Chromosome Mapping in Cattle by Fluorescence

In Situ Hybridization

A dissertation submitted to the
Swiss Federal Institute of Technology - Zurich
for the degree of Doctor of Natural Sciences

presented by

Sabina Solinas Toldo
Dipl. Ing. Agr. ETHZ
born April 28, 1965
citizen of Rovio (TI)

accepted on the recommendation of
Prof. Dr. G. Stranzinger, examiner
PD Dr. R. Fries, co-examiner
Dr. P. Lichter, co-examiner

Zurich 1993
To my mother, Luca, Marta and Emmanuel
Acknowledgements

I'm grateful to Prof. G. Stranzinger, who gave me the enthusiasm to work in the field of cytogenetics. It was his "open door" philosophy, that encouraged me to visit many a time for very useful discussions.

I'm especially grateful to PD Dr. R. Fries, who was not only a very competent teacher, but also provided me with a guideline in understanding the strategy of the research. Moreover, I appreciated his friendship and interesting discussions, which were not only science oriented.

A special thanks is due to Dr. P. Lichter, from German Cancer Research Center, Heidelberg, for agreeing to be co-examiner of this work and also for giving me the possibility to attend an EMBO Course in 1992, where I could improve my knowledge about in situ hybridization.

I also would like to express my gratitude to Dr. F. Kohen from Weizmann Institute, Israel, and S. Hardy for careful reading and commenting on the first draft, and improving the English language used.

The same gratitude is devoted to Prof. U. Gyger, from the University in Pennsylvania, for his helpful comments.

Thanks to all the colleagues and friends at the Institute, particularly André, Andreina, Asoka, Dagmar, Georges, Paul and Roger. They contributed in making a friendly atmosphere in the laboratory.

My thanks are due to the people and their respective laboratories, reported at pages 23 and 24, for providing DNA probes used for the realization of this work.

Recognition is due to the Swiss National Science Foundation and to the Swiss artificial insemination association, who for three years financially supported the research which has been presented here. The Laur Foundation assisted in payment of the printing costs.

I'm greatly indebted to my husband Luca, my sister Marta, and Emmanuel for their patient and continual encouragement, and for many helpful suggestions. Moreover, the wonderful week ends together at home were indispensable for starting a new good week.
# Table of contents

Summary \hspace{1cm} i

Zusammenfassung \hspace{1cm} ii

Riassunto \hspace{1cm} iv

List of figures \hspace{1cm} v

List of tables \hspace{1cm} vii

List of abbreviations \hspace{1cm} viii

1. Introduction \hspace{1cm} 1
   1.1. Gene mapping \hspace{1cm} 2
      1.1.1. Methods for gene mapping \hspace{1cm} 3
      1.1.2. Comparative mapping \hspace{1cm} 4
   1.2. The need for a marker map \hspace{1cm} 5
      1.2.1. DNA marker loci \hspace{1cm} 7
      1.2.2. Different class of DNA markers \hspace{1cm} 8
         1.2.2.1. Restriction fragment length polymorphisms (RFLP) \hspace{1cm} 8
         1.2.2.2. Variable number of tandem repeats (VNTR) \hspace{1cm} 8
            1.2.2.2.1. Microsatellites \hspace{1cm} 10
         1.2.2.3. Random amplified polymorphic DNA (RAPD) \hspace{1cm} 11
      1.2.3. Genetic linkage between two loci \hspace{1cm} 11
   1.3 Gene mapping in cattle \hspace{1cm} 12
   1.4. In situ hybridization (ISH) \hspace{1cm} 13
      1.4.1. Radioactive in situ hybridization \hspace{1cm} 13
      1.4.2. Non isotopic in situ hybridization (NISH) \hspace{1cm} 14
         1.4.2.1. Fluorescence in situ hybridization (FISH) \hspace{1cm} 17
         1.4.2.2. Biotin-avidin system \hspace{1cm} 18
      1.4.2.3. Cosmid approach \hspace{1cm} 20
   1.5. Chromosome identification \hspace{1cm} 21
   1.6. Objectives \hspace{1cm} 22
2. Materials
   2.1. DNA probes
   2.2. Technological demands
      2.2.1. Equipment
      2.2.2. Small material and products
   2.3. Buffers and solutions

3. Methods
   3.1. Preparation of chromosomes
      3.1.1. Chromosomes from leukocytes
      3.1.2. Chromosomes from fibroblasts
      3.1.3. Slide preparation
      3.1.4. QFQ-banding
   3.2. Preparation of cosmid probes
      3.2.1. Source of cosmids
      3.2.2. Growth of bacteria culture
      3.2.3. Qiagen cosmid midi-preparation
      3.2.4. Probe labelling
         3.2.4.1. Nick Translation
         3.2.4.2. Random Priming method
         3.2.4.3. Purification of the labelled DNA
         3.2.4.4. Methods to check the extent of biotin incorporation
   3.3. In situ hybridization
      3.3.1. Chromosomal in situ suppression (CISS)
         3.3.1.1. Preparation of genomic competitor DNA
         3.3.1.2. Preparation of Cot-1 competitor DNA
         3.3.1.3. DNA precipitation and chromosomal in situ suppression
      3.3.2. Hybridization on slides
      3.3.3. Signal detection
      3.3.4. Chromosome counterstaining
      3.3.5. Two probes/one colour in situ hybridization
      3.3.6. Two probes/double colour in situ hybridization
      3.3.7. Determination of the hybridization site
   3.4. Microscopic evaluation
   3.5. Photography
   3.6. Automated Sanger sequencing
3.6.1. Sequence analysis
3.6.1.1. Format conversion
3.6.1.2. Homology search with FASTA
3.6.1.3. Optimal alignment between homologous sequences

4. Results
4.1. Optimization of fluorescence in situ hybridization on bovine chromosomes
4.1.1. Factors that affect the background phenomena
4.1.1.1. Slide preparation
4.1.1.2. RNase treatment
4.1.1.3. Chromosomal in situ suppression
4.1.1.4. Washing steps
4.1.2. Factors that affect the signal strength
4.1.2.1. Probe labelling and analysis of biotin incorporation
4.1.2.2. Probe length
4.1.2.3. Probe concentration
4.1.2.4. Chromosome counterstaining
4.1.2.5. Amplification of the signal intensity
4.1.2.6. Efficiency of fluorescence in situ hybridization
4.2. Assignment of cosmid and phage probes to bovine chromosomes
4.2.1. Map position of loci
4.2.1.1. Double colour fluorescence in situ hybridization with two clones
4.2.2. Distribution of loci on the genome
4.2.3. Centromere specific hybridization of satellite containing cosmids
4.2.4. A chromosome specific satellite for discrimination between chromosome 4 and 6

5. Discussion
5.1. Significance of the physical mapping of the clones
5.2. Centromere specific satellites
5.3. Chromosome specific satellites
5.4. Use of microsatellites as markers
5.5. Completion of the bovine marker map
5.6. Improvements and future developments
Summary

The aim of this work was to contribute to the physical map of the bovine genome. Fluorescence in situ hybridization (FISH) was optimized and used for the assignment of marker loci, using probes which had been cloned in phages or in cosmids. The optimization of the mapping technique consisted mainly of improvement in the chromosome identification and in the biotin labelling of probes, and the establishment of an additional step to suppress signals from the repetitive sequences contained in large insert probes (chromosomal in situ suppression, CISS).

The improved technique was used for the assignment of 41 probes, 30 of which were cosmids containing microsatellites of type (CA)n. Although in other species microsatellites are homogeneously distributed on the genome, in cattle a tendency for clustering on the telomeric and subtelomeric regions was observed.

13 (CA)n containing cosmids produced signals on the centromeres of the autosomes. The partial sequencing of two clones revealed the presence of the known bovine satellites 1.709 and 1.723. A very uncommon and strong hybridization signal on chromosome 6 suggested the presence in this clone of a chromosome specific satellite. This probe is now a very important tool for discrimination between the cytogenetically very similar bovine chromosomes 4 and 6. Another clone gave a particularly strong signal on chromosome 17, and its partial sequencing suggested the presence of a minisatellite sequence.

Besides (CA)n containing anonymous DNA segments, different cosmid and phage clones containing coding sequences were also localized. These assignments revealed interesting results about the evolutionary conservation of the position of genes in mammals.

Since information about the synteny of some probes was also available, new syntenic groups could be assigned. In one case there was a discrepancy between a direct chromosomal assignment by in situ hybridization and the assignment through a previously localized syntenic group. This was a probable case of false synteny in the panel used.
Zusammenfassung


Neben den anonymen (CA)n enthaltenden Cosmidklonen wurden auch verschiedene mit kodierenden Genen assozierte Cosmid- und Phagenklone lokalisiert. Diese Lokalisierungen ergaben zum Teil interessante Hinweise bezüglich der evolutionären Konservierung der Genanordnung bei Säugetieren.

Da für einige der lokalisierten Sonden auch Informationen über die Zugehörigkeit zu bestimmten Syntänegruppen vorlagen, konnten einige bis jetzt noch nicht lokalisierte Syntänegruppen chromosomal zugewiesen werden. In einem Fall wurde eine Diskrepanz zwischen der direkten chromosomalen
Zusammenfassung

Zuweisung mittels *in situ* Hybridisation und der Zuweisung zu einer bereits chromosomal lokализierten Synteniegruppe festgestellt. Es handelte sich dabei möglicherweise um einen Fall von falscher Syntanie.
Riassunto

Questo lavoro si propone come contributo alla mappa fisica del genoma bovino. Ibridazione in situ fluorescente (FISH) fu ottimizzata e utilizzata per la localizzazione di loci marcatori. Le sonde usate erano principalmente clonate in cosmid o fagi. L'ottimizzazione di questa tecnica consistette soprattutto nel miglioramento dell'identificazione cromosomica e nella marcatura delle sonde con biotina. L'uso di lunghi inserti richiese inoltre la soppressione delle sequenze ripetitive in essi contenute (soppressione cromosomale in situ, CISS).

41 sonde, 30 delle quali cosmid contenenti microsatelliti di tipo (CA)n, furono ibridate con tale metodo. Si osservò una tendenza di localizzazione nelle regioni telomeriche e subtelomeriche, contrariamente a quanto noto per altre specie dove tale distribuzione è omogenea.

13 cosmid contenenti (CA)n diedero segnale di ibridazione sui centromeri degli autosomi. L'analisi della loro sequenza rivelò la presenza dei satellite 1.709 e 1.723 già noti nel genoma bovino. Un cosmide produsse un segnale molto forte sul cromosoma 6 e suggerì la presenza in questo clone di un satellite cromosoma-specifico. Questa sonda è ora uno strumento molto importante per la discriminazione fra i cromosomi 4 e 6, che nel bovino sono citogeneticamente molto simili. Un altro clone diede un segnale molto forte sul cromosoma 17; parte della sua sequenza suggerì la presenza di un minisatellite.

Oltre a segmenti di DNA anonimi contenenti (CA)n, furono localizzati diversi cloni di cosmid e fagi con sequenze codificanti. Queste localizzazioni rivelarono interessanti risultati riguardanti la conservazione evoluzionistica della posizione dei geni in mammiferi.

Per alcune sonde erano disponibili informazioni relative alla loro appartenenza ad un gruppo di sintenia; queste permisero la localizzazione di gruppi non ancora assegnati. Si verificò un caso di discrepanza fra la localizzazione cromosomale diretta con ibridazione in situ e quella ottenuta da un gruppo sintenico precedentemente localizzato. Si trattò probabilmente di un caso di falsa sintenia nel pannello utilizzato.
List of figures

Fig. 1: Molecular structure of biotin-11-dUTP 19
Fig. 2: Schematic representation of binding between biotin and avidin, one round of amplification 20
Fig. 3: Polyacrylamide gel showing the size of DNA fragments after nick translation of the probe 35
Fig. 4: Metaphase spreads showing the effect on ISH when different amounts of competitor DNA are added to the probe (CAP2) 52
Fig. 5: 1.5% agarose minigel showing Cot-1 DNA digested with different amounts of S1 nuclease 54
Fig. 6: Dot blot assay of three labelling reactions of the same probe 56
Fig. 7: Model of steric hindrance between biotinylated molecules 57
Fig. 8: Model of steric hindrance between avidin molecules 58
Fig. 9: Metaphase spreads showing the chromosomal assignment of the clones 73
Fig. 10: Standard idiogram (ISCNDA, 1989) showing the localization of the clones 74
Fig. 11: Partial metaphase spread showing double colour FISH with the probes CAP3 and c10BT34, on chromosome 19 77
Fig. 12: Distribution of the (CA)n containing probes on a schematic chromosome 80
Fig. 13: Example of hybridization with a clone containing sequences centromere specific for the autosomes (CAP6) 81
Fig. 14: Sequencing of the cosmid CAP5, (a) with the primer T3, (b) with the primer T7 83
Fig. 15: Sequencing of the cosmid CAP6, (a) with the primer T3, (b) with the primer T7 84
Fig. 16: Sequence homology revealed by BESTFIT, a) between sub2CAP5 sequenced with the primer T3 and the bovine satellite 1.709, b) between sub2CAP6 sequenced with the primer T3 and the bovine satellite 1.723 86
Fig. 17: Interphase nuclei showing the "long hybridization signal" obtained with the probe c10BT33 88
Fig. 18: Metaphase spread showing the simultaneous localization of the probes CAP8 and c10BT33 on chromosome 6 90
Fig. 19: Metaphase spread showing the simultaneous localization of the probe c10BT33 and the NORs genes

Fig. 20: a) Schematic representation of cattle chromosomes 4 and 6
b) QFQ-banded chromosomes 4 and 6 of cattle
List of tables

Table 1: Non radioactive labelling and detection systems used in NISH 16
Table 2: List of the probes available and their origin 23
Table 3: Excitation and emission wavelengths of fluorescent dyes 46
Table 4: Filter blocks used for the detection of fluorescent dyes 46
Table 5: Hybridization results of some randomly considered probes 62
Table 6: Chromosomal assignment of the clones 65
Table 7: Assignment of some probes to syntenic groups 76
Table 8: Distribution of different clones on the genome 79
Table 9: Chromosomes showing a weaker signal than other chromosomes, after hybridization with centromere specific probes 82
Table 10: Comparison between the centromere specific sequences contained in CAP5 and CAP6 and the known bovine satellites 85
List of abbreviations

A adenine
AMCA aminomethylcoumarin acetic acid
ASPS anti-satellite polymerase chain reaction screening
ATP adenosine triphosphate
bp base pair
BCIP 5-bromo-4-chloro-3-indolyl phosphate
BLAD bovine leukocyte adhesion deficiency
BSA bovine serum albumin
C cytosine
Ci curie
CISS chromosomal in situ suppression
cM centimorgan
DNA desoxyribonucleic acid
dpm disintegrations per minute
EMBL European molecular biology laboratory
ETL economic trait locus
FACS fluorescence-activated cell sorter
FISH fluorescence in situ hybridization
FITC fluorescein isothiocyanate
FLcen fractional length centromere
g acceleration of gravity
G guanine
GCG genetic computer group (Madison, WI)
IRS interspersed repetitive sequences
ISH in situ hybridization
kb kilobase
LB Luria-Bertani medium
M molar
MAS marker assisted selection
NBT nitro blue tetrazolium
NISH non isotopic in situ hybridization
PCR polymerase chain reaction
PDME progressive degenerative myeloencephalopathy
PEG polyethylene glycol
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIC</td>
<td>polymorphism information content</td>
</tr>
<tr>
<td>QTL</td>
<td>quantitative trait locus</td>
</tr>
<tr>
<td>RAPD</td>
<td>random amplified polymorphic DNA</td>
</tr>
<tr>
<td>RFLP</td>
<td>restriction fragment length polymorphism</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RT</td>
<td>room temperature</td>
</tr>
<tr>
<td>SAA</td>
<td>syndrome of arachnomelia of arthrogrypose</td>
</tr>
<tr>
<td>SA-AP</td>
<td>streptavidin alkaline phosphatase</td>
</tr>
<tr>
<td>SD</td>
<td>standard deviation</td>
</tr>
<tr>
<td>SMA</td>
<td>spinal muscular atrophy</td>
</tr>
<tr>
<td>SSC</td>
<td>standard saline citrate</td>
</tr>
<tr>
<td>STMS</td>
<td>sequence tagged microsatellite site</td>
</tr>
<tr>
<td>STS</td>
<td>sequence tagged site</td>
</tr>
<tr>
<td>T</td>
<td>thymine</td>
</tr>
<tr>
<td>TE</td>
<td>tris EDTA</td>
</tr>
<tr>
<td>TNE</td>
<td>tris natrium EDTA</td>
</tr>
<tr>
<td>TRITC</td>
<td>tetramethyl rhodamine isothiocyanate</td>
</tr>
<tr>
<td>u</td>
<td>unit</td>
</tr>
<tr>
<td>U</td>
<td>unassigned syntenic group</td>
</tr>
<tr>
<td>U</td>
<td>uracil</td>
</tr>
<tr>
<td>v</td>
<td>volume</td>
</tr>
<tr>
<td>VNTR</td>
<td>variable number of tandem repeats</td>
</tr>
<tr>
<td>YAC</td>
<td>yeast artificial chromosome</td>
</tr>
</tbody>
</table>
1. Introduction

Farm animals contribute in many ways towards the environment and the human beings in providing a source of nutrition, of clothing and of income for the people. To ensure that human populations may also benefit in the future, modern breeding programs require improvements in animal nutrition, reproduction, hygiene and genetics. Improvements of these factors will lead to an increase in productive traits, such as milk and meat production, as well as to animals with special disease resistance. The development and realization of modern breeding programs started from the quantitative genetics and, nowadays, includes a wide range of areas, thereby developing an understanding of the molecular basis of the genome and the principles of inheritance (Stranzinger, 1987). Of major interest is the isolation and characterization of genes which determine economic traits loci (ETL, Smith and Simpson, 1986), specifically quantitative traits loci (QTL, Geldermann, 1975; Beckmann and Soller, 1983, 1988) or loci responsible for hereditary disorders or disease resistance.

Since it is not possible to determine the genotype of an individual by examining the phenotype alone, biological and biometrical methods have been developed to evaluate the genetic variation within a population and to proceed to a proper selection. The breeding success depends in part on the heritability of the trait. Traits such disease resistance and susceptibility, as well as fertility, are traits of low heritability. Carriers of genetic diseases must be recognized early to prevent an increase in gene frequency. Delays in and improper selection could cause huge economical losses. By removing carrier individuals from the breeding population and/or by choosing healthy animals with special interesting characteristics for their progeny, an efficient selection method could be planed.

Breeding programs must be handled very carefully because certain selection of productive traits can indirectly increase the frequency of a defective gene, if it is positively associated (linked) with the selected trait. The genetic defect is then rapidly spread to the progeny population through artificial insemination using semen of untested sires. An example is the bovine Weaver syndrome, a genetic defect causing progressive degenerative myeloencephalopathy, that seems to be genetically linked to high milk performance (Hoeschele and Meinert, 1990). In Switzerland this trait was introduced without knowledge into the Brown-Swiss population in the 1980s by importing semen of an outstanding American bull, that was a carrier for Weaver.
This situation indicates that genetic selection requires the study of intergene relationships to reduce this type of risk.

The identification of single genes and their action is a first important step towards the control of traits by animal breeding. McKusick (1980) has proposed considering the genome as part of the anatomy of an organism. Its study implies the establishment of a map of genes and marker anchor loci. This map will be a useful tool for the identification of economic important traits and/or for their selection. Marker anchor loci which are evenly distributed through the genome allow to consider possible intragene influences and serve as anchor points in the study of linkage maps.

Worldwide efforts of several laboratories are concentrated to the realization of a "bovine marker packet".

This work is a contribution to the physical mapping of suitable markers on the bovine genome for genetic selection of a selected trait.

1.1. Gene mapping

The analysis of loci on the genome implies the knowledge of the order of the genes on the chromosomes, as well as the alignment of other DNA fragments and the relationship between different loci. The joining of such informations forms the gene maps. There are two types of gene maps: genetic and physical, which are distinguished as follows:

• physical maps are based upon the assignment of loci to chromosomes, mainly by the analysis of somatic cell hybrid panels, by in situ hybridization and by comparative mapping. The considerable progress reached today in gene mapping in different species suggests in fact an extensive sequence conservation between different species. Regional assignment of genes has been accomplished also by using naturally occurring reciprocal translocations, or other chromosomal rearrangements. In physical maps the coordinates are the chromosome region or band; the distance between two loci is described as synteny information and in number of base pairs.

• genetic maps are constructed by studying the meiotic linkage relationship between two or more loci through family studies (linkage analysis) or with single sperm typing. Genetic maps do not provide an absolute location of loci but they describe the genetic distance of loci as a function of the frequency of crossing-overs occurring during meiotic recombination. The genetic distance is expressed in units of crossing-over, or centimorgans (cM).
The relationship between physical and genetic maps is not straightforward, due to the fact that crossing-overs do not occur completely random in the genome. The physical distance in kilobases (kb) between loci can however be approximated from the genetic distance in cM, and vice versa. Assuming a constant distribution of crossing-overs, 1 cM contains approximately 1000 kb.

1.1.1. Methods for gene mapping

Different approaches are used to map DNA sequences on the chromosomes: linkage analysis in families, somatic cell hybrid panels and in situ hybridization to fixed chromosomes. These methods are briefly discussed here below:

a) Linkage analysis: a new locus can be assigned to a chromosome through segregation analysis of the new locus and of an already mapped locus in an informative family, that means a family in which at least one parent is heterozygous for the loci to be studied. This analysis, revealing the meiotic recombination frequency between the loci, estimates the distance and the localization of one locus with respect to the other (s. 1.2.3.). Measuring the recombination frequency can also be applied with "single sperm typing" (Li et al., 1988). This method does not require the availability of an informative family and of a large number of individuals, since the recombination frequency between the loci is directly estimated from the count of recombinant meiotic products.

Methods used for linkage analysis provide information about the assignment of a new locus to a chromosome region only if another locus close to the candidate region has been already assigned. Linkage analysis is also very useful to find out the recombination frequency between a marker and a gene or within genes. The relationship between the bovine leukocyte antigen (BoLA) and the prolactin locus (PRL) has been studied using sperm typing (Lewin et al., 1992).

b) Somatic cell hybrid panel: rodent cells (mouse or hamster) are fused with cells of another species when they are cultured together under special conditions. During and after the cell fusion there is preferential loss of the chromosomes of the non rodent species. Different hybrid cell lines retain different chromosome complements of the non-murine parent species. The study of several lines allows the identification of the suitable ones to form a hybrid panel. The presence or
absence of gene products, such as enzymes or proteins, Southern blotting or the amplification of DNA segments with PCR technique through the panel allow the identification of loci which are lost or retained together. These loci define syntenic groups, that is loci located on the same chromosome, which is however not always known. Syntenic groups are designated with an "U" (unassigned syntenic group) followed by a number, ranging from 1 to 29 in the case of cattle and representing the haploid number of autosomal chromosomes.

