The 6th MAFF International Workshop on Genetic Resources

Genetic Diversity and Conservation of Animal Genetic Resources

November 4-5, 1998
National Institute of Agrobiological Resources
Tsukuba, Ibaraki, JAPAN

Sponsored by
Research Council Secretariat of MAFF
National Institute of Agrobiological Resources

in corporation with
National Institute of Animal Industry
National Institute of Sericultural and Entomological Sciences
National Institute of Animal Health
Japan International Research Center for Agricultural Sciences

Contents

Opening address

NAOKI KATSURA
Director General, National Institute of Agrobiological Resources
Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan

First, I would like to extend heartfelt welcome to all participants of this workshop. It is great pleasure for me, particularly, to welcome participants who have traveled a long way to join us. Some invited speakers have traveled more than 10 hours, from Europe or America, and I hope they have overcome the jet lag already. We do appreciate all of you for sparing the time to join in this workshop.

I would like to acknowledge and thank the Agriculture, Forestry and Fisheries Research Council, our sister institutes and committee members who have helped to make this workshop possible.

Conservation of genetic resources of animal should be a global concerned and seen as a way to improve the global food supply. This is needed for the basic security of all people and a most urgent tasks facing humanity. All countries are interdependent when it comes to conserving, evaluating and using genetic resources. Many issues related to bio-diversity are becoming international issues involving government level meetings but these issues have implications for science and scientists. Fostering international cooperation and interaction is one of the results which I hope will emerge from the next few days you all have together in Tsukuba.

Animal genetic resources conservation is the particular focus of your discussions of this workshop. The papers to be presented in this meeting cover a very wide range from genetics to economics. Each field has its own particular interest. However, over the next few days I hope all participants will discuss and exchange their views and be stimulated by the various reports presented. I hope you will find differences among research which will be of interest in the research you are engaged in. Please, seek commonness which will make the broad management of conserving animal genetic resources more efficient.

I would like to introduce the National Institute of Agrobiological Resources (NIAR). NIAR has 5 main research sections:

Genetic Resources,
Molecular Genetics,
Plant Physiology,
Biotechnology,
Radiation Breeding.

I would like to explain the connection between these sections by saying that each aims to provide a foundation for improving agriculture. Conservation of genetic resources as well as, basic research on molecular biology is essential to improve agricultural science and production. We are trying to use fully the advances in biotechnology and information sciences to make scientific progress.

One of the most important facilities is the genebank project for plant, animal, microorganism and DNA genetic resources. For animals, we have more than 700 species or breeds conserved.

However, an institute does not stand alone. All our research depends on partnership both within Japan and abroad. The collaborative linkages we are hopefully strengthen by joint research are an important component of our institutes activity. We are looking forward to enhanced collaboration in Japan and, particularly, with colleagues from overseas in the future.

I sincerely hope that those participating in this workshop from various countries will have something to offer regarding our study for animal genetic resources and will contribute to their future development.

Thank you very much.

Session 1. Improvement and Conservation of Genetic Diversity in Livestock

Effective Population Size and Inbreeding under Selection

TETSURO NOMURA
Department of Biotechnology, Faculty of Engineering,
Kyoto Sangyo University, Kyoto 603-8555, Japan

Abstract

Effective population size is one of the most important parameters in the definition of selection programmes, because the magnitude of this parameter determines the effect of both random genetic drift and inbreeding. In this paper, the basic principle and several extensions of prediction theory of effective population size under selection are presented. The theory is applied to the evaluation of long-term breeding schemes with BLUP selection.

Introduction

The effective population size is one of the most important parameters in the definition of selection programmes, because the magnitude of this parameter determines the effects of both random genetic drift and inbreeding (Falconer, 1989; Caballero, 1994). In the absence of selection, all the individuals of a population have the same expected number of offspring, and differences in family size are due to only random sampling of individuals among families. However, in populations under selection, families do not have equal probabilities of contribution to the next generation because of inherited or noninherited causes. This leads to a reduction in effective population size. Thus, selection inflates the rate of inbreeding and the amount of random genetic drift.

In this paper, the basic principle and several extensions of prediction theory of effective population size under selection are presented. The theory is applied to the evaluation of long-term breeding schemes with BLUP selection.

Basic formula

Wright (1939) derived a formula for the effective population size (N_e) in terms of variance of family size (S_k^2),

$$N_e = \frac{4N}{2 + S_k^2}, \qquad [1]$$

where N is the number of parents. This formula can be applied when selection is acting on a noninherited trait. When selection acts on an inherited trait, the effective size cannot be predicted solely from the variance of the family size at a given generation, because selection affects on the variance of the family size not only in the parental generation but also in more ascended generations (Wray *et al.*, 1990). Thus, selection has a cumulative effect on the effective population size.

The cumulative effect can be understood from the following consideration: Under selection on an inherited trait, more offspring are likely to be selected from a genetically superior parent than from a genetically inferior parent. Thus, the effective size is reduced by an increase in variance of family size (S_k^2 in [1]). However, more grandoffspring are also likely to be selected from a genetically superior grandparent for two reasons. Firstly, the grandparent was genetically superior as a parent and has already contributed more offspring, and therefore has more grandoffspring available for selection. Secondly, the grandoffspring has inherited superior genes from the grandparent, and so are more likely to be selected than their contemporaries.

Following Santiago and Caballero (1995), we formulate the cumulative effect of selection by considering the change in frequency of a neutral allele. A population in which the same number of male and female parents are pair mated to create $N/2$ families each generation is supposed. We consider an autosomal neutral allele unlinked to the selected genes with the initial frequency p_0. Effective population size is computed from the variance of change in gene frequency of this allele among infinite replicates of the population. We first consider effective size after one generation of selection. Later, the asymptotic prediction under continued selection is addressed.

First generation: Changes in gene frequency arise from three independent process. (a) Random association between the neutral allele and families with a selective advantage or disadvantage: For example, if a family with a high selective advantage has the neutral allele with a high frequency, the gene frequency in the next generation is expected to increase, and *vice versa*. (b) Random sampling of individuals among families: Actual variation in contribution of the family to the next generation is caused not only by the differences of selective advantage among families but also random or noninherited causes. (c) Mendelian sampling of the neutral allele in the families in which one or both parents are heterozygotes. Changes in gene frequency due to these three causes after one generation of selection are denoted by S_1, D_1 and H_1, respectively. Thus, the total change in gene frequency is $\Delta p = p_1 - p_0 = S_1 + D_1 + H_1$, and the

variance of the change among infinite replicates is expressed as

$$V(\Delta p) = E(S_1^2) + E(D_1^2) + E(H_1^2), \qquad [2]$$

where E denotes expectation. According to Santiago and Caballero (1995), the three expectations are approximated by

$$E(S_1^2) = \frac{p_0(1 - p_0)}{2N} C^2$$

$$E(D_1^2) = \frac{p_0(1 - p_0)}{2N} \frac{S_k^2}{4}$$

and

$$E(H_1^2) = \frac{p_0(1 - p_0)}{4N},$$

respectively, where C^2 is the variance of selective advantage among families. Substituting the above expressions into [2] and equating the variance of change in gene frequency in an ideal population $p_0 (1- p_0) / 2N_e$, we get the effective size in the first generation

$$N_e = \frac{4N}{2 + S_k^2 + 4C^2}. \qquad [3]$$

Without selection (C^2=0), [3] reduces to [1], as expected. Under phenotypic selection, the variance of selective advantage is approximated by

$$C^2 = i^2 \rho \qquad [4]$$

(Robertson, 1961; Burrows, 1984), where i is the selection intensity and ρ is the intraclass correlation of full-sibs.

