



## Identity-by-descent mapping of recessive traits in livestock: application to map the bovine syndactyly locus to chromosome 15.

C Charlier, F Farnir, P Berzi, et al.

*Genome Res.* 1996 6: 580-589

Access the most recent version at doi:[10.1101/gr.6.7.580](https://doi.org/10.1101/gr.6.7.580)

---

**References** This article cites 23 articles, 5 of which can be accessed free at:  
<http://genome.cshlp.org/content/6/7/580.full.html#ref-list-1>

### License

**Email Alerting Service** Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

---

To subscribe to *Genome Research* go to:  
<https://genome.cshlp.org/subscriptions>

---

Copyright © Cold Spring Harbor Laboratory Press

## RESEARCH

# Identity-by-descent Mapping of Recessive Traits in Livestock: Application to Map the Bovine *Syndactyly* Locus to Chromosome 15

Carole Charlier,<sup>1,2</sup> Frédéric Farnir,<sup>1,2</sup> Paulette Berzi,<sup>1</sup>  
 Pascal Vanmanshoven,<sup>1</sup> Benoît Brouwers,<sup>1</sup> Hans Vromans,<sup>3</sup> and  
 Michel Georges<sup>1,4</sup>

<sup>1</sup>Département de Génétique, Faculté de Médecine Vétérinaire, Université de Liège (B43), 4000-Liège, Belgium; <sup>3</sup>Holland Genetics, 6800 AL Arnhem, The Netherlands

Twelve animals affected with syndactyly or mulefoot were sampled in the Dutch black-and-white cattle population. Analysis of the pedigree data revealed that all of these individuals traced back to a single acknowledged carrier founder individual. Between seven and nine generations separated the founder from its 12 affected descendants. The 12 affected offspring were genotyped for a battery of 213 microsatellites spanning the 29 bovine autosomes. The resulting genotypes were analyzed using a maximum likelihood approach searching for shared homozygous haplotypes among affected individuals. Three candidate regions for the *syndactyly* locus emerged from this initial screening. *syndactyly* was shown to map to one of these candidate regions on chromosome 15 by genotyping 29 additional individuals linking founder and affected offspring and performing a conventional linkage analysis with the LINKAGE programs. This study illustrates the potential of identity-by-descent mapping in livestock populations.

Linkage analysis has proven to be an efficient strategy to determine the localization of genes underlying a variety of inherited phenotypes allowing for their subsequent positional cloning. For a number of inherited traits, especially if genetically complex, the pedigree resources conventionally considered to be appropriate for the application of linkage strategies (i.e., nuclear families characterized by multiple informative offspring) might not exist. This limitation is often encountered when studying inherited disorders in livestock populations. The approach usually implemented to overcome this limitation in domestic species has been to breed informative pedigrees de novo. Obviously this is a very time-consuming and expensive approach.

Alternative and more powerful strategies have been devised recently to compensate for the lack of ideal family material when seeking to map disease genes (Lander and Botstein 1987; Puffenberger et al. 1994; Houwen et al. 1995; Kruglyak and Lander 1995; Kruglyak et al. 1995). These

identity-by-descent (IBD) mapping strategies have in common that they exploit the high density marker maps that are becoming available for an increasing number of organisms and that they look for sharing of marker haplotypes among individuals expressing the trait of interest.

In this paper we describe the application of such a strategy to map the gene causing bovine hereditary syndactyly to chromosome 15. The main feature of hereditary bovine syndactyly, also referred to as mulefoot, is the fusion or non-division of functional digits. The lesions are primarily osteologic in nature, and consist mainly in synostotic phalanges. Detailed examination, however, reveals that more proximal limb structures can be affected as well. In particular, metacarpal and metatarsal III and IV bones show premature fusion in syndactylous fetuses when compared to normal controls. Interestingly, these synostoses are accompanied by an enhanced development of the metacarpal and metatarsal II and V bones, showing well-developed rudimentary digits II and V. The latter symptoms are reminiscent of a polydactyly syndrome. Bovine hereditary syndactyly shows variable expressivity in the sense that the second phalanges are more of-

<sup>2</sup>These authors contributed equally to this work.

<sup>4</sup>Corresponding author.

E-MAIL [michel@stat.fmv.ulg.ac.be](mailto:michel@stat.fmv.ulg.ac.be); FAX 32-41-66.41.22.

ten synostotic than the third and first but mainly due to an interesting right-left and front-rear gradient, meaning that the right front foot is affected more often, followed by the left front, right rear, and left rear foot. Intriguingly, these anatomic defects are accompanied by a susceptibility to hyperthermia induced by elevated ambient temperatures (Eldridge et al. 1951; Leipold et al. 1969; Hart-Elcock et al. 1987).

In cattle, syndactyly has been shown to segregate as an autosomal recessive trait with incomplete penetrance. The corresponding locus is referred to as *SY*, with two alleles *SY*<sup>+</sup> and *sy*. Hereditary syndactyly has been described in a number of cattle breeds, including Angus, Brown Swiss, Chianina, Hariana, Japanese Improved, Simmental, and Spotted Mountain, but occurs at the highest frequency in U.S. Holsteins. It has been argued that the high frequency of the disorder observed in Holsteins might be linked to a superiority of the carriers (*SY*<sup>+</sup>/*sy*) for milk and butterfat production when compared to normal (*SY*<sup>+</sup>/*SY*<sup>+</sup>) controls (Leipold et al. 1973).

At present, diagnosis of the carrier status for syndactyly requires time-consuming and expensive progeny testing (i.e., mating of the tested individuals to affected or known carrier animals). The progeny-testing procedure has been improved in some respect when it was demonstrated that the syndactylous condition could be determined on 60-day-old fetuses (Hart-Elcock et al. 1987). However, there is a definite need for a cheaper and more efficient genetic diagnostic test for the mulefoot condition (as for several other genetic disorders in livestock) that could be applied at any developmental stage. The identification of markers linked to the *syndactyly* (*SY*) locus will allow for marker-assisted selection against this disorder.