The panels currently used consist of up to 20-30 different lines, although theoretically only five hybrid lines consisting of an optimal chromosome complement are needed to assign a locus to cattle chromosomes (Fries and Ruddle, 1986). The method of choice to assign a new locus to a defined syntenic group is now done by PCR amplification of the DNA fragment in each hybrid line. If the syntenic group has already been assigned to a chromosome, the locus is indirectly assigned to the same chromosome. Due to this fact, the physical localization of a locus which belongs to a syntenic group also localizes the whole syntenic group to the same chromosome.

A drawback is that only genes of which the products can be analyzed in the cell culture or for which cloned DNA sequences or sequence information is available can be assigned by this approach. Moreover, regional chromosomal localization can only be achieved if the hybrid panel contains chromosomal rearrangements.

c) **In situ hybridization:** it is the most direct way to localize a DNA sequence to a chromosome and it allows the localization of a locus with high resolution. The DNA probe is hybridized on metaphase chromosomes and the hybridization locus is visualized immediately. This method is main topic of the work presented here and is therefore extensively described in 1.4.

### 1.1.2. Comparative mapping

Homologies in the genome between human beings and other mammals have been observed (Nadeau and Taylor, 1984; Nadeau, 1989; Stranzinger and Hediger, 1990; Dietz et al., 1992). The criteria defining homologies between species are reported in the reports of the Committee on comparative gene mapping (e.g. O'Brien and Graves, 1991), which also assures uptodate information.

The bovine genome is also included in these comparative studies, indicating not only very strong homologies with other Bovidae (Hediger et al.,
Introduction

1991; Gallagher and Womack, 1992; Johnson et al., 1993), but a surprisingly high chromosome conservation between bovine and human (Womack and Moll, 1986; Threadgill and Womack, 1990a; Threadgill and Womack, 1990b). This similarity of chromosomal loci arrangement exceeds that observed between human and mouse.

Comparative mapping is based on the observations that extensive genome regions, mainly coding sequences, are conserved between different species. The identification of conserved chromosomal segments provides a preliminary information about the location of a locus in one species if the homologue of such a locus has been mapped in another species. We could consider such predictions as a method for provisional locus assignments.

Knowledge of genome regions that are conserved among different species allows the use of mapping data in a species to select new probes for identifying polymorphic markers with predictable spacing in other species (Womack, 1987; Moore et al., 1991).

Comparative mapping often answers important questions concerning genome organization (Stranzinger, 1990) and evolution, and it is particularly useful in the study of related species (O'Brien and Seuanez, 1988; Threadgill et al., 1991). Very recently, O'Brien et al. (1993) reported a list of 321 reference anchor loci suitable for comparative genome mapping in mammals.

1.2. The need for a marker map

Today many efforts are focused on studies of genetic defects in mammalian species, therefore, the next sections will be focused onto the need for markers in disease control. In cattle, genetic disorders such as spinal muscular atrophy (SMA), (El-Hamidi et al., 1989; Nielsen et al., 1990), Weaver (PDME, progressive degenerative myeloencephalopathy) (Stuart and Leipold, 1985), Spiderleg (SAA, syndrome of arachnomelia arthrogrypose) (König et al., 1987), bovine leukocyte adhesion deficiency (BLAD) (Shuster et al., 1992) and others, have been already described in different breeds, and cause huge economical losses. Therefore, a selection against carrier animals is required.

These inherited diseases are usually monogenic autosomal recessive disorders caused by mutations in the gene. In monogenic diseases the identification of the mutation is a relatively realistic goal, although not always straightforward. In such cases it is possible to perform a screening of the population to find individual carriers of the mutated gene. When the candidate
locus is not known, or in the case where diseases are controlled from several genes (polygenic), it is helpful to know a marker linked to the pathological phenotype.

Markers are polymorphic "reference points" in the phenotype or genotype of a species. They can be used to diagnose a genetic disease, to map new genes or to select for major genes that determine economically important traits. Informative markers are highly polymorphic, i.e. they present several variants. With several variants occurring in a pedigree, it will be more likely to find parental individuals heterozygous at different marker loci; heterozygous individuals are informative for linkage analysis. The linkage between a marker variant and the locus of interest can be established for example among the progeny in segregating families by correlating the inheritance of a disease trait with the inheritance of a specific genotype/phenotype information. A complete map of polymorphic markers is a prerequisite to find a marker linked to every locus of interest. To saturate the genome, a different minimal number of polymorphic markers is required for each domestic animal species, according to the genome size, the chromosome number and the desired density of markers. The application of a marker map in selective breeding is described under the name of marker assisted selection (MAS) (Smith and Simpson, 1986; Beckmann and Soller, 1987; Georges and Massey, 1991). Such a map is also a tool for a marker based diagnosis of genetic diseases and for reverse genetics (s. 1.2.1.).

Thus far, several types of markers have been developed. The first generation was confined to biochemical serological variations of enzymes, serum proteins, blood group factors, blood and milk proteins. These biochemical markers do not provide any information about their localization in the genome. Therefore, a homogeneous coverage of the genome is not accomplishable. With the advent of recombinant DNA technologies, DNA markers have been discovered and are described in more detail below.

The establishment of marker maps became a major occupation of molecular genetics in the last few years. Almost complete linkage maps spanning the whole genome have recently been published for the human and for the mouse genome, by Weissenbach et al. (1992) and Dietrich et al. (1992) respectively.

Beckmann and Soller (1983, 1987) were the first to point out the need of molecular markers in farm animals as well.
1.2.1. DNA marker loci

DNA markers are based on polymorphisms detected at the DNA level, that means that one locus is the site of several alleles. These markers are tested for linkage relationships and are used to construct linkage maps (s. 1.2.3.). Segregation studies of the marker in a family reveal which allele is inherited together with the gene of interest.

Marker loci are evaluated according to their polymorphism information content, abbreviated as PIC. The PIC of a marker is calculated from allele frequencies in the population and is related to the repeat length. A PIC value greater than 0.5 is considered highly informative, a PIC value between 0.25 and 0.5 indicates a reasonably informative marker, while markers with PIC smaller than 0.25 are only slightly informative (Botstein et al., 1980).

DNA markers allowed the development of a new approach, called reverse genetics (Orkin, 1986) or positional cloning (Collins, 1992). In this approach the isolation and cloning of a gene is not referred to biological or biochemical information about the corresponding protein, but it is deduced from the map location of the marker linked to it (Wicking and Williamson, 1991) or from its own map position. In human genetics several genes involved in hereditary diseases such as chronic granulomatous disease (Royer-Pokora et al., 1986) and cystic fibrosis have been cloned in this way (Riordan et al., 1989).

Due to the possibility to directly determine their chromosomal position, DNA markers can be defined as anchor loci for further analysis of the genome. O’Brien (1991) discriminates between two types of anchor loci:

- Type I anchor loci: they are coding gene loci conserved among mammalian species and mapped at least in man and mouse. They are often monomorphic or slightly polymorphic, mostly with just 2 alleles (RFLPs). These loci contribute considerably to the construction of comparative maps between related species.
- Type II anchor loci: they are highly polymorphic (PIC ≥ 0.6) DNA segments, as for instance VNTRs (minisatellites and microsatellites, s. 1.2.2.2.). These loci are usually not conserved among different species and are therefore less useful for comparative evolutionary studies. On the other hand, they give helpful informations to map other loci of interest or are used for parentage control.
1.2.2. Different class of DNA markers

1.2.2.1. Restriction fragment length polymorphisms (RFLP)

Botstein et al. (1980) and Wyman and White (1980) were the first to propose restriction fragment length polymorphisms as DNA markers.

RFLPs are observed when a restriction enzyme recognizes a variable restriction site, due to insertions, deletions or point mutations. This variation affects the length of restriction fragments, defining different alleles, which are visualized by Southern blot techniques. RFLPs are inherited in a codominant mode, and they usually have only two alleles, indicating whether the restriction site is present or absent. The probability that a parent is heterozygous at one locus is usually less than 50%; they are therefore not informative for linkage analysis in many families and in populations without selection.

In human genetics, RFLP techniques have been used to directly identify DNA sequence changes in genes of interest (Geever et al., 1981) and to detect specific alleles linked to the mutant gene. Genetic defects such as Huntington's disease (Gusella et al., 1983), Duchenne's muscular dystrophy (Monaco et al., 1986; Murray et al., 1982), cystic fibrosis (Tsui et al., 1985, White et al., 1985) and others have been mapped through linkage found between an RFLP and the disease.

RFLPs have been found at many loci in coding and non-coding sequences in man, plants, insects and also in domestic animals.

Today, however, the interest has moved to the study of VNTRs, which present more appropriate characteristics to be used as genetic markers.

1.2.2.2. Variable number of tandem repeats (VNTR)

VNTRs are based on the variation in the number of tandem repeats of a defined sequence motif at a defined locus. This variation is visualized by cleavage with a restriction enzyme, which generates fragments of different length and determines their polymorphism. Heterozygosity of 90% is common, and since they provide linkage information in almost all families, they represent a very suitable type of DNA markers (Nakamura et al., 1987a). Jeffreys et al. (1985a) noted that in human several VNTRs share a homologous "core sequence" of about 10 nucleotides. Nakamura et al. (1987a) observed a similar property but concerning a slightly different "core sequence". This core element is similar to the Chi sequence of E.coli and is suspected to help to generate satellite sequences. This occurs by promoting an initial tandem duplication of unique DNA and subsequent stimulating unequal crossing over during meiotic
recombination or by DNA slippage during replication. The effect of this phenomenon is the high degree of polymorphism (Jeffreys et al., 1985a).

These repeats are found at many loci in the mammalian genome and cross-hybridization of VNTRs between different species has been observed and described by Georges et al. (1988). The core sequence in farm animals, however, seems to stem from an ancestor artiodactyla genome (Jeffreys, 1987).

Different types of VNTRs are distinguished, according to the length of the repeat unit and the repetition number:

• **Satellites**: they have a repeat unit of few hundreds bp to several kilobases, many repeats form up to thousands of kilobases. Satellite sequences have been found mainly in centromeric regions of mammalian chromosomes.

• **Midisatellites**: consist of repeating units of an average of 40 bp and clustered in a locus to form a 250-500 kb long repetitive DNA (Nakamura et al., 1987b).

• **Minisatellites**: the repeat unit shows a length of about 20 to 60 bp. Minisatellites have already been used as genetic markers and have been found prevalently on proterminal regions of human and mouse chromosomes (Royle et al., 1988). Their polymorphic character is visualized in Southern blot hybridization with "single locus probes" (monolocus) or "multilocus probes".

Single locus probes recognize locus-specific VNTRs, generating two bands if the individual is heterozygous and one if homozygous. They are especially suited for forensic analyses. Multilocus probes recognize simultaneously a large number of hypervariable minisatellites dispersed throughout the genome and they provide a DNA profile of high complexity with many polymorphic bands. The obtained pattern represents an individual-specific "DNA fingerprint" (Jeffreys et al., 1985b), also called "genetic bar code". Relatedness is deduced from the number of bands that two individuals have in common. Multilocus probes are less sensitive than monolocus, requiring more DNA. The interpretation of eventually missing bands is very difficult, and it is often due to laboratory failures, as well as the detection of foreign DNA contaminations. However, DNA fingerprinting is a powerful technique used in paternity testing and forensic medicine (Jeffreys et al., 1985b; 1985c) or in mapping of the genome through linkage analysis. Georges et al. (1990) proposed the use of DNA fingerprints in linkage studies in cattle.

• **Microsatellites**: due to their growing interest in genetic projects and due to the fact that they are main topic of this work, microsatellite markers are described in a separate paragraph (s. 1.2.2.2.1.).
1.2.2.2.1. Microsatellites

Microsatellites, described at the same time by three groups (Weber and May, 1989, Litt and Luty, 1989, Tautz, 1989), are considered a special type of VNTRs, consisting of tandemly repeated sequences with many repeats of di- to hexanucleotides. The most common microsatellites are the poly (CA-GT) repeats. The variation of the number of copies of the repeat determines the polymorphism, which is very high for these sequences. The higher the number of repeats, the higher the polymorphism (Weber, 1990). The stretch of tandem repeats is often not longer than 100 to 200 nucleotides. Therefore, individual genotyping can be performed with PCR amplification, the length of the amplified product indicating the number of repeat units (Weber and May, 1989; Litt and Luty, 1989). Since their polymorphism is characterized with PCR, microsatellites are represented in sequence tagged site format (STS) (Olson et al., 1989), a site which is characterized by the two primers chosen for PCR amplification. Beckmann and Soller (1990) defined the STMS as sequence tagged microsatellite site. This property is very useful for single sperm typing, since PCR amplification is possible from single cells.

Two different approaches for the detection of microsatellites are used: experimental search by screening a genomic library with a (dC-dA)n(dG-dT)n oligonucleotide probe or computational search in a GenBank database (Moore et al., 1992). The microsatellites used in this work had been developed by using the first method.

The function of microsatellites is still unknown, although they may play a role in genetic recombination (Pardue et al., 1987) or in the enhancement of transcriptional activity (Hamada et al., 1984). CA-repeats seem to play a role in condensing and decondensing DNA in eukaryotic chromosomes during the cell cycle (Stallings et al., 1991).

Until now these sequences seem to occur abundantly and at random throughout most eukaryotic genomes. Beckmann and Weber (1992) calculated an average distribution of microsatellites every 6 kb in the human genome. Initial studies found the particular type of microsatellites poly (CA-GT) every 30 kb in human (Stallings et al., 1991; Beckmann and Weber, 1992), every 20 kb in mouse and rat (Stallings et al., 1991) and only every 180 kb in cattle (Steffen et al., 1993). Polymorphic microsatellites in the bovine genome are described for example by Fries et al. (1990). Since primers for a microsatellite designed for one species often work well on related species, they can be efficiently used to
identify new markers in different species (Swarbrick et al., 1992). Practically, this is exploited within Bovidae.

Due to the advantages mentioned above, today microsatellites are the marker of choice for the realization of marker maps in eukaryotes. In human microsatellites have already contributed to the localization of new genes.

1.2.2.3. Random amplified polymorphic DNA (RAPD)

First described by Williams et al. (1990), RAPD are polymorphic DNA sequences randomly amplified with PCR from genomic DNA by using short primers of about 10 bp of arbitrary sequence. Since no prior sequence information is required, amplification can be performed with little knowledge of the genome being studied. By using a multiplex primer approach, which means a combination of arbitrary primers, or by working with a relatively low stringency, a large number of PCR products can be generated. The characterization of the amplified fragments will reveal which will be useful as markers.

A potential disadvantage of this method is that the primers are selected without regard to the biological function of the sequences they contain, so that DNA sequences of an unknown function are characterized.

However, RAPD are useful to detect polymorphisms between individuals, which show mutations in priming sites or other mutations that change the size of the amplified segment. RAPDs have been used for instance to determine the genetic relationship between different wheat types (Vierling and Nguyen, 1992).

1.2.3. Genetic linkage between two loci

By definition, two loci are genetically linked if they are on the same chromosome and if their recombination rate \( \Theta \) is lower than 50%. Contrary to syntenic groups, where the only information available is the belonging of different loci to the same chromosome, in linkage groups the recombination rate is known and defines the distance between the loci. The recombination occurs during meiosis and \( \Theta \) represents the proportion of recombined gametes among those which are produced, that means the frequency of crossing-overs between homologous chromosomes. One factor which influences the probability of crossing-overs is in fact the physical distance between the loci.

For the purpose of genetic analysis and to keep the number of animals needed for recombination studies as small as possible, a marker should be no more than 20 cM (corresponding to 20% of recombination) away from another locus (Soller and Brody, 1976). For the realization of a complete marker map this
means that polymorphic marker loci cannot be more distant than 40 cM apart from each other to allow detection of linkage with a reasonable number of individuals.

Testing linkage between loci is equivalent to test $\Theta = 50\%$ against the alternative $\Theta < 50\%$. Morton (1955) proposed the calculation of this probability with the lod score test (logarithm of odds test) first developed by Wald (1947). The lod score test is based on the method of maximum likelihood and it is defined as follows:

$$L = \log_{10} \frac{\text{likelihood of observation of linkage (when } \Theta<0.5)}{\text{likelihood of observation of no linkage (when } \Theta=0.5)}$$

Lod scores of +3 or higher mean that the likelihood of linkage is at least 1000 times greater than the likelihood of no linkage, and are considered to be strong evidence for linkage. Lod scores of -2 or less mean that the linkage hypothesis must be rejected.

Algorithms for multilocus linkage analysis have been developed and computer programs such as LINKAGE (Lathrop and Lalouel, 1984) allow the estimation of genetic distance between several loci simultaneously.

1.3. Gene mapping in cattle

The bovine karyotype consists of 29 acrocentric to telocentric autosomal pairs and two sex chromosomes, presenting huge difficulties in the identification. After the first gene assignment by ISH in this species (Fries et al., 1986) scientists have applied a range of different techniques which have contributed towards the gaining of knowledge about the bovine genome. Today the 29 bovine syntenic groups (U1-U29) as well as the X-chromosome are represented from a panel of 20 hybrid cell lines established in the laboratory of J. Womack (Texas A & M University) (Womack and Moll, 1986). Through this and other panels and with in situ hybridization, 260 Type I anchor loci and 104 Type II anchor loci have been localized on the bovine genome. Eighteen syntenic groups are assigned to chromosomes (Fries et al., 1993), some of them have been confirmed by the localization of more than one locus. The bovine gene map ranks fourth, behind the maps of man, mice and rat.

The bovine genome has been estimated to have a length of about 3000 cM. In order to establish a relationship between all the loci and a marker, 3
markers are required for each chromosome of average length of 100 cM: two about 20 cM from the centromere and telomere, respectively, and one in the middle (Fries et al., 1989). Practically however, the longer chromosomes should carry 4 markers and the shorter only two. The complete bovine marker map would theoretically consist of 90 markers. Since marker spacing cannot be directed by using a random approach such as that adopted by this work, more than 90 markers have to be found, characterized and mapped.

1.4. In situ hybridization (ISH)

In situ chromosomal mapping is based on the pioneer work of Gall and Pardue (Gall and Pardue, 1969; Pardue and Gall, 1969).

This technique provides a means of localizing DNA or RNA sequences on tissues, interphase nuclei or metaphase chromosomes. Initially it allowed only the localization of highly repeated sequences, but later improvements in the hybridization procedure made possible the localization of single copy genes.

ISH is based upon the principle of the molecular hybridization between nucleic acids. A single stranded labelled nucleic acid probe is hybridized with the previously denatured specimen, resulting in a new double stranded molecule if the two reaction partners have a complementary base sequence. This reaction usually takes place between the probe in suspension and immobilized chromosomes or interphase nuclei.

Two methods are common practice today, radioactive ISH and non isotopic ISH (NISH), and they are discussed in detail below.

1.4.1. Radioactive in situ hybridization

In this technique labelling of the probe is carried out by incorporating one or more radiolabelled nucleotides, such as \(^{35}\text{S}-\text{dNTP}\), \(^{125}\text{I}-\text{dNTP}\) or \(^{3}\text{H}-\text{dNTP}\) into the DNA by using different labelling methods. Tritiated probes are usually preferred due to their higher resolution power and long term stability. The detection of hybridized sequences is performed by autoradiography with a photographic emulsion. Autoradiographic silver grains are scored on several metaphases, grains located on chromosomes are plotted on a standardized idiogram representing the haploid genome. A simple statistical approach evaluates the distribution of silver grains, indicating a significant accumulation (peak) over a chromosome region as the specific hybridization signal. This
Introduction

statistical approach requires the analysis of up to 100 metaphases, and it is therefore very laborious.

Improvements of this methodology over the last two decades allow now the detection of single copy sequences of only a few hundred base pairs in length (Harper et al., 1981). The hybridization of labelled viral RNA to cellular preparations was also demonstrated (Harper et al., 1986).

The resolution obtained with ISH depends on the chromatin contraction in the chromosomes and on the resulting banding pattern. Ruddle (1981) estimated that a resolution of 5-10 cM can be achieved. A limiting factor in radioactive ISH is the fact that, due to the photographic emulsion, the silver grains are spatially positioned at a certain distance from the hybridization site.

The limitations of radioactive ISH are the relatively poor resolution, the background noise of silver grain deposition, the long time required for autoradiography and the disadvantages associated with radioactive probes (waste accumulation, health hazards,...). In addition, the laborious statistical analysis of the silver grain accumulation is particularly tedious when it concerns the identification of quite difficult chromosomes, as in the case of cattle, dog or fish.

To overcome these limitations innovative non isotopic detection techniques have been studied and developed.

1.4.2. Non isotopic in situ hybridization (NISH)

In recent years alternative approaches to detect specific nucleotide sequences have been developed and very quickly became popular due to their several advantages. Non isotopic in situ hybridization is also found in a wide range of applications in clinical cytogenetics, i.e. for the detection of numerical (Lichter et al., 1988a; Pinkel et al., 1988) and structural (Lichter et al., 1991; Cremer et al., 1988) chromosome aberrations in metaphase and in interphase nuclei (Cremer et al., 1990).

Two types of nonradioactive in situ hybridization methods are distinguished:

- direct procedures: the probe is labelled with nucleotides modified with a reporter molecule, which after hybridization can be directly detected (Dirks et al., 1990; Wiegant et al., 1991; Koch et al., 1992). This method is used mainly when large regions are targeted or when chromosome painting is performed.
- indirect procedures: they are based on modified labelled probes, which are detected using sandwich techniques. This is carried out by exploiting the
interaction between the modification and a compound which is carrying the reporter molecule, and is directed against the modification. The nucleotides used to label the probe are conjugated with biotin (Langer et al., 1981), digoxigenin (De Frutos et al., 1990; Zhang et al., 1990), acetyaminofluorene (Tchen et al., 1984; Landegent et al., 1984; 1985), mercury (Hopman et al., 1986; 1987) or other compounds. When the probe is labelled with biotin, the detection can be achieved through a reporter molecule linked to an antibody or to avidin (s.1.4.2.2.). The reporter groups include (i) enzymes, like horseradish peroxidase and alkaline phosphatase, and their respective insoluble substrates, (ii) fluorescent molecules, such as fluorescein and rhodamine, and (iii) colloidal gold particles, visualized under an electron microscope (Hutchison et al., 1982).