Asymptotic prediction: Although the term C^2 in [3] reflects the effect of selection in offspring on change in gene frequency, this change partially persists in descendants. All of the persisted changes are ascribed to the random association between the neutral allele and families with a selective advantage or disadvantage, existed in generation 0. Thus, we should incorporate the persisted changes into [3]. This problem was first addressed by Robertson (1961). He showed that the cumulative effect of selection on the change in gene frequency increases up to a limiting value of Q times the initial effect. This means that to obtain the asymptotic prediction, the term arising from selection (C^2 in [3]) should be multiplied by Q^2. Thus, the prediction equation for the asymptotic effective size is

$$N_e = \frac{4N}{2 + S_k^2 + 4Q^2C^2}. \qquad [5]$$

Since the average selective advantage (and, therefore, the changes of gene frequency of

the neutral allele) is expected to be reduced by one-half each generation in descendants by segregation and recombination, Robertson (1961) expressed Q by a geometric series $Q = 1 + 1/2 + 1/4 + 1/8 + \cdots = 2$. This expression overpredicts the cumulative effect of selection, because the reduction of genetic variance due to selection (Bulmer effect; Bulmer, 1980) is not accounted for. Santiago and Caballero (1995) derived a more precise expression

$$Q = 1 + \left[\frac{G}{2}(1+r)\right] + \left[\frac{G}{2}(1+r)\right]^2 + \cdots$$
$$= \sum_{i=0}^{\infty}\left[\frac{G}{2}(1+r)\right]^i = \frac{2}{2 - G(1+r)},$$

where G is the remaining proportion of genetic variance in the selected individuals, and r is the correlation between the expected selective advantages of male and female parents. Under truncation selection on a normally distributed trait controlled by an infinitesimal model of gene effects (Bulmer, 1980), $G = 1 - kh^2$, where $k = i\,(i - x)$, x is the truncation point of the standardized normal distribution, and h^2 is the heritability.

The above derivation has been based on the variance of change in gene frequency (variance effective size). Using the rate of inbreeding (inbreeding effective size), Wray and Thompson (1990a) developed a method to approximate the effective size of selected populations, as a function of the mean and variance of the contributions of ancestors in the first generation, to descendants in the limit. Their method, however, requires complex recurrence computations. To overcome this problem, Woolliams *et al.* (1993) derived equations to predict the mean and variance of the contributions of ancestors to descendants. Although the prediction equation derived by Woolliams *et al.* (1993) is slightly different from [5], simulation studies have shown that both equations give good predictions (Woolliams and Thompson, 1994; Santiago and Caballero, 1995).

Several extensions

Nonrandom mating and different numbers of male and female parents: Suppose a population in which N_m males are mated with N_f females ($N_m < N_f$). The number of females mated with one male is an integer number of N_f / N_m. The prediction equation for the effective size derived by Santiago and Caballero (1995), but including some terms arising from selection and nonrandom mating (Wang, 1996, 1998; Nomura,

1997a), is

$$\frac{1}{N_e} = \frac{1}{16N_m}\left[\left(\frac{1}{\mu_{mm}} + \frac{1}{\mu_{mf}}\right)\left(1-\alpha_{I,m}\right)\right.$$
$$\left.+\left(\frac{S_{mm}^2}{\mu_{mm}^2} + \frac{2S_{mm,mf}}{\mu_{mm}\mu_{mf}} + \frac{S_{mf}^2}{\mu_{mf}^2} + 4Q^2C_m^2\right)\left(1+\alpha_{I,m}+4\alpha_0\right)\right]$$
$$+\frac{1}{16N_f}\left[\left(\frac{1}{\mu_{fm}} + \frac{1}{\mu_{ff}}\right)\left(1-\alpha_{I,f}\right)\right. \qquad [6]$$
$$\left.+\left(\frac{S_{fm}^2}{\mu_{fm}^2} + \frac{2S_{fm,ff}}{\mu_{fm}\mu_{ff}} + \frac{S_{ff}^2}{\mu_{ff}^2} + 4Q^2C_f^2\right)\left(1+\alpha_{I,f}\right)\right],$$

where μ_{su} ($= N_u / N_s$) and S_{su}^2 are the mean and variance of offspring number of sex u ($= m$ or f) from parents of sex s ($= m$ or f), respectively, $S_{sm,sf}$ is the covariance between the numbers of male and female offspring from parents of sex s, α_0 is the correlation between genes of mated parents, $\alpha_{I,s}$ is the deviation from Hardy-Weinberg proportions in parents of sex s, and C_s^2 is the variance of selective advantage of parents of sex s.

Overlapping generations: In most domestic animals, the generations are not discrete but are overlapping. Nomura (1996) formulated the cumulative effect of selection in populations with overlapping generations by applying the gene-flow technique (transition matrix approach), that was originally proposed by Hill (1974) to describe the dissemination of genetic gain achieved in a given year. Under random mating, the prediction equation for the effective size per generation is

$$\frac{1}{N_e} = \frac{1}{16N_{m,1}L}\left[\left(\frac{1}{\mu_{mm}} + \frac{1}{\mu_{mf}}\right)\left(1-\alpha_{I,m}\right)\right.$$
$$\left.+\left(\frac{S_{mm}^2}{\mu_{mm}^2} + \frac{2S_{mm,mf}}{\mu_{mm}\mu_{mf}} + \frac{S_{mf}^2}{\mu_{mf}^2} + 4Q^2C_m^2\right)\left(1+\alpha_{I,m}\right)\right]$$
$$+\frac{1}{16N_{f,1}L}\left[\left(\frac{1}{\mu_{fm}} + \frac{1}{\mu_{ff}}\right)\left(1-\alpha_{I,f}\right)\right. \qquad [7]$$
$$\left.+\left(\frac{S_{fm}^2}{\mu_{fm}^2} + \frac{2S_{fm,ff}}{\mu_{fm}\mu_{ff}} + \frac{S_{ff}^2}{\mu_{ff}^2} + 4Q^2C_f^2\right)\left(1+\alpha_{I,f}\right)\right],$$

where $N_{s,l}$ is the number of parents of sex s entering the population each year and L is the generation interval. The variance and covariance terms in [7] is same as [6] but should be computed based on the lifetime progeny numbers of parents. In comparing [7] and the corresponding equation for discrete generations (equation [6]), it is found that the effective size of a selected population with overlapping generations is the same as that for a discrete-generation population having the same number ($N_{s,l}$ L) of individuals entering the population each generation, which is in accordance with the corresponding result from nonselected populations (Hill, 1979).

Hierarchical structure: Animal breeding programmes frequently show a hierarchical structure, in which genetic improvement is made in a nucleus herd, and the improvement is transferred by migration to base or commercial. Extending the transition matrix approach, Nomura (1997b) derived a prediction equation for the effective size of populations with hierarchical structure.

Sex-linked locus: With a sex-linked gene, the effect of selection is accumulated through more complex process, because the heterogametic sex has only one copy of the sex-linked gene. Nomura (1997c) described the cumulative effect of selection on the change in frequency of a sex-linked neutral gene, by applying the transition matrix method. For species in which males are heterogametic, the coefficients accounting for the cumulative change in gene frequency from parents of sex s to offspring of sex u, Q_{su}, are expressed as

$$Q_{mf} = Q_{ff} = \frac{2 + G_m(1+r)}{2 - \{G_f(1+r)/2\}\{1 + G_m(1+r)/2\}}$$

$$Q_{fm} = \frac{2}{2 - \{G_f(1+r)/2\}\{1 + G_m(1+r)/2\}},$$

where G_s is the remaining proportion of the additive genetic variance after selection of individuals of sex s. Using these coefficients, the effective size of population under random mating is predicted by

$$\frac{1}{N_e} = \frac{1}{9N_m}\left[1 - \alpha_{I,f} + 2\left\{\frac{S_{mf}^2}{\mu_{mf}^2} + Q_{mf}^2 C_{mf}^2\right\}\right]$$

$$+\frac{1}{9N_f}\left[1 - \alpha_{I,f} + \left\{\frac{S_{fm}^2}{\mu_{fm}^2} + \frac{2S_{fm,ff}}{\mu_{fm}\mu_{ff}} + \frac{S_{ff}^2}{\mu_{ff}^2} + \left(Q_{fm}C_{fm} + Q_{ff}C_{ff}\right)^2\right\}\left(1 + \alpha_{I,f}\right)\right],$$

where C_{su}^2 the variance of the relative selective advantage of parents of sex s in their

contribution to offspring of sex *u*. The equation extended to nonrandom mating populations is given by Wang (1998).