## RESULTS

### A Genome-Wide Scan Reveals Three Candidate Regions for the *SY* Locus Using IBD Analysis

To map the *SY* locus blood samples were collected from 12 affected individuals from the Dutch black-and-white Friesian cattle population. Analysis of the pedigree data (Fig. 1) revealed that all 12 individuals traced back to a single acknowledged mulefoot carrier by two pathways characterized by seven to nine meioses. Therefore, it could be postulated that these 12 individuals would likely be homozygous for a

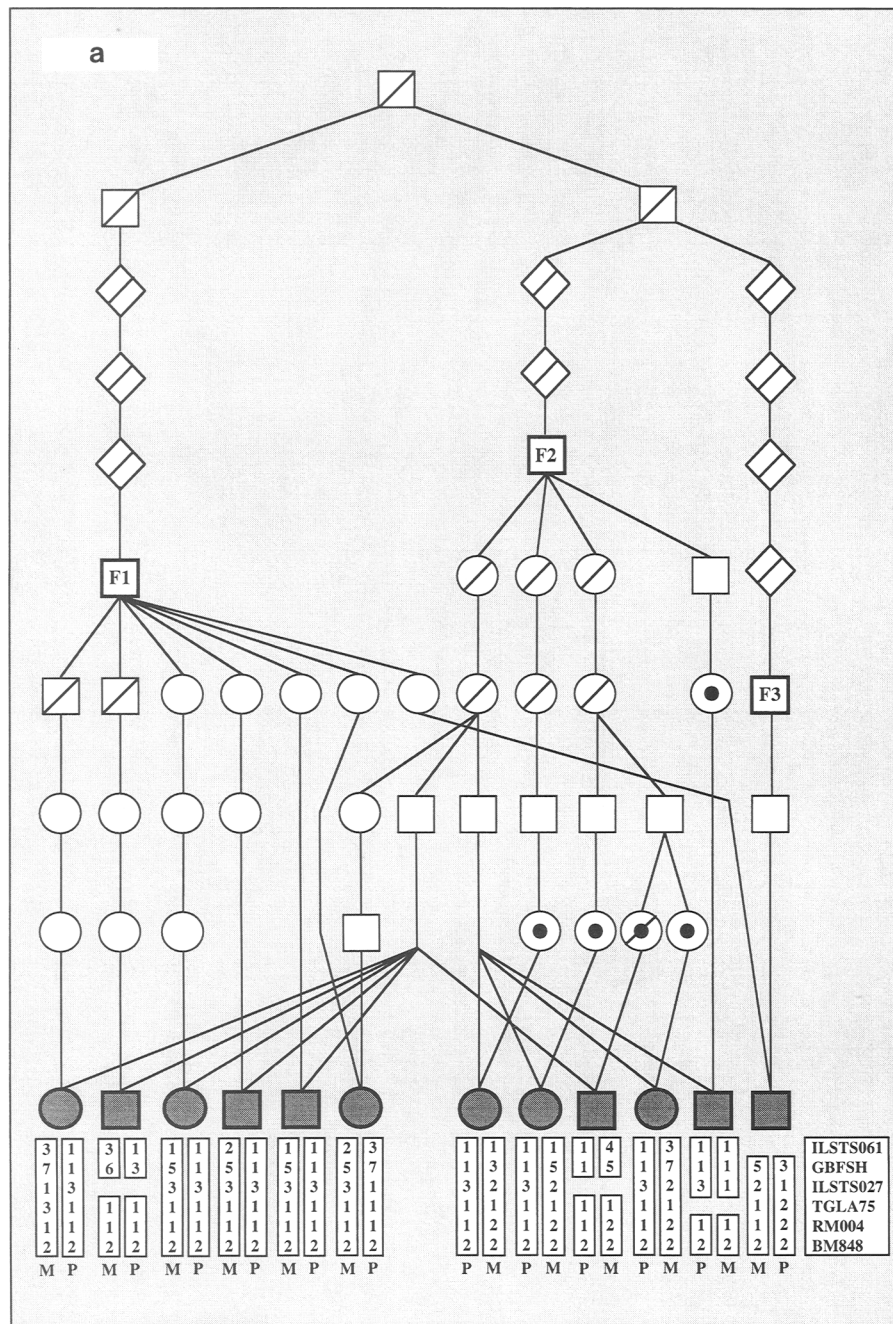
common chromosome segment flanking the *SY* locus (Houwen et al. 1995).

A genome scan was undertaken using 213 microsatellite markers compiled from three published marker maps (Barendse et al. 1994; Bishop et al. 1994; Georges et al. 1995) and spanning the 29 bovine autosomes. On the basis of published map data we estimated that the selected marker set provided a coverage of ~2565 cM (Kosambi) as bracketed chromosome segments (average ~14 cM), which might represent as much as 80% from the bovine genome.

Marker genotypes were determined for the 12 affected individuals, as well as an equimolar DNA pool of 10 unrelated individuals from the same breed. The resulting marker genotypes were analyzed with the ANIBD program (F. Farnir and M. Georges, in prep.) to identify chromosomal regions where marker haplotypes were shared among affected individuals more often than expected by chance (see Methods). The microsatellite patterns obtained from the DNA pool were analyzed by visual examination to estimate marker allele frequencies required to perform the likelihood calculations. The results were expressed as multipoint lod score curves, and are represented for the entire genome in Figure 2.

Analysis of the lod score patterns reveals unusually high lod scores (i.e., >3) for most chromosomes. At least two factors could underlie this inflation of lod scores. (1) Linkage between a marker and the disease locus will manifest itself primarily by a difference in the marker allele frequencies between a control population and the sample of affected individuals. Therefore, inaccurate estimation of control allelic frequencies is expected to generate spurious positive lod scores. Given the crude method applied for the estimation of marker allele frequencies, this effect could be expected in this analysis. (2) The algorithm used (see Methods) assumes that the pathways connecting the affected descendants and the common founder are independent. Examination of the pedigree data (Fig. 1) shows that this assumption is violated for several offspring.

