas. Cambridge University Press, UK, p. 173.

Stanchi F., Corso V., Scannapieco P., Ievolella C., Negrisolo E., Tiso N., Lanfranchi G. & Valle G. (2000) TUBA8: a new tissue-specific isoform of alpha-tubulin that is highly conserved in human and mouse. *Biochemical and Biophysical Research Communications* **270**, 1111-8.

Tonoki H., Saitoh S. & Kobayashi K. (1998) Patient with del(12)(q12q13.12) manifesting abnormalities compatible with Noonan syndrome. *American Journal of Medical Genetics* **75**, 416–8.

Tsuda T., Yoshida K., Ohashi S., Yanase T., Sueyoshi M., Kamimura S., Misumi K., Hamana K., Sakamoto H. & Yamakawa M. (2004) Arthrogryposis, hydranencephaly and cerebellar hypoplasia syndrome in neonatal calves resulting from intrauterine infection with Aino virus. *Veterinary Research* **35**, 531-8.

Tuggle C., Green J., Fitzsimmons C., Woods R., Prather R., Malchenko S., Soares B., Kucaba T., Crouch K., Smith C., Tack D., Robinson N., O'Leary B., Scheetz T., Casavant T., Pomp

D., Edeal B., Zhang Y., Rothschild M., Garwood K. & Beavis W. (2003) EST-based gene discovery in pig: virtual expression patterns and comparative mapping to human. *Mammalian Genome* **14**, 565–79.

Van Laere A.-S., Nguyen M., Braunschweig M., Nezer C., Collette C., Moreau L., Archibald A., Haley C., Buys N., Tally M., Andersson G., Georges M. & Andersson L. (2003) A regulatory mutation in *IGF2* causes a major QTL effect on muscle growth in the pig. *Nature* **425**, 832-6.

Vignal A., Milan D., Sancristobal M. & Eggen A. (2002) A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution* **34**, 275-305.

Vincent A., Newland C., Brueton L., Beeson D., Riemersma S., Huson S.M. & Newsom-Davis J. (1995) Arthrogryposis multiplex congenita with maternal autoantibodies specific for a fetal antigen. *Lancet* **346**, 24-5.

Vögeli P., Bolt R., Fries R. & Stranzinger G. (1994) Co-segregation of the malignant hyperthermia and the Arg615-Cys615 mutation in the skeletal muscle calcium release channel protein in five European Landrace and Pietrain pig breeds. *Animal Genetics* **25**, 59-66.

Whittem J.H. (1957) Congenital abnormalities in calves: arthrogryposis and hydrocephaly. *Journal of Pathology* **73**, 375-87.

Yamamoto F. & Yamamoto M. (2001) Molecular genetic basis of porcine histo-blood group AO system. *Blood* **97**, 3308-10.

Yerle M., Pinton P., Robic A., Alfonso A., Palvadeau Y., Delcros C., Hawken R., Alexander L., Beattie L.B., Milan D. & Gellin J. (1998) Construction of a whole genome radiation hybrid panel for high-resolution gene mapping in pigs. *Cytogenetics and Cell Genetics* **82**, 182-8.

Yueh Y.G., Gardner D.P. & Kappen C. (1998) Evidence for regulation of cartilage differentiation by the homeobox gene Hoxc-8. *Proceedings of the National Academy of Sciences* **95**, 9956-61.

Appendix

A-1 Submitted porcine sequences

Table A-1 gives the new porcine sequences, derived from cross-species PCR from the putative corresponding human genes that were submitted to GenBank.

Table A-1: Porcine sequences submitted to GenBank.
Gene symbol, gene name, the corresponding human region and the assigned GenBank accession number are shown.

Gene symbol	Gene name	Human region	GenBank number
SCN8A	Voltage-gated sodium channel, alpha subunit 8	Exon 1	DQ180322
KCNA1	Potassium voltage-gated channel, subfamily A, member 1	Exon 1 (only one exon)	DQ180323
KCNA1	Potassium voltage-gated channel, subfamily A, member 1	Exon 1 (only one exon)	DQ355786
KIF21A	Kinesin family member 21A	Exon 18, Intron 18 and Exon 19	DQ180324
CNTN1	Contactin 1	Exon 7, Intron 7 and Exon 8	DQ180325
CNTN1	Contactin 1	Exon 20	DQ180326
ABCD2	ATP-binding cassette, sub-family D, member 2	Exon 1	DQ338453
ABCD2	ATP-binding cassette, sub-family D, member 2	Exon 1	DQ338454
ABCD2	ATP-binding cassette, sub-family D, member 2	Exon 10	DQ338455