Alternative methods, already used in other applications but still in development for ISH, are the chemiluminescence (Ekins et al., 1989) and the time resolved fluorescence with lanthanides (Dahlén, 1987; Verwoerd et al., 1992).

Due to the higher detection sensitivity, the most widely used procedures are the biotin-avidin and digoxigenin-antidigoxigenin systems, combined with fluorochromes. Biotin incorporation is the method of labelling chosen for this work and will be discussed later. The use of digoxigenin was introduced in 1987 by Boehringer Mannheim Biochemicals. Digoxigenin, a steroid isolated from the plant Digitalis purpurea, is present only in the blossoms and leaves of these plants. Therefore, it is totally unlikely that non specific binding of antibodies to digoxigenin will occur with biological materials other than Digitalis purpurea plants. Antibodies anti-digoxigenin are conjugated to reporter molecules.

The most common systems used for NISH are summarized in Table 1. (Any other modified nucleotide can be used to label the DNA instead of dUTP).
Table 1: Non radioactive labelling and detection systems used in NISH.
FL: fluorochrome, AP: alkaline phosphatase, ab: antibody to..., AAF: acetylaminofluorene, BrdU: bromodeoxyuridine

<table>
<thead>
<tr>
<th>Direct procedures</th>
<th>Indirect procedures</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA modification</td>
<td>DNA modification</td>
</tr>
<tr>
<td>fluorescein-dUTP</td>
<td>biotin-dUTP</td>
</tr>
<tr>
<td>rhodamine-dUTP</td>
<td></td>
</tr>
<tr>
<td>resorufin-dUTP</td>
<td>digoxigenin-dUTP</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>hydroxycoumarin-</td>
<td>acetylaminofluorene-</td>
</tr>
<tr>
<td>dUTP</td>
<td>dUTP</td>
</tr>
<tr>
<td></td>
<td>bromodeoxyuridine-</td>
</tr>
<tr>
<td></td>
<td>dUTP</td>
</tr>
</tbody>
</table>

Compared to conventional methods, NISH offers several important advantages, such as:

a) higher signal resolution: the detection of the hybridized DNA takes place exactly at the site where the probe is hybridized,
b) it is possible to amplify the signal intensity by using immunological techniques (s. table 1 and 1.4.2.2.),
c) when fluorescent ISH is applied, this technique has the potential to combine several different fluorochromes to detect simultaneously different target regions (s. 1.4.2.1.),
d) increased speed of the procedure,
e) avoidance of radioactive waste accumulation,
f) stability of the labelled probe for several months at -20°C.

Compared to radioactive ISH, the sensitivity limitation is a disadvantage of NISH. Single copy genes of a few kb in length cannot be routinely assigned, but the use of high copy number probes or large insert probes such as cosmids (s. 1.4.2.3.) or yeast artificial chromosomes (YACs) improves the success rate. However, the localization of unique sequences of 900 bp, 1 kb and 2.3 kb of
length have been described respectively by Viégas-Péquignot et al. (1991), by Garson et al. (1987) and by Viégas-Péquignot et al. (1989).

New developments, such as specialized camera systems (CCD camera, charged coupled device) or confocal laser scanning microscopes, have improved the detection sensitivity of non isotopic methods, allowing more sophisticated evaluations, and therefore the visualization of very weak signals. Single-copy DNA fragments of about 1 kb have been successfully mapped by evaluation with image processing systems.

The flexibility of NISH and its wide application in cell biology and genetics has been reported by Lichter and Ward (1990).

1.4.2.1. Fluorescence in situ hybridization (FISH)

In FISH the probe can be labelled with a fluorochrome and directly visualized, or the detection can take place after hybridization by binding of a fluorescent molecule conjugated to an antibody or to avidin/streptavidin.

The fluorochrome most commonly used is fluorescein isothiocyanate (FITC), which gives a yellow/green signal. Over the years many other fluorochromes have been developed, for example tetramethyl rhodamine isothiocyanate (TRITC, red) or aminomethylcoumarin acetic acid (AMCA, blue).

In order to become fluorescent, fluorochromes must be excited by light of a defined wavelength range (excitation spectrum) and the emitted fluorescent light is detected under a slightly longer wavelength (emission spectrum).

Fluorescence is generally more suitable than other detection systems, because of better spatial resolution, less background and the great potential for simultaneous visualization of several target sequences on the same specimen (Nederlof et al., 1989; 1990; Ried et al., 1992). This approach, called multicolor FISH, consists in the labelling of one probe for example with biotin and of a second one with digoxigenin; the first one is detected by binding of avidin conjugated to FITC and the second with an antibody anti-digoxigenin conjugated to TRITC. The detection is achieved by simultaneously exciting the two fluorochromes with a double band-pass filter, which has recently become available. Alternatively, two or more pictures of the same metaphase are taken separately with different filters and are then overlapped. Overlapping can be performed with conventional photography or with image processing systems.

These properties, together with the high resolution power, are very suitable when two or more neighboring probes must be ordered relative to each other along the chromosome (Lawrence et al., 1988; 1990). FISH to
chromosomes can resolve two probes that are separated by approximately 1 Mb, the detection of sequences within interphase nuclei, where the chromatin is less condensed, allows a resolution of up to 50 kb (Trask, 1991; Trask et al., 1991), while in pronuclei probes which are separated by 40 to 50 kb can be consistently resolved (Brandriff et al., 1991). Heng et al. (1992) demonstrated the resolution of sequences spaced 21 kb apart on "free chromatin". Van den Engh et al. (1992) proposed a model for the calculation of the spacing between sequences along the linear DNA molecule from interphase distance measurements. Such a model is required due to the different possible configurations of the chromatin fibre. Recently, Lawrence et al. (1992) reviewed the development of interphase mapping.

FISH is widely used for many applications, such as gene mapping on metaphase chromosomes and interphase nuclei, localization of viral integration sites, studies on chromosomal aberrations, analysis of somatic cell hybrid panels, etc.

1.4.2.2. Biotin-avidin system

Some years ago it was observed that by feeding animals with raw egg white as their sole source of protein, they developed a nutritional deficiency characterized by neuromuscular disorders and called "egg-white injury". Investigations revealed that the injury was due to deficiency of the vitamin biotin and it was shown that egg white contains a glycoprotein, later termed avidin, which binds biotin with extremely high affinity. The history of these investigations has been well summarized by György (1954). The discovery of this exceptional complex resulted in the development of a new and sensitive assay as described below.

Biotin belongs to the vitamin B-group and it is also called vitamin B7 or vitamin H. It has a molecular weight of 244.

Avidin, reviewed by Green (1975), has a molecular weight of 68000 daltons and an affinity constant with biotin of $10^{15}$/mol. Another biotin-binding compound was isolated in 1963 from culture filtrates of Streptomyces avidinii and was called streptavidin (Chaiet and Wolf, 1964). Both avidin and streptavidin have four binding sites for biotin and similar binding affinity, but a different amino acid sequence. Streptavidin is not glycosilated and due to its structure it shows less binding with other components than biotin, resulting in less background (Wilchek and Bayer, 1988). The three dimensional structure of the biotin-avidin complex has been reported very recently by Livnah et al. (1993).
In 1981 Langer et al. proposed the synthesis of a new molecule to label nucleic acids: a biotin molecule covalently linked to the C-5 position of the pyrimidine ring of dUTP (fig.1). Langer and Ward (1981) reported one of the first uses of biotin for in situ gene mapping.

![Molecular structure of biotin-11-dUTP](image)

Due to their specific interaction, after hybridization of the labelled probe to the specimen, avidin or streptavidin can be bound to biotin. Avidin is conjugated with a colorimetric reporter molecule or with a fluorochrome. Very often it is necessary to amplify the intensity of the hybridization signal and this is obtained by adding a layer of biotinylated antibody anti avidin, followed by a second layer of avidin-FITC. Pinkel et al. (1986) calculated that by using this system with FITC each amplification step enhances 6 fold the intensity of the signal. Amplification of the signal is not confined to the biotin-avidin system, and can be performed also in other FISH methods (s. table 1).

The principle of the detection of the biotin-avidin system is illustrated in fig. 2.
Some years ago, it already had been predicted that the biotin-avidin system would find many applications in molecular biology (Bayer and Wilchek, 1980). Some years later the same authors (Wilchek and Bayer, 1989) reviewed the exceptional development of this technology, which has wide application in several different disciplines, as localization and separation of antigens, immunotherapy, affinity chromatography, immunoassay and hybridization studies.

1.4.2.3. Cosmid approach

Because of the limited sensitivity of NISH, a logical strategy to produce strong hybridization signals with unique sequences is to use large probes. Landegent et al. (1987) demonstrated that the use of cosmid vectors in non radioactive ISH resulted in an improved detection limit. Lichter et al. (1990) were able to order numerous cosmids on human chromosome 11 based on measurements of signal position relative to total chromosome length from digitalized images.

Cosmid vectors are hybrids, derived from plasmids and lambda phages, containing the "cos site" of the lambda phage. They allow the accommodation of genomic DNA fragments ranging in size from 35 to 45 kb.
In this work microsatellites cloned in cosmids were chosen as probes for NISH. The microsatellite and sufficient unique sequences to yield a strong hybridization signal are contained in the large insert.

Due to its large size and due to the interspersed repetitive sequences (IRS) scattered also in the bovine genome, cosmid probes contain repeats which can obscure the specific signal. This problem and the way to circumvent it are discussed in the chapter of methods, 3.3.1.

1.5. Chromosome Identification

The unambiguous identification of the chromosomes is the preliminary requirement for the cytogenetic analysis of chromosomes and for gene mapping in every species. Stranzinger and Fechheimer (1988) described the need to develop a worldwide common identification system standardized in international ideograms.

The precise description and identification of chromosomes is possible with staining methods summarized for farm animals by Gustavsson (1980).

The "Proceedings of the First International Conference for the Standardization of Banded Karyotypes of Domestic Animals" discussed in Reading, was the first official standard available to identify chromosomes in domestic species (Reading Conference, 1976). G-banded (Giemsa stained) karyotypes of *Bos taurus, Ovis aries, Capra ircus, Sus scrofa, Equus caballus, Felis catus* and *Oryctolagus cuniculus* are therein described. Since the similarity between G-banding and other banding patterns, such as that obtained with Quinacrine staining, is high, the first gene assignments on cattle were based on this standard. In the following years, several other chromosome banding techniques were developed, and it was necessary to standardize karyotypes with different banding patterns.

The Second International Conference on Standardization of Domestic Animal Karyotypes was carried out in 1989 in Paris. The result of this meeting was the "International System for Cytogenetic Nomenclature of Domestic Animals" (ISCNDA, 1989). In this new standard a GTG-banded, a QFQ-banded and two types of RBA-banded cattle karyotypes with exact description and numbering of bands are compared. Besides that of cattle, karyotypes of goat and sheep are presented. A combined haploid karyotype of RBG-banded chromosomes of goat, sheep and cattle is also proposed. These studies
document the homology of banding patterns within the *Bovidae* family (Hediger et al., 1991; Hayes et al., 1991; Gallagher and Womack, 1992).

1.6. Objectives

Besides providing a contribution to the analysis of the bovine genome, the specific aims of this thesis were:

- the establishment and optimization of fluorescence *in situ* hybridization of DNA markers on the bovine genome
- the physical mapping of several polymorphic microsatellite markers and of other loci
- the identification of chromosome-specific satellites
- the characterization of centromere specific satellites
- the analysis of loci distribution on the bovine genome
- the development and proposal of chromosome specific markers as tool in the cytogenetic identification of the chromosomes.
2. Materials

2.1. DNA probes

As outlined in the introduction, the main topic of this work was the localization of cosmids containing microsatellites and other large insert probes. The probes were either developed in our laboratory as described in 3.2.1, or they were provided by other laboratories. For this reason, the probes used show heterogeneity in their characteristics, being not only cosmids and not only clones containing necessarily microsatellites. Probes containing sequences candidate for coding genes, were in fact also available. Informations about the probes used in this work and their origin are reported in table 2.

Table 2: List of the probes available and their origin

<table>
<thead>
<tr>
<th>Probes</th>
<th>Type of clone</th>
<th>Vector</th>
<th>Provided by</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAP1, CAP2, CAP3,</td>
<td>cosmid</td>
<td>SuperCos1 (Stratagene)</td>
<td>Library constructed and screened in our laboratory (Eggen, 1992 and Steffen, 1992)</td>
</tr>
<tr>
<td>CAP4, CAP5, CAP6,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP7, CAP8, CAP9,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP10, CAP11, MSI/2,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSI/3, 3R-100-9,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAP1B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c10BT18, c10BT21,</td>
<td>cosmid</td>
<td>Triple Helix</td>
<td>Olsaker, I., Dept. of Animal Genetics, Norwegian College of Vet. Medicine, 0033 Oslo, Norway</td>
</tr>
<tr>
<td>c10BT33, c10BT34,</td>
<td></td>
<td>(Stratagene)</td>
<td></td>
</tr>
<tr>
<td>c10BT93, c10BT133,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c10BT134, c10BT145,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c10BT146, c10BL26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C5, C7, C8, M10/13,</td>
<td>cosmid</td>
<td>SuperCos1 (Stratagene)</td>
<td>Bishop, M., USDA, Agricultural Research Service, Clay Center, 68933 Nebraska, USA</td>
</tr>
<tr>
<td>M14/1, 12A, 24A, 17B,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>99-1, 99-2, 99-3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Materials

<table>
<thead>
<tr>
<th>Probes</th>
<th>Type of clone</th>
<th>Vector</th>
<th>Provided by</th>
</tr>
</thead>
<tbody>
<tr>
<td>μ35Cos9</td>
<td>cosmid</td>
<td>pWE15 (Stratagene)</td>
<td>Vaiman, D., Lab. de Génétique Biochimique, INRA-CRJ, 78350 Jouy-en-Josas, France</td>
</tr>
<tr>
<td>60-21, 59-11, 65</td>
<td>phage</td>
<td>EMBL</td>
<td>Franke, W.W., Inst. of Cell and Tumor Biology, German Cancer Research Center, 6900 Heidelberg, Germany</td>
</tr>
<tr>
<td>plgR</td>
<td>phage</td>
<td>EMBL SP6/T7</td>
<td>Kulseth, M.A., Dept. of Animal Science, Agricultural University, 1432 Ås, Norway</td>
</tr>
</tbody>
</table>

#### 2.2. Technological demands

##### 2.2.1. Equipment

- **β-Counter**: Betamatic II (Kontron)
- **Centrifuges**: Biofuge B (Heraeus)
- **Developing machine**: Agfaprint BW 37-50
- **Enlarger machine**: Durst Laborator 900
- **Incubator**: Heraeus
- **Lab Shaker**: Kühner AG
- **Microscope equipped with fluorescence**: Leitz Diaplan
- **Mini gels**: Mini-Sub Cell (BioRad)
- **Sonicator**: Micro Tip Limit
- **Spectrophotometer**: UV-VIS 550 (Perkin-Elmer Cetus)
- **Water baths**: Julabo, Digitana
2.2.2. Small material and products

Acetic acid glacial 100%  
Alpha-medium  
Agarose type I Low EEO  
Agefix fixativ  
Agfa Brovira Speed paper  
Agfa Ortho 25 professional films  
Ampicillin  
Antibiotic-antimycotic  
Avidin-FITC  
Bacto agar  
Bacto-yeast extract  
Bacto-tryptone  
Biotin-11-dUTP  
Biotin-16-dUTP  
BluGENE Kit  
BSA  
Chloroform  
Colcemid  
Coverslips 18x18 mm  
Coverslips 24x60 mm  
DABCO (1,4-diazo-bicyclo-2,2,2-octane) FLuka 33480  
D-19 developer  
Dextran sulphate  
DMEM (Dulbecco’s modified Eagle’s Medium)  
Dnase I  
EDTA (Ethylenediaminetetraacetic acid) Sigma E-5134  
Ethanol  
Falcon tubes  
Fetal calf serum  
Formamide  
G-50 sephadex Columns  
Goat anti-avidin  
\(^{3}\)H-dATP
<table>
<thead>
<tr>
<th>Material</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCl 32%</td>
<td>Merck 319</td>
</tr>
<tr>
<td>Isopropanol</td>
<td>Fluka 59300</td>
</tr>
<tr>
<td>Kodak Technical Pan 2415 films</td>
<td>Kodak</td>
</tr>
<tr>
<td>L-glutamine</td>
<td>Seromed K 0280</td>
</tr>
<tr>
<td>Methanol</td>
<td>Merck 6009</td>
</tr>
<tr>
<td>Microscope slides</td>
<td>Knittel</td>
</tr>
<tr>
<td>NaCl</td>
<td>Merck 6404</td>
</tr>
<tr>
<td>NaOH</td>
<td>Fluka 71690</td>
</tr>
<tr>
<td>Na-pyruvat</td>
<td>Fluka 15990</td>
</tr>
<tr>
<td>Neutol</td>
<td>NE Agfa</td>
</tr>
<tr>
<td>Nick translation labeling kit</td>
<td>Boehringer Mannheim 976 776</td>
</tr>
<tr>
<td>Nitrocellulose filter</td>
<td>Schleicher &amp; Schuell 401 116</td>
</tr>
<tr>
<td>Nuclease S1</td>
<td>Boehringer Mannheim 818 330</td>
</tr>
<tr>
<td>$^{32}$PdATP</td>
<td>Rahn PB 10204</td>
</tr>
<tr>
<td>PBS Dulbecco</td>
<td>Seromed L 182-10</td>
</tr>
<tr>
<td>PEG 1000 (Polyethylene glycol)</td>
<td>Fluka 81190</td>
</tr>
<tr>
<td>Phenol</td>
<td>Fluka 77610</td>
</tr>
<tr>
<td>Pokeweed mitogen</td>
<td>Seromed M 5060</td>
</tr>
<tr>
<td>Propidium iodide</td>
<td>Fluka 81845</td>
</tr>
<tr>
<td>Qiagen-tips 100 column</td>
<td>Qiagen 10044</td>
</tr>
<tr>
<td>Quinacrine Mustard</td>
<td>Sigma Q-2876</td>
</tr>
<tr>
<td>Random primed labeling kit</td>
<td>Boehringer Mannheim 1004 760</td>
</tr>
<tr>
<td>RNase A</td>
<td>Fluka 83832</td>
</tr>
<tr>
<td>Rodinal</td>
<td>Agfa</td>
</tr>
<tr>
<td>RPMI 1640 medium</td>
<td>Seromed F1215</td>
</tr>
<tr>
<td>Scintillationcocktail</td>
<td>Optiphase &quot;Hisafe 3&quot; LKB Wallac</td>
</tr>
<tr>
<td>SDS (Sodium dodecyl sulfate)</td>
<td>Fluka 71729</td>
</tr>
<tr>
<td>Silica gel</td>
<td>Fluka 85340</td>
</tr>
<tr>
<td>SSDNA</td>
<td>Sigma D-1626</td>
</tr>
<tr>
<td>Sterile vials</td>
<td>Venoject VT-100SH</td>
</tr>
<tr>
<td>T25, T75 tissue culture flasks</td>
<td>Falcon</td>
</tr>
<tr>
<td>Trypsin</td>
<td>Seromed L 2123</td>
</tr>
<tr>
<td>Tween 20</td>
<td>Fluka 93773</td>
</tr>
<tr>
<td>Urea</td>
<td>Fluka 51459</td>
</tr>
<tr>
<td>Xylol</td>
<td>Merck 8681</td>
</tr>
</tbody>
</table>
2.3. Buffers and solutions

Antifade solution 2.3 % of DABCO (1,4-diazo-bicyclo-2,2,2-octane) in 20mM TrisHCl pH 8.0 and 90% glycerol.

AP 7.5 buffer  0.1 M Tris-HCl pH 7.5, 0.1 M NaCl, 2 mM MgCl₂
AP 9.5 buffer  0.1 M Tris-HCl pH 9.5, 0.1 M NaCl, 50 mM MgCl₂

Chloroform/Isoamylalkohol
  Substitute 40 ml of chloroform in a 1 l bottle with 40 ml of Isoamylalkohol

LB medium (Luria-Bertani Medium)
  Bacto tryptone 10 g/l, Bacto yeast extract 5 g/l,
  NaCl 10 g/l, Bacto agar 15 g/l

P1 buffer  100 μg/ml RNase A in 50 mM Tris/HCl, 10 mM EDTA, pH 8.0
P2 buffer  200 mM NaOH, 1% SDS
P3 buffer  2.55 M KAc, pH 4.8

Phenol/Chloroform
  Dissolve 250 g of phenol with 250 ml of chloroform/i soamylalkohol.
  Add 0.25 g of 8-hydroxy-chinolin and fill up with TE pH 8.0 to saturation

QBT buffer  750 mM NaCl, 50 mM MOPS, 15% EtOH, pH 7.0, 0.15% Triton X-100
QC buffer  1.0 M NaCl, 50 mM MOPS, 15% EtOH, pH 7.0
QF buffer  1.25 M NaCl, 50 mM MOPS, 15% EtOH, pH 8.2

Sørensen's phosphate buffer (pH 6.8)
  1/15 M KH₂PO₄ Merck 4873 and 1/15 M Na₂HPO₄ x 2 H₂O Merck 6580

TE pH 8.0  10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0
TNE pH 8.0  10 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA pH 8.0
Trypsin/EDTA  1.5 ml 1% EDTA in 100 ml Trypsin
10 x TBE  0.9 M Tris-borate, 0.02 M EDTA
20 x SSC  3M NaCl, 0.3 M Na-Citrate
3. Methods

3.1. Preparation of chromosomes

Chromosomes were prepared with standard methods from blood cells or from tissue as described below.

3.1.1. Chromosomes from leukocytes

Peripheral blood was obtained from male cattle (*Bos taurus* L.) at the ETH research station Chamau. The blood was dispensed into 10 ml sterile vials containing 150 u.s.p.u. sodium heparin and kept at RT during transportation to the laboratory.