To evaluate the impact of both these factors on the generated lod scores, we simulated the segregation of a chromosome segment of 50 cM unlinked to the disease locus in a pedigree comprising 12 individuals tracing back to a common ancestor through two pathways. For each descendent, the length of both pathways was chosen to mimic the situation in the actual *SY* pedigree (Fig. 1). The chromosome segment was subdivi-



**b** Markers Allelic frequencies

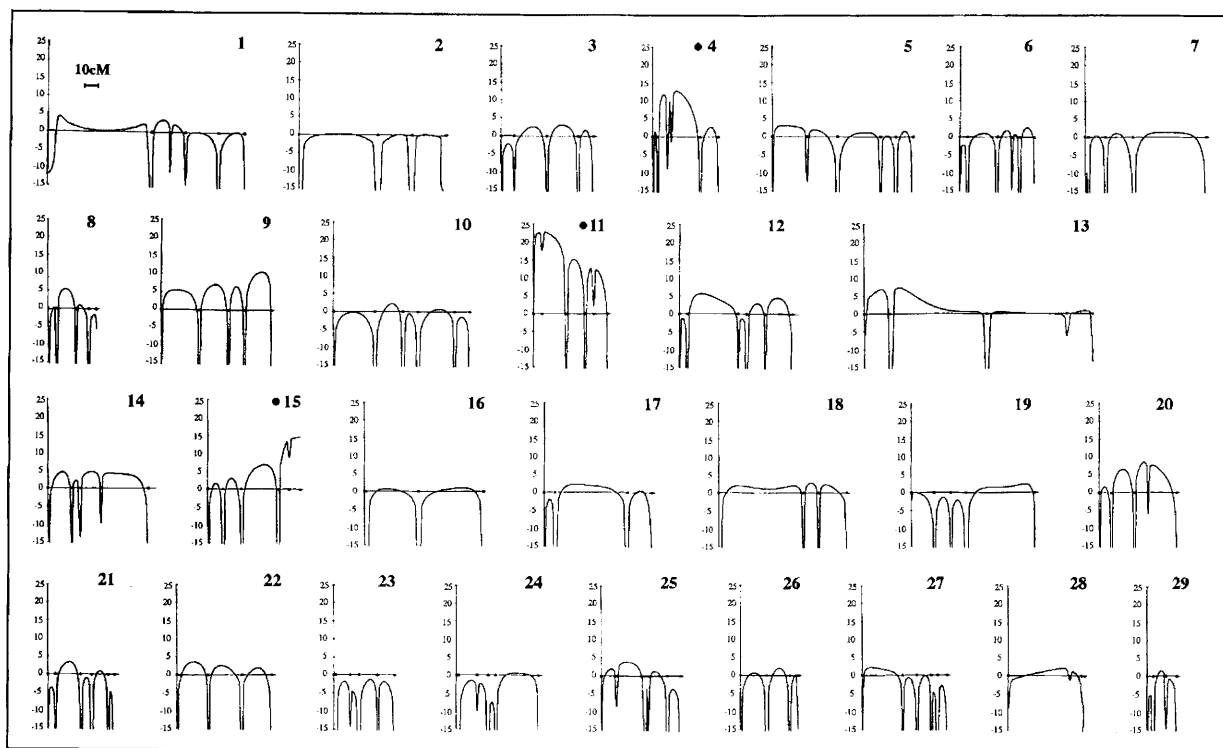
Markers	Allelic frequencies						
	1	2	3	4	5	6	7
ILSTS061	.50	.13	.24	.13			
GBFSH	.10	.10	.10	.13	.10	.13	.34
ILSTS027	.53	.26	.21				
TGLA75	.38	.12	.29	.21			
RM004	.24	.38	.10	.04	.24		
BM848	.47	.53					

**Figure 1** (a) Familial relationships between the 12 individuals affected with syndactyly (filled symbols). Missing individuals are marked with a diagonal line. The microsatellite genotypes of the affected individuals are shown for the six distal chromosome 15 markers. Alleles are sorted (boxed) by parental haplotype (M, maternal; P, paternal), whenever this information could be inferred unambiguously from available parental genotypes. Conventional linkage analysis was performed by exploiting the gene flow from known carrier individuals  $F_1$ ,  $F_2$ , and  $F_3$  to the affected offspring as described in Results, requiring the cutting of loops at the individuals marked by a dot. (b) Maximum likelihood marker allele frequencies for the six distal chromosome 15 microsatellites as estimated with the genotypes of the available members of the SY pedigree.

maternal; P, paternal), whenever this information could be inferred unambiguously from available parental genotypes. Conventional linkage analysis was performed by exploiting the gene flow from known carrier individuals  $F_1$ ,  $F_2$ , and  $F_3$  to the affected offspring as described in Results, requiring the cutting of loops at the individuals marked by a dot. (b) Maximum likelihood marker allele frequencies for the six distal chromosome 15 microsatellites as estimated with the genotypes of the available members of the SY pedigree.

vided by five polymorphic markers into four segments of 20, 15, 10, and 5 cM, respectively, corresponding to intervals comparable to the actual ones. It was assumed that each of these markers was characterized in the population of interest by four alleles with randomly drawn frequencies adding to one. This model yields an average heterozygosity of 67% very similar to the actual heterozygosity of 61% found for the collection of used microsatellites in the Dutch black-and-white cattle population. In a first simulation all pathways linking founder and descendants were considered to be independent. To test the effect on the obtained lod scores of inaccuracies in the estimation of marker allele frequencies, we compared the distribution of lod scores obtained along the map when using the true marker allele frequencies with that obtained when using frequency estimates. To mimic the experimental conditions as accurately as possible, the estimates were obtained from a simulated random sample of 10 unrelated individuals and rounded out to the closest multiple of 10%. To test the effect of the nonindependence of pathways linking founder and affected offspring, the same simulated chromosome segment was now allowed to segregate in a tree with the actual topology of our SY pedigree as represented in Figure 1.

Three thousand simulations were performed un-



**Figure 2** Multipoint lod scores obtained with the IBD analysis for the 29 bovine autosomes. The dots on the x-axis (cM) represent the positions of the available microsatellite markers for the respective chromosomes. The size of a 10-cM segment is shown. Lod scores are given on the y-axis on a scale from +25 to -15.