A-2 Products and equipment

A-2.1 Reagents and chemicals

Product	Firm/locality (Article number)
• ABI Prism™ Dye Terminator v3.0 Cycle Sequencing Ready Reaction Kit	Perkin Elmer, Applied Biosystems, CH-Rotkreuz (4390242)
• Acrylamide:bisacrylamide 29:1	Biorad laboratories AG, CH-Reinach (1610121)
• Agar	Gibco, Invitrogen AG, CH-Basel (30391-023)
• Agarose	Gibco, Invitrogen AG, CH-Basel (1551-027)
• APS	Fluka Chemie GmbH, CH-Buchs (09915)
• Ampicillin	Fluka Chemie GmbH, CH-Buchs (10044)
• β-mercaptoethanol	Fluka Chemie GmbH, CH-Buchs (63689)
• Boric acid	Fluka Chemie GmbH, CH-Buchs (15660)
• 100 bp-ladder	Amersham Pharmacia Biotech Europe GmbH, CH-Dübendorf (27-400101)
• 50 bp-ladder	Amersham Pharmacia Biotech Europe GmbH, CH-Dübendorf (27-400501)
• Bromophenol blue	Fluka Chemie GmbH, CH-Buchs (18030)
• BSA	Sigma-Aldrich Chemie, CH-Buchs (A-9647)
• Chloramphenicol	Fluka Chemie GmbH, CH-Buchs (23275)
• Cy3 monoreactive dye pack	Amersham Pharmacia Biotech Europe (PA23001)
• Cy5 monoreactive dye pack	Amersham Pharmacia Biotech Europe (PA25001)
• DEPC	Fluka Chemie GmbH, CH-Buchs (32490)
• dNTPs (dATP, dCTP, dGTP and dTTP)$_{100\,mM}$	Amersham Pharmacia Biotech Europe GmbH, CH-Dübendorf (27-2025-01)
• D(+)-Sucrose	Fluka Chemie GmbH, CH-Buchs (84100)
• Ethanol (EtOH absolute)	Merck Eurolab AG, CH-Dietikon (1.00983)
• Ethidium bromide	Sigma-Aldrich Chemie, CH-Buchs (E-8751)
• EDTA	Sigma-Aldrich Chemie, CH-Buchs (E-5134)
• Fairplay labeling kit	Stratagene, La Jolla, USA (252002)
• *FOK*I (+ 1x NEBuffer 4)	Bioconcept, New England Biolabs, CH-Allschwil (R0109S)
• Formaldehyde	Fluka Chemie GmbH, CH-Buchs (47629)
• Formamide	Fluka Chemie GmbH, CH-Buchs (47670)
• Glycerol (87%)	Fluka Chemie GmbH, CH-Buchs (49770)
• Half big dye (*half*BD™)	Genetix GmbH, D-München-Dornach (K1110)
• HCl (32%)	Merck Eurolab AG, CH-Dietikon (1.00319)
• IPTG	Eurobio, Chemie Brunschwig AG, CH-Basel (018025)
• Isopropanol	Fluka Chemie GmbH, CH-Buchs (59300)
• λ-DNA, non methylated	Amersham Pharmacia Biotech Europe GmbH, CH-Dübendorf (27-4118-01)
• $MgCl_2$	Fluka Chemie GmbH, CH-Buchs (63064)
• Methylene blue	Fluka Chemie GmbH, CH-Buchs (66720)
• MOPS	Fluka Chemie GmbH, CH-Buchs (69949)
• NaAc	Fluka Chemie GmbH, CH-Buchs (71190)
• NaCl	Fluka Chemie GmbH, CH-Buchs (71380)
• NaOH	Fluka Chemie GmbH, CH-Buchs (71690)
• Nucleobond AX100 kit	Macherey-Nagel, D-Düren (740579)
• Orange G (crystalline)	Chroma Gesellschaft GmbH, D-Münster (1A116)
• Phenol	Fluka Chemie GmbH, CH-Buchs (77610)
• Primers	Microsynth, CH-Balgach
• Prism™ Genescan-350 Tamra	Applied Biosystems, CH-Rotkreuz (401736)
• Proteinase K	Sigma-Aldrich Chemie, CH-Buchs (P-0390)
• Reverse transcription system	Promega, Catalys AG, CH-Wallisellen (A3500)
• RNeasy Maxi extraction kit	Qiagen AG, CH-Basel (75162)
• SDS	Fluka Chemie GmbH, CH-Buchs (71729)
• Taq DNA Polymerase and 10x PCR buffer	Sigma-Aldrich Chemie, CH-Buchs (D1806)

