VELOGENETICS, OR THE SYNERGISTIC USE OF MARKER ASSISTED SELECTION AND GERM-LINE MANIPULATION

M. GEORGES AND J.M. MASSEY

Genmark Inc., 417 Wakara Way, Salt Lake City, Utah 84108

INTRODUCTION

Until recently, artificial selection has relied on the biometrical evaluation of individual breeding values from an animal's own performance and from performance of its relatives. This biometrical strategy is based on relatively simple genetic premises, operating within a "black box". Briefly, the majority of economically important traits are so-called complex or quantitative traits, meaning that the phenotype of an animal is determined by both environment and a large number of genes with individually small, additive effects. The proportion of the phenotypic variation observed in a given population that is genetic in nature is the heritability of the trait. Substantial genetic progress has been obtained using this approach. One of the powers of the biometrical approach is that it obviates the need for any detailed molecular knowledge of the underlying genes or Economic Trait Loci (ETL).
However, it is believed that the molecular identification of these ETLs should allow for an increased genetic response by affecting both time and accuracy of selection, through a procedure called Marker Assisted Selection (MAS)(1,2). Moreover, we propose to use a scheme that we call "velogenetics", or the combined use of Marker Assisted Selection and germ-line manipulations aimed at shortening the generation interval of domestic species (especially cattle), which would allow the efficient introgression of mapped Economic Trait Loci between genetic backgrounds.

1. MAPPING ECONOMIC TRAIT LOCI:

One strategy towards the isolation of ETLs is called "reverse genetics" (3). It relies on the use of DNA Sequence Polymorphisms (DSP) as genetic markers in linkage studies.

1.a. DNA Sequence Polymorphism (DSP):

The typical mammalian genome is composed of an approximately $3x10^9$ base pairs long stretch of DNA, divided over a species-specific number of chromosomes, and containing all the information required for the proper development and functioning of a normal being. Each individual has two copies of this message: one paternal and one maternal in origin. Although their overall architecture and content is virtually identical, the paternal and maternal DNA sequences exhibit subtle "allelic" differences, referred to as "DNA Sequence Polymorphisms (DSP)".
DSP encountered in a given population find their origin in mutational events occurring in the germline and escaping the DNA repair machinery. The fate of these germline mutations in the population is dominated by two kinds of effects: stochastic effects ("random drift") and deterministic effects (negative, positive and balancing selection).
One can recognize three types of DSP: 1. single base pair polymorphisms: substitutions (transitions and transversions) or the insertion/deletion of a single base pair; 2. DNA sequence rearrangements, such as the insertion or deletion of a stretch of DNA, DNA sequence inversions and duplications; mobile genetic elements play a key role in this kind of DSP; 3. expansion-contraction type polymorphism.
The latter type deserves special attention. It affects a peculiar class of sequences widely termed "satellite DNA". Satellite DNA is characterized by the head-to-tail or tandem repetition of a defined sequence motif. Both the length of the repeated motif and the repetition number differ a lot, allowing one to distinguish macro-, midi-, mini- and microsatellites. The function of satellite sequences, if any, is essentially unknown. An important feature of all satellite sequences is that the maintenance of their tandemly repeated organization is dependent on the concerted evolution of the repeats. This concerted evolution is thought to result from subsequent rounds of unequal crossing-over (or any other mechanisms fitting the "card deck"

model (4)), which are favored by the tandemly repeated structure itself. The proposed unequal crossing over mechanism, whether happening between sister chromatids or homologous chromosomes, explains the substantial degree of length polymorphism, here referred to as "expansion-contraction polymorphism", characterizing those sequences. Moreover, the ensuing shuffling of slightly divergent repeat units or Minisatellite Variant Repeats (MVRs)(5) within the satellite generates additional internal site polymorphism. Hence, the majority of individuals will have inherited different alleles from their respective parents, with heterozygosities ≥ 90% being common. This value must be compared with the maximum heterozygosity of 50% (more often much less) that can be obtained with the two other diallelic types of DSP. These peculiar properties of satellite sequences has made them an invaluable source of highly informative genetic markers both in the human and in domestic species (reviewed in (6)). The length variation characterizing mini- and microsatellites in particular, is visualized as DNA Fragment Length Polymorphisms either by Southern Blot Hybridization, or after PCR (Polymerase Chain Reaction) amplification of the corresponding loci.