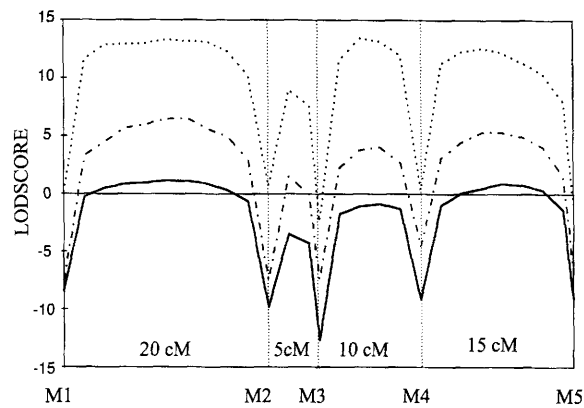
der each assumption. Figure 3 shows the distribution of the 99.9% lod score percentiles as obtained for these different conditions along the hypothetical map, corresponding to the significance threshold associated with a 5% type 1 error rate. Indeed, to account for the fact that we performed a whole genome scan spanning ~2500 cM (Kosambi), which is 50 times more than the simulated 50-cM chromosome segment, we applied a Bonferroni correction  $[(1 - \alpha_{5\%}) = (1 - \alpha_{0.1\%})^{50}]$  shifting the threshold associated with a type 1 error of 5% from the 95% to the 99.9% percentile. Analyzing genotypic data simulated in a pedigree with independent pathways connecting founder and affected offspring using estimated rather than true marker allele frequencies shifts the required lod score threshold from 1.1 to 6.5. When analyzing genotypes simulated in a pedigree with the structure of the actual one (Fig. 1) and using estimated allele frequencies, the lod score threshold moves up to 13.5.

As can be seen from Figure 2, three chromosomes yielded lod scores superior to this threshold value of 13.5 in our quest for the *SY* locus: chromosome 4 ( $Z_{\max} = 13.6$ ), chromosome 11

( $Z_{\max} = 23.1$ ), and chromosome 15 ( $Z_{\max} = 15.1$ ). Therefore, the corresponding chromosome regions were considered good candidates and were selected for further analysis.

#### Conventional Linkage Analysis Confirms the Localization of the *SY* Locus on Bovine Chromosome 15

To verify the role of the three respective candidate regions in the determination of syndactyly, we collected samples from an additional 29 individuals belonging to the pedigree represented in Figure 1. These individuals were genotyped for the microsatellites flanking the highest lod scores on chromosomes 4, 11, and 15, and the resulting genotypes used in a conventional linkage analysis using the LINKAGE and FASTLINK programs (Lathrop and Lalouel 1984; Cottingham et al. 1993). The numerous marriage and inbreeding loops, however, as well as the remaining high proportion of missing data, preclude the analysis of the complete pedigree data using these programs. This illustrates their limitations and the



**Figure 3** The 99.9% lod score percentile obtained with the ANIBD programs from  $3 \times 3000$  simulated segregations of a 50-cM chromosome unlinked to the disease locus: The solid line indicates a pedigree characterized by 12 affected individuals tracing to a common founder through independent pathways (length of the pathways as in Fig. 1) with the real marker allele frequencies in the likelihood calculations; the dashed-dotted line indicates the same pedigree but with marker allele frequencies estimated from a sample of 10 unrelated individuals in the likelihood calculations, and the dotted line indicates the pedigree represented in Fig. 1 with marker allele frequencies estimated from a sample of 10 unrelated individuals in the likelihood calculations. The positions of the markers on the simulated chromosome are marked by M1, M2, ..., M5, bounding four intervals of 5, 10, 15, and 20 cM, respectively.

need to develop alternative strategies. Nevertheless, to exploit the additional genotypes, the pedigree was split into two subpedigrees, A and B, as described previously (Georges et al. 1993). Subpedigree A dealt with the gene flow between 6 of the affected individuals and a proven mulefoot carrier sire referred to as  $F_2$  in Figure 1, whereas subpedigree B addressed the corresponding gene flow between the 12 affected individuals and 2 other proven mulefoot carrier sires,  $F_1$  and  $F_3$ . To avoid extracting redundant information from the six affected offspring tracing back to  $F_2$  by pathway  $\alpha$  and either  $F_1$  or  $F_3$  by pathway  $\beta$ , and therefore included in both subpedigrees, the relationships between the parents located in subpedigree B were ignored when analyzing subpedigree A and vice versa. The ancestral relationships between individuals  $F_1$ ,  $F_2$ , and  $F_3$  were ignored as no genotypic information was available for the corresponding individuals. Subpedigree A was still characterized by five loops that were broken at the individuals marked by dots in Figure 1 to perform the analysis. lod scores were calculated using maximum likelihood marker allele frequencies as obtained from the SY pedigree using the LINKAGE programs. The resulting two-point lod scores are reported in Table 1.

None of the two-point lod scores obtained with the LINKAGE programs on chromosomes 4 and 11 were significant, contrary to the results

**Table 1. Maximum Likelihood Two-point lod Scores Between Candidate Markers and the SY Locus**

Chromosome	Marker	LINK	IBD1	IBD2
4	TGLA159	0.73 (24%)	6.16 (9%)	0.05 (26%)
4	TGLA215	0.98 (26%)	5.80 (4%)	0.00 (50%)
11	TGLA436	1.27 (13%)	14.49 (4%)	3.00 (6%)
11	TGLA77	0.67 (0%)	1.34 (4%)	0.00 (50%)
15	TGLA75	4.75 (6%)	7.50 (2%)	6.32 (2%)
15	BM848	4.67 (0%)	6.92 (0%)	6.34 (0%)