Index selection: In animal breeding practice, selection on an index including family information is often carried out. To consider the effect of index selection on the effective size, we suppose that selection is based on an index of an individual's phenotype (P) and the mean of its full sib family ($\overline{P}_D$, including the individual):

$$I = \beta_1 (P - \overline{P}_D) + \beta_2 \overline{P}_D .$$

The optimal index weights are given by

$$\beta_1 = \frac{1 - r}{1 - t} h^2$$

$$\beta_2 = \frac{1 + (n-1)r}{1 + (n-1)t} h^2$$

(Falconer, 1989), where r is the genetic correlation among full-sibs ($r = 1/2$), t is the phenotypic intraclass correlation among full-sibs ($t = h^2/2 + \sigma_C^2/\sigma_P^2$: σ_C^2 = the common environmental variance; σ_P^2 = the phenotypic variance) and n is the number of full-sibs per family. The intraclass correlation of full-sibs for the index value is expressed as

$$\rho = \left[-\beta_1^2 \frac{1}{n}\left(\sigma_{Aw}^2 + \sigma_E^2\right) + \beta_2^2 \left\{ \sigma_{Ab}^2 + \sigma_C^2 + \frac{1}{n}\left(\sigma_{Aw}^2 + \sigma_E^2\right) \right\} \right] \Big/ \sigma_I^2 , \qquad [8]$$

where σ_{Aw}^2 is the genetic variance within family, σ_{Ab}^2 is the genetic variance between family, σ_E^2 is the environmental variance, and σ_I^2 is the variance of index. With this correlation, the variance of selective advantage C^2 is computed from [4]. Thus, the effective size after one generation of selection is predicted from [3].

Under phenotypic selection ($\beta_1 = \beta_2 = h^2$) in the absence of common environmental variance ($\sigma_C^2 = 0$), [8] reduces to $\rho = \sigma_{Ab}^2/\sigma_P^2 = h^2/2$, meaning that ρ arises only from the breeding values of parents. Since the cumulative effect of selection on effective size is caused by the transmission of breeding values of parents to their descendants, full portion of the differences of selective advantages among families (σ_{Ab}^2) have the cumulative effect on the effective size. Thus, as already seen, the asymptotic effect is predicted by $Q^2 C^2 = Q^2 i^2 \rho$.

Under index selection, [8] is rewritten as

$$\rho = \beta_2^2\sigma_{Ab}^2/\sigma_I^2 + \left[\beta_2^2\sigma_C^2 + \left(\beta_2^2 - \beta_1^2\right)\left(\sigma_{Aw}^2 + \sigma_E^2\right)/n\right]/\sigma_I^2$$
$$= \rho_A + \rho_E,$$

where ρ_A is the component of ρ due to the breeding values of parents (which has the cumulative effect on N_e), and ρ_E is the component of ρ due to within-family variance and common environmental variance (which has no cumulative effects on N_e). Thus, the asymptotic prediction of the cumulative effect is written as

$$i^2\rho + i^2\left(Q^2 - 1\right)\rho_A = i^2\rho_E + i^2Q^2\rho_A \qquad [9]$$

(Wray *et al.*, 1994; Caballero *et al.*, 1996). Replacing Q^2C^2 by [9], we get a prediction equation for the effective size

$$N_e = \frac{4N}{2 + S_k^2 + 4i^2\rho + 4i^2\left(Q^2 - 1\right)\rho_A}$$
$$= \frac{4N}{2 + S_k^2 + 4i^2\rho_E + 4i^2Q^2\rho_A}. \qquad [10]$$

Evaluation of long-term breeding schemes with BLUP selection

The theory of selection index has been developed to enable simultaneous prediction of an individual's breeding value and fixed effects, such as year and season, using information from all the available relatives. This is known as best linear unbiased prediction (BLUP) of breeding values with an animal model (Henderson, 1984). BLUP selection maximizes the response after one generation. However, through the increase in the probability of coselection of close relatives, BLUP selection reduces the effective size and, then, inflates the rate of inbreeding. As long-term selection response is highly dependent on the effective size (Robertson, 1960), BLUP selection applied to long-term breeding schemes may be less efficient than simple phenotypic selection.

Wray and Hill (1989) have developed a method for predicting asymptotic rate of response to BLUP selection using an index (pseudo-BLUP) of an individual's record, the mean of the records of its full- and half-sibs and the estimated breeding values (EBVs) of its sire, its dam and all the mates of its sire. In this section, this method is used to predict cumulative selection response and inbreeding under BLUP selection.

Model and assumptions: The trait under selection is assumed to be determined by an infinite number of additive loci, each with infinitesimal effect (the infinitesimal model; Bulmer, 1980). A population with discrete generations is supposed, in which N_m males

are randomly mated with N_f females. The number of females mated with one male is an integer number of $d = N_f / N_m$. Subscript s (= m or f) is used to specify sex. Each female produces n progeny consisting of n_f progeny of each sex ($n = 2n_f$). Each male has $n_m = dn_f$ progeny of each sex. Among the progeny, the best N_m males and the best N_f females are selected on the basis of their EBVs from BLUP with an animal model. The proportion of selected for sex s is $p_s = 1/n_s$ and the selection intensity is $i_s = z_s / p_s$, where z_s is the normal ordinate corresponding the truncation point of selection (x_s). The variance reduction factor due to the Bulmer effect (Bulmer, 1980) is $k_s = i_s(i_s - x_s)$.

EBVs from BLUP are approximated by the pseudo-BLUP (Wray and Hill, 1989). The selection index assumed in the pseudo-BLUP is

$$I = \beta_{1(t)}\left(P - \bar{P}_D\right) + \beta_{2(t)}\left(\bar{P}_D - \bar{P}_H - \hat{A}_f/2 + \bar{\hat{A}}_f/2\right) \qquad [11]$$
$$+\beta_{3(t)}\left(\bar{P}_H - \hat{A}_m/2 - \bar{\hat{A}}_f/2\right) + \beta_{4(t)}\left(\hat{A}_m/2 + \hat{A}_f/2\right),$$

where P is the individual's record, $\bar{P}_D$ is the mean of records of n (including the individual) full-sib records, $\bar{P}_H$ is the mean of nd (including the n full-sibs) half-sib records, $\hat{A}_m$ and $\hat{A}_f$ are EBVs of the sire and dam, respectively, and $\bar{\hat{A}}_f$ is the mean of EBVs of all dams mated to the sire. The selection index weights are updated and optimized each generation using the equations given in Appendix 1.

Prediction of effective size and inbreeding: The asymptotic genetic parameters, which are required for the prediction of asymptotic effective size and rate of inbreeding, are obtained in the following way. The total phenotypic variance at generation t is

$$\sigma^2_{P(t)} = \sigma^2_{A(t)} + \sigma^2_C + \sigma^2_E,$$

where $\sigma^2_{A(t)}$ is the additive genetic variance, σ^2_C is the variance due to the common environment of full-sibs, and σ^2_E is the environmental variance. The additive genetic variance is further decomposed into between-sire family (σ^2_{Am}), between-dam family (σ^2_{Af}) and within-full-sib family (σ^2_{Aw}) components:

$$\sigma^2_{A(t)} = \sigma^2_{Am(t)} + \sigma^2_{Af(t)} + \sigma^2_{Aw}.$$

These components are obtained from

$$\left.\begin{aligned}\sigma^2_{Am(t)} &= \frac{1}{4}\left(1-k_m r^2_{(t-1)}\right)\sigma^2_{A(t-1)}\\ \sigma^2_{Af(t)} &= \frac{1}{4}\left(1-k_f r^2_{(t-1)}\right)\sigma^2_{A(t-1)}\\ \sigma^2_{Aw} &= \frac{1}{2}\sigma^2_{A(0)}\end{aligned}\right\} \qquad [12]$$

where $r_{(t-1)}$ is the accuracy of selection in generation t-1. The accuracy of selection is computed as $r_{(t)} = \sigma_{I(t)}/\sigma_{A(t)}$, where $\sigma_{I(t)}$ is the standard deviation of the index. The genetic parameters are computed each generation by using recurrently equations [12]. Preliminary analysis showed that all the genetic parameters reach essentially steady values after several generations. Thus, the values after 10 generations are used as the asymptotic values (asterisks will denote asymptotic values).