1.b. Linkage Strategies:

Two loci are said to be genetically linked if, during meiosis, they recombine at a significantly lower than 50% rate, i.e. generate significantly more parental than recombinant gametes. The recombination rate between loci reflects the frequency of occurrence of an uneven number of crossing-overs between the loci. Because the probability for crossing-over is proportional to the distance separating the loci, the recombination rate can be used as a unit of chromosomal length. This length unit is known as the Morgan (M), 1 cM corresponding to the distance separating two loci exhibiting a 1% recombination rate. For small distances (≤30cM) the relation with recombination rate is essentially linear; for longer distances, however, the relation is more complex, depending on the frequency of double cross-overs, itself affected by eventual interference. Parental and recombinant gametes will only be distinguishable for doubly heterozygous individuals, hence the need for highly polymorphic markers.
Recently, and due to the advent of the PCR, it has been possible to directly determine the genotype of individual gametes (7). However, most of the time, the gametic contribution is inferred from the genotype of the offspring and linkage studies are performed within families. Most modern linkage studies use the lodscore test for evaluation of linkage: a sequential test based on the method of maximum likelihood (8). The lodscore corresponds to $\log_{10}$(LR), where LR corresponds to the ratio: likelihood of observation under the alternative hypothesis of linkage ($\theta \leq 0.5$) / likelihood of observations under null hypothesis of no linkage ($\theta = 0.5$). In human genetics, a lodscore > 3 is accepted as significant evidence for linkage. The prior probability of linkage between two loci has been used to justify this stringent critical value. Note that 2ln(LR) can be used as well, having a chi-square distribution with one degree-of-freedom under the null hypothesis of no linkage.
Recently, algorithms for multilocus linkage analysis have been developed, allowing the estimation of the most likely gene orders and genetic distances between several loci simultaneously (9,10,11).

1.c. Linkage Maps:

Using this linkage approach, combined with alternative mapping strategies such as "in situ" hybridization, the use of somatic cell hybrid panels and comparative mapping, the map location of a large set of DSP can be determined in order to build a genetic marker map (see for instance 12, 13, 14). Assuming a total map length of 30M as for the human, and a desirable maximum distance of 20cM between markers, a set of 150 DSP could cover the entire genome. However, many more markers will be needed to generate a reasonable map, and this essentially for two reasons: 1. most of the time we have no a priori information on the location of the characterized markers. Hence, some chromosomal regions will initially be over-represented in our map, others under-represented. This problem is expected to become critical in the later stages of the development of a map. Comparative data will then become critical, allowing to search for markers whose location can be predicted from other species. 2. An individual will only be informative for the markers for which he is heterozygous; parts of his genome won't thus be explorable, because he will be homozygous for the corresponding markers. To compensate for this, one will have to identify more markers, the number required being inversely proportional to their heterozygosity. Hence the importance of highly informative systems.

Once such a map is available, however, any gene for which the appropriate segregating family material is available, can be located on the map. Assuming a maximum marker-target gene distance of 10cM, the expected lodscore (ELOD) for a doubly informative, phase-known meioses approximates 0.16 (15). Therefore, 20 such meioses are theoretically sufficient to establish linkage with a lodscore of 3. In practice, however, the number of individuals to analyze will be higher, function among other factors of the quality of the marker, expressed as its Polymorphism Information Content (PIC)(16).
The efficiency of this approach has been illustrated by the recent mapping of a large number of genes involved in human single gene disorders (see for instance 17). The identification of DNA markers for a defined gene can be the first step towards its molecular cloning. The feasibility of the isolation of a gene based on its map location has recently been demonstrated spectacularly by the isolation of the cystic fibrosis and other genes (see for instance 18).