For each animal several T25 culture flasks containing 10 ml of RPMI 1640 medium supplemented with 20% fetal calf serum, 1% L-glutamine, 1% pokeweed mitogen and 1% antibiotic/antimycotic were inoculated with 1 ml of blood. The flasks were incubated at 37°C and in a 5% CO₂ atmosphere. Half of them were harvested after 72 hours and half after 96 hours. In the first experiments colcemid was added to a final concentration of 0.02 µg/ml during the last 1.5 h of incubation. Since we observed that colcemid was responsible for shortening of the chromosomes, later the cells were harvested without adding colcemid. The cultures were agitated to ensure that the cells were in suspension before their transfer into 15 ml centrifuge tubes. After centrifugation at 160-g for 10 min at RT, the supernatant was removed using a Pasteur pipette. The hypotonic treatment was performed by resuspending the pellet in 10 ml of 0.075 M KCl which had been prewarmed to 37°C and by subsequent incubating for 30 min at 37°C. The tube was then centrifuged at 300-g at RT, the supernatant was removed by leaving about 0.5 ml of fluid to resuspend the pellet. This suspension was transferred into a new 15 ml centrifuge tube containing 10 ml of freshly prepared ice-cold fixative (methanol:glacial acetic acid, 3:1). The cell suspension was kept on ice for at least 10 min, allowing further lysis of the erythrocytes. The fixation step was repeated 6 to 8 times by rounds of centrifugation, removal of supernatant, resuspension of pellet and addition of new fixative. The thoroughness of fixation was partly responsible for the quality of the final preparations. A high number of fixations resulted in cytoplasm free specimens, a very important requirement for ISH. After the last fixation the cells were
resuspended in a small volume of fixative and dropped onto slides as described in 3.1.3.

3.1.2. Chromosomes from fibroblasts

Primary fibroblasts, established from tissue of male bovine fetuses, were cultured in 5 ml of improved alpha-medium supplemented with 10% fetal calf serum, 1% L-glutamine and 1% Na-pyruvat. The cell cultures were incubated in T25 culture flasks at 37°C with 5% CO₂ for some days. Confluent cells in one T25 flask were split in 2-4 T75 culture flasks containing 12 ml of supplemented alpha-medium. Cells were harvested for chromosome preparations when the number of round mitotic cells peaked.

After removing the medium with a Pasteur pipette, the cells were rinsed with 3 ml of PBS and removed with 1 ml of Trypsin/EDTA. 9 ml of new medium were added and the cell suspension was transferred into 15 ml centrifuge tubes. After centrifugation at 70-g for 8 min at RT, the supernatant was removed and the pellet resuspended in 10 ml of 0.075 M KCl. Hypotonic incubation was for 30 min at 37°C. The cells were then centrifuged again at 110-g for 10 min and fixed several times with methanol:acetic acid, 3:1, as described in 3.1.1.

In order to obtain longer chromosomes, a shorter but stronger hypotonic treatment was in some cases performed. By reducing the time of the hypotonic treatment it was possible to consistently obtain longer chromosomes.

Instead of 0.075 M KCl, 10 ml of 0.02 M or 0.05 M KCl were added to the cell suspension, which was incubated for a total of 13 min at 37°C. After 8 min of incubation, 1 ml of fixative was added, so that the cells could proceed for other 5 min of hypotonic treatment in presence of fixative. This improved the quality of the fixation in the subsequent step.

3.1.3. Slide preparation

Microscope slides of 26 x 76 mm were cleaned by leaving them overnight in a solution of 2N HCl and subsequent rinsing for 5 to 10 min with running deionized H₂O, then they were stored in ddH₂O at 4°C.

Using a micropipette, 15-18 μl of cell suspension were dropped from about 3 cm of distance onto cleaned and wet slides, held by forceps at an angle of 45°. Slides were left for short time on a rack at RT, and before drying completely they were rinsed with fresh fixative. The concentration of cells was checked under a phase contrast microscope and adjusted by diluting the cell
suspension with a little more fixative if it was too high or resuspending the pellet after centrifugation in a smaller volume of fresh fixative if only few cells were visible. Unused cell suspension was stored at -20°C in a closed vial. Before using the cells at a later date they were spun down and resuspended in fresh fixative.

The slides were kept overnight at RT and then they were stored at -20°C in sealed boxes containing silica gel until use. Preservation and hardening of the nuclear material was essential for successful ISH and was optimally accomplished by storing the preparations in frozen state.

3.1.4. QFQ-banding

QFQ-banding is an abbreviation for "Q-bands by fluorescence using Quinacrine" and indicates a method which was introduced by Caspersson and coworkers (Caspersson et al., 1968; 1969; 1970) and is the oldest of the modern banding techniques, as well as one of the simplest and most reproducible methods.

Q-bands are produced by the fluorescent dyes Quinacrine (Q) or Quinacrine Mustard (QM). They produce an identical pattern. However, the histochemical background is not fully understood. A review about historical studies performed to understand the Q-banding mechanism is given by Sumner (1990).

There are two separate considerations when interpreting the banding: the mechanism by which the dye binds to the chromosomes and the factors affecting their fluorescence. The primary mode of binding of QM to DNA is by intercalation, the second covalent by alkylation of guanine. Selander and de la Chapelle (1973) showed that QM could also react with different histones, and speculated that this could be a factor in causing the bands. Moreover, there is evidence for more efficient excitation of fluorescence in some chromosomal regions than in others: the fluorescence is brighter when the dye is bound to AT-rich DNA than when it is bound to GT-rich DNA.

QFQ-banding was chosen here because of its high reproducibility and its low interference with subsequent ISH.

Q-bands were done prior to ISH by dipping the slides in a solution of 0.005% Quinacrine Mustard in sterile ddH₂O for 30 sec at RT. The slides were then thoroughly rinsed in deionized H₂O and air dried for a short time. Few drops of Sørensen’s phosphate buffer at pH 6.8 were used to overlay the slides that were covered with a coverslip of 24 x 60 mm. Well spread metaphases, with
distinctive banding pattern and non-overlapped chromosomes, were photographed with a fluorescence microscope as described in 3.5. After photography the chromosome preparations were washed by dipping the slides for 2 min in xylol. Thereafter, they were ready to be processed for ISH.

3.2. Preparation of cosmid probes

3.2.1. Source of cosmids

A partial bovine cosmid library of recombinant E. coli (XL1-Blue strain) was constructed (Eggen, 1992). Cosmids were from type SuperCos 1 (Stratagene); instead of one, this vector contains two cos sites allowing cloning of very small amounts of nonsized selected DNA and eliminating the need to dephosphorylate the cosmid. The vector contains the primers T3 and T7.

Cosmids containing (CA)n-microsatellites were identified by screening the library with a synthetic (dC-dA)n(,dG-dT)n oligonucleotide probe (Pharmacia). Positive clones were selected, multiplied and stored at -80°C until use (Steffen, 1992).

3.2.2. Growth of bacteria culture

Agar medium was prepared with autoclaved LB medium (s. 2.3.). After cooling to 48°C, the medium was supplemented with ampicillin to a concentration of 50 mg/l, about 30 ml were cast in a Petri dish and allowed to gelify at RT. Petri dishes were stored at RT and used within few days.

Bacteria containing recombinant cosmids, stored at -80°C, were streaked on the agar plate with a loop under sterile conditions and grown up overnight at 37°C.

Single colonies were picked with a sterile toothpick and inoculated in glass tubes containing 5 ml of LB medium supplemented with ampicillin to a final concentration of 50 mg/l. Precultures were shaken at 250 rpm for about 12 hours at 37°C. When the precultures appeared turbid, indicating that the bacteria had grown to saturation, 1 ml of each preculture was transferred in 100 ml of LB medium, without ampicillin, in a 500 ml-Erlenmeyer flask and shaken overnight at 37°C.

3.2.3. Qiagen cosmid midi-preparation

Cosmid preparations were carried out using the columns Qiagen-tip 100, according to the protocol of the manufacturer (Qiagen, 1990). This method was
used to prepare DNA for ISH purposes, but such DNA was not suitable for sequencing.

Briefly, the volume of 100 ml of culture preparation was divided in two 50 ml Falcon tubes and centrifuged at 2500-g for 15 min at 4°C. Each bacterial pellet was resuspended in 2 ml of P1 buffer containing 100 µg/ml RNase A. The tubes were kept on ice. Breaking of the bacterial cell walls was achieved by adding 4 ml of P2 buffer to each Falcon tube, by mixing gently and incubating for 5 min at RT. After addition of 4 ml of P3 buffer containing 2.55 M KAc, the content was mixed immediately but gently. The tubes were centrifuged at 15000-g for 40 min at 4°C and the chromosomal DNA precipitated. The supernatant was removed and kept on ice until loading on the column. In the meantime the tip 100 column was equilibrated with 3 ml of QBT buffer, allowing it to empty by gravity flow. The supernatant was loaded on the column and allowed to enter the resin. A washing step followed with 10 ml of buffer QC. The DNA was eluted with 5 ml of buffer QF. It was then precipitated with 3.5 ml of isopropanol at RT and centrifuged at 3500-g for 30 min at 4°C. The DNA pellet was washed with 5 ml of 70% ethanol, air dried for 20 min, redissolved in 100 µl of TE (pH 8.0) and stored at -20°C.

The DNA yield from a 100 ml culture amounted from 20 to 200 µg.

Cosmid preparation for sequencing was according to a protocol supplied with the sequencing kit (Applied Biosystems Inc.).

3.2.4. Probe labelling

Several procedures like PCR labelling, end labelling or photoreactions are available to label the DNA, but labelling by Nick Translation and by Random Priming are the most efficient and routinely used methods for ISH.

3.2.4.1. Nick Translation

Nick translation is a frequent method for labelling probes to be used in non isotopic in situ procedures. The protocol is a modification of the original one developed by Rigby et al. (1977). It is based upon the activities of the DNase I and of the E. coli DNA polymerase I. Pancreatic DNase I introduces scissions or nicks into the DNA molecules. The exonuclease activity of polymerase I acts at the sites of the nicks and removes nucleotides in the 5'→3' direction. At the same time, using the 3'-hydroxyl group of the terminal nucleotide of the nick as a primer, the polymerase activity of polymerase I replaces the pre-existing
unlabelled nucleotides with labelled nucleotides. The simultaneous activities of the polymerase I result in a movement of the nick along the DNA, therefore the designation "nick translation".

In a standard reaction, 250-300 ng of template DNA were labelled with biotin-11-dUTP or biotin-16-dUTP in 20 μl of final volume, by using a Nick Translation Kit (Boehringer Mannheim). Labelling with biotin-11-dUTP (Sigma) was possible without addition of dTTP, but biotin-16-dUTP (Boehringer Mannheim) required additional dTTP in the reaction. This topic is discussed in more detail in 4.1.2.1.

A standard reaction was as follows:

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>ddH₂O</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>dATP (0.4 mM)</td>
<td>3 μl</td>
<td>(60 μM)</td>
</tr>
<tr>
<td>dGTP (0.4 mM)</td>
<td>3 μl</td>
<td>(60 μM)</td>
</tr>
<tr>
<td>dCTP (0.4 mM)</td>
<td>3 μl</td>
<td>(60 μM)</td>
</tr>
<tr>
<td>bio-11-dUTP (1 mM)</td>
<td>1.2 μl</td>
<td>(60 μM)</td>
</tr>
<tr>
<td>10 x buffer</td>
<td>2 μl</td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>250-300 ng</td>
<td></td>
</tr>
<tr>
<td>enzyme mixture</td>
<td>2 μl (DNA polymerase I at 1 Ku/ml, DNase I at 0.8 u/ml)</td>
<td></td>
</tr>
<tr>
<td>DNase I</td>
<td>1 μl (0.008 u/μl)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20 μl</td>
<td></td>
</tr>
</tbody>
</table>

or instead of bio-11-dUTP: bio-16-dUTP (1 mM) 1 μl (50 μM) + dTTP (0.4 mM) 0.5 μl (10 μM)

This mixture was incubated for 90 min in a 14°C water bath. For enzyme inactivation 2 μl of 0.5 M EDTA were added and heating to 65°C for 10 min was performed. Purification of the probe followed as described in 3.2.4.3. The labelling efficiency was checked either with a radioactive tracer or by dot blot assay (s. 3.2.4.4.).

Advantages of this technique are an uniform labelling through the DNA molecule and the possibility to adjust the fragment size by varying the amount of DNase I. A requirement for successful ISH is in fact the correct size of the probe after the labelling reaction. Lawrence and Singer (1985) suggested that the
labelled probe should be between 200 and 400 nucleotides in length. Biotinylated probes larger than 500 nucleotides have difficulty penetrating the specimen and have a tendency to stick to both the glass and the cellular material, resulting in high background (P. Lichter, personal communication).

Control of the fragment size is particularly important when large probes such as cosmids are labelled. In order to optimize the labelling protocol for these probes, three Nick Translation reactions were carried out by adding $3.2 \times 10^{-3}$ u of DNase I in the first, $8 \times 10^{-3}$ u in the second and $16 \times 10^{-3}$ u in the third experiment, respectively. $18 \mu$Ci of $^{32}$P-dATP, corresponding to 6 pmoles, were added as a tracer. After incubation, the three labelled fractions were run on a sequencing polyacrylamide gel, and the fragment size was estimated by running in parallel a M13 sequence ladder. In the first reaction the low amount of enzyme resulted in DNA fragments of about 350 bp of length. DNase I in the concentration of $16 \times 10^{-3}$ u caused degradation of the DNA. Fragments of about 180 bp were generated with the use of $8 \times 10^{-3}$ u (fig. 3). This last amount of DNase I, although generating fragments lightly to short according to the literature, was used in all the subsequent labelling reactions, because ISH experiments revealed them to be the most appropriate.
3.2.4.2. Random Priming method

First described by Feinberg and Vogelstein (1983; 1984) the random priming DNA labelling method was applied mainly when very small amounts of probe were available or when probes other than cosmids were to be labelled.

The sample was labelled in the following standard reaction by using a Random Primed Kit (Boehringer Mannheim):

Fig. 3: Polyacrylamide gel showing the size of DNA fragments after nick translation of the probe. Left: M13 plasmid run as size marker. Lanes a,b,c) nick translation with the addition of $3.2 \times 10^{-3}$ u of DNase I, three different amounts of DNA are loaded on the gel. d,e,f) in analogous manner, with $8 \times 10^{-3}$ u. g,h,i) with $16 \times 10^{-3}$ u.
ddH₂O  
X

<table>
<thead>
<tr>
<th>ddATP (0.5 mM)</th>
<th>2 µl (50 µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>dGTP (0.5 mM)</td>
<td>2 µl (50 µM)</td>
</tr>
<tr>
<td>dCTP (0.5 mM)</td>
<td>2 µl (50 µM)</td>
</tr>
<tr>
<td>bio-11-dUTP (1 mM)</td>
<td>2 µl (50 µM)</td>
</tr>
<tr>
<td>reaction mixture</td>
<td>2 µl</td>
</tr>
<tr>
<td>DNA</td>
<td>20-30 ng</td>
</tr>
<tr>
<td>kleenow enzyme</td>
<td>2 µl (1u/µl)</td>
</tr>
</tbody>
</table>

20 µl

As for nick translation, biotin-11-dUTP could be substituted from:

<table>
<thead>
<tr>
<th>bio-16-dUTP (1 mM)</th>
<th>0.8 µl (40 µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>dTTP (0.5 mM)</td>
<td>0.4 µl (10 µM)</td>
</tr>
</tbody>
</table>

The DNA, first mixed with the requested amount of H₂O, was denatured by heating for 10 min in boiling H₂O and following cooling on ice. The other reagents were added on the same tube on ice, to a final volume of 20 µl. The tube was left overnight at RT and the following morning the reaction was stopped with 2 µl of 0.5 M EDTA and by heating to 65°C for 10 min. The probe was purified as described under 3.2.4.3. The labelling efficiency was checked either with a radioactive tracer or by dot blot assay (s. 3.2.4.4.).

3.2.4.3. Purification of the labelled DNA

Prior to its use as hybridization probe, the DNA has to be separated from unincorporated nucleotides, otherwise considerable levels of background would occur.

Labelled DNA was separated from unincorporated nucleotides through sephadex G-50 columns (Pharmacia). The columns were rinsed and equilibrated with 3 ml of TNE, and then loaded with 20 µl of labelling reaction. By elution with 400 µl of TNE for each fraction, 5 fractions were collected in microcentrifuge tubes.
3.2.4.4. Methods to check the extent of biotin incorporation

a) with radioactive tracer

During the optimization of the protocols, the labelling efficiency was determined from the measurement of the incorporation of 6.5 μCi of ³H-dATP (100 pmol), added to the reaction. Since ³H-dATP was supplied in an ethanol solution, which interferes in enzymatic reactions, the nucleotide had to be lyophilized in a vacuum centrifuge and the other reagents were added to the lyophilized pellet. When ³H-dATP was added to the reaction, cold dATP was usually not included or a only small trace was added to the reaction in the molarity ratio as described for dTTP and biotin-16-dUTP (s. 3.2.4.1.).

After sephadex separation (s. 3.2.4.3.), 1 μl of each fraction was added to a tube containing 3 ml of scintillation-cocktail and the dpm value (disintegrations per minute) of the different fractions was measured with a β-Counter. The distribution of the radioactivity content of the fractions was analyzed; the absence of a clear peak denoted an unsuccessful labelling reaction, while a clear high peak in the second fraction indicated an accumulation of suitably labelled DNA fragments.

This method, beyond the known disadvantages related to the use of radioactivity, is an indirect evaluation, assuming an efficient biotin incorporation only when efficient incorporation of a hot nucleotide is determined. However, this correlation is not always true, as for instance in the case of degraded hot nucleotides, or when to little DNA is present in the reaction for the incorporation of both radioactivity and biotin.

b) with dot blot analysis

Although laborious, this approach is more suitable because it visualizes directly the quantity of biotin molecules present in the DNA probe. It consists of an enzymatic assay based upon the binding between biotin and streptavidin conjugated to alkaline phosphatase (SA-AP). The substrate for the enzyme is BCIP (5-bromo-4-chloro-3-indolyl phosphate) combined with NBT (nitro blue tetrazolium). The enzyme catalyzes the reaction by producing the precipitation of the reaction product, which is blue.

The test was performed using the BluGENE Kit (BRL) according to the protocol recommended by the supplier. Standard biotinylated DNA was prepared at concentrations of 3, 8, 15 and 30 pg/μl in 6xSSC containing 0.1 mg/ml sheared salmon sperm DNA. Test DNA was diluted in the same fashion.
and 1 µl of each dilution was spotted on a nitrocellulose filter. Fixation of DNA to the filter was done by baking it for 1 hour at 80°C. Afterwards the filter was rehydrated in AP 7.5 buffer for 1 min and blocking was performed for 1 hour at 37°C in blocking buffer (3% BSA in AP 7.5 buffer). A solution of SA-AP at a concentration of 1 µg/ml in AP 7.5 buffer was prepared, in which the filter was incubated for 30 min at 37°C. After washing 2 x 5 min in AP 7.5 buffer and 1 x 10 min in AP 9.5 buffer, the detection solution was prepared by adding 33 µl of NBT to 7.5 ml of AP 9.5 buffer and by mixing carefully, followed by 25 µl of BCIP again mixed gently to avoid precipitation. The filter was incubated in this solution in the dark until color development was satisfactory. Usually the reaction could be stopped after 15 min by washing the filter in TE pH 7.0 and air drying. It was assumed that if the probe was at least half as intense as the standard DNA at the corresponding dilution, it was suitable for ISH.

3.3. In situ hybridization

3.3.1. Chromosomal in situ suppression (CISS)

When in situ hybridization is to be performed with large insert clones or chromosome specific libraries, it must be considered that the DNA probe contains, in addition to unique sequences, interspersed repetitive elements (IRS). The application of a standard procedure results in hybridization signals distributed over the whole chromosomes, due to the ubiquitous presence of IRS throughout the genome. This phenomenon is called "cross hybridization", interfering often very strongly with the hybridization efficiency. The portion of the signal caused by IRS within a probe must be suppressed, to allow the unequivocal detection of the specific signal produced by the target site specific single copy sequences. The suppression hybridization technique, called Chromosomal In Situ Suppression (CISS) (Landegent et al., 1987; Lichter et al., 1988b), utilizes a partial reannealing step to reduce the number of labelled repetitive sequences which are available to hybridize to the specimen. Labelled probe and an excess of unlabelled competitor DNA, highly enriched in repetitive sequences, are denatured together. Thereafter, preannealing is allowed: IRS of the probe hybridize rapidly with IRS of the competitor DNA, while unique sequences remain single stranded, available to hybridize to the chromosomes (Sealey et al., 1985). The larger the complexity of the probe (chromosome specific library>YAC>cosmid>phage) the greater the amount of competitor needed.
The application of this procedure prevents non specific signals mainly through two events:
1) preannealed double stranded IRS DNA cannot hybridize to the chromosomes
2) not preannealed IRS of the competitor, which is in large excess and not labelled, hybridize with IRS of the fixed chromosomes and prevent hybridization of remaining single stranded IRS elements of the probe.

Although every DNA containing sufficient repetitive sequences can be used, today two types of competitor DNA are used: total genomic competitor DNA and Cot-1 DNA. Total genomic DNA is easy to prepare and gives usually satisfactory results, but when big amounts of competitor are needed, in addition to IRS, site specific signals could be suppressed as well. To avoid this risk, Cot-1 DNA is an alternative.

Some particularly interesting applications of this new technique are reviewed by Cremer et al., 1988; Pinkel et al., 1988 and Lichter et al., 1990.

3.3.1.1. Preparation of genomic competitor DNA
A big amount of unlabelled genomic DNA can serve as competitor if sheared in small fragments.

This type of competitor was prepared from calf thymus in the following way.

A small piece of frozen thymus was powderized in liquid N2 and resuspended in 10 ml of TNE at pH 8.0. Addition of 500 μl of 10% SDS disrupted the cell membranes, protein digestion was done by adding 3 times 40 μl of 20 mg/ml proteinase K and by incubating from 1 hour to overnight at 50°C. Phenol/chloroform extraction and ethanol precipitation were then performed. Contamination of RNA was eliminated by treating the DNA with 100 μl of RNaseA at 10 mg/ml for 30 min at 37°C. Proteinase K digestion and phenol extraction were repeated, the DNA was then ethanol precipitated and resuspended in 2-3 ml of TE at pH 8.0. The DNA was sonicated 9 times for 10 sec at speed 6.5 of a Micro Tip Limit sonicator. Between bursts of sonication, the tube was placed on ice to cool. The DNA size was checked on a 1% agarose gel to estimate if the bulk of the fragments was between 200 and 400 bp of length. Sonication was repeated if necessary. Alternatively, shearing of the DNA was performed by passing it few times through a 17-gauge hypodermic needle.