The asymptotic effective size of sex s is predicted by the equation combined [6] and [10]

$$N_{es} = 4N_s \Big/ \left[\frac{1}{\mu_{sm}} + \frac{1}{\mu_{sf}} + \frac{S^2_{sm}}{\mu^2_{sm}} + \frac{S^2_{sf}}{\mu^2_{sf}} + 4\rho^*_s i^2 + 4\rho^*_{As} i^2\left(Q^2-1\right)\right], \qquad [13]$$

where ρ^*_s is the asymptotic intraclass correlation of families from parents of sex s, ρ^*_{As} is component of ρ^*_s due to the breeding values of the parents, and i [$= (i_m + i_f)/2$] is the unweighted mean of selection intensities for the two sexes. The expressions of the correlations ρ^*_s and ρ^*_{As} are given in Appendix 2. The effective size of the population and the asymptotic rate of inbreeding are obtained by

$$N_e = \frac{4N_{em}N_{ef}}{N_{em}+N_{ef}} \qquad [14]$$

and

$$\Delta F = \frac{1}{2N_e} = \frac{1}{8N_{em}} + \frac{1}{8N_{ef}}. \qquad [15]$$

Prediction of selection response: In the calculation of the selection response, the reduction of the additive genetic variance due to the inbreeding is taken into account by using the equation suggested by Wray and Thompson (1990) and Wray and Goddard (1994),

$$\sigma^2_{A(t)} = \frac{1}{4}\left[\left(1 - k_m r^2_{(t-1)}\right) + \left(1 - k_f r^2_{(t-1)}\right)\right]\sigma^2_{A(t-1)} \frac{1 + F_{t-1} - 2F_t}{1 - F_{t-1}} + \frac{1}{2}\sigma^2_{A(0)}\left(1 - F_{t-1}\right). \quad [16]$$

The inbreeding coefficient at generation t is predicted by

$$F_t = F_{t-1} + \frac{1}{2N_e}\left(1 - 2F_{t-1} + F_{t-2}\right) \quad [17]$$

(Wright, 1931), where N_e is the asymptotic effective size computed from [13] and [14]. The response to selection in generation t is computed from $R_{(t)} = ir_{(t)}\sigma_{A(t)}$. Then the cumulative response over t generations is obtained by

$$CR_{(t)} = \sum_{j=1}^{t} R_{(j)} = \sum_{j=1}^{t} ir_{(j)}\sigma_{A(j)}. \quad [18]$$

Simulation: Stochastic simulation was carried out to check the prediction equations. Breeding values of parents in generation 0 were randomly chosen from an N(0, 1) distribution. Thus, the initial additive genetic variance, $\sigma^2_{A(0)}$ was 1. On the basis of the inifinitesimal model of gene effect (Bulmer, 1980), the breeding value (A_i) of an individual in the following generation was generated as

$$A_i = \frac{A_m + A_f}{2} + \phi_i\sqrt{\frac{1}{2}\sigma^2_{A(0)}\left(1 - \frac{F_m + F_f}{2}\right)},$$

where A_m and A_f are the breeding values of the sire and dam, respectively, F_m and F_f are the inbreeding coefficients of the sire and dam, respectively, and ϕ_i is a random number taken from an N(0, 1) distribution. Assuming that there are no common environmental factors ($\sigma_c^2 = 0$), the phenotypic value (P_i) was obtained as

$$P_i = A_i + \varphi_i\sigma_E,$$

where φ_i is a random number taken from an N(0, 1) distribution. The population with $N_m = 10$, $N_f = 60$ and $n_f = 3$ was simulated. The initial heritabilities (h_0^2) were assumed to be 0.1, 0.2, 0.4, 0.6 or 0.8. The number of simulated generations were 8 and the number of replicated runs were 50. Selection was based on BLUP with an animal model. The rate of inbreeding was obtained from the average inbreeding coefficients between 5 and 8 generations

$$\Delta F = \frac{1}{4}\sum_{i=5}^{8} \frac{F_i - F_{i-1}}{1 - F_{i-1}}.$$

For comparison, simple phenotypic selection was also simulated. The predictions for

phenotypic selection were made by putting $\beta_{1(t)} = \beta_{2(t)} = \beta_{3(t)} = \beta_{4(t)} = h_t^2$.

The inbreeding coefficients observed in simulation are shown in Figure 1. As seen from this figure, the rate of inbreeding under BLUP selection is much higher than under phenotypic selection, especially when the selected trait has a low heritability. As the heritability becomes low, the rate of inbreeding under BLUP selection tends to increase. This is because for a trait with low heritability, BLUP places larger weights on family information and, then, the probability of coselection of close relatives increases.

Table 1 gives the rate of inbreeding and effective size observed (sim) in simulation and predicted (pre) from equations [13]-[15]. On the whole, the equations give a good predictor, though it slightly over-predicts the rate of inbreeding and under-predicts the effective size.

Figure 2 shows the inbreeding coefficients observed in simulation and predicted from [17]. Although the prediction was made with the asymptotic effective size, it gave a good predictor of initial inbreeding under phenotypic selection. For BLUP selection, the presented method over-predicts the initial inbreeding. The presented method may give a conservative prediction of the effective size and rate of inbreeding under BLUP selection.

The cumulative selection responses observed in simulation and predicted from equations [16] and [18] are compared in Figure 3. In all the cases studied, there was a good agreement between the observed and predicted values.

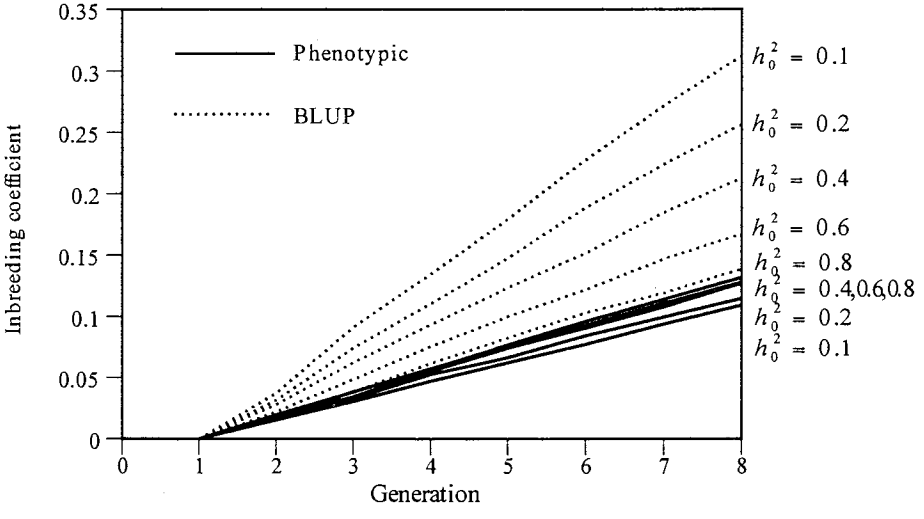


Figure 1. Observed inbreeding coefficients under phenotypic and BLUP selection with different initial heritabilities.

Table 1. Observed (sim) and predicted (pre) effective population size and rate of inbreeding under phenotypic and BLUP selection with different initial heritabilities.