1.d. Progress towards a primary DNA marker map in cattle:

Several laboratories around the world are now involved in the development of markers and the construction of genetic maps for our main domestic species, especially cattle, pigs and poultry.
Our laboratory has focused in the last two years in the development of a primary DNA marker map for cattle. We have now developed close to 200 highly polymorphic DNA markers of three types:

1. Variable Number of Tandem Repeat Markers (VNTR):
Hypervariable minisatellites are known to show significant cross-hybridization between species (19,20). We have exploited this to isolate over 50 bovine VNTR sequences revealing hypervariable, locus-specific patterns in cattle (21). The mean heterozygosity for all markers was estimated at 59%. 60% of these bovine VNTRs are giving polymorphic, locus-specific patterns in sheep as well. Several clones were mapped by in situ hybridization in both cattle and sheep demonstrating perfect conservation of chromosomal location. Hence, a significant part of the data generated in cattle will be readily usable in other Bovidae, including sheep.

2. Multisite haplotypes:
The relatively low PIC of diallelic DSP compared to expansion-contraction type DSP, can be compensated for if identifying several such closely spaced markers. The theoretical number of allelic combinations or haplotypes then equals 2^n, where n corresponds to the number of identified DSP. Although linkage disequilibrium reduces the number of haplotypes actually present in a given population, the heterozygosity obtained when considering those haplotypes is much higher than when considering the individual DSP.
To exploit this, random cosmid clones were probed to genomic DNA, digested with 14 restriction enzymes, in order to identify RFLPs (22). Out of 110 cosmids studied, 85% were showing polymorphism within the Holstein breed. A mean of 2.5 polymorphic events were identified per cosmid, generating multisite haplotypes. The mean heterozygosity for the identified multisite haplotypes was estimated at 52%. The average heterozygosity per nucleotide site or nucleotide diversity, π, was estimated from the RFLPs attributed to point mutations, giving a value of ≈ 0.0006. This value is ≈ 3 times lower than in the human and confirms our previous observations (23, 24). The lower π value is however compensated by a high frequency of RFLPs due to insertion/deletion events, which constitute 27.7% of the polymorphisms in our sample. The mutation rate at cytosines followed by guanines was estimated ten times higher than for other nucleotides, as expected from results in the human (25).

3. Microsatellites:
More than 100 bovine (TG)n repeats (26, 27, 28) have been isolated and characterized by sequencing. Their frequency in the bovine genome has been estimated at $\geq$ 150,000. 50 of them have been amplified using the Polymerase Chain Reaction, all of them showing a substantial degree of genetic polymorphism when analyzing the product on high-resolution, denaturing polyacrylamide gels. The mean heterozygosity for these systems was estimated at $\approx$ 55%. Surprisingly more than half of the bovine system work in sheep as well (29). Magnetic solid phase DNA sequencing procedures are used for the massive generation of sequence information and multiplex approaches are explored for genotypic collection.

The relative location of the markers is determined by linkage analysis in MOET generated pedigrees. To

assign linkage groups to specific chromosomes, highly polymorphic "anchor markers" are mapped using somatic cell hybrids in collaboration with Jim Womack (Texas A&M), and by in situ hybridization in collaboration with Rudy Fries (ETH, Zurich).
We anticipate that a primary DNA marker map will soon be available for cattle, allowing us to systematically explore the bovine genome in linkage studies.

1.e. Mapping Quantitative Trait Loci:

The majority of traits dealt with in animal production are so-called quantitative traits, characterized by continuous variation. The phenotype of an animal with respect to a particular trait is the resultant of the effect of a several "polygenes" known as Quantitative Trait Loci (QTLs), combined with environmental effects. The number of polygenes involved is essentially unknown. Classically, it is considered very large, each gene contributing for a very small part of the genetic variation. However, there is evidence both from the plant and the animal world, that QTLs with significant effects are common (30, 31, 32). The most likely model, is to assume that their are indeed a large number of genes involved, but that their is a broad distribution of effects, substantial for some of them. Polygenes with extreme effects, whose segregation in a population may cause skewness and bi- or trimodality are known as "major genes". Examples in animal breeding are "double muscling" genes in both cattle and pigs, the "White Shorthorn" gene in the determinism of "White Heifer Disease" and the fertility "Booroola" gene in sheep (33). Even with significant effects on the trait of interest, however, their contribution to the total genetic variation may be limited in case of low population frequency.
When dealing with quantitative traits, direct determination of genotype for the corresponding QTLs is impossible. Nevertheless, strategies have been designed to map QTLs by linkage analysis. Within segregating populations, which is usually the case for our domestic species, QTL mapping can be performed both within families and at the population level.