The DNA concentration was determined by measuring its optical density at 260 nm.
3.3.1.2. Preparation of Cot-1 competitor DNA

The Cot-1 DNA is a DNA highly enriched for repetitive sequences. Its preparation is based upon the renaturation kinetic properties first described by Subirana and Doty (1966). In 1968 Britten and Kohne proposed "Cot" as a measure for the rate of DNA reassociation.

The reassociation of two DNA strands is a function of the initial DNA concentration and of the time of renaturation:

\[ \text{Cot-1} : \ Co \times t = 1; \]

\( C_0 \) = concentration of the DNA in moles of nucleotides per liter at the time 0
\( t \) = time of reassociation in seconds

The preparation of the Cot-1 competitor DNA is based upon the principle that the repetitive sequences reassociate faster than the unique sequences. After the appropriate time, the remaining single stranded DNA is digested with S1 nuclease, a single-strand specific endonuclease which cleaves DNA and RNA to yield 5'-phosphate mono- or oligonucleotides while double stranded DNA is relatively resistant (Ando, 1966; Vogt, 1973).

For the preparation of Cot-1 competitor, total genomic DNA was extracted from calf thymus as described in 3.3.1.1. Few mg of DNA were sheared in pieces of 500 to 2000 bp of length by sonicating for 90 sec at speed 6.5 with a Micro Tip Limit sonicator and by testing the fragment size on a 1% agarose gel stained with ethidium bromide and photographed under UV-light. The DNA was then diluted with S1 nuclease buffer to a concentration of 0.25 \( \mu \text{g}/\mu\text{l} \). Denaturation was done by heating the DNA in boiling \( \text{H}_2\text{O} \) for 15 min. Afterwards, reassociation of the repetitive sequences was allowed by incubating the DNA at 63°C during 22 min, as shown by the following calculation:

\[ 0.25 \text{ mg DNA/ml} = 7.57 \times 10^{-4} \text{ M} \]

\[ \text{Cot-1: } 1 = 7.57 \times 10^{-4} \times \text{ x sec} \]

\[ \text{sec} = 1321 \text{ (= 22 min)} \]

The temperature of the tube was lowered to about 37°C and digestion of remaining single stranded DNA was performed for 2 min with 0.08 unit of S1 nuclease per each \( \mu \text{g} \) of DNA. The reaction was stopped with 10% (v/v) of 0.5 M EDTA and the tube was put on ice. To separate the double stranded repetitive
DNA from digested fragments, the DNA was phenol extracted and ethanol precipitated.

Since bovine genomic DNA contains about 40\% of repetitive sequences, the amount of remaining DNA, measured at the spectrophotometer, should represent approximately this percentage. Variations were explained as variabilities due to the laboratory manipulations.

3.3.1.3. DNA precipitation and chromosomal in situ suppression

When a new cosmid probe had to be handled, different conditions of CISS were tried. The total amount of 400 \( \mu \text{l} \) of labelled probe (s. 3.2.4.3.), corresponding to about 200 ng of DNA, was sufficient for 4 hybridizations. This amount of DNA was therefore divided in 4 tubes. DNA from an unrelated species, salmon sperm DNA, was added to reduce nonspecific binding of labelled probe to chromatin, cell debris and glass. The other components were added as follows:

<table>
<thead>
<tr>
<th>Conditions</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>probe in TNE</td>
<td>100 ( \mu \text{l} ) *</td>
<td>100 ( \mu \text{l} ) *</td>
<td>100 ( \mu \text{l} ) *</td>
<td>100 ( \mu \text{l} ) *</td>
</tr>
<tr>
<td>competitor DNA</td>
<td>0 ( \mu \text{g} )</td>
<td>1.5 ( \mu \text{g} )</td>
<td>3.0 ( \mu \text{g} )</td>
<td>6.0 ( \mu \text{g} )</td>
</tr>
<tr>
<td>salmon sperm DNA</td>
<td>10 ( \mu \text{g} )</td>
<td>8.5 ( \mu \text{g} )</td>
<td>7.0 ( \mu \text{g} )</td>
<td>4.0 ( \mu \text{g} )</td>
</tr>
<tr>
<td>100% ethanol, cold</td>
<td>400 ( \mu \text{l} )</td>
<td>400 ( \mu \text{l} )</td>
<td>400 ( \mu \text{l} )</td>
<td>400 ( \mu \text{l} )</td>
</tr>
</tbody>
</table>

\* 100 \( \mu \text{l} \) = 50 ng DNA

The tubes were put at -80\(^\circ\)C for at least 2 hours. They were centrifuged at 12500·g for 30 min at 4\(^\circ\)C, the supernatant was removed and the pellet washed with 300 \( \mu \text{l} \) of 70\% ethanol. The pellet was allowed to dry at RT and was then resuspended in 10 \( \mu \text{l} \) of hybridization mixture containing 10\% dextran sulphate, 2xSSC, 50\% formamide and 1\% Tween 20. The labelled probe in the hybridization mixture could be stored at -20\(^\circ\)C for few days. For longer storage resuspension in TE pH 8.0 (50 ng DNA/2 \( \mu \text{l} \) TE) was more suitable. The DNA in 2 \( \mu \text{l} \) of TE was later combined with 8 \( \mu \text{l} \) of hybridization mixture.
3.3.2. Hybridization on slides

Prior to denaturation the slides were treated with 150 μl of RNase A (100 μg/ml in 2xSSC) under a coverslip for 1 hour at 37°C. Thereafter, they were washed 3 x 5 min in 2xSSC, dehydrated in a series of 70%, 80% and 95% ethanol, 3 min each at RT and were allowed to dry.

The chromosome preparations were denatured for 2 min in a bath of 70% formamide and 2xSSC at 70°C. After denaturation the slides were immersed immediately in ice-cold 70% ethanol, followed by 80%, 90% and 100% ethanol baths for 3 min each, then they were allowed to dry.

The probe, dissolved in 10 μl of hybridization mixture, was denatured for 5 min at 75°C and reannealed for 12 min at 37°C. It was then applied to the slides under a 18 x 18 mm coverslip in the previously photographed region, as soon as possible. Usually no more than 30 min were left between denaturation of the chromosomes and hybridization.

Hybridization was carried out overnight at 37°C in a moist chamber sealed with parafilm.

3.3.3. Signal detection

After hybridization overnight, the coverslips were removed in the first of three washes in 50% formamide, 2xSSC at 43°C, 5 min each. Three subsequent washes in 2xSSC at 43°C, 5 min each, were also performed. Beyond this step the slides were never allowed to dry.

The slides were immersed for a minimum of 15 min in 4xSSC, 0.05% Tween 20 to equilibrate them to a lower stringency. This solution could be also used for storing the slides and for further washing steps, therefore it will be recalled as wash buffer.

Immediately after this equilibration, the slides were preincubated in 3% BSA, 4xSSC and 1% Tween 20 (preincubation buffer) for 45 min. After draining, the detection of the hybridization signal was performed by applying to each slide 60 μl of avidin-FITC (avidin conjugated with fluorescein isothiocyanate) at 5 μg/ml in preincubation buffer. The slides were covered with a 24 x 60 mm coverslip and incubated in a moist chamber for 45 min at 37°C. To prevent a possible decrease of the fluorescence, from here forwards a direct exposure of the slides to the light was avoided by wrapping the Coplin jars containing the slides with alu-foil.

After the first incubation the slides were washed 3 x 5 min in the wash buffer at RT or at 42°C, depending on the background to be removed.
For all probes the intensity of fluorescence was amplified by applying to each slide 60 µl of biotinylated goat anti-avidin antibody at a concentration of 5 µg/ml in preincubation buffer and by incubating them at 37°C for 45 min. After washing as before, this step was followed by a second layer of avidin-FITC and another washing step. This "sandwich" approach, has been first described by Pinkel et al. (1986). They estimated that each amplification step produced a 6-fold increase in signal intensity. In some cases more than one round of amplification was necessary for an optimal intensity of the hybridization signal.

### 3.3.4. Chromosome counterstaining

After the first amplification, the slides were carefully rinsed with deionized H₂O, allowed to dry and then stained for 5 min in a solution of propidium iodide at a concentration of 200 ng/ml in 2xSSC. Again, rinsing and drying followed.

To prevent fading of the signal, preparations were sealed with three drops of antifade solution (Johnson and Araujo, 1981) under a coverslip. Antifade was prepared as described in 2.3. and aliquots were stored at -20°C.

### 3.3.5. Two probes/one colour in situ hybridization

In few cases two probes were hybridized simultaneously on the same specimen to confirm or exclude the belonging of the probes to the same chromosome. In particular, this was done with the probes CAP8 + c10BT33 and with c10BT33 + a probe for the NORs.

The three probes were labelled with biotin as already described.

CAP8 and c10BT33 were ethanol precipitated with competitor DNA, the NORs gene without competitor as following:

<table>
<thead>
<tr>
<th></th>
<th>CAP8</th>
<th>c10BT33</th>
<th>NORs genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>probe in TNE</td>
<td>200 µl *</td>
<td>200 µl *</td>
<td>200 µl *</td>
</tr>
<tr>
<td>competitor DNA</td>
<td>6 µg</td>
<td>6 µg</td>
<td>-</td>
</tr>
<tr>
<td>salmon sperm DNA</td>
<td>14 µg</td>
<td>14 µg</td>
<td>25 µg</td>
</tr>
<tr>
<td>100% ethanol, cold</td>
<td>800 µl</td>
<td>800 µl</td>
<td>800 µl</td>
</tr>
</tbody>
</table>

* 200 µl = 100 ng DNA
Each pellet was then resuspended in 10 µl of hybridization mixture. 5 µl of CAP8 + 5 µl of c10BT33 were combined in a tube and very well mixed. Denaturation and reannealing followed as usually. Hybridization was carried out under a 18 x 18 mm coverslip.

5 µl of c10BT33 in one tube and 5 µl of NORs in another tube were denatured separately. The probe c10BT33 was then allowed to reanneal for 12 min at 37°C, while the NORs probe was quickly put on ice after denaturation, to avoid the reassociation of the repetitive sequences. Then the two probes were mixed and hybridized on the slide as above.

**3.3.6. Two probes/double colour in situ hybridization**

During an EMBO Course in Heidelberg, the two probes CAP3 and c10BT34 were hybridized simultaneously in a double colour hybridization experiment. Since the clones had been previously localized on the same chromosome very close to each other, the aim was to exactly determine their position relative to the telomere.

CAP3 was labelled with biotin-11-dUTP as described in 3.2.4.1. and c10BT34 was labelled by nick translation with digoxigenin-11-dUTP as follows:

| DNA | 2 µg |
| 10 x buffer | 10 µl |
| 0.1 M -mercapto | 10 µl |
| dNTP | 10 µl |
| DNase I (1:1000) | 3 µl |
| DNA polymerase I | 2 µl |
| ddH2O | X |

\[ \text{100 µl} \]

dNTP: Digoxigenin-11-dUTP (1 mM) 12.5 µl (125 µM)  
dTTP (10 mM) 3.8 µl (380 µM)  
dATP, dCTP, dGTP (10 mM each) 5 µl (500 µM)  
\[ 78.7 \text{ µl} \]  
\[ \text{ddH2O} \]  
\[ 100 \text{ µl} \]
This mixture was incubated for 2 hours in a 14°C water bath. Enzyme inactivation was performed by adding 3 μl of 0.5 M EDTA, 1 μl of 10% SDS and by heating to 68°C for 15 min.

The labelled probe was separated from unincorporated nucleotides as described in 3.2.4.3.

DNA precipitation, CISS and hybridization were as described for the other probes (s. 3.3.), except that 60 ng of labelled CAP3 were mixed with 60 ng of labelled c10BT34 and hybridized together under the same 18 x 18 mm coverslip. The detection of the hybridization signals was performed in three steps. The first step consisted in a layer of avidin-FITC, which bound to the biotinylated probe CAP3. After washing, the second layer was the amplification of the fluorescent signal as already described (s. 3.3.3.). The third step consisted in a mixture of avidin-FITC (1:200) and of sheep anti-digoxigenin antibody conjugated with TRITC (1:75). Avidin-FITC enhanced the signal intensity of CAP3 and the conjugated antibody bound to c10BT34.

After washing, counterstaining of the chromosomes followed in a solution of DAPI at a concentration of 50 ng/ml in 2xSSC, by agitation for 15 min at RT. Rinsing, drying and mounting were as already described.

The evaluation of the ISH was performed in this particular case with a Zeiss microscope. The signal of each probe was detected separately with the appropriate filter and the two images saved, overlapped and processed with a CCD-camera (s. 4.2.1.1.).

3.3.7. Determination of the hybridization site

To identify the chromosome pair carrying the hybridization signal, duplicate photographs of QFQ-banded and hybridized metaphases were compared and the labelled chromosome identified according to ISCNDA, 1989.

The distance between the centromere and the hybridization signal, as well as the total length of the chromosome were measured. The ratio between both of these measures gave the "fractional length centromere" (FL<sub>cen</sub>), a value to express the position of the probe relative to the centromere.

3.4. Microscopic evaluation

A Leitz Diaplan microscope equipped with a fluorescent Hg lamp of 50W was used for most of the microscopic analyses, with exception of 3.3.6.
Methods

Unstained chromosome preparations were evaluated with a phase contrast objective of 20x magnification.

Fluorescent metaphases were observed and photographed with a 63x/1.30 oil objective. Table 3 lists excitation and emission wavelengths of the fluorescent dyes which have been used.

Table 3: Excitation and emission wavelengths of fluorescent dyes

<table>
<thead>
<tr>
<th>Fluorochrome</th>
<th>Excitation max (nm)</th>
<th>Emission max (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quinacrine Mustard</td>
<td>455</td>
<td>495</td>
</tr>
<tr>
<td>Propidium Iodide</td>
<td>330 and 520</td>
<td>620</td>
</tr>
<tr>
<td>Fluoresceinisothiocyanate</td>
<td>490</td>
<td>520</td>
</tr>
</tbody>
</table>

The filter block E3 allowed the detection of quinacrine mustard, the I 2/3 excited propidium iodide and FITC. The filter block K3 was selective for FITC but did not detect the counterstain propidium iodide. The main characteristics of these filters are summarized in table 4.

Table 4: Filter blocks used for the detection of fluorescent dyes

<table>
<thead>
<tr>
<th>Leitz filterblock</th>
<th>E3</th>
<th>I2/3</th>
<th>K3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band pass filter (BP)</td>
<td>436/7</td>
<td>450/490</td>
<td>470/490</td>
</tr>
<tr>
<td>Reflection short-pass filter (RKP)</td>
<td>475</td>
<td>510</td>
<td>510</td>
</tr>
<tr>
<td>long pass filter (LP)</td>
<td>490</td>
<td>515</td>
<td>515</td>
</tr>
</tbody>
</table>

3.5. Photography

Photography was carried out using a Leitz Orthomat E camera system with a 35 mm camera.

For QFQ-banded chromosomes it was important to detect the different staining intensities between bands. This was achieved with the high contrast black and white films Kodak Technical Pan 2415. Good results were obtained by 25 to 27 DIN settings, corresponding to exposure times of approximately 10 to 20 sec. Technical Pan 2415 were developed with D19, 4 min at RT, and fixed 4 min with Agefix 1+7.
Ektachrome 160 ASA films were used to photograph metaphases carrying the hybridization signal on color slides, films Kodak Gold 100 were used for colour prints and Agfa Ortho 25 professional for black and white prints. Agfa Ortho 25 has a very high resolving power and extremely fine grain. It is a red insensitive film, allowing an optimal detection of the FITC-yellow/green signal but without detecting eventual differences in the red staining of the chromosomes, caused by propidium iodide. Exposition times of up to 3-4 min were required. Agfa Ortho films were developed 5 min with Rodinal 1+10 at RT and fixed 4 min with Agefix 1+7.

To make prints, negatives were enlarged with a Durst Laborator 900 enlarger. Agfa Brovira Speed paper of different hardness was used according to the contrast required; usually paper of hardness number 2 was chosen for QFQ-banding and of hardness number 3 for propidium iodide/FITC pictures. Paper was developed with an Agfaprint BW 37-50 machine using Neutol 1+7 developer and Agefix 1+3.

The microscope enlargement of the chromosomes on the negative was calculated in the following way:

\[
\text{enlargement} = \text{objective magnific.} \times \text{ocular magnific.} \times \text{tube factor} \times \text{camera factor}
\]

The Leitz Diaplan microscope has a ocular magnification of 12.5 and a tube factor of 0.8. The camera system has a factor of 0.32.

3.6. Automated Sanger sequencing

Automated sequencing was performed with the ABI 373 A DNA sequencer (Applied Biosystems Inc.). The protocol was based on the dideoxy termination method (Sanger et al., 1977) and according to a manual supplied with the Taq cycle sequencing Kit (Applied Biosystems, 1992).

The templates were recombinant pBluescript II SK+ (Steffen, 1992) subcloned from cosmids. Primers specific for the T3 and T7 promoter sequences of pBluescript were used for sequencing.
3.6.1. Sequence analysis

3.6.1.1. Format conversion

In order to use the sequence in conjunction with the GCG sequence analysis software package version 7.2 (Genetic Computer Group, Madison, WI, Devereux et al., 1984), the sequence had to be transferred onto a VAX 9000 computer and converted to the SEQ format, by using the program REFORMAT of the GCG package.

3.6.1.2. Homology search with FASTA

A systematic search for homologous sequences was performed through the whole GenEMBL database release 34.0 with over 78703 sequence entries by using the program FASTA (Pearson and Lipman, 1988).

3.6.1.3. Optimal alignment between homologous sequences

An optimal alignment of the best segment of similarity between pairs of sequences was performed by running the program BESTFIT (Devereux et al., 1984). A gap weight of 5.00 and gap length weight of 0.30 was used. Optimal alignment was found by inserting gaps to maximize the number of matches.
4. Results

The localization of several different cosmid or phage clones on bovine chromosomes required the development and establishment of a reliable technique. In the present study different experimental parameters were evaluated and they lead to the optimized FISH protocol described in the methods.

By using large probes, it was usually possible to obtain clear and strong hybridization signals without noticeable background. However, it must be mentioned that the results varied depending on the probes used. The sequence of the probe and the ratio between unique sequences and repetitive sequences, as well as the purity of the DNA influenced the results of ISH. In some cases the experiment had to be repeated by varying the ratio between the concentration of the probe and of competitor DNA (conditions of CISS).

The size, shape and intensity of the hybridization signal allowed to conclude which type of sequence was contained in the clones (s. 4.2.4.).

The first part of this chapter describes the optimization of the technique, in the second part the mapping results are presented.

4.1. Optimization of fluorescence in situ hybridization on bovine chromosomes

The quality of the hybridization is generally evaluated through qualitative and quantitative parameters.

Qualitative parameters are for instance the sharpness of the specific signal, its intensity, its position, the ratio between specific signal and background, the type and amount of background. An optimal signal is represented from a distinct and bright double dot positioned on each of both homologous chromosomes, perpendicular to their long axis. The double dot is composed of two small fluorescent spots (twin spots), one on each chromatid. The ratio between specific signal and background should be as high as possible. A satisfactory ratio can be obtained by increasing the intensity of the signal and/or by decreasing the amount of background.

The number of metaphases carrying the specific signal compared to those without signal, as well as the chromosomes carrying two signals compared to those carrying one are considered quantitative parameters.
The background is defined as the non-specific signal visible on the specimen after hybridization. It results from the non-specific binding of the probe, of the avidin-FITC and of the antibodies to the chromosomal DNA, to proteins, glass, etc. The background observed was mainly of two types: a cytoplasm associated background and a chromosome associated background, the latter also being referred to as "cross-hybridization".

Cytoplasm associated background appeared as a layer of very small fluorescent dots evenly distributed between the metaphases and on the chromosomes. It was distinguishable from the specific hybridization signal due to its relatively weak brightness and occurrence on a slightly different focal plane. However, in some cases it hampered the unequivocal detection of the hybridization signal.

"Cross-hybridization" consisted in the binding on chromosomal DNA of repetitive sequences contained within the probe. The type of signal produced presented the same shape as the specific one, but instead of being localized at a specific site on both homologues it was distributed along all the chromosomes. The cytoplasm between the metaphases was not associated with this signal. Depending on the probe, the use of more competitor DNA was needed in order to suppress cross-hybridization.

The lowering of non specific binding allowed to perform two or three rounds of amplification (three to four layers of avidin-FITC) without noticeable increase of background and therefore allowed the improvement of sensitivity of the FISH approach.

The factors which influence the background and the intensity of the signal are discussed in the following sections.

4.1.1. Factors that affect the background phenomena

4.1.1.1. Slide preparation

Slides of good quality were prerequisite for successful ISH. In some cases the unequivocal detection of the specific hybridization signal was hindered by slides of lesser quality.

The way of dropping the cell suspension on well cleaned slides, as well as the number of chromosome fixations were determining factors. An insufficient number of fixations resulted in specimens carrying an accumulation of cytoplasm which caused unspecific binding of the DNA probe, avidin-FITC and antibodies. Several chromosome fixations were thus performed, 6 to 8 being the minimal number required to assure good results.
Surface-tension forces during drying after dropping the chromosomes on glass appeared important for obtaining well spread metaphases. In case of too slow drying due to high air humidity the slide was placed close to a warm plate and in case of too little humidity in the air, very wet slides were used. Rinsing the slides with fixative before drying completely contributed to getting excellent preparations.

4.1.1.2. RNase treatment

Digestion with RNase of the chromosome preparations prior to hybridization improved the specific signal relative to the background, by avoiding binding between the probe and RNA traces on the slide. Moreover, when stained with propidium iodide, the contour of the treated chromosomes was sharper.

4.1.1.3. Chromosomal in situ suppression

An amount of 3 µg of competitor DNA was usually needed to obtain clean and clear specific hybridization signals when used in conjunction with 50 ng of probe. Nevertheless, some probes required the addition of 2 or 3 times more competitor. The effect of varying the concentration of competitor DNA is also discussed in 4.1.2.6. and shown in table 5.

Only in particular cases, namely the cosmids c10BT33 and CAP4, very satisfactory results were obtained without the addition of competitor DNA. Cosmid c10BT33 produced an uncommonly strong signal under all experimental conditions. Further investigations indicated that this cosmid contains a satellite sequence (s. 4.2.4.). CAP4 provided very clear hybridization signals without traces of cross-hybridization. Preliminary sequence information indicated numerous (GT)n repeats interrupted by stretches of G and T residues (data not shown), suggesting the presence of a minisatellite sequence.

Centromere specific sequences were insensitive to variations in the amount of competitor DNA, and gave always a very strong hybridization signal on the centromeres (s. 4.2.3.).