	Rate of inbreeding				Effective population size			
	Phenotypic		BULP		Phenotypic		BULP	
h^2_0	sim	pre	sim	pre	sim	pre	sim	pre
0.1	0.0166	0.0171	0.0558	0.0572	30.07	29.17	8.96	8.74
0.2	0.0167	0.0190	0.0438	0.0523	29.90	26.92	11.42	9.56
0.4	0.0195	0.0207	0.0347	0.0412	25.59	24.20	14.42	12.14
0.6	0.0204	0.0207	0.0257	0.0315	24.46	24.18	19.42	15.85
0.8	0.0192	0.0195	0.0211	0.0235	26.05	25.62	23.64	21.31

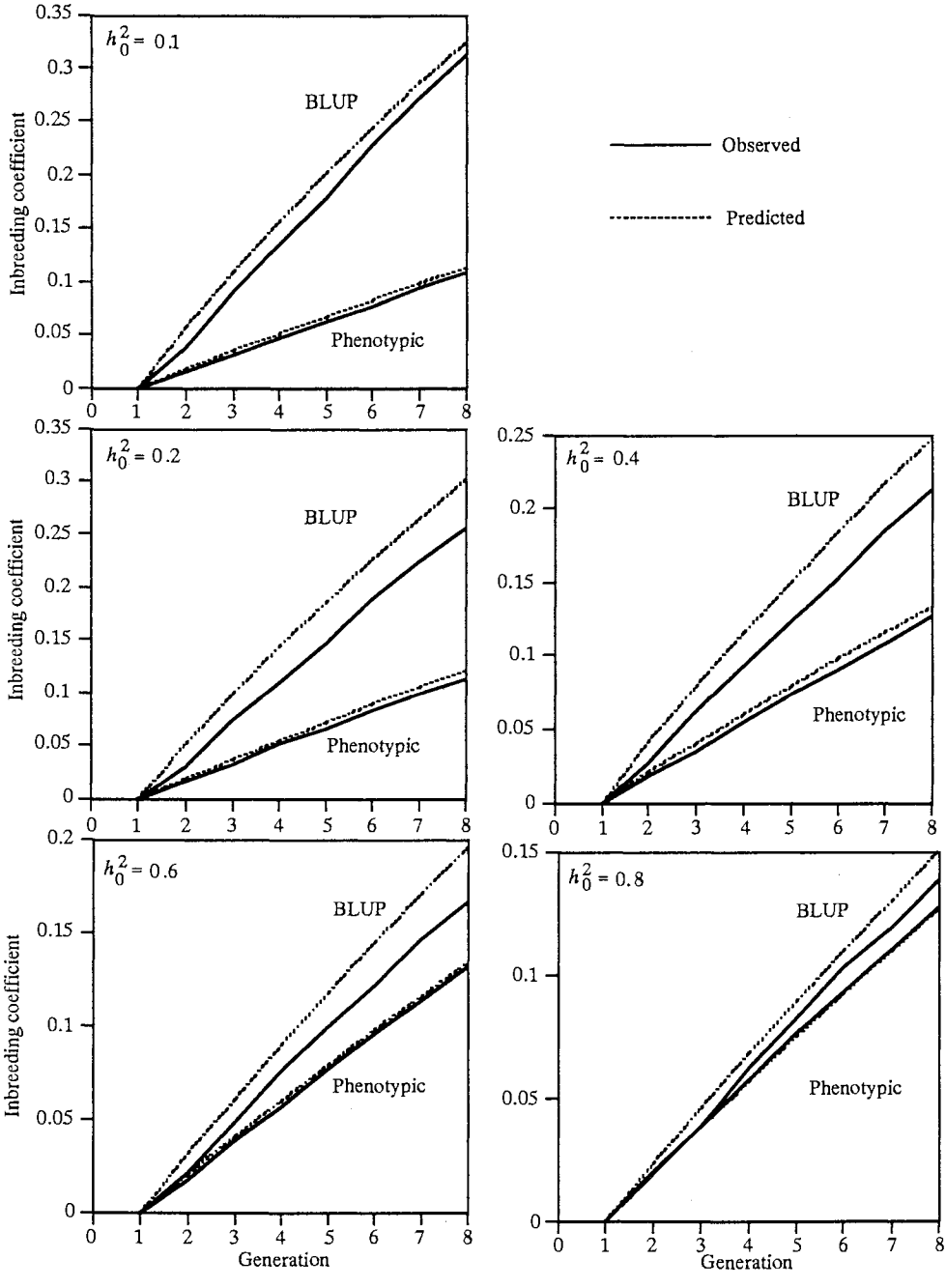


Figure 2. Observed and predicted inbreeding coefficients under phenotypic and BLUP selection with different initial heritabilities.

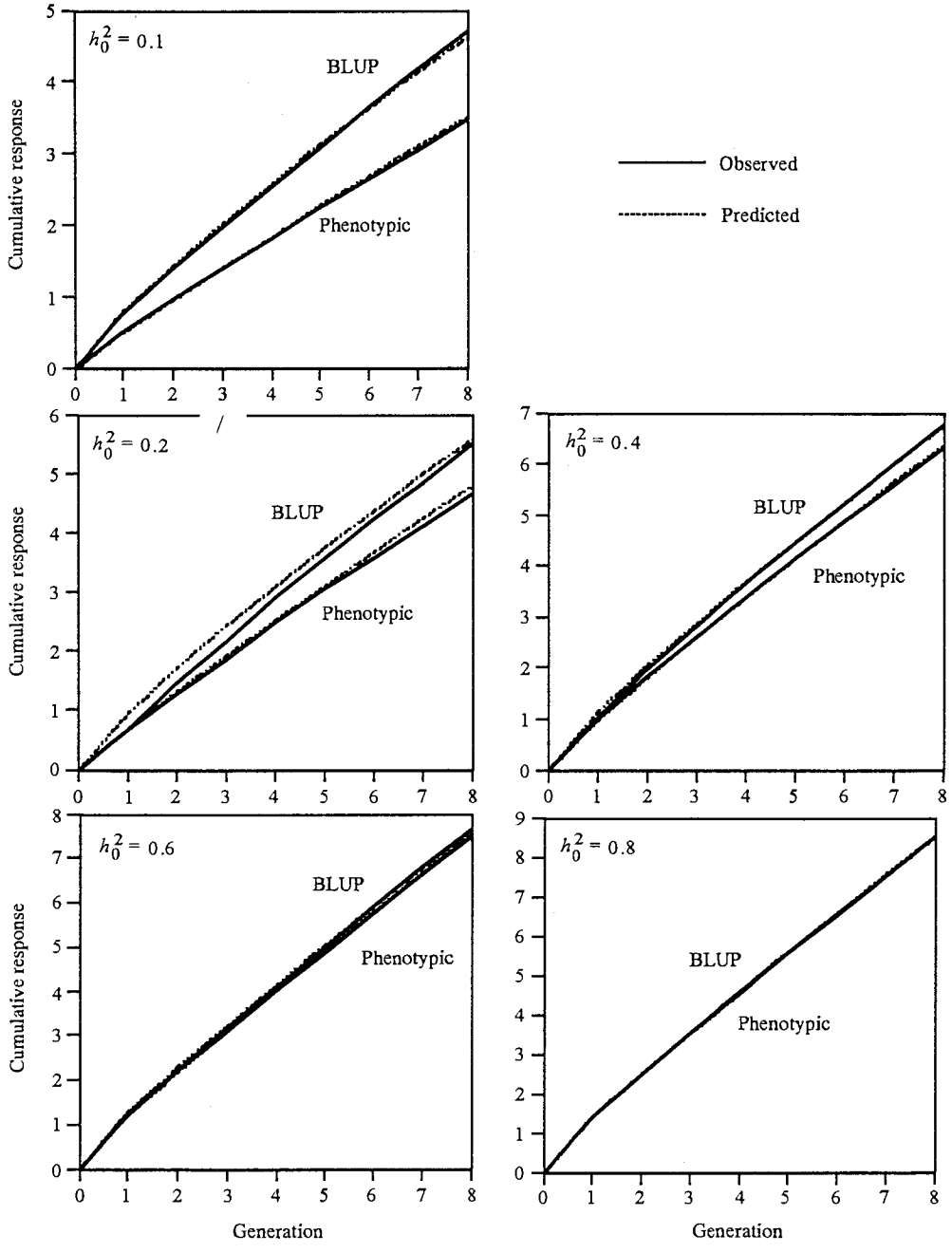


Figure 3. Observed and predicted cumulative responses under phenotypic and BLUP selection with different initial heritabilities.