• TEMED	Bioconcept, CH-Allschwil (0761)
• Trizma base	Sigma-Aldrich Chemie, CH-Buchs (T-1503)
• Trizma HCl	Sigma-Aldrich Chemie, CH-Buchs (T-3253)
• *Tsp*45I (+ 1x NEBuffer 1)	Bioconcept, New England Biolabs, CH-Allschwil (R0583S)
• Urea	Fluka Chemie GmbH, CH-Buchs (51459)
• XCFF	Fluka Chemie GmbH, CH-Buchs (95600)
• X-Gal	Boehringer Mannheim, Roche Diagnostics AG, CH-Rotkreuz (651745)

A-2.2 Labware

Product	Firm/locality (Article number)
• Cryotube 1.8 ml	Nunc, Life Technologies AG, CH-Basel
• Disposable gloves latex L	SAFE EX (41026105)
• Kitchen towel wiper roll	Kimberly-Clark, Synmedic AG, CH-Zurich (7100)
• MicroAmp® optical tubes	Applied Biosystems, CH-Rotkreuz (N801-0933)
• Micropipets	Gilson Medical Electronics AG, FR-Villers-le-Bel
• Pasteur pipet	WU Mainz, BRD
• Petri dishes (10 cm)	Falcon, Becton Dickinson Biosciences, CH-Basel
• Pipet tips	
◦ 1-10 µl	Omikron-Labotec, CH-Reinach (RC10)
◦ 100-1000 µl	Treff, Fischer Scientific AG, CH-Wohlen (96.1702.6.02)
◦ 20-200 µl	Treff, Fischer Scientific AG, CH-Wohlen (96.1701.4.02)
• Polypropylene tubes	
◦ 14 ml	Falcon, Becton Dickinson Biosciences, CH-Basel (2059)
◦ 50 ml	Falcon, Becton Dickinson Biosciences, CH-Basel (9150)
◦ 0.5 ml (PCR)	Omikron-Labotec GmbH, CH-Reinach (710912)
◦ Microtube 1.5 ml	Treff, Fischer Scientific AG, CH-Wohlen (96.7246.9.01)
• Optical caps	Applied Biosystems, CH-Rotkreuz (4323032)
• Quartz cuvettes 10 mm	Hellma GmbH, D-Müllheim/Baden (104.002-QS)
• RNA 6000 Nano LabChip kit	Agilent Technologies, Cheshire, UK (5065-4476)
• Serological pipette	
◦ 10 ml	Falcon, Becton Dickinson Biosciences, CH-Basel (7534)
◦ 5 ml	Falcon, Becton Dickinson Biosciences, CH-Basel (7543)
• Spin columns	Millipore AG, CH-Volketswil
• Sterile vials (+ EDTA K3, blood collection)	Venoject, Cosanum-Ancilla AG, CH-Schlieren
• Ultracentrifuge tubes, polyallomer	
◦ 13.5 ml	Beckman Coulter Int., CH-Nyon (32-6814)
◦ 38.5 ml	Kontron Instr. AG, CH-Zurich (9190196)

A-2.3 Instruments

Instrument	Firm/locality (Article number)
• ABI PRISM 377 DNA Sequencer	Applied Biosystems, CH-Rotkreuz
• ABI PRISM 7700 Sequence Detector	Applied Biosystems, CH-Rotkreuz
• Agilent 2100 Bioanalyser	Agilent Technologies, Cheshire, UK (G2940CA)
• Autoclave	O. Kleiner AG, CH-Wohlen (Typ 23)
• Centrifuges	
◦ Cryofuge 8000	Cryofuge, Heraeus, Kendro Laboratory Products AG, CH-Zurich
◦ Ultra	Kontron Instruments AG, CH-Zurich (centrikon T-2070)
◦ Tabletop	Biofuge, Heraeus, Kendro Laboratory Products AG, CH-Zurich
• GeneTAC Hybridization Station	Genomic Solutions Ltd, Cambridgeshire, UK
• Homogenizer Ultraturrax ® T25 basic	IKA Labortechnik, D-Staufen (T25)