1. QTL mapping within families:
Traditionally one proceeds as follows: offspring from an individual heterozygous for both marker and QTL are grouped according to which allele at the marker locus they inherited; a statistically significant difference between the phenotypic means of the two groups indicates linkage between marker and QTL. Test for statistical significance is done by linear regression (i.e. one-way analysis of variance) under the assumption of normally-distributed residual environmental variance. Classically, markers are tested one at the time for possible linkage with a QTL affecting the trait of interest. One of the drawbacks of this approach is that it is impossible to unequivocally estimate both map location of the QTL with respect of the marker, and its effect on the considered trait; no distinction can be made between a closely linked QTL with small effect and a loosely linked QTL with major effect.
Recently, the lodscore method has been improved, allowing to deal with quantitative and other complex traits and fully exploiting the power of the nearly complete marker maps which become available for different organisms; an approach known as interval mapping. Not only does interval mapping solves the problem of simultaneous estimation of location and effect, but because of its increased power, it reduces the number of individuals to test to detect linkage with a QTL of given effect (34).
Assuming that the marker is the QTL, the number of individuals to test in order to detect an effect of given amplitude, δ, can be estimated from:

$$n \geq \frac{4(t_o + t_1)^2 . s^2}{\delta^2}$$

where n gives the required sample size, s^2 is an estimate of the residual variance, t_o is the t value associated with Type I error, and t_1 is the t value associated with Type II error; t_1 equals tabulated t for probability 2(1-P) where P is the required probability of detecting δ if such a difference exists (35).
For dairy production for instance, and if performing the linkage analysis using the "daughter yield deviations" (DYD; σ^2_{DYD}≈600lb) from paternal half-sibs ("granddaughter design" (36)), one would have to study respectively 1500, 378 and 168 individuals to detect QTLs with differences of 200lb, 400lb and 600lb between alternate alleles. Assuming a phenotypic variance of $(2500lb)^2$, such effects correspond to 0.08, 0.16 and

0.24 standard deviations respectively. These estimates assume a Type I error of 5%, a Type II error of 10% and absence of recombination between marker and QTL.
If the tested marker and the QTL recombine at a rate e, the number of individuals to test increases by a factor $1/(1-2e)^2$ for single marker analysis, by a factor $\approx(1-\tau)/(1-2e)^2$ in case of interval mapping, τ corresponding to the recombination rate between the flanking markers (34).

In view of the costs and time involved in genotyping, it is important to minimize the required sample size. This can be achieved by various ways (34):

- identify the individuals most likely to be heterozygous, hence informative for the studied QTLs;
- selective genotyping of the extreme progeny;
- decreasing environmental variance via progeny testing;
- reducing genetic noise by studying several genetic regions simultaneously, or "simultaneous" search;
- exploiting "tagged QTLs": The direct effect of selection for a production trait will be to increase the frequency of the favorable alleles at the segregating QTLs. However, this selection pressure may indirectly affect loci in linkage disequilibrium by so-called "hitch-hiking". This is probably what happened to the genetic defect causing progressive degenerative myeloencephalopathy, or Weaver in Brown Swiss, shown to be linked to a major gene for milk production (32). Because of the deleterious effect of the Weaver causing gene, it is the heterozygous "carrier" genotype which is selectively most advantageous, generating a "balanced polymorphism", the Weaver causing allele being maintained in the population at a relatively high frequency. We are exploiting this to map the corresponding QTL by going through the relatively easy exercise (compared to QTL mapping) of finding a marker linked to this single gene disorder.

QTLs for a variety of polygenic traits have been identified, both in plants and animals. Using complete DSP maps in tomato, Paterson et al. (30) identified at least 6 genes controlling fruit mass, 4 controlling soluble solids and 5 controlling fruit pH, accounting for respectively 58%, 44% and 48% of the phenotypic variance. Martin et al. (31), using a similar approach, identified at least 3 tomato genes controlling water use efficiency. In cattle, Geldermann et al. (37) found significant effects on milk yield (+ 200 kgs) and fat content (+ 1%), especially for the β-lactoglobulin locus; while Cowan et al. (38) recently demonstrated significant effects on predicted difference milk (282.93kgs), predicted difference dollars, cheese yield dollars and protein dollars using a prolactin DSP as marker.