Application: Quinton *et al.* (1992) suggested that it is important to compare selection programmes for response at the same level of inbreeding. This can be done graphically, relating the cumulative selection response with the level of inbreeding for a range in the number of sire used. As an application of the presented method, breeding programmes with BLUP and phenotypic selection are compared at the same level of inbreeding. A population with $h_0^2 = 0.3$, $N_f = 120$ and $n_f = 3$ was supposed. The number of sires was either N_m = 5, 10, 15, 20, 30, 40 or 60.

Figure 4 shows the cumulative responses over 10 and 20 generations of selection relative to the cumulated inbreeding. Except the case of N_m = 5 and 20 generations of selection, where the effect of selection and inbreeding on the cumulative response is most severe, BLUP selection always gives a higher response than phenotypic selection, for a given number of sires selected. However, for lower levels of inbreeding (e.g. $F_{10} < 10\%$ and $F_{20} < 20\%$), the cumulative response at the same inbreeding is greater with phenotypic selection. This suggests that BLUP selection may not be always optimal, especially when inbreeding is a limiting factor in the choice of selection programmes.

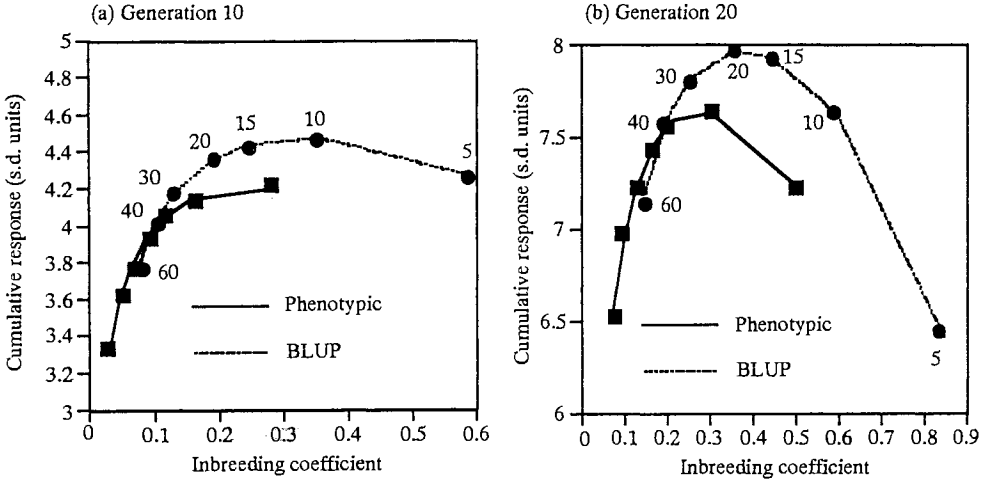


Figure 4. Cumulative response and inbreeding from 10 (a) or 20 (b) generations of phenotypic and BLUP selection. The number of sires: 5, 10, 15, 20, 30, 40 or 60 (the marked points on the graphs). The number of dams: 120

Final remark

This paper has investigated the prediction theories of effective population size and inbreeding under selection. All the theories are based on the infinitesimal model of gene effects, in which traits are assumed to be determined by an unlinked system of additive loci. This model is, however, inappropriate to make predictions of the effect of marker-assisted selection (MAS) on effective size, because quantitative trait loci (QTLs) and selected markers are linked and, therefore, they tend to be passed together over generations. Due to the hitchhiking effect, the drift process of genes linked with markers is expected to be larger than that predicted by the infinitesimal model (Santiago and Caballero, 1998). The prediction theory for effective size and inbreeding under MAS should be developed in future studies.

Appendix 1. Index weights in the pseudo-BLUP

Optimum weights for the pseudo-BLUP (equation [11]) are obtained by:

$$\beta_{1(t)} = \frac{\sigma^2_{Aw}}{\sigma^2_{Aw} + \sigma^2_E}$$

$$\beta_{2(t)} = \frac{\sigma^2_{Af(t)} + \sigma^2_{Aw}/n - V(\hat{A}_f)/4}{\sigma^2_{Af(t)} + \sigma^2_C + (\sigma^2_{Aw} + \sigma^2_E)/n - V(\hat{A}_f)/4}$$

$$\beta_{3(t)} = \frac{\sigma^2_{Am(t)} + \sigma^2_{Af(t)}/d + \sigma^2_{Aw}/dn - [V(\hat{A}_m) + V(\hat{A}_f)/d]/4}{\sigma^2_{Am(t)} + (\sigma^2_{Af(t)} + \sigma^2_C)/d + (\sigma^2_{Aw} + \sigma^2_E)/dn - [V(\hat{A}_m) + V(\hat{A}_f)/d]/4}$$

$$\beta_{4(t)} = 1,$$

where $V(\hat{A}_m)$ and $V(\hat{A}_f)$ are the variance of estimated breeding values of sires and dams, respectively. These variances are

$$V(\hat{A}_m) = r^2_{(t-1)}(1 - k_m)\,\sigma^2_{A(t-1)}$$

and

$$V(\hat{A}_f) = r^2_{(t-1)}(1 - k_f)\,\sigma^2_{A(t-1)}.$$

Appendix 2. Asymptotic intraclass correlation of families for the value of pseudo-BLUP

Let ρ^*_D and ρ^*_H be the asymptotic intraclass correlations of full- and half-sibs,

respectively. These are expressed as

$$\rho_D^* = \left\{-\beta_1^{*2}\left(\sigma_{Aw}^2 + \sigma_E^2\right)/n + \beta_2^{*2}(1-1/d)\left[\sigma_{Af}^{*2} + \sigma_C^2 - V^*\left(\hat{A}_f\right)/4 + \left(\sigma_{Aw}^2 + \sigma_E^2\right)/n\right]\right.$$
$$+\beta_3^{*2}\left[\sigma_{Am}^{*2} + \left(\sigma_{Af}^{*2} + \sigma_C^2\right)/d + \left(\sigma_{Aw}^2 + \sigma_E^2\right)/dn - \left[V^*\left(\hat{A}_m\right) + V^*\left(\hat{A}_f\right)/d\right]/4\right]$$
$$\left.+\beta_4^{*2}\left[V^*\left(\hat{A}_m\right) + V^*\left(\hat{A}_f\right)\right]/4\right\}/\sigma_I^{*2}$$

and

$$\rho_H^* = \left\{-\beta_2^{*2}\left[\sigma_{Af}^{*2} + \sigma_C^2 + \left(\sigma_{Aw}^2 + \sigma_E^2\right)/n - V\left(\hat{A}_f\right)/4\right]/d\right.$$
$$+\beta_3^{*2}\left[\sigma_{Am}^{*2} + \left(\sigma_{Af}^{*2} + \sigma_C^2\right)/d + \left(\sigma_{Aw}^2 + \sigma_E^2\right)/dn - \left(V^*\left(\hat{A}_m\right) + V^*\left(\hat{A}_f\right)/d\right)/4\right]$$
$$\left.+\beta_4^{*2}\,V^*\left(\hat{A}_m\right)/4\right\}/\sigma_I^{*2},$$

where asterisks denote the asymptotic values. With these correlations, the intraclass correlation of families from parents of sex s is obtained by

$$\rho_s^* = \left(1 - \frac{N_s}{N_f}\right)\rho_H^* + \frac{N_s}{N_f}\rho_D^*$$

The asymptotic intraclass correlations of full- and half-sibs due to the breeding values of parents ($\rho_{A,D}^*$ and $\rho_{A,H}^*$, respectively) are expressed as

$$\rho_{A,D}^* = \left\{\beta_2^{*2}(1-1/d)\left[\sigma_{Af}^{*2} - V^*\left(\hat{A}_f\right)/4\right]\right.$$
$$+\beta_3^{*2}\left[\sigma_{Am}^{*2} + \sigma_{Af}^{*2}/d - \left[V^*\left(\hat{A}_m\right) + V^*\left(\hat{A}_f\right)/d\right]/4\right]$$
$$\left.+\beta_4^{*2}\left[V^*\left(\hat{A}_m\right) + V^*\left(\hat{A}_f\right)\right]/4\right\}/\sigma_I^{*2}$$

and

$$\rho_{A,H}^* = \left\{-\beta_2^{*2}\left[\sigma_{Af}^{*2} - V\left(\hat{A}_f\right)/4\right]/d\right.$$
$$+\beta_3^{*2}\left[\sigma_{Am}^{*2} + \sigma_{Af}^{*2}/d - \left(V^*\left(\hat{A}_m\right) + V^*\left(\hat{A}_f\right)/d\right)/4\right]$$
$$\left.+\beta_4^{*2}\,V^*\left(\hat{A}_m\right)/4\right\}/\sigma_I^{*2}$$

With these correlations, the component of ρ_s^* to the breeding values of parents is obtained by

$$\rho_{As}^* = \left(1 - \frac{N_s}{N_f}\right)\rho_{A,H}^* + \frac{N_s}{N_f}\rho_{A,D}^*.$$

References

Bulmer, M.G. 1980. The Mathematical Theory of Quantitative Genetics. Clarendon Press, Oxford.