The effect of the variation of competitor DNA added to the probe prior to hybridization is shown in fig. 4.
Fig. 4: Metaphase spreads showing the effect on ISH when different amounts of competitor DNA are added to the probe (CAP2). a) 0 μg comp./50 ng probe, b) 1.5 μg comp./50 ng probe, c) 3 μg comp./50 ng probe, d) 6 μg comp./50 ng probe. The arrow indicates the specific signal on chromosome 5.
Large differences between the use of genomic competitor DNA and Cot-1 DNA were not observed, both competitors gave very satisfactory results when used in proper amounts.

For the preparation of Cot-1 DNA many parameters had to be optimized in order to obtain a suitable competitor.

The successful preparation of Cot-1 DNA was evaluated based on the yield of DNA and on the size of the fragments, separated by electrophoresis on a 1.5% agarose minigel. Since the repetitive sequences of the bovine genome are estimated to represent about 40% of the genome, the desired yield of Cot-1 DNA was 40% of the total genomic DNA used. However, this percentage was not consistently reproducible, and differences were observed in the yield although the procedure was not modified. This was probably due to the phenol/chloroform extraction and the ethanol precipitation, which could result in loss of varied amounts of DNA, when working with small quantities.

The definitive criterion used to judge the suitability of each new preparation of competitor, was its capacity to suppress the repetitive sequences in the ISH experiments. This test was performed with a probe which had been previously localized, using a known amount of competitor DNA and therefore could be considered as a standard.

Different conditions were studied for the optimization of this protocol:

- **Time of reannealing**: the time granted for the DNA to reassociate ranged from 22 min (theoretical minimum, s. 3.3.1.2.) to 2 hours. Its variation did not show any change in the yield or size of the competitor DNA obtained. This suggests that in 22 min the repetitive sequences of DNA at a concentration of 0.25 µg/µl can reassociate, and more than 2 hours are needed to allow reassociation of the unique sequences as well. Therefore, the time of reassociation was not considered as a decisive parameter.

- **Amount of S1 nuclease**: according to Vogt (1980), 1 unit of S1 nuclease is "the enzyme activity which releases 1 µg of acid soluble deoxynucleotides from denatured DNA per 1 min at 37°C". When working with a total of 10 µg of DNA, the recommended 6 units of enzyme were used to digest 6 µg of unique sequence DNA. However, since this amount of enzyme consistently caused degradation of the DNA, 0.1 or 0.08 units of S1 nuclease per µg of DNA were used for digestions. A considerable difference in the quality of the DNA was then observed. In fact this DNA showed a homogeneous smearing of 100/200 bp long
DNA on the gel, while degraded DNA appeared as an accumulation of very short DNA pieces (fig. 5).

Fig. 5: 1.5% agarose minigel showing Cot-1 DNA digested with different amounts of S1 nuclease. Lanes a) λ HindIII marker, b) reannealed DNA, not digested, c) and e) 1u of S1 nuclease/μg DNA, d) and f) 0.1 u of S1 nuclease/μg DNA.

- *Time of S1 nuclease digestion:* by increasing the time of digestion with S1 nuclease from 1 to 20 min, the yield of competitor DNA seemed to decrease, although the results were not entirely consistent. For practical reasons, a time of 2 min was chosen and kept in all the experiments, although Vogt (1980) recommended a digestion time of 1 min.
4.1.1.4. Washing steps

Labelled probes can hybridize non-specifically to sequences which bear homology but are not identical to the sequence of the probe. This phenomenon results in undesired background, such as cross-hybridization. These hybrids are less stable than perfectly matched hybrids, and therefore they can be dissociated by performing washes at higher stringency. The stringency of washing can be modified by varying the formamide concentration, the salt concentration and the temperature of the solution. The lower the salt concentration and the higher the formamide concentration and wash temperature, the more stringent the wash becomes.

After hybridization with cosmid and phage probes, the slides were washed using the methods previously described. Increasing the temperature of the wash solution (increase of the stringency) to reduce the background resulted in the simultaneous disappearance of the specific signal. However, washes with decreased salt concentration (0.1 x SSC at 65°C instead of 2 x SSC at 43°C), were ideal for centromere specific probes.

The ratio salt/formamide in the wash solution was never changed to avoid the influence of too many different factors on the stringency applied to a probe.

Washing steps between layers of avidin-FITC and antibodies were first performed at RT, but increasing the temperature to 42°C allowed essential improvements in the ratio signal/background.

4.1.2. Factors that affect the signal strength

4.1.2.1. Probe labelling and analysis of biotin incorporation

The two biotin compounds used to label the probes were provided from two different manufacturers and differed in the length of the arm between the biotin core and the dUTP. Biotin-11-dUTP (Sigma) has an arm of a carbon chain with 11 atoms and biotin-16-dUTP (Boehringer Mannheim) a carbon chain of 16 atoms. This difference seems not to influence the extent of biotin incorporation in the DNA strand during the nick translation, but is due to patent regulations.

It was found that the use of biotin-16-dUTP without the addition of unlabelled dTTP in the reaction (s. 3.2.4.1.), did not allow the same biotin incorporation as that obtained when using biotin-11-dUTP with the same molarity. Based upon the criterion of a labelled probe's suitability for ISH (s. 3.2.4.4., b)), the following dot blot assay shows that the probe labelled with biotin-16-dUTP and without unlabelled dTTP is not suitable (fig. 6 d). In fact, the hybridization with such a probe only rarely gave a specific hybridization signal.
The addition of 10 nM of dTTP to biotin-16-dUTP in the labelling reaction (fig. 6 c) produced a considerable improvement in the intensity of the hybridization signal. A positive difference was observed by labelling with biotin-11-dUTP (fig. 6 b).

![Dot blot assay of three labelling reactions of the same probe](image)

**Fig. 6:** Dot blot assay of three labelling reactions of the same probe
a) marker, b) with biotin-11-dUTP, c) with biotin-16-dUTP + dTTP, d) with biotin-16-dUTP without dTTP

Analogous observations were obtained when DNA probes are labelled with digoxigenin-11-dUTP or with FITC-12-dUTP, where the addition of dTTP is a prerequisite for successful ISH. A model based on steric hindrances is proposed to explain this observation.

Biotin-16-dUTP from Boehringer Mannheim is a highly purified product and does not contain any other nucleotides. On the other hand, Sigma’s biotin-11-dUTP seems to contain traces of unlabelled dUTP and dTTP. Due to the long chain of the biotinylated compounds, the molecule can assume different spatial conformations and this results in a relatively cumbersome molecule. When too
many of these collide, they hinder each other's accommodation on the DNA strand. An appropriate portion of unlabelled dTTP molecules in the reaction facilitates the incorporation of biotins, because some nicks in the DNA strand are filled up with dTTP and a suitable amount of other nicks remains free for the biotin molecules (fig. 7).

Fig. 7: Model of steric hindrance between biotinylated molecules a) an excess of biotinylated dUTP molecules is present in the labelling reaction b) unlabelled dUTP and dTTP are also added.
A second argument to explain our observations is the correct amount of biotins incorporated in the DNA strand. Compared to biotin, avidin is a very big molecule, and it is foreseeable that when too many biotins are present in the probe, a steric hindrance would occur between the avidin molecules which normally bind to the biotin. This model is represented in fig. 8.

Fig. 8: Model of steric hindrance between avidin molecules a) too many biotins are incorporated in the DNA strand, b) an optimal proportion of biotins are incorporated in the DNA strand
4.1.2.2. Probe length

Large probes, such as cosmids, are not sufficiently cut during the nick translation reaction and thus have difficulties in penetrating the dense matrix of the chromosomes; therefore, the probe size of the fragments after the labelling has to be adjusted, as described in detail in the paragraph 3.2.4.1. Good results were obtained by generating DNA fragments of about 180 bp of length.

When labelling recombinant phages, the amount of enzyme present in the reaction was already sufficient, and fragments of optimal length could be obtained without addition of DNase I in the nick translation. When hybridizing with phages, the percentage of metaphases carrying a visible hybridization signal was lower than that shown by the cosmid clones. The hybridization with plasmid probes was even more difficult than work completed using phages. This difficulty is explained by two factors: the number of biotin molecules incorporated in the probe and the time required for the probe to hybridize on the specimen. It is estimated that during the nick translation one molecule of biotin is incorporated in the DNA strand approximately every 20 bp. Therefore, in one labelled cosmid with an average length of 40 kb there are theoretically 2000 biotins ready to be bound with avidin-FITC and then to produce a fluorescent signal, whereas in a phage of 15 kb there are only 750 biotins to be found. It must also be taken into account that when hybridizing shorter probes, such as phages or plasmids, the size of the target DNA is smaller, and the probe requires a longer period of time to find the homologous strand on the specimen. This suggests the danger of rapid reannealing of chromosomal DNA, which may limit the hybridization efficiency. This was avoided by increasing the amount of DNA on the same surface of specimen, as discussed in 4.1.2.6. Moreover, "inefficiency" of shorter probes was circumvented by performing two rounds of signal amplification.

4.1.2.3. Probe concentration

The probe concentration affects the rate at which the first few base pairs hybridize (nucleation reaction), the adjacent base pairs are formed afterwards. The nucleation reaction is a limiting step in the hybridization.

Satisfactory results with cosmids were achieved only when the concentration of the probe was increased to 50 ng/10 μl of hybridization mixture, hybridized on a surface of 18 x 18 mm.

Phages and plasmids required from 100 to 200 ng of DNA per 10 μl of hybridization mixture under a 18 x 18 mm coverslip.
By farther increasing the DNA amount, the degree of background increased. Therefore, when the amount of DNA probe was increased, the quantity of competitor DNA had to be adjusted accordingly.

4.1.2.4. Chromosome counterstaining

By lowering the stringency of the staining solution, a major improvement in the detection of the hybridization signal could be achieved, allowing the detection of even faint signals. Instead of staining with propidium iodide dissolved in H2O, better results were obtained by staining the chromosomes with propidium iodide which was dissolved in 2xSSC. Staining in diluted solution for a longer time also improved the detection of the small double dot signal. The optimal staining was obtained following the procedure described in 3.3.4.

4.1.2.5. Amplification of the signal intensity

In order to obtain spots on both chromatids of both homologues, one round of amplification was necessary for all the cosmids, except for those containing satellite sequences. Phages often were already showing a sufficient signal on some metaphases immediately after the first amplification (two layers of avidin-FITC). However, two amplifications were usually performed to obtain more spreads carrying the four hybridization signals. The fluorescent signals were, in fact, considerably increased by using the amplification method (s. 3.3.3.). The results obtained, however, are not linear: while the signal intensity is increasing, the ratio signal/background is decreasing. Signals on slides not affected by background after the first amplification, could be amplified a second time without producing too high levels of background. In the cases where no specific signal was observed after two rounds of amplification, a third amplification was superfluous, since it would have caused a very strong background. In these cases the lack of signal was due to other reasons.

4.1.2.6. Efficiency of fluorescence in situ hybridization

Although this technique offers a very high potential to localize any cosmid or phage probe on chromosomes, differences were observed in the hybridization results by using different probes. "Efficient probes" are designated as those clones that showed 4 specific hybridization signals on almost all the metaphases, without noticeable background. In such cases, the chromosomal localization of the cosmid could be determined by identifying the labelled chromosomes of one or two good metaphases. However, for each probe 15-20
metaphases were usually carefully analyzed, because some chromosomes are particularly difficult to identify and the shape of the chromosomes and banding pattern were not always very typical. In addition, there were also incomplete signals, which showed up either as a single signal on one of the two homologues, or as a single spot on only one of the chromatids.

The two signals on the two chromatids of a chromosome were approximately equal in intensity, whereas the variation among chromosomes was wide.

In Table 5 the results observed with 9 probes and different amounts of competitor DNA are summarized. After the first amplification, metaphases were counted without signal, with signal on 1 chromosome or signal on 2 chromosomes, as well as metaphases carrying remarkable cross-hybridization. The same evaluation was made after the second amplification.

It is obvious that by carrying a second amplification, the proportion of metaphases carrying the specific signal on both homologous chromosomes increases significantly. Metaphases that do not show any signal after the first amplification, likely bear signals after the second amplification. In some cases, however, i.e. for the probes "59-11" and "60-21", a disturbing increase in background occurs instead.

By adding more competitor DNA, the percentage of metaphases showing signals on two chromosomes augments, while the metaphases affected from background (cross-hybridization) decrease (s. CAP11 and MAP1B). When the added amount is too big, as for instance in the case of CAP11 with 9 μg of competitor DNA, the specific signal can be suppressed as well.

Interphase nuclei also present clear hybridization signals, frequently seen as doublets due to the labelling of both chromatids in late S-phase and G-2 phase cells.
Table 5: Hybridization results of some randomly considered probes. The number of metaphases showing none, one, two signals and/or background is reported, according to the amount of competitor DNA used, after 1 and 2 amplifications of the signal.

<table>
<thead>
<tr>
<th>Probe</th>
<th>Comp. DNA (µg)</th>
<th>tot metap</th>
<th>1. Amplification</th>
<th>2. Amplification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>no signal</td>
<td>signal on 1 chr.</td>
<td>signal on 2 chr.</td>
</tr>
<tr>
<td>CAP11</td>
<td>1.5</td>
<td>14</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50%</td>
<td>43%</td>
</tr>
<tr>
<td></td>
<td>3.0</td>
<td>14</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>43%</td>
<td>36%</td>
</tr>
<tr>
<td></td>
<td>6.0</td>
<td>10</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>60%</td>
<td>30%</td>
</tr>
<tr>
<td></td>
<td>9.0</td>
<td>9</td>
<td>9</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100%</td>
<td>-</td>
</tr>
<tr>
<td>MAP1B</td>
<td>0</td>
<td>18</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>55.5%</td>
<td>22.5%</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>11</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>36%</td>
</tr>
<tr>
<td></td>
<td>3.0</td>
<td>16</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>31%</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>6.0</td>
<td>18</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50%</td>
<td>33%</td>
</tr>
<tr>
<td>Probe</td>
<td>Comp. DNA (µg)</td>
<td>tot metap</td>
<td>no signal</td>
<td>signal on 1 chr.</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------</td>
<td>-----------</td>
<td>-----------</td>
<td>------------------</td>
</tr>
<tr>
<td>59-11</td>
<td>3.0</td>
<td>18</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>55.5%</td>
<td>28%</td>
</tr>
<tr>
<td></td>
<td>6.0</td>
<td>11</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>33%</td>
<td>17%</td>
</tr>
<tr>
<td>60-21</td>
<td>3.0</td>
<td>19</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>16%</td>
<td>42%</td>
</tr>
<tr>
<td></td>
<td>6.0</td>
<td>17</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>35%</td>
<td>-</td>
</tr>
<tr>
<td>C8</td>
<td>3.0</td>
<td>14</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>57%</td>
<td>21.5%</td>
</tr>
<tr>
<td>M14/1</td>
<td>6.0</td>
<td>16</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>37.5%</td>
<td>12.5%</td>
</tr>
<tr>
<td>M10/13</td>
<td>3.0</td>
<td>16</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25%</td>
<td>-</td>
</tr>
<tr>
<td>99-1</td>
<td>3.0</td>
<td>16</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>75%</td>
<td>25%</td>
</tr>
<tr>
<td>24A</td>
<td>3.0</td>
<td>15</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>53%</td>
<td>20%</td>
</tr>
</tbody>
</table>

**Results**
4.2. Assignment of cosmid and phage probes to bovine chromosomes

4.2.1. Map position of loci

The location of each probe on the respective chromosomes could be roughly determined by comparing the chromosome pair carrying the hybridization signal with the photograph of the same metaphase previously banded. However, in order to define the exact position of the hybridization signal, a minimum of 6 to 30 chromosomes was carefully analyzed and the FL\text{cen} value determined as described in 3.3.7.

The chromosomal assignment of the clones, the FL\text{cen} and the standard deviation (SD) of FL\text{cen} are reported in table 6.

The probes are classified in groups according to the type of library screened to obtain them. The same table shows whether the probe contains a microsatellite repeat, indicated as (CA)\text{n}, or candidate sequences for a gene.

The accuracy of the assignment depends on the degree of chromosome condensation and on the banding pattern. Usually the probes could be assigned to a single band with a relatively high certainty. An average band in the cattle karyotype encompasses 7500 kb, corresponding to about 8 cM. Therefore, it is concluded that a resolution of roughly 8 cM was routinely achieved. In humans, cosmids have been mapped on chromosomes with a resolution of more than 3 cM (Trask, 1991). This resolution could also be achieved in cattle, using elongated and/or prometaphasic chromosomes.

Metaphase spreads showing the physical mapping are shown in fig. 9. For each clone a QFQ-banded partial metaphase prior to ISH is compared to the same metaphase spread after hybridization. All the localizations are then summarized on a standard idiogram (ISCNDA, 1989) in fig 10.
Table 6: Chromosomal assignment of the clones.

1) microtubule associated protein gene, 2) candidate gene for calpastatin, 3) candidate gene for growth hormone receptor, 4) desmocollin gene, 5) polyimmunoglobulin receptor gene

FL cen: "fractional length centromere", see text, SD: standard deviation

<table>
<thead>
<tr>
<th>Probe</th>
<th>Assignment</th>
<th>FL cen</th>
<th>SD</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAP1 (CA)n</td>
<td>21q17</td>
<td>0.55</td>
<td>0.043</td>
<td>in the middle of band q35, almost telomeric</td>
</tr>
<tr>
<td>CAP2 (CA)n</td>
<td>5q35</td>
<td>0.91</td>
<td>0.031</td>
<td>telomeric</td>
</tr>
<tr>
<td>CAP3 (CA)n</td>
<td>19q23</td>
<td>0.97</td>
<td>0.031</td>
<td>telomeric</td>
</tr>
<tr>
<td>CAP4 (CA)n</td>
<td>17q26</td>
<td>0.99</td>
<td>0.013</td>
<td>distal portion of band q26, telomeric, very strong signal</td>
</tr>
<tr>
<td>CAP5 (CA)n</td>
<td>centrom. aut.</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>CAP6 (CA)n</td>
<td>centrom. aut.</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>CAP7 (CA)n</td>
<td>13q21/q22</td>
<td>0.71</td>
<td>0.049</td>
<td>between bands q21 and q22</td>
</tr>
<tr>
<td>CAP8 (CA)n</td>
<td>6q35</td>
<td>0.94</td>
<td>0.034</td>
<td>almost telomeric. This chromosome is number 6 according to the standard of Reading 1976 (s. 4.2.4.)</td>
</tr>
<tr>
<td>CAP9 (CA)n</td>
<td>11q27</td>
<td>0.92</td>
<td>0.033</td>
<td></td>
</tr>
<tr>
<td>CAP10 (CA)n</td>
<td>5q25</td>
<td>0.55</td>
<td>0.012</td>
<td></td>
</tr>
<tr>
<td>CAP11 (CA)n</td>
<td>16q21</td>
<td>0.62</td>
<td>0.032</td>
<td>proximal portion of band q21</td>
</tr>
<tr>
<td>MSI/2 (CA)n</td>
<td>19q17</td>
<td>0.61</td>
<td>0.037</td>
<td>proximal portion of band q17</td>
</tr>
<tr>
<td>MSI/3 (CA)n</td>
<td>8q21</td>
<td>0.64</td>
<td>0.028</td>
<td>distal portion of band q21</td>
</tr>
<tr>
<td>3R-100-9 (CA)n</td>
<td>centrom. aut.</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>MAP1B gene 1)</td>
<td>20q14</td>
<td>0.32</td>
<td>0.042</td>
<td>proximal portion of band q14</td>
</tr>
<tr>
<td>c10BT 18 (CA)n</td>
<td>no signal</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>c10BT 21 (CA)n</td>
<td>centrom. aut.</td>
<td>-</td>
<td>-</td>
<td>it did not work under any condition</td>
</tr>
<tr>
<td>c10BT 33 (CA)n</td>
<td>6q12-q15</td>
<td>0.11 to 0.33</td>
<td>0.020</td>
<td>it produced an unusually large signal, in the subcentromeric region</td>
</tr>
<tr>
<td>c10BT 34 (CA)n</td>
<td>19q22</td>
<td>0.87</td>
<td>0.069</td>
<td>in the middle of band q22</td>
</tr>
<tr>
<td>Probe</td>
<td>Assignment</td>
<td>FL&lt;sub&gt;cen&lt;/sub&gt;</td>
<td>SD</td>
<td>Observations</td>
</tr>
<tr>
<td>---------------</td>
<td>------------</td>
<td>-----------------</td>
<td>-----</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>c10BT 93 (CA)n</td>
<td>centrom. aut.</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>c10BT 133 (CA)n</td>
<td>centrom. aut.</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>c10BT 134 (CA)n</td>
<td>centrom. aut.</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>c10BT 145 (CA)n</td>
<td>centrom. aut.</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>c10BT 146 (CA)n</td>
<td>centrom. aut.</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>c10BL 26 (CA)n</td>
<td>1q45</td>
<td>0.96</td>
<td>0.018</td>
<td>distal portion of band q45, almost telomeric</td>
</tr>
<tr>
<td>C 5 (CA)n</td>
<td>19q14/q15</td>
<td>0.42</td>
<td>0.062</td>
<td>between bands q14 and q15</td>
</tr>
<tr>
<td>C 7 (CA)n</td>
<td>centrom. aut.</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>C 8 (CA)n</td>
<td>25q15</td>
<td>0.49</td>
<td>0.064</td>
<td></td>
</tr>
<tr>
<td>M10/13 (CA)n</td>
<td>22q24</td>
<td>0.97</td>
<td>0.034</td>
<td>distal portion of band q24, telomeric</td>
</tr>
<tr>
<td>M14/1 (CA)n</td>
<td>18, 19q22/q23, and 22</td>
<td>-</td>
<td>-</td>
<td>chimeric cosmid showing signals on different chromosomes, however mainly on 18, 19 and 22</td>
</tr>
<tr>
<td>12 A (CAL) ²</td>
<td>centrom. aut.</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>24 A (CAL)</td>
<td>3q37</td>
<td>0.99</td>
<td>0.011</td>
<td>distal portion of band q37, telomeric</td>
</tr>
<tr>
<td>17 B (CAL)</td>
<td>no signal</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>99-1 (GHR) ³</td>
<td>7q24</td>
<td>0.63</td>
<td>0.027</td>
<td>proximal portion of band q24</td>
</tr>
<tr>
<td>99-2 (GHR)</td>
<td>centrom. aut.</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>99-3 (GHR)</td>
<td>no signal</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>μ35Cos9 (CA)n</td>
<td>centrom. aut.</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>&quot;60-21&quot;(DSC) ⁴</td>
<td>24q21/q22</td>
<td>0.53</td>
<td>0.056</td>
<td>between bands q21 and q22</td>
</tr>
<tr>
<td>&quot;59-11&quot; (DSC)</td>
<td>24q21/q22</td>
<td>0.54</td>
<td>0.060</td>
<td>between bands q21 and q22</td>
</tr>
<tr>
<td>&quot;65&quot; (DSC)</td>
<td>24q21/q22</td>
<td>0.56</td>
<td>0.045</td>
<td>between bands q21 and q22</td>
</tr>
<tr>
<td>PIGR ⁵</td>
<td>16q13</td>
<td>0.28</td>
<td>0.041</td>
<td>proximal portion of band q13</td>
</tr>
</tbody>
</table>
Results

CAP4 (CA)n

CAP7 (CA)n

CAP8 (CA)n
MAP1B gene

MSI/2 (CA)n

MSI/3 (CA)n

MAP1B gene
Results

c10BT 34 (CA)n

c10BL 26 (CA)n

C 5 (CA)n
Fig. 9: Metaphase spreads showing the chromosomal assignment of the clones. Left: QFQ-banded partial metaphase prior to ISH. Right: the same metaphase spread after ISH, stained with propidium iodide.
Fig. 10: Standard idiogram (ISCNDA, 1989) showing the localization of the clones
In order to confirm some of the localizations, primers specific for 8 probes were synthesized (Steffen, 1992) and made available to H. Neibergs and J. Womack (Texas A & M University) for PCR analysis of a hybrid panel. The synteny assignments are reported in table 7 and they are important for the interpretation of the physical mapping (s. 5.1.).