2. QTL mapping within populations:
One can expect to find an effect of marker alleles linked to QTLs also outside of a family context, i.e at the population level, if the two loci are in linkage disequilibrium. As reported by Hanset (39), and assuming a diallelic marker (alleles M1 and M2 with respective frequencies p1 and p2) linked to a diallelic QTL, the phenotypic difference between the respective homozygotes at the marker loci equals:

$$\delta = 2a \cdot \frac{D}{p1.p2}$$

with D measuring the linkage disequilibrium and 2a corresponding to the phenotypic difference between the two homozygotes for the QTL.
Markers for which a priori expectation for linkage disequilibrium is highest are the so-called "candidate genes": genes expected from their physiological role to be likely candidates for the QTL itself. DSPs at those loci, even selectively neutral by themself, can be expected to exhibit linkage disequilibrium with the hypothetical functional mutations because of their very tight linkage.

2. "VELOGENESIS" OR GENERATION SKIPPING:

Ruminant oogenesis and folliculogenesis have been recently reviewed by Russe (40) and Betteridge et al. (41). Briefly, oogenesis begins with the formation of primordial germ cells in the region of the allantois. These precursor cells migrate to the developing gonads where after a period of mitotic proliferation, they enter meiosis. Meiosis is arrested at the diplotene stage of prophase I by the poorly understood "meiotic

division I arrest system". Small numbers of primary oocytes are successively recruited for oocyte growth, required for the acquisition of meiotic competence. Resumption of meiosis and oocyte maturation are stimulated by gonadotropins acting through the follicular cells. Oocytes maturation seems to be triggered by a decrease in intracellular cyclic AMP, which leads to the dephosphorylation and concomitant activation of the M-phase promoting factor (MPF), a complex of a $p34^{cdc2}$ protein kinase subunit with a cyclin (42). The maturing oocyte completes the first meiotic division and enters the second (becoming a secondary oocyte) which will be arrested as well at metaphase II until fertilization. This "meiotic division II arrest system" is thought to reflect the stabilization of MPF mediated by the kinase activity of $pp39^{mos}$ on either a cyclin protease or on cyclin itself. Fertilization relieves this block, by increasing the intracellular Ca^{2+} concentration, triggering calcium-dependent protease activity (42).

Although spontaneous oocyte maturation and ovulation do not begin until puberty, waves of oocyte growth are seen in prepubertal animals, even in utero. As summarized by Betteridge et al. (41) for cattle, activation of primordial follicles starts at 140 days of gestation, while secondary and tertiary follicles appear at respectively 210 and 230 days of gestation.

These observations raise the theoretical possibility to growth, mature and fertilize prepubertal oocytes in vitro. Such a procedure, which we refer to as "velogenesis", could reduce the generation interval of cattle to, maybe as little as 3 to 6 months (41). As a matter of fact offspring have been obtained from gonadotropin-stimulated calve oocytes, transferred to postpubertal animals (43). Also encouraging is the development of culture systems supporting the growth of rodent immature oocytes, their maturation, fertilization and preimplantation development.

The direct impact of "velogenesis" on breeding programs has been discussed by Betteridge et al. (41). By using in-vitro fertilization of fetal oocytes by selected, progeny-tested sires, annual responses in milk yield could be doubled compared to conventional progeny testing.

3. VELOGENETICS OR THE SYNERGISTIC USE OF MARKER ASSISTED SELECTION AND GERM-LINE MANIPULATION.

The value of a marker for a QTL, is usually considered within the biometrical framework, as an additional source of information for the evaluation of individual breeding values. Stated otherwise, by including information on mapped QTLs, selection accuracy will increase and, hence, genetic response as well (1,2). For most economically important traits, however, and because of their relatively high heritability, the gain to be made is only modest in view of the invested effort, unless most of the genetic variation can be explained by mapped QTLs. For typical polygenic traits, the latter situation is highly unlikely in the near future. Marker Assisted Selection is expected to be more useful for low heritability traits such as disease resistance. Unfortunately, the corresponding QTLs are also the most difficult ones to map by linkage analysis.