• Incubator	Thermo instruments, D-Haake (DC10-P5)
• Mini-shaker	A. Kühner AG, CH-Birsfelden
• pH-Meter	Metrohm Ltd., CH-Herisau (E 512)
• Power supply (agarose gels), Desagatronic 3x 500/1000	Desaga GmbH, D-Heidelberg
• Power supply (polyacrylamide gels), PowerPac 3000	Bio-Rad laboratories AG, CH-Reinach (165-60)
• Scales	
• <3 g	Mettler Toledo, CH-Greifensee (AG 204)
• >3 g	Mettler Toledo, CH-Greifensee (PE 3600)
• Scan Array 5000 XL	GSI Lumonics, Watertown, USA
• Spectrophotometer Lambda Bio UV/VIS 550	Perkin Elmer, Applied Biosystems, CH-Rotkreuz
• Thermocycler	
• PCR Express	Hybaid, MWG-Biotech AG, CH-Münchenstein
• Thermocycler	Hybaid, MWG-Biotech AG, CH-Münchenstein
• Tissue grinder with teflon pestle	Kimble Kontes, Brechbuhler AG, CH-Zurich (886000)
• Ultracentrifuge rotors	
• TST 28-38	Kontron Instruments AG, CH-Zurich
• TFT 65-13	Kontron Instruments AG, CH-Zurich
• Ultrapure water purification systems	Millipore AG, CH-Volketswil (Milli-Q synthesis A10)
• UV-Transilluminator	MWG-Biotech AG, CH-Münchenstein
• Vacuum concentrator (centrifuge)	Bachofer, D-Reutlingen (BA-VC-300H)
• Vortex	Bender and Hobein AG, CH-Zurich (K-550-GE)

A-2.4 Media and solutions

Media and stock solutions were kept at room temperature except when it is specified.

• Acrylamide:bisacrylamide 49:1 solution 40%	39.2% (*w/v*)	Acrylamide
	0.8% (*w/v*)	N, N'-methylene-bisacrylamide
	Protected from direct light, stored at 4°C	
• Agarose gel	0.8-1.75% (*w/v*)	Agarose
	in 0.5x	TBE
	warmed up to complete dissolution of agarose	
	cooled at ~60-70°C	
	0.1 µg/ml	ethidium bromide
• DEPC-treated H_2O (to use with RNA)	0.2% (v/*v*)	DEPC
	hardly shaked, autoclaved after 30 min	
• DNA double loading dye	0.25% (*w/v*)	XCFF
	0.25% (*w/v*)	Bromophenol blue
	70%	Glycerol
• DNA loading buffer (genescan and sequencing)	50mg/ml	Blue dextran
	25 mM	EDTA (pH 8)
	stored at 4°C	
• DNA triple loading dye	0.25% (*w/v*)	XCFF
	0.25% (*w/v*)	Bromophenol blue
	0.25% (*w/v*)	Orange G
	40% (*w/v*)	D(+)-sucrose
• Ethidium bromide plate	0.8% (*w/v*)	Agarose
	in 0.5×	TBE
	warmed up to complete dissolution of agarose	
	cooled at ~60-70°C	
	1 µg/ml	ethidium bromide (added under hoove)
	Poured into mini Petri dishes, protected from direct light and stored upside down at 4°C	