Table 7: Assignment of some probes to syntenic groups (Solinas Toldo et al., 1993; * H. Neibergs, personal communication)

<table>
<thead>
<tr>
<th>probe</th>
<th>syntenic group</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAP2</td>
<td>U3</td>
</tr>
<tr>
<td>CAP3</td>
<td>U20</td>
</tr>
<tr>
<td>CAP8</td>
<td>U15</td>
</tr>
<tr>
<td>CAP9</td>
<td>U16</td>
</tr>
<tr>
<td>CAP10</td>
<td>U3</td>
</tr>
<tr>
<td>CAP11</td>
<td>U1</td>
</tr>
<tr>
<td>MSI/2</td>
<td>U21</td>
</tr>
<tr>
<td>MAP1B *</td>
<td>U20</td>
</tr>
</tbody>
</table>

CAP3 has been localized on chr. 19 and MAP1B on chr. 20. Therefore, they cannot belong to the same syntenic group. Moreover, CAP3 and MSI/2, which both hybridize on chr. 19, should belong to the same syntenic group, but according to the somatic hybrid panel the result does not agree. These disagreements are discussed in 5.1.

4.2.1.1 Double colour fluorescence in situ hybridization with two clones

Due to the closeness of the probes CAP3 and c10BT34, both assigned on the telomeric region of chr.19, a further experiment was performed in which the two probes were simultaneously hybridized as is described in the methods (s. 3.3.6.). This experiment facilitated the unequivocal positioning of the probes relative to the telomere. Fig. 11 shows one of the metaphases analyzed, where
CAP3 is detected with avidin-FITC (yellow/green signal) and c10BT34 with antidigoxigenin-TRITC (red signal).

Fig. 11: Partial metaphase spread showing double colour FISH with the probes CAP3 and c10BT34, on chromosome 19. CAP3 is detected with fluorescein isothiocyanate (yellow/green signal) and c10BT34 with tetramethyl rhodamine isothiocyanate (red signal). The signal of these probes is detected with a CCD camera and computer processed.

By considering the coordinates of the signals on the X and Y axis, a further determination of the exact location of the two probes could be ascertained. In fig. 11 the two spots of the probe CAP3 were recorded on position X=490; Y=307 and X=496; Y=298. The spots of c10BT34 were on X=502; Y=307 and X=505; Y=300. Thus the coordinates on the X axis for the probe c10BT34 indicate a more distal position of this probe.

This result is not in agreement with the assignments of the two probes done by measuring their FLcen values, which are reported in table 6. However, the standard deviation (SD) of FLcen for c10BT34 is relatively high (0.069) and indicates remarkable variability in the contraction of the chromosomes used for
Leer - Vide - Empty
such localization and in the position of the signal. In the case of close proximity, the FL\textsubscript{cen} values are not sufficient to determine the order of probes. Multicolor FISH experiments can indisputably show the exact position of two probes on a chromosome, even when close to each other, as fig. 11 demonstrates. Linkage analysis of probes on the same chromosome would be an alternative possibility.

4.2.2. Distribution of loci on the genome

Many cosmids were found to be centromere specific for the autosomes, but not for the two sex chromosomes. Results concerning these probes are reported in 4.2.3. Moreover, site specific cosmids were found on 16 of 29 autosomal chromosome arms. More than one microsatellite hybridized on chr. 5 and on chr. 19, one microsatellite and one gene hybridized on chr. 16. Only few probes could not be assigned under any conditions of hybridization, but after further analysis, probes were found to contain no insert.

A tendency of clustering of microsatellite containing probes in the telomeric region of chromosomes was observed. In order to support this observation and to evaluate the distribution of the clones on the genome, the chromosomes were divided into 5 intervals, corresponding to a FL\textsubscript{cen} from 0 to 0.2, from 0.2 to 0.4, from 0.4 to 0.6 etc. The number of probes localized to each interval has been counted and compared. These results, together with the number of centromere specific probes and those which did not produce any signal are reported in table 8. Fig. 12 shows the distribution of the non centromeric \((CA)_n\) containing probes.

Table 8: Distribution of different clones on the genome. Centromere specific probes are exclusively located on the autosomes. FL\textsubscript{cen}: "fractional length centromere".

<table>
<thead>
<tr>
<th></th>
<th>total probes</th>
<th>centro-meric</th>
<th>FL\textsubscript{cen} 0-0.2</th>
<th>FL\textsubscript{cen} 0.2-0.4</th>
<th>FL\textsubscript{cen} 0.4-0.6</th>
<th>FL\textsubscript{cen} 0.6-0.8</th>
<th>FL\textsubscript{cen} 0.8-1</th>
<th>no signal or chimeric</th>
</tr>
</thead>
<tbody>
<tr>
<td>((CA)_n)</td>
<td>30</td>
<td>11</td>
<td>1</td>
<td>-</td>
<td>4</td>
<td>4</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>candid. genes</td>
<td>11</td>
<td>2</td>
<td>-</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>total</td>
<td>41</td>
<td>13</td>
<td>2</td>
<td>7</td>
<td>5</td>
<td>9</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>
Fig. 12. Distribution of the (CA)\textsubscript{n} containing probes on a schematic chromosome, divided in 5 equal intervals, from FL\textsubscript{cen} 0 to FL\textsubscript{cen} 1.0

The $\chi^2$ Test is used to statistically analyze whether the distribution of the probes on the genome follows the hypothesis of a random distribution or not. In a negative case, it supports the observation of a tendency of the microsatellites to be localized on the telomeric region of the chromosomes.

The following calculation is done according to the values reported in table 8:

$$\chi^2 = \sum_{i}^{k} \left( \frac{(f_i - F_i)^2}{F_i} \right) = 11.53$$

$k$ : number of intervals into which each chromosome has been divided ($k = 5$)

$f_i$ : number of microsatellites localized on the interval $i$

$F_i$ : theoretical average number of microsatellites on the interval $i$ ($F_i = 3.4$)

With 4 degrees of freedom, a $\chi^2$ of 11.53 was calculated. This indicates that the clustering in the centromeric regions is significant ($p=0.025$, Geigy tables, 1980).

Although caution is required due to the relatively small number of probes analyzed, a first indication of a telomeric tendency of microsatellites in the bovine genome is reported.
4.2.3. Centromere specific hybridization of satellite containing cosmids

As already shown in table 8, a relatively high portion of the (CA)$_n$ probes, namely 13 among 41 probes, resulted in being centromere specific for the autosomes. The centromeres were very strongly labelled, as shown in fig. 13. Chromosomes X and Y did not show any trace of positive signal, suggesting that centromeres of autosomes and sex chromosomes in cattle do not carry the same sequences.

CAP6 hybridized with all the centromeres of the autosomes at the same intensity of fluorescence. On the other hand, by hybridizing with CAP5 and all the other centromere specific clones, some chromosomes showed a weaker signal. In table 9 these chromosomes are reported for each probe analyzed.

Fig. 13: Example of hybridization with a clone containing sequences centromere specific for the autosomes (CAP6)
Table 9: Chromosomes showing a weaker signal than other chromosomes, after hybridization with centromere specific probes

<table>
<thead>
<tr>
<th>Probe</th>
<th>Chromosomes showing a weaker signal</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAP5</td>
<td>2, 3, 6, 7, 10, 11, 17, 21</td>
</tr>
<tr>
<td>c10BT21</td>
<td>2, 3, 6, 7, 10, 11, 17, 21</td>
</tr>
<tr>
<td>c10BT93</td>
<td>2, 3, 6, 7, 10, 11, 17, 21, (5), (12)</td>
</tr>
<tr>
<td>c10BT133</td>
<td>1, 2, 3, 6, 7, 10, 11, 17, 21, (5), (12)</td>
</tr>
<tr>
<td>c10BT134</td>
<td>2, 3, 6, 7, 10, 11, 17, 21</td>
</tr>
<tr>
<td>c10BT145</td>
<td>1, 2, 3, 6, 7, 10, 11, 17, 21</td>
</tr>
<tr>
<td>c10BT146</td>
<td>(2), (7), 11, 17, 21, (5), (12), (8)</td>
</tr>
</tbody>
</table>

The majority of centromere specific probes was screened from a library of vector Triple Helix (s. table 2) which had been amplified. Amplification leads to a higher representation of satellite sequences, which give centromere specific signals.

CAP5 and CAP6 are only 2 centromere specific clones, among a total of 13 (CA)n-containing clones screened from the same library of SuperCos 1. The choice of this vector, which involves the dephosphorylation of the genomic DNA during the construction of the cosmid library, excludes the possibility of chimera formation. Therefore, the localization of CAP5 and CAP6 on the centromeres has to be discussed (s. later).

Some of the centromeric clones were partially sequenced. The results of the sequencing of CAP5 and CAP6 with the primers specific for T3 and T7 are shown in fig. 14 and fig. 15. The sequence signals started directly after the priming site. The first bases, reported in *italic*, are part of the vector. For each sequence, subsequences are defined, from the first base to the first N or to another N. These are chosen arbitrarily by trying to find a compromise between a sequence with few N, but also long enough to be used in homology studies.
Results

a)  

1  NCCCCCTAAAN  ACCCACCCCC  ATTTTNCTGN  CAAGTGGAGG  TCCACCGCGG  
51  TGGCGGCCGCG  TCTAGAAACTA  GTGAGACTTTG  CTTGGTGTGGT  TGGTGCTTTGGT  
101  TGGTGCTTTGGT  TGGTGCTTTGGT  TGCTGCGCGG  CGGGGGAGGG  GAGGGGCAGCT  
151  ACTATCAGCTG  ATAGGTATGGT  NCACCAATCT  GTAAGTAGCC  AATATAAAAG  
201  TCAATNTCTTG  GTTGCTGAGG  GGAAGTNTAT  GACTTGGNAT  AAAANAAAGG  
251  TATGTTGCAT  AAAATGCAGAT  TTCTGAAGGG  AAAAAAGCGCA  CGTNTGCTNT  
301  TCCAGGTGNC  ANANACANNA  TCNNAGAACG  AAANNAATNG  CTNCCCCNT  
351  CCGATCTGAA  NNACCAAAANC  CTTGCCCCAN  AAAAAAGCCCA  AACCNGGCNT  
401  CNTAGANCAT  CNGNACNCA  NAATCCCAT  TTGTCNCTTA  NCAACCCTCC  
451  ANNNTANAG  GGCAANNACA  CCCCCTACNC  CCCCCAAAAA  ANNGAAATNC  
501  CCTCCANNAN  CTNCACGCTT  CGACNANCC  NCAANANAAA  NGANNNCNCC  
551  CCCTCCTCCA  NANCNCCANC  GNATNNANAN  CCCCCCTCCA  N

Subsequence 1, Subsequence 2

b)  

1  CACCCANNNN  GNNNNNATTT  TNNNCGGCTT  GGGTACCGGG  CCCCCCTCGG  
51  AGGTGACCGG  TATCGATAAGG  CTTGATACGG  AATCCTGCA  GCCCCGGGGA  
101  TCTTTTTCGG  GATGTTCACA  TCTTCTATTT  AGAGTTAGT  CATCAATTAG  
151  ACCAAGGTAA  TTCTGTGTPTG  AATAGCGATA  GGCACGCGTA  TTCACTGAA  
201  TACTGGTTTC  TCCGAAATAC  TTGCACGGG  AAACAGCGA  ACTTTGCCC  
251  TGGTATGGRTT  TATCCACGCG  GATACGAACA  TTGGAATGAG  TCACATCGAA  
301  AGAAATACGA  CCAACTGAGA  TATCTAGACC  AAGACTCAAC  AAAAAACCGG  
351  GATNGTTATTT  ATCCGGTATT  CTGAATACGG  GTATNACCCG  CCAGGATAGNT  
401  ATTCAGGGCG  NAGTAGTAGG  TTGGCTGGAA  ATNAAATTAGG  TCAACCNGA  
451  ATATCATCTT  TTAAGACCCG  ACCGGNNTAA  TATCCACGATG  CCAACGCAGG  
501  GGGTNAATTTG  GACAACCGGTG  AGAAGGCTNG  GAAATTCCCN  NTNT

Subsequence 1

Fig. 14: Sequencing of the cosmid CAP5, a) with the primer T3, b) with the primer T7. In a) the microsatellite (GT)₂₀ found in the first 125 bases is outlined.
The sequence reading length was between 500 and 600 bases. The sequencer generated sharp peaks on the chromatogram (data not shown), deterioration of the peaks occurred generally only after 350-400 bp. In the case of CAP5, sequenced with T3, deterioration occurred already after the base 145.
This is probably due to the presence of a G-rich stretch after the microsatellite (GT)$_{20}$ found in the first 125 bp of the sequence. Deterioration of the sequence represents positions where the sequencer was unable to determine a base and placed an N. The definition of subsequences allows to obtain shorter but more reliable sequences suitable for the analysis and alignment within the GenEMBL database.

The results of similarity with bovine satellites DNA are reported in the following table.

Table 10: Comparison between the centromere specific sequences contained in CAP5 and CAP6 and the known bovine satellites.

<table>
<thead>
<tr>
<th>Subsequence</th>
<th>Sequence in database</th>
<th>Name of satellite</th>
<th>Similarity</th>
<th>Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub1CAP5T3</td>
<td>Btsat1</td>
<td>1.709</td>
<td>98.8%</td>
<td>84</td>
</tr>
<tr>
<td>Sub2CAP5T3</td>
<td>Btsat1</td>
<td>1.709</td>
<td>98.3%</td>
<td>117</td>
</tr>
<tr>
<td>Sub1CAP5T7</td>
<td>none sigif.</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sub1CAP6T3</td>
<td>Btsata</td>
<td>1.723</td>
<td>95.7%</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>Btsatb</td>
<td>1.723</td>
<td>97.1%</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>Btsatc</td>
<td>1.723</td>
<td>96.6%</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>Btrep17</td>
<td>sat II</td>
<td>95.7%</td>
<td>70</td>
</tr>
<tr>
<td>Sub2CAP6T3</td>
<td>Btsata</td>
<td>1.723</td>
<td>93%</td>
<td>171</td>
</tr>
<tr>
<td></td>
<td>Btsatb</td>
<td>1.723</td>
<td>94.7%</td>
<td>171</td>
</tr>
<tr>
<td></td>
<td>Btsatc</td>
<td>1.723</td>
<td>87.7%</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>Btrep17</td>
<td>sat II</td>
<td>93.3%</td>
<td>149</td>
</tr>
<tr>
<td>Sub1CAP6T7</td>
<td>none sigif.</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Btsat1: bovine genomic fragment for 1.709 satellite DNA
Btsata: bovine 680 bp repeat unit of 1.723 satellite DNA
Btsatb: bovine 1.723 satellite DNA
Btsatc: bovine 1.723 satellite DNA
Btrep17: bovine satellite II DNA repeat unit
The best segments of similarity were found between the bovine genomic fragment for satellite 1.709 (btsat1, repeat length 3800 bp) (Skowronska et al., 1984) and sub2CAP5T3 (98.3% of similarity on a stretch of 117 bp), and between the bovine 680 bp repeat unit of 1.723 (btsatb) (Plucienniczak et al., 1985) and sub2CAP6T3 (94.7% of similarity on a stretch of 171 bp). Sequences were matched for maximum homology and the sequence comparisons are shown in fig. 16.

a)

```
14 TCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGCGTGCGTGCGGGGG 63
3259 TGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGCGTGCGTGCGGGGG 3210
64 AGGGGAGGGGCGCTACTATACACTGATAGGTATGTNCACCAATCTGTGCAAA 113
3209 AGGGGAGGGGCGCTACTATACACTGATAGGTATGTTCACCAATCTGTGCAAA 3160
114 TGCCAATATAAAAGTCA 130
3159 TGCCAATATAAAAGTCA 3143
```

b)

```
1 GATCATGGGCTGGTCCCCTGCGCTGCGGNTGTGCTAGGGAGNCCAAACG 50
51 CCCAGGGGAAGCCCGCACAGCACATCCACCTGCCTCTAGGCGGGAGGACA 100
101 GNTCCAGGGAAGGGGGCTGTNCAGGCGGTTCCTCTGCAAAACGGGAGC 150
151 AACACTCTNGCTGTGGAGCTG 171
```

Fig. 16: Sequence homology revealed by BESTFIT, a) between sub2CAP5 sequenced with the primer T3 and the bovine satellite 1.709, b) between sub2CAP6 sequenced with the primer T3 and the bovine satellite 1.723.
The (CA)$_n$ repeat found by sequencing CAP5T3 is known to be present in the satellite 1.709. Screening with the (CA)$_n$ repeat must therefore result in the detection also of cosmids containing this satellite.

On the other hand, no dinucleotide repeat was detected after sequencing of about 500 bp of CAP6. The published sequence of the 680 bp repeat basic unit of the satellite 1.723 does not contain a (CA)$_n$ repeat. Hybridization of the (CA)$_n$ probe with CAP6 resulted therefore from a yet unknown dinucleotide repeat associated with the satellite 1.723. Another possibility is the isolation of a new sequence, which is partly homologous to the satellite 1.723.

The sequencing of c10BT21, c10BT93, c10BT133 and c10BT145 (collaboration with I. Olsaker) and of μ35Cos9 (collaboration with D. Vaiman) revealed the presence of the satellite 1.709 in all of them.

An experiment was carried out by using μ35Cos9 as the probe and by performing the CISS with the satellite 1.709 as competitor DNA. The in situ hybridization again revealed labelling on the centromeres of the autosomes. The presence of more than one satellite in the clone was therefore assumed and later confirmed (D. Vaiman, personal communication).

4.2.4. A chromosome specific satellite for discrimination between chromosome 4 and 6

The microsatellite containing cosmid clone c10BT33 was obtained after screening a library constructed in Triple Helix (Stratagene). This clone, instead of displaying a typical double spot assigned to a unique chromosome band, revealed a very uncommon hybridization signal, which encompassed the subcentromeric region of only one chromosome pair, later identified as chromosome 6. The signal was localized on 6q12-q15, and was much too intense to be due to a single 40 kb copy sequence, therefore indicating the presence of a satellite sequence in the cosmid.

Analysis of the interphase nuclei often revealed long, punctuated and diffuse signals, suggesting an interspersed organization of the repetitive sequence, instead of a tandem repetition (fig. 17). This hypothesis must of course be verified by characterizing the sequence in more detail.
The cosmid was then analyzed further, and in collaboration with I. Olsaker it was subcloned into pBluescript vectors.

FISH experiments performed with one of the subclones of 4.3 kb of length, p10BT33, gave the same intense large signal as the whole cosmid. Measurements of the signal position on the chromosome revealed the plasmid to be in the exact same region of chr. 6 as the cosmid. Thus, it appears that the plasmid contains at least part of the repetitive element producing the very strong signal. In experiments carried out with these two clones, 100% of the metaphases showed a strong hybridization signal on both homologous chromosomes 6.

Previously, the cosmid CAP8 had been assigned to chromosome 6. However, due to the difficulty in identifying this chromosome, a further experiment with the simultaneous hybridization of the two probes was carried out. Fig. 18 shows that both probes are localized on the same chromosome, so that the chromosome identification done for one probe confirms that done for the other probe. As shown in table 7, CAP8 belongs to syntenic group U15. Threadgill and Womack (1990c) assigned the genes for the caseins to syntenic
Results

group U15 and by *in situ* hybridization to chromosome 6q31-33. Therefore, the probe c10BT33 can be applied as chromosome specific marker for the identification of the "chromosome of the caseins", for instance in ISH experiments with simultaneous localization of more probes.
Fig. 18: Metaphase spreads showing the simultaneous localization of the probes CAP8 (small arrow) and c10BT33 (big arrow) on chromosome 6. a) QFQ-banding prior to hybridization, b) FITC signal on counterstained chromosomes
The identification of the chromosome carrying the probe c10BT33 and the bovine caseins corresponds to the assignment of the casein β on sheep chromosome 6 (Ansari et al., 1992). The bovine chr. 6 is in fact cytogenetically homologous to ovine chr. 6; the same is valid for chr. 4. By using ISH, Ansari co-localized the caseins and the genes for NORs on sheep and found that these genes are not located on the same chromosome. By identifying chromosomes according to the standard of Reading (1976), he assigned the NORs on 4 pairs of chromosomes, one of these being chr. 4 (personal communication).

A FISH experiment with the clone c10BT33 and the NORs genes was carried out as an additional contribution for the identification of the labelled chromosome and the two probes did not hybridize on the same chromosome. The NORs were localized on chr. 2, 3, 4, 11 and 28. In fig. 19 a partial metaphase spread with both chromosomes 6 carrying c10BT33 and some chromosomes 2, 3, 4 and 11 carrying the NORs is shown.
Fig. 19: Metaphase spreads showing the simultaneous localization of the NORs genes (small arrow) and of the probe c10BT33 (big arrow).
a) QFQ-banding prior to hybridization, b) FITC signal on counterstained chromosomes
Other publications reported different assignments of the NORs genes, indicating first misidentifications between chr. 4 and 6, and later differences which depended on whether the nomenclature of Reading or that of ISCNDA had been used to identify the chromosomes.