It should be noted also, that the increase in accuracy of selection is subordinate to the accurate estimation of the QTL effects. This may require larger samples than the ones needed for the detection of linkage. Once a QTL mapped by within-family linkage studies, it may be more effective to identify supplementary flanking markers and to accurately determine the effect of the thus generated haplotypes at the population level. Selection can then focus on the best haplotype instead of spending initial selection efforts on intermediate ones.

Moreover, this classical Marker Assisted Selection approach is limited to the exploitation of genetic variation preexisting within the commercial breed of interest, and only if present in a "high merit" genetic background. Favorable mutations appearing within a mediocre background would be difficult to exploit, even with markers.

"Velogenesis", however, may offer new possibilities for Marker Assisted Selection: Marker Assisted Engineering or "Velogenetics". By substantially reducing the generation interval using "Velogenesis", mapped Economic Trait Loci (ETL) could be introgressed very quickly into new genetic backgrounds by repeated backcrossing, while the retention of the ETL of interest would be monitored using the linked markers. The main advantage of markers, namely the possibility to monitor the segregation of mapped ETLs without having to rely on phenotypic expression, and hence at any developmental stage, would be fully exploited.

Contrary to transgenesis, the organization and chromosomal localization of the manipulated genes is entirely respected, eliminating aberrant expression patterns. Moreover, segregation and hence engineering of

several genes simultaneously is perfectly feasible with "velogenetics". It is noteworthy, however, that "velogenetics" could be used for the efficient introgression of transgenes between backgrounds. Importantly, this approach offers the possibility to efficiently exploit "exotic" genetic variation identified outside the breed of interest. This consideration is particularly attractive because the use of "exotic germplasm" also greatly facilitates the mapping of the genes of interest.

It should be noted that Marker Assisted Selection is already applied in specific instances: in cattle, preimplantation embryos are sexed with Y-specific probes (reviewed in 44,45); in pigs, Marker Assisted Selection is already used to reduce the frequency of the major gene causing Porcine Stress Syndrome (PSS). Susceptibility to PSS seems to correlate with Halothane sensitivity or Malignant Hyperthermia. This condition has been mapped to a linkage group on pig chromosome 6, encompassing the following markers: S(A-O)-GPI-Hal-H-A1BG-PGD (reviewed in 46). These markers are used for the Marker Assisted Selection against the PSS condition. Recently, the ryanodine receptor gene has been identified as a good candidate for the Malignant Hyperthermia or Hal gene (47).

We have no doubt that Marker Assisted Selection will have its place in future breeding program, not only because it enhances the power of the present-day breeding strategies, but because, in combination with the manipulation of the germ-line, it offers entirely new possibilities as well.

REFERENCES:

1. SOLLER, M. & BECKMAN, J.S. (1982). Restriction Fragment Length Polymorphisms and Genetic Improvement. Proceedings of the 2nd World Congress on Genetics Applied to Livestock Production. Madrid **6**:396-404.
2. SMITH, C. & SIMPSON, S.P. (1986). The use of genetic polymorphism in livestock improvement. J. Anim. Breedg. Genet. **103**:205-217.
3. ORKIN, S.H. (1986). Reverse Genetics and Human Disease. Cell **47**:845-850.
4. DOOLITTLE, W.F. (1985). RNA mediated gene conversion? Trends in Genetics 1 (3),64,65.
5. JEFFREYS, A.J.; NEUMANN, R.; WILSON, V. (1990). Repeat unit sequence variation in minisatellites: a novel source of DNA polymorphism for studying variation and mutation by single molecule analysis. Cell **60**:473-485.
6. GEORGES, M. (1990). Hypervariable minisatellites and their use in animal breeding. In: Gene mapping: strategies, techniques and applications (ed. by L.B. Schook, Lewin and McLaren). Marcel Dekker, New York. to be published.
7. LI, H.; GYLLENSTEN, U.B.; CUI, X.; SAIKI, R.K.; ERLICH, H.A.; ARNHEIM, N. (1988). Amplification and analysis of DNA sequences in single human sperm and diploid cells. Nature **335**:414.
8. MORTON, N.E. (1955). Sequential tests for the detection of linkage. American Journal of Human Genetics **7**:277-318.
9. LATHROP, G.M.; LALOUEL, J.M.; JULIER, C.; OTT, J. (1984). Strategies for multilocus linkage analysis in humans. Proc. Natl. Acad. Sci. USA **101**:3443-3446.
10. LATHROP, G.M.; LALOUEL, J.M.; JULIER, C.; OTT, J. (1985). Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. American Journal of Human Genetics **37**:482-498.
11. LANDER, E.S.; GREEN, P. (1987). Construction of multilocus linkage maps in humans. Proceedings of the National Academy of Sciences USA **84**:2363-2367.
12. WHITE, R.; LEPPERT, M.; BISHOP, D.; BARKER, D.; BERKOWITZ. J.; BROWN, C.; CALLAHAN, P.; HOLM, T.; JEROMINSKI, L. (1985). Construction of linkage maps with DNA markers for human chromosomes. Nature **313**:101-105.
13. DONIS-KELLER, H. et al. (1987). A genetic linkage map of the human genome. Cell **51(2)**:319-337.
14. FRIES, R.; BECKMANN, J.S.; GEORGES, M.; SOLLER, M.; WOMACK, J. (1989). The bovine gene map. Animal Genetics **20**:3-29.
15. LANDER, E.S.; BOTSTEIN, D. (1986). Mapping complex genetic traits in humans: new methods using complete RFLP linkage maps. Cold Spring Harbor Symposia on Quantitative Biology **51**:49-62.
16. BOTSTEIN, D.; WHITE, R.L.; SKOLNICK, M.; DAVIS, R.W. (1980). Construction of a Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms. American Journal of Human Genetics **32**:314-