Maximum likelihood two-point lod scores and associated recombination rates obtained with the chromosome 4, 11, and 15 markers flanking the highest location scores as obtained by multipoint IBD analysis using the ANIBD programs (see Fig. 2). (LINK) lod scores obtained with the FASTLINK programs (Cottingham et al. 1993) used as described in Results, including the genotypes of 29 individuals from the intermediate generations of the pedigree represented in Fig. 1 and using maximum likelihood estimates of marker allele frequencies given the SY pedigree data. (IBD1) Two-point lod scores obtained with the ANIBD programs, using genotypes of affected individuals only, and marker allele frequencies as estimated from a DNA pool of 10 unrelated individuals (see Methods). (IBD2) Two-point lod scores obtained with the ANIBD programs, using genotypes of affected individuals only and maximum likelihood estimates of marker allele frequencies given the SY pedigree data.

obtained when performing a whole-genome scan with the ANIBD algorithm. The corresponding two-point lod scores obtained with the ANIBD programs using allelic frequencies estimated from the DNA pools are given in Table 1 for comparison, clearly showing highly significant lod scores for TGLA159 and TGLA215 on chromosome 4 and TGLA436 on chromosome 11. Examination of the genotypes shows that these high lod scores are attributable to the high frequency among affected offspring of a specific allele at the corresponding markers. The frequency of these alleles was estimated from the DNA pool to be very low in the population of interest, making it apparently very likely for these shared alleles to be IBD. Analysis of the intermediate generations, however, clearly indicated their identity-by-state rather than by descent status, causing the lod scores to drop when taking this information into account as was done when using the LINKAGE programs.

Altogether, these results clearly indicate that the high lod scores initially obtained on these chromosomes were fallacious and at least in part attributable to inaccurate estimation of marker allele frequencies. Genotyping the 10 individuals composing the DNA pool individually rather than mixed, resulted in very similar estimates (data not shown), excluding unequal amplification of microsatellite alleles as the primary cause of poor frequency estimates. As the lod score thresholds adopted in the IBD analysis (Fig. 3) were chosen to account for sampling errors, these false-positive results might indicate some level of population stratification and illustrate the difficulty in identifying a well-matched control sample.

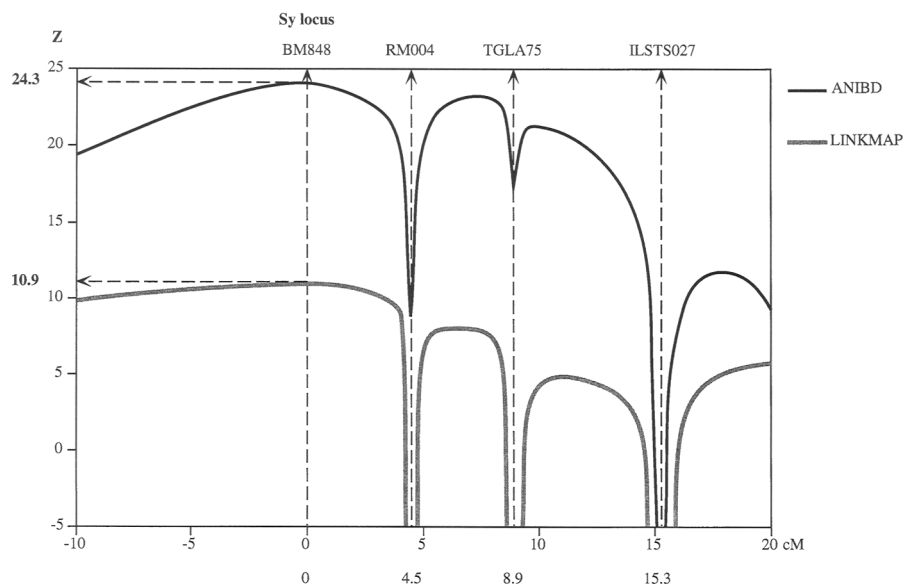
When using the ANIBD programs with the maximum likelihood allelic frequencies obtained from the syndactyly pedigree rather than the estimates from the DNA pool, the lod scores dropped to zero for TGLA159, TGLA215, and TGLA77 (Table 1). This confirmed that inaccurate frequency estimates were responsible for the lod score inflation noted with these markers. For marker TGLA436, however, this approach still yielded a lod score of 3. Despite the considerable drop from the previous lod score of 14.49, this value remains noticeably high. Analysis of the TGLA436 genotypes showed that two sires having jointly fathered 10 of the 12 affected offspring (Fig. 1) were both homozygous for the same allele. Therefore, this allele appeared unusually abundant among affected offspring, espe-

cially if ignoring the half-sib status of these individuals, as is the case with the ANIBD algorithm. The high lod scores initially obtained with TGLA436, therefore, illustrate how incorrect marker allele frequencies and ignored relationships can generate spurious results.

Contrary to the results obtained on chromosomes 4 and 11, the chromosome 15 markers consistently yielded lod scores superior to three in the two-point analysis (Table 1), supporting the genuine nature of the signal obtained on this chromosome in the IBD analysis. It is noteworthy that allelic frequency estimates obtained from the DNA pool or from the SY pedigree differed only marginally for all corresponding markers. Three additional microsatellite markers known from published data to be located on the telomeric half of chromosome 15 were identified: RM004 (Barendse et al. 1994), ILSTS061, and ILSTS027 (Ma et al. 1996). The International Bovine Reference Panel (IBRP) (Barendse et al. 1994) was genotyped for these three markers as well as a relevant subset of the markers that were part of the initial chromosome 15 selection: FSHB, TGLA75, and BM848 (Barendse et al. 1994; Bishop et al. 1994; Georges et al. 1995). The resulting genotypes were used to construct the corresponding most likely sex-averaged map using the LINKAGE programs, yielding the following order and recombination rates between adjacent markers: ILSTS061 (9%)–FSHB (6%)–ILSTS027 (5%)–TGLA75 (4%)–RM004 (4%)–BM848. All available members of the SY pedigree were genotyped for these six markers and the resulting data used in conjunction with the established map to estimate the most likely position of the SY locus. The multipoint lod score curve obtained with the LINKMAP option of the FASTLINK programs with respect to the four most distal markers held in fixed positions is shown in Figure 4. A very significant lod score of 10.9 was obtained at 0% recombination from BM848. The marker allele frequencies used in these analyses were obtained by maximum likelihood approach from the SY pedigree. Use of the same marker map and frequency estimates with the ANIBD programs yielded a maximum lod score of 24.3 at the position of BM848 (Fig. 4). Altogether, these results allowed us to confidently assign the bovine SY locus to the telomeric end of chromosome 15.