Burrows, P.M. 1984. Inbreeding under selection from unrelated families. Biometrics 40: 357-366.

Caballero, A. 1994. Developments in the prediction of effective population size. Heredity 73: 657-679.

Caballero, A., Wei, M., Hill, W.G. 1996. Survival rates of mutant gene under artificial selection using individual and family information. Journal of Genetics 75: 63-80.

Falconer, D.S. 1989. Introduction to Quantitative Genetics. 3rd edn. Longman, Harlow, Essex.

Henderson, C.R. 1984. Application of Linear Models in Animal Breeding. University of Guelph, Canada.

Hill, W.G. 1974. Prediction and evaluation of response to selection with overlapping generations. Animal Production 18: 117-139.

Hill, W.G. 1979. A note on effective population size with overlapping generations. Genetics 92: 317-322.

Nomura, T. 1996. Effective size of selected populations with overlapping generations. Journal of Animal Breeding and Genetics 113: 1-16.

Nomura, T. 1997a. On the effective size of nonrandom mating populations under selection. Japanese Journal of Biometrics 18: 1-11.

Nomura, T. 1997b. Prediction of effective population size in open nucleus breeding systems. Journal of Animal Breeding and Genetics 114: 333-347.

Nomura, T. 1997c. Effective population size for a sex-linked locus in populations under selection. Mathematical Biosciences 142: 79-89.

Quinton, M., Smith, C., Goddard, M.E. 1992. Comparison of selection methods at the same level of inbreeding. Journal of Animal Science 70: 1060-1067.

Robertson, A. 1960. A theory of limits in artificial selection. Proceedings of the Royal Society of London B153: 234-249.

Robertson, A. 1961. Inbreeding in artificial selection programmes. Genetical Research 2: 189-194.

Santiago, E., Caballero, A. 1995. Effective size of populations under selection. Genetics 139: 1013-1030.

Santiago, E., Caballero, A. 1998. Effective size and polymorphism of linked neutral loci in populations

under directional selection. Genetics 149: 2105-2117.

Wang, J. 1996. Inbreeding and variance effective sizes for nonrandom mating populations. Evolution 50: 1786-1794.

Wang, J. 1998. Effective size of populations under selection: some extensions. Proceedings of the 6th World Congress on Genetics Applied to Livestock Production 26: 13-16.

Woolliams, J.A., Wray, N.R., Thompson, R. 1993. Prediction of long-term contributions and inbreeding in populations undergoing mass selection. Genetical Research 62: 231-242.

Woolliams, J.A., Thompson, R. 1994. A theory of genetic contributions. Proceedings of the 5th World Congress on Genetics Applied to Livestock Production 19: 127-134.

Wray, N.R., Goddard, M.E. 1994. Increasing long term response to selection. Genetics, Selection, Evolution 26: 431-451.

Wray, N.R., Hill, W.G. 1989. Asymptotic rates of response from index selection. Animal Production 49: 217-227.

Wray, N.R., Thompson, R. 1990a. Prediction of rates of inbreeding in selected populations. Genetical Research 55: 41-54.

Wray, N.R., Thompson, R. 1990b. Advances in selection theory. Proceedings of the 4th World Congress on Genetics Applied to Livestock Production XIII: 167-176.

Wray, N.R., Woolliams, J.A., Thompson, R. 1990. Methods for predicting rates of inbreeding in selected populations. Theoretical and Applied Genetics 80: 503-512.

Wray, N.R., Woolliams, J.A., Thompson, R. 1994. Prediction of rates of inbreeding in populations undergoing index selection. Theoretical and Applied Genetics 87: 878-892.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16: 97-159.

Wright, S. 1939. Statistical Genetics in Relation to Evolution. Exposes de Biometrie et de Statistique Biologique, Herman & Cie, Paris, France.

Genetic Improvement in Dairy Cattle and its Consequence on Effective Population Size

MRODE RA, SWANSON GJT and LINDBERG CM
Animal Data Centre Ltd, Fox Talbot House,
Greenways Business Park, Chippenham SN15 1BN UK

Abstract

Genetic progress that has accompanied the globalization of Holstein cattle was investigated by examining the rate of genetic change in protein yield in several countries. Current trends in effective population size (Ne) were studied by examining the number of sires of sons used each year and the percentage in common use across several countries using Interbull test results. The influence of rates of inbreeding in several Holstein populations and the genetic correlation (r_g) among countries on Ne was examined. The results were compared with those for the Jersey breed with a less globalized breeding programme. Substantial rates of genetic progress of about 2% per year in milk traits have been observed in most Holstein Friesian populations in the last 5 years. This however has been accompanied by a decreasing trend in effective population size. The number of bull sires used is decreasing with time while the percentage in common usage is increasing. The average relationship among bulls, rates of inbreeding and percentage of inbred cows all show an increase with years. International evaluations by the Interbull Centre account for r_g among countries and therefore offers opportunities to select from a wide range of top bulls. This, in addition to national selection objectives which include other traits related to profit in addition to milk traits, should ensure different countries sample different bulls for progeny testing and hence increase Ne. The use of selection using indices with restrictions on family information which substantially reduce levels of inbreeding while maintaining present rate of genetic gain should be exploited. The current rates of inbreeding could increase substantially due to advances in reproductive technology. The adoption of special breeding schemes such as factorial mating designs, equalizing family sizes and deliberate choice of less closely related animals would become necessary to reduce the rate of inbreeding to conserve available genetic diversity.

Introduction

Selection based on progeny testing remains the major tool for genetic improvement in dairy cattle. Traditionally improvement has resulted from selection applied on the usual four pathways: bulls to breed bulls (BB), bulls to breed cows (BC), cows to breed bulls (CB) and cows to breed cows (CC).

Prior to 1970, this was mainly within the several strains of Holstein and

Friesian (black and white) dairy cattle existing in different countries around the world. Comparisons of the various strains of the black and white (B&W) cattle carried out in the mid-seventies indicated the superiority of the North American Holstein for milk production (Jasiorowski *et al.*, 1988). This stimulated large scale importation of Holsteins by many countries resulting in the replacement of most strains of European Friesian dairy cattle by the North American Holsteins. In the 1980's, due to improvements in reproductive technology, this process was accelerated. This resulted in bulls with similar parentage being progeny tested simultaneously in several countries and began the creation of a single global population for B&W dairy cattle. This globalization process was accompanied by substantial rates of genetic progress due to the genetic lift in importing countries from the between strain selection, opportunities for higher intensity of selection provided by a large global population and a certain amount of heterosis from crossing the North American Holstein with the local Friesian genotypes.