Very recently the caseins have been assigned on ovine chr. 4, according to ISCNDA (Hayes et al., 1992). Although the possibilities of small chromosome rearrangements during the evolution in Bovidae should be considered, we believe that the discrepancies observed are due to different interpretation of the two standards taken as reference: Reading Conference, 1976 or ISCNDA, 1989.

When Threadgill and Womack (1990c) reported the location of the casein genes on chr. 6, they stated to have used the ISCNDA standard to analyze their results. However, a careful observation of the two standards suggests probably an exchange of chromosomes: chr. 4 of ISCNDA (1989) corresponds to chr. 6 of Reading (1976) and chr. 6 corresponds to chr. 4. Because of this exchange, c10BT33 and CAP8 have been localized according to the old standard, that of Reading, while the other chromosomes have been identified according to ISCNDA.

With the availability of a marker and due to the genes already localized on this chromosome denoted as chr. 6, we propose a change in the standard ISCNDA, 1989, where the number 6 will be attributed to chr. 4 and the number 4 to chr. 6, as originally proposed in the Reading Standard.

However, whatever number the scientific community will attribute to the "chromosome of the caseins", c10BT33 can be proposed as chromosome specific marker to assist in the identification of these two chromosomes, which are very difficult to differentiate.

The identifying features of chromosomes 4 and 6 are proposed below, whereby only the typical Q-band landmarks are shown. Some typical QFQ-banded chromosomes follow as example.
Fig. 20:  

a) Schematic representation of cattle chromosomes 4 and 6 (according to Reading, 1976), the most important characteristics to distinguish these chromosomes are stressed. 
b) QFQ-banded chromosomes 4 and 6 of cattle
Chromosome 4 shows a large dark proximal band below the centromere, a distal band, which resembles the very typical band of chromosome 6 but it is more distal, and a very terminal (telomeric) band, which indicates the end of the chromosome.

Chromosome 6 resembles chr. 4 but lacks the terminal band. Moreover, the distal band is less distal than on chr. 4. Close to the centromere, the chromosome 6 often shows a narrowing. This is very typical of this chromosome, as drawn in fig. 20 a) and as visible in some chromosomes of fig. 20 b).
5. Discussion

The results obtained in this thesis have shown that the technique described is a useful tool for the development of a marker map in the bovine genome. The heterogeneous group of probes available revealed different types of assignments, chromosome arm specific microsatellites, centromeric clones, chromosome specific satellites. Their significance, their application and new perspectives are discussed.

5.1. Significance of the physical mapping of the clones

The localization of CAP11 to chromosome 16 and its assignment to the syntenic group U1 allow the assignment of the whole syntenic group to chromosome 16. Five DNA segments and 6 genes belong to this syntenic group, among those also the polyimmunoglobulin receptor gene (PIGR, Kulseth et al., 1992). In agreement with the assignment of CAP11, the phage clone for PIGR hybridizes on 16q13.

CAP9 is syntenic with a marker for the formerly unassigned syntenic group U16. The loci of U16 can therefore be assigned to chr. 11. This assignment is confirmed by a recent chromosomal mapping of another member of U16 (Hayes and Petit, 1993).

CAP2 and CAP10, which hybridize on chr. 5, have been assigned to U3. This result is in agreement with the previous assignment of the syntenic group, based on other probes.

With the progress in establishing the bovine map, the results confirm each other and eventual discrepancies can be found and corrected. The combined use of in situ hybridization and analysis of somatic hybrid cell lines allows the rapid development of the gene maps of any species.

In the case of CAP8 and c10BT33, which are localized on a difficult chromosome, the knowledge of the previous localization of the syntenic group to which they have been assigned was of great help in the unambiguous identification of the chromosome, as discussed in 4.2.4.

24A, CAP7, MAP1B, M10/13 and C8 represent the first assignments on chr. 3, 13, 20, 22 and 25, respectively. Therefore, it will be very important to develop their primers and to assign them to a syntenic group. In this way new syntenic groups will be assigned to chromosomes. However, primers designing
was not always possible, either because the probes belonged to other laboratories or because sequencing to design primers did not provide the necessary informations.

Of the three candidates for the calpastatin gene, only the probe 24A probably represents the real sequence of the gene. The two other probes could contain false positives.

The probe 99-1, a candidate for the growth hormone receptor, hybridized on chr. 7q24, while in human this gene has been assigned to chr. 5p14-p12. The gene for the RAS p21 protein activator (RASA), which is localized on human chr. 5 as well (5q13), has also been assigned to bovine chr. 7q24-pter (Eggen et al., 1992). Moreover, a more distal large portion of human chr. 5q is known to be conserved on bovine chr. 7 (O’Brien et al., 1993). Therefore, we can conclude that a portion of the short arm of human chromosome 5 is also conserved in bovine chr. 7 and that the probe 99-1 is likely to be the right candidate for the growth hormone receptor. This assignment represents a new information supporting the homology described.

The combination of more techniques can result in the discovery of incongruities, as the assignment of CAP3 and MAP1B to syntenic group U20 demonstrated. U20 is in fact known to be on chromosome 23, while the results of in situ hybridization clearly show that CAP3 is on chr. 19 and MAP1B on chr. 20. By contrast, MSI/2, which has also been localized on chr. 19, gave a positive signal on the somatic hybrid panel on syntenic group U21. U21 has, in fact, previously been assigned to chr. 19. Such results indicate conflicts in the assignment of this syntenic group by using the hybrid panel available. The discrepancy can be explained by a non entirely random chromosome segregation in the hybrid clones of the panel. The need to analyze an independent panel to resolve this group into at least two independent syntenic groups requires international collaboration between laboratories where hybrid panels are available.

Moreover, particular interest should be devoted to chr. 19, to which already 4 microsatellites have been assigned. This chromosome carries genes, such as the growth hormone (Hediger et al., 1990), the homeobox region 2 (Gunawardana and Fries, 1992), the cytokeratin (Fries et al., 1991) and others. Therefore, linkage studies between these genes and the closest marker could further characterize the bovine chr. 19.
5.2. Centromere specific satellites

Two different kinds of repeated sequences in the mammalian genome are distinguished: tandemly repeated DNA and reiterated segments interspersed among other genomic sequences. Tandem repetitions, widely termed satellites, are based on repetition of a monomer repeat unit and occur usually at the centromeres and telomeres of the chromosomes. Reiterated segments are LINEs (Long Interspersed repetitive Elements) and SINEs (Short Interspersed repetitive Elements). The function of repeated DNA is still unclear, the most part of these sequences being non coding. However, there are repeated sequences known to be parts of genes, as for instance ribosomal RNA genes.

Repetitive sequences in mammalian genomes have been reviewed by Singer (1982), the bovine genome has been analyzed by Macaya et al., (1978). Eight different centromeric satellites have been described in bovine DNA. Compared to other eukaryotic genomes, these satellites constitute a larger proportion, that is over 23% of the total DNA. The bovine repeat units are a complex set of interrelated sequences, studies on these sequences suggest the development from a common ancestral sequence at least for some of them (Plucienniczak et al., 1982; 1985). However, the sequences of bovine 1.723 and 1.709 (s. results) are not related to the other bovine satellites.

The satellite 1.709 consists of a 3800 bp long repeat and represents 4.6% of the bovine genome. The satellite 1.723, with a repeat length of about 700 bp represents 0.05% of the genome (Macaya et al., 1978). Therefore, the probability to obtain a DNA fragment containing a centromeric satellite sequence in a cosmid vector is noticeable. Moreover, 1.709 contains a \((CA)_n\) repeat, so that screening with a poly \((CA\cdot GT)_n\) easily provides positive satellites. This explains the high proportion of centromere specific cosmids. The second possibility is the formation of chimeras, favoured by particular vectors, as mentioned in 4.2.3. Although this is usually a phenomenon occurring when working with Yeast Artificial Chromosomes (YACs) (Selleri et al., 1992), it likely also occurs with cosmids.

Due to the localization of clones containing different satellite sequences, such as CAP 5 (satellite 1.709) and CAP 6 (satellite 1.723) on the bovine centromeres, there is evidence for the presence of different satellites in the centromeres of the autosomes. This observation suggests a similarity to what was found on the human centromeres, which appears as complex assemblies of numerous different satellite sequences. The variation in the intensity of
hybridization from one chromosome to another observed by hybridizing with a clone containing the satellite 1.709 (s. table 9) indicates either that the number of repeat units varies greatly or that subsequences are present for this satellite. Moreover, due to the absence of hybridization signal on X and Y chromosomes, there is evidence for a difference in the centromeres between autosomes and sex chromosomes.

A possible approach to exclude clones carrying satellite sequences prior to ISH is the PCR amplification of the DNA with primers developed from the more common centromere specific sequences. In the case of positive amplification, the clone must be eliminated, while absence of amplification indicates that the probe does not contain that type of sequences. Recently, Vaiman et al. (1992) designed the ASPS method (anti-satellite polymerase chain reaction screening) to overcome this problem, specifically for the satellite 1.709.

The need to routinely include such a step in searching for (CA)n containing cosmids, prior to ISH, has been already envisaged and is more frequently practiced.

5.3. Chromosome specific satellites

The sequences contained in the clones c10BT33 and CAP4 are the first chromosome specific bovine satellites reported to date, aside from the repetitive sequence on the Y-chromosome (Perret et al., 1990).

An autosome specific large repeat unit very similar to c10BT33 has been found on mouse chromosome 8 (Boyle and Ward, 1992). This clone on mouse is very species specific. By analogy, the satellite found on the bovine chromosome 6 could be species specific as well. To test this hypothesis, FISH experiments on the chromosomes of other Bovidae and also of other animal species should be performed. Hybridization on sheep chromosomes revealed clear hybridization signals on the heterochromatic subcentromeric region of chr. 6 (data not shown). This region is the homologue of the assignment found on bovine chr. 6. Thus, this satellite is localized on the same region as in cattle. Moreover, experiments to estimate the number of copies of the repeat in the bovine genome could be performed and compared to the 60 to 80 copies found in the mouse (Boyle and Ward, 1992).

The sequence analysis of CAP4 suggests the presence of minisatellite sequences, which are in this case specific for the telomere of chr. 17.
At this point it would be very interesting to study whether each chromosome contains some sort of chromosome specific repeat unit and if homology between these repeats can be found. Their study generates a first interest in ascribing possible functions and in better understanding the amplification and evolution of such genomic sequences.

Another important role to be mentioned is the function of such repeat elements as chromosome specific markers for cytogenetic purposes, as described in 4.2.4. Due to the difficulties in the identification of the bovine chromosomes and to the need to confirm chromosomal assignments, especially for the small chromosomes, it would be very convenient to have an "identification kit". The "identification kit" could consist of a set of probes, one specific for each bovine chromosome, suitable to be used in cohybridization experiments, by producing strong signals on all the metaphases.

5.4. Use of microsatellites as markers

The frequency of (CA)$_n$ microsatellites in the bovine genome has been estimated to be much lower than in other mammalians (s. 1.2.2.2.1., Solinas Toldo et al., 1993). Therefore, this type of markers is probably still sufficient for a low resolution marker map, but it could be difficult to develop high resolution maps with (CA)$_n$ microsatellites.

As shown in fig. 12, the probes considered in this work suggest a clear preferential location in the telomeric and subtelomeric regions of the bovine chromosomes, although microsatellites seem to occur randomly in other species (Hearne et al., 1992). Therefore, the establishment of a marker map consisting of randomly found dinucleotide repeats on the bovine genome will require the placement of more loci than in other species.

Beyond the physical localization on the genome through in situ hybridization or the use of a hybrid panel, microsatellites, either obtained by a library screening with a poly (CA-GT)$_n$ oligonucleotide probe or from a GenBank database, need to be characterized. The characterization of markers consists of sequencing, determination of their polymorphism information content value (PIC, s. 1.2.1.) and PCR analysis on genomic DNA samples from a reference family, which reveals the number of alleles. Eight of the probes used in this work have been characterized further and 7 revealed polymorphism (Solinas Toldo et al., 1993). Moreover, a correlation between number of dinucleotide repeats and number of alleles has been observed, as previously reported by Weber (1990).
Only a map of suitable polymorphic markers can be used to test linkage relationship with relevant genes.

5.5. Completion of the bovine marker map

A physical map of genetic markers requires microsatellites covering the whole genome and evenly distributed along the chromosomes. As already discussed in 1.3., an average of three polymorphic markers are required for each cattle chromosome.

The search for the markers used in this work has been done randomly. The clones have in fact been randomly sorted from a cosmid or a phage library. Therefore, it is evident that after the first chromosomal localizations, saturation of some chromosomes or restricted regions of chromosomes will occur, while other regions of the genome will be still uncovered. This has already been occurred for instance on bovine chromosome 19, which carries 4 microsatellites, while others are still empty.

Consequently, since marker spacing cannot be directed, much more than 90 optimally situated markers need to be mapped to cover the bovine genome and alternative strategies should be developed as completion of the proposed approach.

An interesting technique to isolate clones carrying sequences from a specific chromosome region involves the microdissection and microcloning of regions where no markers have yet been localized. The chromosome region is cut out, the DNA content purified and amplified by PCR. The PCR product is labelled and then used either directly for in situ hybridization (Lengauer et al., 1991; Hampton et al., 1991), or it is first used to screen a library and isolate new clones. Lüdecke et al., (1990) and Fiedler et al. (1991) propose this approach for the direct development of markers linked to genetic syndroms. A set of markers with the proper marker spacing could then be set up for efficient linkage studies. In cattle, however, a major disadvantage of this technique is the impracticable identification of the chromosomes directly under the microscope, due to the acrocentric nature of all autosomes, thereby not allowing the use of the centromere position as help in the identification. Banding procedures, which do not interact with a subsequent PCR amplification of the DNA, could be used for the identification of the chromosomes during the microdissection. A more careful method, although quite laborious, would be the banding and subsequent photographic documentation of good metaphases, as described in the methods.
Subsequently, with the photograph, the microdissection of the desired chromosome should be feasible.

Alternatively, markers on a predetermined chromosome can be generated from chromosome specific libraries, prepared from flow-sorted chromosomes. Again, difficulties are encountered when working with cattle chromosomes; their sorting is not straightforward, due to their large number and uniformity. The exploitation of chromosome translocations or the use of hybrid cell lines is of assistance in their sorting. Dixon et al. (1992) investigated flow karyotypes of domestic animals with a fluorescence-activated cell sorter (FACS) and demonstrated that cattle-mouse hybrids are an important tool to resolve the poor separation of chromosomes in cattle.

FACS separation has already been used for making chromosome-specific libraries and chromosome-specific markers in humans (Van Dilla et al., 1986; Cotter et al., 1989).

In order to have landmarks on the genome which facilitate the identification of genes, it would be convenient to develop CpG islands as gene markers. CpG islands are short dispersed regions of non-methylated DNA with a high frequency of CpG dinucleotides, which have been found to be associated with genes. In human, almost all housekeeping genes carry a CpG island and 40% of tissue specific genes are associated with islands (Larsen et al., 1992). The exploitation of such knowledge permitted for instance the isolation of a genomic sequence close to the cystic fibrosis locus, which was a strong candidate for the CF gene (Estivill et al., 1987). By knowing this approach already applied in the human genome, CpG islands should be considered in the search for new markers also in the bovine genome.

5.6. Improvements and future developments

As already outlined, the FISH technique is a relatively new powerful method which allows to rapidly and precisely localize DNA probes for instance on metaphase chromosomes. The advantages of FISH compared to other techniques are undisputed, and have been demonstrated in this work, but in order to exploit its full potential new developments are under way.

An important part of FISH is the photographic documentation of the results. Fluorescent signals are generally detectable using a conventional epifluorescence microscope with the appropriate filters and equipped for photomicroscopy. In the case of weak signals, high quality optics and filters with
very specific wavelengths for the fluorochromes used are required. This type of
documentation demands the use of special photographic films, which are not
always easy to handle. They often require long exposures of the fluorescent
signal to UV-light, which is associated with a rapid fading of the signal.
Moreover, in the case of multicolor FISH, the documentation of hybridization
results cannot be routinely performed with only a conventional microscope,
because multiple exposures of films do not adequately resolve images from
metaphases differently labelled. This restriction is overcome with digital imaging
camera systems, such as charge coupled device (CCD) camera. This device
allows one to collect gray scale images of each fluorophore and to perform
powerful image processing. Different images can for instance be overlapped
and pseudocolours are attributed to the fluorochromes. Since there is the
possibility for digital filtering and enhancing the ratio between signal and
background, improvements in the detection of the signal are also achieved. In
some cases, signals which are not visible with traditional equipment can be
detected with these new instruments.
6. Conclusion

The aim of this work was to establish a reliable and efficient technique to map highly polymorphic marker loci on the bovine genome, which serve as reference loci along the chromosomes. The physical mapping by fluorescence in situ hybridization of clones, such as cosmids, phages, markers and genes demonstrated the versatility of this technique. The relatively consistent appearance of the hybridization signal on both homologous chromosomes required on average the analysis of less than 20 metaphases for each assignment, rendering fluorescent in situ hybridization much more efficient than radioactive in situ hybridization.

The clones assigned to specific chromosomes can be used as landmarks on these chromosomes or as tool for their identification. The discovery of chromosome specific satellite is particularly interesting. These probes hybridize very efficiently on all the metaphases and are very suitable chromosome specific markers, especially when located on cytogenetically difficult chromosomes.

Since a considerable part of the bovine genome consists of repetitive sequences, some of them with a (CA)_n repeat, the probability to obtain a DNA fragment containing a centromeric satellite in a cosmid vector must be considered. An anti-satellite detection step could be a necessary additional step in the search for microsatellite sequences.

Although tendentiously located more in the subtelomeric regions of the chromosomes and showing wider spacing in the bovine genome than in other mammalian species, microsatellites are still good candidates to develop maps of polymorphic markers.
7. Probe nomenclature

For the development of a bovine gene map which is understandable and internationally available to everybody, each probe assigned, especially anonymous loci, should be named according to a conventional nomenclature. Based on the last up-to-date of the bovine genome map (Fries et al., 1993), the following nomenclature is adopted.

<table>
<thead>
<tr>
<th>Probe</th>
<th>DNA Segment</th>
<th>Chromosome</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAP1</td>
<td>DNA segment</td>
<td>D21S8</td>
</tr>
<tr>
<td>CAP2</td>
<td>DNA segment</td>
<td>D5S2</td>
</tr>
<tr>
<td>CAP3</td>
<td>DNA segment</td>
<td>D19S2</td>
</tr>
<tr>
<td>CAP4</td>
<td>DNA segment</td>
<td>D17S2</td>
</tr>
<tr>
<td>CAP7</td>
<td>DNA segment</td>
<td>D13S1</td>
</tr>
<tr>
<td>CAP8</td>
<td>DNA segment</td>
<td>D6S3</td>
</tr>
<tr>
<td>CAP9</td>
<td>DNA segment</td>
<td>D11S5</td>
</tr>
<tr>
<td>CAP10</td>
<td>DNA segment</td>
<td>D5S3</td>
</tr>
<tr>
<td>CAP11</td>
<td>DNA segment</td>
<td>D16S5</td>
</tr>
<tr>
<td>MSI/2</td>
<td>DNA segment</td>
<td>D19S3</td>
</tr>
<tr>
<td>MSI/3</td>
<td>DNA segment</td>
<td>D8S1</td>
</tr>
<tr>
<td>C10BT33</td>
<td>DNA segment</td>
<td>D6Z1</td>
</tr>
<tr>
<td>C10BT34</td>
<td>DNA segment</td>
<td>D19S4</td>
</tr>
<tr>
<td>C10BL26</td>
<td>DNA segment</td>
<td>D1S12</td>
</tr>
<tr>
<td>C5</td>
<td>DNA segment</td>
<td>D19S5</td>
</tr>
<tr>
<td>C8</td>
<td>DNA segment</td>
<td>D25S1</td>
</tr>
<tr>
<td>M10/13</td>
<td>DNA segment</td>
<td>D22S1</td>
</tr>
<tr>
<td>24A</td>
<td>DNA segment</td>
<td>D3S1</td>
</tr>
<tr>
<td>99-1</td>
<td>DNA segment</td>
<td>D7S3</td>
</tr>
</tbody>
</table>

In this nomenclature "D" indicates a DNA segment. "D" is followed from the number of the chromosome on which it is located. "S" means segment, followed by its number.

D-numbers are also assigned to sequences obtained after screening of libraries for specific genes as long as their identity is not confirmed (C10BL26, supposed lysozyme coding, 99-1 growth hormone receptor, 24A calpastatin).
8. References


Boyle, A.L. and Ward, D.C. (1992) Isolation and initial characterization of a large repeat sequence element specific to mouse chromosome 8, <i>Genomics</i>, 12, 517-525.


Eggen, A., Solinas-Toldo, S., Dietz, A.B., Womack, J.E., Stranzinger, G., Fries, R. (1992) RASA contains a polymorphic microsatellite and maps to bovine syntenic group U22 on chromosome 7q2.4-qter, Mammalian Genome, 3, 559-563.


References


References


References


derived microsatellite markers as anchor loci on bovine chromosomes, *Mammalian Genome, 4, in press.*


Wilchek, M. and Bayer, E.A. (1989) Avidin-biotin technology ten years on: has it lived up to its expectations?, *Trends in Biological Sciences*, 14, 408-412.


Leer - Vide - Empty
Curriculum Vitae

Sabina Solinas Toldo
born in Milan on 28th April 1965
citizen of Rovio, TI, Switzerland
married

1971-1976  Primary School in Milan (Italy)
1976-1981  Literargymnasium in Mendrisio, TI (Switzerland)
1981-1984  High School degree, Matura Type C (scientific) with Latin and Greek in Mendrisio
1984-1989  Study of Agriculture Engineering, Animal Sciences at the Swiss Federal Institute of Technology (ETH), Zurich
1989      Diploma thesis by Prof. G. Stranzinger and Dr. R. Fries
1990 (4 months)  Stage at the NIH, NCI, laboratory of biology, Bethesda (USA), by Dr. J. DiPaolo and Dr. N. Popescu
                Collaboration with the USDA, reproduction laboratory, Beltsville (USA) with Dr. V. G. Pursel
1992 (2 weeks)  EMBO course on "Non-isotopic in situ hybridization to metaphase chromosomes and interphase nuclei", in Heidelberg, by Dr. P. Lichter