## DISCUSSION

Intense selection allied with reproductive meth-



**Figure 4** Multipoint lod score curves obtained for the SY locus with respect to microsatellite markers on distal chromosome 15, using (1) the LINKMAP option of FASTLINK programs (Cottingham et al. 1993) (shaded line), and (2) the ANIBD programs (solid line), as described in Results and Methods.

ods, such as artificial insemination, has led to a considerable reduction in the effective population size of most domestic breeds. This has been accompanied by sudden increases in the frequency of breed-specific recessive defects attributable to genetic drift sometimes assisted by indirect selection (hitchhiking) caused by linkage disequilibrium between the defective gene and loci influencing the selected traits, or a direct selection caused by the pleiotropic effect of the defective gene on a production trait. The frequency of progressive degenerative myeloencephalopathy in Brown Swiss possibly linked to an associated increase in milk production (Hoeschele and Meinert 1990) illustrates the former mechanism, whereas porcine malignant hyperthermia associated with improved carcass characteristics likely illustrates the latter (Fuji et al. 1991). Ignoring the added influence of selection, therefore, domestic breeds show some similarity with human isolates (e.g., Hästbacka et al. 1992), which have proven particularly suitable for the IBD mapping of genes underlying simple and complex genetic traits. Given the population structure, it is reasonable to predict that affected individuals within a breed will share IBD disease-causing alleles. Actually, analysis of the genealogical records—scrupulously documented for most domestic species—will very often reveal common

acknowledged carrier “founders” in the ancestry of affected individuals. This has been verified for syndactyly in this study as well as for progressive degenerative myeloencephalopathy in a previous study (Georges et al. 1993). In this study we illustrate how this can be exploited to efficiently locate disease genes with the available livestock maps, even when family material conventionally considered suitable for linkage studies—both in terms of numbers and relationships—is not available, therefore, obviating the need to generate such pedigrees by directed breeding. The initial screening phase required  $13 \times 213 = 2769$  microsat-

ellite genotypes representing less than a month of work for one person. For most inherited disorders observed in livestock populations it is relatively straightforward to collect samples from small numbers of affected individuals within a given breed. For several such defects, fixed tissue samples of affected individuals have been collected routinely by a specialized pathologist over time and are therefore readily available. Although the corresponding individuals very seldom compose nuclear families optimal for conventional linkage analysis, they are expected to share common haplotypes in the chromosomal area carrying the defective gene.

This study points toward the need to consider the results of an IBD mapping experiment with much caution. In particular, we illustrate how inaccurate estimates of marker allele frequencies and ignored relationships between affected individuals can inflate lod scores at loci unlinked to the disease locus.

Although individuals from intermediate generations in this study could be studied to sort among the candidate loci resulting from the IBD mapping, such samples might not be obtainable in most cases. The availability of denser marker maps in the future, however, will allow the replacement of this second screening based on the typing of additional individuals by a second

screening based on the typing of additional markers in the regions of interest. The more markers compose the haplotypes shared among affected individuals, the more discriminating power one has to distinguish identity by state versus IBD. This is illustrated to some extent in this study by the increase in lod score observed on chromosome 15 after the addition of three distal microsatellite markers.

The development of the vertebrate limb is a subject of intensive research. Several genes known to be essential for normal limb development have been identified, including bone morphogenetic protein-2a (*Bmp2a*) and 4 (*Bmp4*), cellular retinol-binding protein (*Rbp*), engrailed-1 (*En1*), fibroblast growth factor-1 (*Fgf1*) and -2 (*Fgf2*), fibroblast growth receptor-1 (*Fgfr1*) and -2 (*Fgfr2*), homeo box cluster A (*Hoxa*), homeo box complex C (*Hoxc*) and homeo box D (*Hoxd*), insulin-like growth factor-1 (*Igf1*), *msh*-like homeo box 1 (*Msx1*) and 2 (*Msx2*), retinoic acid receptor  $\beta$  (*Rarb*), sonic hedgehog homolog (*Shh*), and wingless-related MMTV integration site 7A (*Wnt-7A*). Syndactyly or related symptoms have been reported resulting from mutations in a number of these genes (Gilbert 1994; Roberts and Tabin 1994; Mulvihill 1995; Tickle 1995). Comparative mapping data, however, allow one to question if not to exclude most of these genes as being responsible of bovine syndactyly. Alignment of the human and bovine maps using anchored reference loci (O'Brien et al. 1993), as well as recent data from comparative chromosome painting (zoo-fish) (Solinas-Toldo et al. 1995), indicate that bovine chromosome 15 is homologous to part of human chromosome 11, the remaining of this human chromosome corresponding to bovine chromosome 25. The only gene from the cited list that maps to human chromosome 11 is *FGF-4*, mapping to 11q13.3 together with *FGF-3*. The latter gene has been assigned previously to bovine chromosome 25 (O'Brien et al. 1993), and we have demonstrated that the same holds for *FGF-4* using somatic cell hybrids (data not shown). Comparing the mouse and bovine maps points toward mouse chromosomes 2, 7, 9, and 19 as being potentially homologous in part to bovine chromosome 15. A number of mouse mutants exhibiting limb deformities as part of the corresponding syndrome have been reported on these chromosomes, including *oligodactyly* (*ol*) and *extra toes spotting* (*Xs*). However, more refined comparative mapping is required to evaluate the potential relevance of this information.