• IPTG/X-Gal	50 µl	IPTG 2.5% (*w/v*)
	50 µl	X-Gal 2.5% (*w/v*) in dimethylformamide
• LB-agar plate	1.5% (*w/v*)	Agar
	in LB medium	
• LB medium	1% (*w/v*)	Bacto tryptone
	0.5% (*w/v*)	Yeast extract
	170 mM	NaCl
	pH 7 with NaOH	
• Lysis buffer A (DNA isolation from blood)	0.32 M	D(+)-saccharose
	10 mM	Tris-HCl (pH 7.5)
	5 mM	$MgCl_2$
	0.94% (*w/v*)	Triton X-100
	100 µg/ml	Proteinase K
	stored at -20°C	
• Lysis buffer B (DNA isolation from tail)	100 mM	Tris-HCl (pH 8)
	5 mM	EDTA (pH 8)
	0.2%	SDS
	200 mM	NaCl
• Methylene blue	0.02%	Methylene blue
	10 mM	Tris-HCl (pH 7.5)
• MOPS 10×	0.4 M	MOPS
	0.1 M	NaAc
	10 mM	EDTA (pH 8)
	pH 7 with NaOH	
• NaAc (pH 5.2)	408.24 g/l	NaAc x $3H_2O$
	500 ml	ddH_2O
• NEBuffer 1 (1x, pH 7)	10 mM	Bis-Tris-Propane-HCl
	10 mM	$MgCl_2$
	1 mM	Dithiothreitol
	Stored at -20°C	
• NEBuffer 4 (1x, pH 7.9)	20 mM	Tris-acetate
	50 mM	Potassium acetate
	10 mM	Magnesium acetate
	1 mM	Dithiothreitol
	Stored at -20°C	
• PCR buffer 10× (standard PCR)	100 mM	Tris-HCl (pH 8.3)
	500 mM	KCl
	15 mM	$MgCl_2$
	0.01%	gelatin
	Stored at -20°C	
• PCR buffer turbo (DNA isolation from blood)	50 mM	KCl
	10 mM	Tris-HCl (pH 8.3)
	0.1% (*w/v*)	Gelatin
	0.45% (*w/v*)	Nonidet P40
	0.45% (*w/v*)	Tween-20
	Stored at -20°C	
• Polyacrylamide gel (4.5%) (genescan and sequencing)	18 g	Urea
	23 ml	ddH_2O
	5 ml	TBE 10x
	7.5 ml	Acrylamide:bis (29:1), 30%
	15 µl	TEMED
	350 µl	APS 10% (*w/v*)
• Proteinase K (stock solution)	20 mg/ml	Proteinase K in ddH_2O

- RNA loading buffer

1x	MOPS (pH 7)
16.6% (v/*v*)	Formaldehyde
50% (*v*/*v*)	Formamide

incubated 15 min at 60°C, cooled on ice before adding loading dye

- RNA loading dye

0.04%	Bromophenol blue
1 mM	EDTA (pH 8)
50% (*w*/*v*)	Glycerol

Stored at 4°C

- Reverse transcription buffer 1x

10 mM	Tris-HCl (pH 9)
50 mM	KCl
0.1%	Triton X-100

Stored at -20°C

- TBE 10×

0.9 M	Trizma base
0.88 M	Boric acid
20 mM	EDTA (pH 8)

pH 8.3 without correction

- TE 1x (pH 7.5-8)

10 mM	Tris-HCl (pH 7.5-8)
1 mM	EDTA (pH 8)

- Tris HCl

121.16 g/l	1 M Trizma base
157.6 g/l	1 M Trizma HCl

adjust to required pH by mixing the two products

CURRICULUM VITAE

PERSONAL DATA

Name:	Sem Genini
Date of birth:	July 1, 1976
Place of birth:	Locarno, TI, Switzerland
Nationality:	Swiss (citizen of Cresciano, TI)

EDUCATION

2002 – 2006	PhD study and research assistant in the group of Prof. Dr. G. Stranzinger, Breeding Biology, Institute of Animal Science, Swiss Federal Institute of Technology Zurich (ETHZ)
2002 – 2006	Departmental Conference, delegate for assistants and PhD students, Department of Agricultural and Food Sciences, ETHZ
2002 – 2006	Departmental Teaching Committee, delegate for assistants and PhD students, Department of Agricultural and Food Sciences, ETHZ
2005	Visiting student, ARK-Genomics, Roslin, UK
1995 – 2002	Graduation as Dipl. Ing. Agr. ETHZ with specialization in Agrobiotechnology
2000 – 2002	Certificate in Didactics from the Department of Agricultural and Food Sciences, ETHZ
1999 – 2000	Research assistant in the group of Prof. Dr. J. Nösberger, Crop Science, Institute of Plant Science, ETHZ
1999	Research assistant in the group of Prof. Dr. C. Candolfi-Vasconcelos, Viticulture group, Department of Horticulture, Oregon State University, Corvallis, OR, USA
1991 – 1995	High school degree, Matura type C, Liceo Lugano 1, Lugano (TI)

PUBLICATIONS

Genini S., Nguyen T.T., Malek M., Talbot R., Gebert S., Rohrer G., Nonneman D., Stranzinger G. & Vögeli P. (2006) Radiation hybrid mapping of 18 positional and physiological candidate genes for porcine arthrogryposis multiplex congenita (AMC) on chromosome 5. Accepted for publication in *Animal Genetics*.