331.
17. WHITE, R.; CASKEY, C.T. (1988). The human as an experimental system in molecular genetics. Science **240**:1483-1488.
18. RIORDAN, J.R.; ROMMENS, J.M.; KEREM, B.-S.; ALON, N.; ROZMAHEL, R.; GRZELCZAK, Z.; ZIELENSKI, J.; LOK, S.; PLAVSIC, N.; CHOU, J.L.; DRUMM, M.L.; IANUZZI, M.C.; COLLINS, F.S.; TSUI, L.-C. (1989). Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science **245**:1066-1073.
19. GEORGES, M.; LEQUARRÉ, A.S.; CASTELLI, M.; HANSET, R.; VASSART, G. (1988a). DNA fingerprinting in domestic animals using four different minisatellite probes. Cytogenet Cell Genet **47**:127-131.
20. GEORGES, M.; LATHROP, M.; HILBERT, P.; MARCOTTE, A.; SCHWERS, A.; SWILLENS, S.; VASSART, G.; HANSET, R. (1990). On the use of DNA fingerprints for linkage studies in cattle. Genomics **6**, 461-474.
21. GEORGES, M.; GUNAWARDANA, A.; THREADGILL, D.; LATHROP, M.; OLSAKER, I.; HOLM, T.; MISHRA, A.; SARGEANT, L.; STEELE, M.; TERRY, Ch.; ZHAO, X.; FRIES, R. & WOMACK, J. Characterization of a set of VNTR-markers for Bovidae. submitted.
22. STEELE, M.; GEORGES, M. *Generation of bovine multisite haplotypes using random cosmids. submitted.*
23. GEORGES, M.; LEQUARRE, A.S.; HANSET, R.; VASSART, G. (1987). Genetic variation of the bovine thyroglobulin locus studied at the DNA level. Animal Genetics **18**:41-50.
24. HILBERT, P.; MARCOTTE, A.; SCHWERS, A.; HANSET, R.; VASSART, G.; GEORGES, M. (1989). Analysis of genetic variation in the Belgian Blue Cattle Breed using DNA sequence polymorphism at the growth hormone, low density lipoprotein receptor, α-subunit of glycoprotein hormones and thyroglobulin loci. Animal Genetics **20**, 383-394.
25. BARKER, D.; SCHAFER, M.; WHITE, R. (1984). Restriction sites containing CpG show a higher frequency of polymorphism in human DNA. Cell **36**, 131-138.
26. WEBER, J.L.; MAY, P.E. (1989). Abundant class of human DNA polymorphisms which can be typed using the Polymerase Chain Reaction. American Journal of Human Genetics **44**:388-396.
27. LITT, M.; LUTY, J.A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am. J. Hum. Genet. **44**:397-401.
28. TAUTZ, D. (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucl Acids Res **17** (16):6463-6471.
29. MOORE, S.S.; SARGEANT, L.; MATTICK, J.S.; KING, T.J.; HETZEL, D.J.S., GEORGES, M. The conservation of DNA microsatellite sequences allows the use of heterologous pcr primer pairs in cattle and sheep. submitted.
30. PATERSON, A.H.; LANDER, E.S.; HEWITT, J.D.; PETERSON, S.; LINCOLN, S.E.; TANKSLEY, S.D. (1989). Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature **335**:721-726.
31. MARTIN, B.; NIENHUIS, J.; KING, G.; SCHAEFFER, A. (1989). Restriction fragment length polymorphisms associated with water use efficiency in tomato. Science **243**:1725-1727.
32. HOESCHELE, I.; MEINERT, T.R. (1990) Association of genetic defects with yield and type traits: a major production gene is linked to Weaver. Journal of Dairy Science: in the press.
33. HANSET, R. (1982). *Major genes in animal production, examples and perspectives: cattle and pigs.* Proceedings of the 2nd World Congress on Genetics applied to Livestock Production (madrid) **6**:439-453.
34. LANDER. E.S.; BOTSTEIN, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics **121**:185-199.
35. SOLLER, M.; BRODY, T.; GENIZI, A. (1976). On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. Theoretical and applied genetics **47**:35-39.
36. WELLER, J.I.; KASHI, Y.; SOLLER, M. (1990). Power of "daughter" and "granddaughter" designs for genetic mapping of quantitative traits in dairy cattle using genetic markers. in the press.
37. GELDERMANN, H.; PIEPER, U.; ROTH, B. (1985). Effects of marked chromosome sections on milk performance in cattle. Theoretical and Applied Genetics **70**:138-146.
38. COWAN, C.M.; DENTINE, M.R.; AX, R.L.; SCHULER, L.A. (1990). Structural variation around prolactin gene linked to quantitative traits in an elite Holstein sire family. Theoretical and Applied Genetics **79**:577-582.
39. HANSET, R. (1989). Les perpectives d'une selection assistee par marqueurs. Annales de Medecine Veterinaire **133**:465-488.
40. RUSSE, I. (1983). Oogenesis in cattle and sheep. Bibliotheca anatomica **24**:77-92.

41. BETTERIDGE, K.J.; SMITH, C.; STUBBINGS, R.B.; XU, K.P.; KING, W.A. (1989). Potential genetic improvement of cattle by fertilization of fetal oocytes in vitro. Journal of Reproduction and Fertility Suppl. **38**:87-98.
42. HUNT, T. (1989). Under arrest in the cell cycle. Nature **342**:483-484.
43. SEIDEL, G.E. (1981). Superovulation and embryo transfer in cattle. Science **211**:351-358.
44. HERR, C.M.; MATTHAEI, K.I.; BRADLEY, M.P.; REED, K.C. (1990). Rapid, accurate sexing of livestock embryos. Proceedings of the 4th world congress on genetics applied to livestock production, **XVI**: 334-343.
45. PERRET, J.; SHIA, Y-C.; FRIES, R.; VASSART, G.; GEORGES, M. (1990). A polymorphic satellite sequence maps to the pericentric region of the bovine Y chromosome. Genomics **6**:482-490.
46. ARCHIBALD, A.L.; IMLAH, P. (1985). The halothane sensitivity locus and its linkage relationships. Animal Blood Groups and Biochemical Genetics **16**:253-263.
47. MacLENNAN, D.H.; DUFF, C.; ZORZATO, F.; FUJII, J.; PHILLIPS, M.; KORNELUK, R.G.; FRODIS, W.; BRITT, B.A.; WORTON, R.G. (1990). Ryanodine receptor gene is a candidate for predisposition to malignant hyperthermia. Nature **343**:559-561.22