## METHODS

### Marker Genotyping

A battery of 213 microsatellite markers distributed over the 29 bovine autosomes and compiled from three partially overlapping and recently published bovine maps (Barendse et al. 1994; Bishop et al. 1994; Georges et al. 1995), was chosen to allow for a systematic scanning of the bovine genome. Microsatellite genotyping was performed essentially as reported previously (Georges et al. 1995). Briefly, two to four different microsatellite systems were amplified simultaneously in 10- $\mu$ l reaction volumes, from 25 ng of each template DNA. Reagent concentrations were 75 mM KCl, 15 mM Tris-HCl (pH 8.4), 2.25 mM MgCl<sub>2</sub>, 0.02% gelatin, 0.3 mM of dATP, dGTP, and dTTP, 0.03 mM dCTP, 0.5  $\mu$ M of each primer, 0.05 units of AmpliTaq per microliter, and 0.05  $\mu$ Ci [ $\alpha$ -<sup>32</sup>P]dCTP/microliter. PCR reactions were set up with a Biomek 1000 robotic station (Beckman Instruments, Palo Alto) and carried out in Techne MW2 devices (Cambridge, UK). Samples were denatured at 95°C for 5 min and cycled 30 times under the following conditions: 93°C for 1 min, 60°C for 1 min, and 72°C for 1 min. After addition of one volume of formamide dye and denaturation at 95°C for 5 min, 5  $\mu$ l of each product was electrophoresed on a 6% acrylamide gel containing 32% formamide, 5.6 M urea, 135 mM Tris, 45 mM boric acid, and 2.5 mM EDTA. The gels were autoradiographed for 2 hr to overnight.

### Linkage Analysis

IBD mapping was performed with the ANIBD programs (F. Farnir and M. Georges, in prep.). Briefly, the algorithm used by these programs assumes that all affected individuals trace back to a common carrier founder individual through a pathway  $\alpha$  characterized by "a" meioses and a pathway  $\beta$  characterized by "b" meioses. Therefore, affected individuals are expected to share marker haplotypes linked to the disease locus at a higher frequency than expected by chance alone for markers unlinked to the disease locus. The pathways connecting affected individuals to the founder are assumed to be unbranched (i.e., only one possible route) and independent. The genotypes of the diseased gamete inherited by the founder [hereafter referred to as S(ource) gamete] and the gametes inherited by the affected individuals [hereafter referred to as D(estination) gamete] through the  $\alpha$  and  $\beta$  pathways, respectively, are recorded in binary format assuming two possible allelic states for each locus: identical by descent or not identical by descent. The S gamete is always in the all-identical-by-descent state, whereas the D gametes will be compatible with a subset of all possible 2<sup>n</sup> gametic configurations, where n corresponds to the number of marker loci plus the disease locus. The subset of compatible D gametes depends on the actual genotype of the S and D gametes considered. If, for a given locus, the allele carried by the D gamete is the same as for the S gamete, these can be either identical by descent or not identical by descent (that is identical by state); if different, they can obviously only be not identical by descent. The probability for an affected offspring to have inherited a given D gamete given the founder's S gamete is computed using hidden Markov chain theory from the r-step transition matrix  $P^{(r)} = P^r$ , where  $r = a$  or  $b$

## CHARLIER ET AL.

for the  $\alpha$  and  $\beta$  pathways, respectively. The entries  $p_{ij}$  in the transition matrix  $P$  are easily computed as functions of the recombination rates between adjacent markers as well as allelic frequencies for all loci. The likelihood of the pedigree data are computed as

$$\sum_{i=1}^x p_i \prod_{k=1}^o \sum_{l=1}^{2^n} p_{i \rightarrow kl\alpha}^{(a)} \times p_{i \rightarrow kl\beta}^{(b)}$$

where

$\sum_{i=1}^x =$  The sum over all  $x$  possible disease-carrying gametes inherited by the founder, or  $S$  gametes.

$p_i =$  The probability of  $S$  gamete  $i$ , computed as the product of the population frequencies of the considered alleles at the  $n$  marker loci.

$\prod_{k=1}^o =$  The product over all  $o$  affected offspring.

$\sum_{l=1}^{2^n} =$  The sum over all possible linkage phases from affected offspring  $k$  ( $2^n$  rather than  $2^{n-1}$  phases are considered to account for the fact that each phase can arise in two ways depending on which gamete derives from pathways  $\alpha$  or  $\beta$ ).

$p_{i \rightarrow kl\alpha}^{(a)} =$  The  $a$ -step transition probability from  $S$  gamete  $i$  to  $D$  gamete  $kl\alpha$ .

$p_{i \rightarrow kl\beta}^{(b)} =$  The  $b$ -step transition probability from " $S$ " gamete  $i$  to " $D$ " gamete  $kl\beta$ .

For each chromosome, the putative disease locus is slid through the available marker map, and for each position the likelihood of the data is calculated as described. Results are expressed as multipoint lod score curves.

Marker order, as well as recombination rates between adjacent markers, was deduced from published maps. Marker allele frequencies, required for the likelihood calculations, were estimated by visual examination of the microsatellite patterns obtained from an equimolar DNA pool of 10 unrelated individuals from the same population. The frequency of the syndactyl  $sy$  allele in the population excluding this pedigree was set at 0%.

Conventional linkage analyses were performed with the 5.1 versions of the LINKAGE package (Lathrop and Lalouel 1984) and with the 2.3P (June 1995) version of the FASTLINK programs (Cottingham et al. 1993). The penetrance for syndactyl was assumed to be unity in these analyses.

## ACKNOWLEDGMENTS

We are grateful to Ir. Jos Koopman for his help in collecting the samples composing the  $SY$  pedigree. We are indebted to Dr. Womack for providing us with panel of somatic cell hybrids. We sincerely thank Dr. Beattie, Dr. Hetzel, Dr. Leveziel, Dr. Teale, and Dr. Womack for providing us with DNA samples from the (IBRP) reference families. This work was funded by Holland Genetics and Livestock Improvement Corporation. The commercial use of the DNA markers linked to the  $SY$  locus is subject to patent protection. The continuous support of Dr. Nanke den Daas and Dr. Brian Wickham is gratefully acknowledged.