Genini S., Stranzinger G. & Vögeli P. (2005) Arthrogryposis Multiplex Congenita (AMC) beim Schwein als Beispiel einer im Bluttypisierungslabor untersuchten Erbkrankheit. *Suisseporcs-Information* **10**, 14-6.

Nguyen T.T., Genini S., Ménétrey F., Malek M., Vögeli P., Goe M.R. & Stranzinger G. (2005) Application of bovine microsatellite markers for genetic diversity analysis of Swiss yak (*Poephagus grunniens*). *Animal Genetics* **36**, 484-9.

Nguyen T.T., Ménétrey F., Genini S., Nguyen V.L., Vögeli P., Nguyen B.X. & Stranzinger G. (2005) Application of bovine microsatellite markers on Saola (*Pseudoryx nghetinhensis*). *Journal of Animal Breeding and Genetics* **122**, 195-8.

Genini S., Malek M., Špilar Š., Nguyen T.T., Ménétrey F., Gebert S., Hagger C., Neuenschwander S., Kadarmideen H., Stranzinger G. & Vögeli P. (2004) Arthrogryposis Multiplex Congenita (AMC), a

hereditary disease in swine, maps to chromosome 5 by linkage analysis. *Mammalian Genome* **15**, 935-41.

Genini S., Gebert S., Stranzinger G. & Vögeli P. (2003) Arthrogrypose beim Schwein: Erste Ergebnisse des Forschungsprojekts und Bekämpfungsstrategie. *Suisseporcs-Information* **8**, 12-3.

Mattiacci L., Rudelli S., Ambühl-Rocca B., Genini S. & Dorn S. (2001) Systemically induced response of cabbage plants against a specialist herbivore, *Pieris brassicae*. *Chemoecology* **11**, 167-73.

CONFERENCES AND POSTERS

Genini S., Nguyen T.T., Joller D., Talbot R., Malek M., Gebert S., Rohrer G., Nonneman D., Stranzinger G. & Vögeli P. (2006) Porcine Arthrogryposis Multiplex Congenita (AMC), a disease model for human medicine. First European Conference on Pig Genomics, Parco Tecnologico Padano, Lodi, Italy (oral presentation).

Genini S., Nguyen T.T., Joller D., Talbot R., Malek M., Gebert S., Rohrer G., Nonneman D., Stranzinger G. & Vögeli P. (2006) Radiation hybrid mapping to SSC5 of 18 positional and functional candidate genes for Arthrogryposis Multiplex Congenita (AMC). First European Conference on Pig Genomics, Parco Tecnologico Padano, Lodi, Italy (poster).

Nguyen T.T., Eggen A., Genini S., Schelling C., Bui L.C., Renard J.P., Nguyen B.X. & Stranzinger G. (2005) Cytogenetics studies on Saola (Pseudoryx nghetinhensis, 2n=50) and embryos derived from interspecies nuclear transfer. 2nd Asian Reproductive Biotechnology Conference, Bangkok, Thailand (poster).

Genini S., Downing A., Malek M., Talbot R., Stranzinger G. & Vögeli P. (2005) DNA microarray analysis of differentially expressed genes associated with Arthrogryposis Multiplex Congenita (AMC) in swine. 3rd International Symposium on Genetics of Animal Health, Iowa State University, Ames, IA, USA (poster).

Genini S., Špilar Š., Nguyen T.T., Ménétrey F., Gebert S., Hagger C., Stranzinger G. & Vögeli P. (2004) Arthrogryposis Multiplex Congenita (AMC), eine Erbkrankheit beim Schwein, liegt auf Schweine-Chromosom 5. Conference of the Swiss Association for Animal Production, Swiss Federal Institute of Technology (ETH), Zurich (oral presentation).

Genini S., Malek M., Špilar Š., Nguyen T.T., Ménétrey F., Gebert S., Hagger C., Neuenschwander S., Kadarmideen H., Stranzinger G. & Vögeli P. (2004) Arthrogryposis multiplex congenita (AMC) a hereditary disease in swine, maps to chromosome 5 by linkage and QTL analyses. 29th International Conference on Animal Genetics, Tokyo, Japan (poster).

Ménétrey F., Le Gal F., Nguyen T.T., Genini S., Saner R. & Stranzinger G. (2004) Molecular test to screen cytogenetic defects on preimplatation bovine embryos. Conference of the Swiss Association for Animal Production, Swiss Federal Institute of Technology (ETH), Zurich (poster).