The publication costs of this article were defrayed in part by payment of page charges. This article must there-

fore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

## REFERENCES

- Barendse, W., S.M. Armitage, L.M. Kossarek, A. Shalom, B.W. Kirkpatrick, A.M. Ryan, D. Clayton, L. Li, H.L. Neibergs, N. Zhang, W.M. Grosse, J. Weiss, P. Creighton, F. McCarthy, M. Ron, A.J. Teale, R. Fries, R.A. McGraw, S.S. Moore, M. Georges, M. Soller, J.E. Womack, and D.J.S. Hetzel. 1994. A genetic linkage map of the bovine genome. *Nature Genet.* **6**: 227–235.
- Bishop, M.D., S.M. Kappes, J.W. Keele, R.T. Stone, S.L.F. Sunden, G.A. Hawkins, S. Solinas-Toldo, R. Fries, M.D. Grosz, J. Yoo, and C.W. Beattie. 1994. A genetic linkage map for cattle. *Genetics* **136**: 619–639.
- Cottingham, R.W., R.M. Idury, and A.A. Schäffer. 1993. Faster sequential genetic linkage computations. *Am. J. Hum. Genet.* **53**: 252–263.
- Eldridge, F.E., W.H. Smith, and W.M. McLeod. 1951. Syndactylism in Holstein-Friesian cattle. *J. Hered.* **42**: 241–250.
- Fuji, J., K. Otsu, F. Zorzato, S. Deleon, V.K. Khanna, J.E. Weiler, P.J. O'Brien, and D.H. MacLennan. 1991. Identification of a mutation in the porcine ryanodine receptor associated with malignant hyperthermia. *Science* **253**: 448–451.
- Georges, M., M. Lathrop, A.B. Dietz, A. Lefort, F. Libert, A. Mishra, D. Nielsen, L.S. Sargeant, M.R. Steele, X. Zhao, H. Leipold, and J.E. Womack. 1993. Microsatellite mapping of the gene causing weaver disease in cattle will allow the study of an associated QTL. *Proc. Natl. Acad. Sci.* **90**: 1058–1062.
- Georges, M., D. Nielsen, M. Mackinnon, A. Mishra, R. Okimoto, A.T. Pasquino, L.S. Sargeant, A. Sorensen, M.R. Steele, X. Zhao, J.E. Womack, and I. Hoeschele. 1995. Mapping quantitative trait loci controlling milk production by exploiting progeny testing. *Genetics* **139**: 907–920.
- Gilbert, S.F. 1994. Development of the tetrapod limb. In *Developmental biology*, 4th ed., pp. 690–715. Sinauer Associates, Inc., Sunderland, MA.
- Hart-Elcock, L., H.W. Leipold, and R. Baker. 1987. Hereditary bovine syndactyl: Diagnosis in bovine fetuses. *Vet. Pathol.* **24**: 140–147.
- Hästbacka, J., A. de la Chapelle, I. Kaitila, P. Sistonen, A. Weaver, and E. Lander. 1992. Linkage disequilibrium mapping in isolated founder populations: Diastrophic dysplasia in Finland. *Nature Genet.* **2**: 204–211.
- Hoeschele, I. and T.R. Meinert. 1990. Association of genetic defects with yield and type traits: The Weaver locus effect on yield. *J. Dairy Sci.* **73**: 2503–2515.

- Houwen, R.H.J., S. Baharloo, K. Blankenship, P. Raeymaekers, J. Juyn, L.A. Sandkuijl, and N.B. Freimer. 1995. Genome screening by searching for shared segments: Mapping a gene for benign recurrent intrahepatic cholestasis. *Nature Genet.* **8**: 380–386.
- Kruglyak, L., M.J. Daly, and E. Lander. 1995. Rapid multipoint linkage analysis of recessive traits in nuclear families, including homozygosity mapping. *Am. J. Hum. Genet.* **56**: 519–527.
- Kruglyak, L. and E. Lander. 1995. High-resolution genetic mapping of complex traits. *Am. J. Hum. Genet.* **56**: 1212–1223.
- Lander, E.S. and D. Botstein. 1987. Homozygosity mapping: A way to map human recessive traits with the DNA of inbred children. *Science* **236**: 1567–1570.
- Lathrop, M. and J.M. Lalouel. 1984. Easy calculation of lod scores and genetic risk on small computers. *Am. J. Hum. Genet.* **36**: 460–465.
- Leipold, H.W., K. Huston, D.M. Trotter, S.M. Dennis, and M.M. Guffy. 1969. Anatomy of hereditary bovine syndactylism. I. Osteology. *J. Dairy Sci.* **52**: 1422–1431.
- Leipold, H.W., S.M. Dennis, and K. Huston. 1973. Syndactyly in cattle. *Vet. Bull.* **43**: 399–403.
- Ma, R.Z., J.E. Beever, Y. Da, C.A. Green, I. Russ, C. Park, D.W. Heyen, R.E. Everts, S.R. Fisher, K.M. Overton, A.J. Teale, S.J. Kemp, H.C. Hines, G. Guerin, and H.A. Lewin. 1996. A male linkage map of cattle (*Bos taurus*) genome. *J. Hered.* (in press).
- Mulvihill, J.J. 1995. Craniofacial syndromes: No such thing as a single gene disease. *Nature Genet.* **9**: 101–103.
- O'Brien, S.J., J.E. Womack, L.A. Lyons, K.J. Moore, N.A. Jenkins, and N.G. Copeland. 1993. Anchored reference loci for comparative genome mapping in mammals. *Nature Genet.* **3**: 103–112.
- Puffenberger, E.G., E.R. Kauffman, S. Bolk, T.C. Matisse, S.S. Washington, M. Angrist, J. Weissenbach, K.L. Garver, M. Mascari, R. Ladda, et al. 1994. Identity-by-descent and association mapping of a recessive gene for Hirschprung disease on human chromosome 13q22. *Hum. Mol. Genet.* **3**: 1217–1225.
- Roberts, D.J. and C. Tabin. 1994. The genetics of human limb development. *Am. J. Hum. Genet.* **55**: 1–6.
- Solinas-Toldo, S., C. Lengauer, and R. Fries. 1995. Comparative genome map of human and cattle. *Genomics* **27**: 489–496.
- Tickle, C. 1995. Vertebrate limb development. *Curr. Opin. Genet. Dev.* **5**: 478–484.

Received February 16, 1996; accepted in revised form May 10, 1996.