The major disadvantage of the emergence of the global breeding structure of B&W dairy cattle has been a reduction in Ne resulting in the loss of genetic diversity. In addition, as a result of several changes in breeding practices in the last decade such as the availability of more accurate genetic evaluations through improved statistical procedures, an even higher intensity of selection is possible especially in the BB pathway. This could result in the use of fewer top sires world wide thereby increasing the rate of inbreeding and further reducing Ne. Further globalization of dairy cattle breeding programmes is intensifying due to international progeny testing and across border mergers of major companies involved in progeny testing bulls. It is therefore important that current trends in rates of inbreeding and Ne are examined in the B&W dairy cattle population and the implications in terms of conserving genetic diversity assessed. Where appropriate, results are compared with those from the Jersey breed which is a numerically smaller population and with a less globalized breeding programme.

Initially, this paper examines the rates of genetic progress in several countries that have accompanied the globalization of the B&W and Jersey breeds. Subsequently, the consequences of genetic improvement in terms of rates of inbreeding and Ne and the implications in terms of conservation of genetic diversity are discussed.

Genetic Improvement Scheme in dairy cattle

Currently genetic improvement in dairy cattle in most developed countries is based on progeny testing (PT), see Figure 1. While initially it was thought that nucleus breeding schemes utilizing Multiple Ovulation and Embryo Transfer (MOET) with selection based on parent and sib information (Nicholas and Smith, 1983) would replace progeny testing schemes, this has not been so in practice. The potential advantages of the nucleus schemes in terms of higher rate of genetic gain were greatly reduced after the reduction in genetic variance due to the Bulmer effect and levels of inbreeding were taken into account (Dekkers, 1992). Generally these schemes have operated in conjunction with conventional PT in countries where they exist rather than replace it. In a national breeding programme based on PT and selection in the four traditional pathways, the rate of response per year (ΔG) based on the expression of Rendel and Robertson, (1950) is:

$$\Delta G = \frac{I_{BB} + I_{BC} + I_{CB} + I_{CC}}{L_{BB} + L_{BC} + L_{CB} + L_{CC}}$$

Where the L_{jk} is the mean generation interval for the pathway jk and I_{jk}, the selection index for the jk^{th} pathway is:

$$I_{jk} = r_{\hat{G}G}\ i\ \sigma_G$$

With $r_{\hat{G}G}$ being the correlation between the predicted and true genetic value (the accuracy of selection), i is the standardized selection differential and σ_G is the genetic standard deviation. These are the major factors that determine the rate of genetic progress.

Schmidt and Van Vleck (1974) indicated that the proportions of genetic gain from the four pathways are approximately: I_{CC} 2%, I_{BC} 27%, I_{CB} 32% and I_{BB} 39%. These values indicate that most of the genetic improvement (about 71%, (BB + CB)) would be due to matings to produce young bulls for testing in the next generation. This proportion is likely to be higher with the availability of more accurate estimates of breeding values (EBV) from Best Linear Unbiased Prediction (BLUP) as the tendency in most PT schemes is to use sires with the highest EBV to breed bulls (BB) in the next generation. In addition selection of cows on the basis of EBV would result in most dams (CB) selected to breed young bulls being daughters and granddaughters of the elite bulls (BB) of the previous generation. Consequently as Goddard and Smith (1990) indicated, the bull breeding part of the population is like a dispersed nucleus herd with the same

bulls used in the BB and CB paths in the nucleus section. The system therefore effectively reduces to a two-path system rather than the conventional four path system. While this could result in substantial rates of genetic progress, it could also markedly increase the rate of inbreeding and hence reduce Ne.

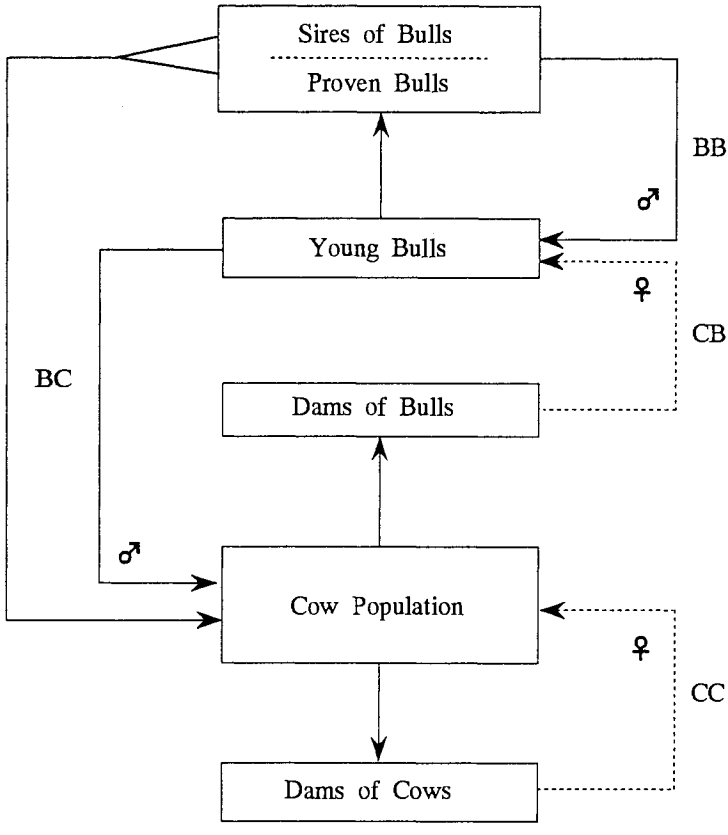


Figure 1. Structure of conventional progeny-testing shcemes, identifing the four paths of selection: sires of bulls (BB), sires of cows (BC), dams of bulls (CB) and dams of cows (CC).

Current Rates of Genetic improvement

Theoretically PT schemes can result in rates of genetic improvement of 2% to 3% per year (Van Vleck, 1981). Presented in Tables 1 and 2 are current rates of genetic progress for protein yield for Holstein bulls and cows in some countries. These rates have been expressed where possible as a percentage of the mean protein yield of cows born in 1990.

Table 1. Genetic trends for Holstein Friesian bulls for protein yield in several countries

Country	Last 10 year period	Kg / year	% of mean*	Last 5 year Period	Kg/ year	% of mean
Canada	1982-92	3.9	1.4	1988-92	4.8	1.7
Holland	1980-92	3.6	1.6	1988-92	4.5	2.0
UK	1981-90	2.7	1.5	1986-90	3.4	1.9
France	1981-91	4.4	1.5	1986-91	4.8	1.8
Italy	1983-92	5.1	1.9	1988-92	5.1	1.9

*Mean of protein yield of cows born in 1990.

Table 2. Genetic trends for Holstein Friesian cows for protein yield in several countries

Country	Last 10 year period	Kg/ year	% of mean*	Last 5 year Period	Kg/ year	% of mean
Canada	1984-94	3.8	1.3	1989-94	4.3	1.5
USA	1986-95	3.8	1.3	1991-95	3.7	1.3
Holland	1984-94	4.1	1.9	1990-94	4.1	1.9
UK	1984-94	2.3	1.3	1990-94	3.8	2.1
France	1988-93	3.6	1.3	1988-93	4.8	1.8
Italy	1986-95	5.02	1.9	1991-95	5.2	1.9

*Mean of protein yield of cows born in 1990.

In percentage terms the rate of genetic progress is similar across countries ranging from 1.5% to 2.5% for bulls and 1.3% to 2.0% for cows over the last 10 year period. The similarity in the rate of change is further confirmed by Figure 2, which indicates the genetic trends for bulls for several countries relative to the United Kingdom (UK) base from Interbull results in February, 1998. In the Jersey breed, the rate of genetic gain for protein yield varies from 0.7% to 1.5%/year for cows in the last 10 years and about 1.8% per year in Canada and the UK in the last 5 years. In most countries studied, the rates of change over the last 5 years are higher than for the last 10 year period in both the B&W and Jersey populations. This is indicative of a higher intensity of selection in recent years and the increasing influence of animals of higher genetic merit from North America in countries such as the UK with a high rate of importation.

These results indicate that the annual change in genetic progress in milk traits in the global Holstein population and Jersey populations are going up at an increasing rate. What then are the implications of such change in terms of effective population size or rates of inbreeding? These are examined subsequently in this paper.

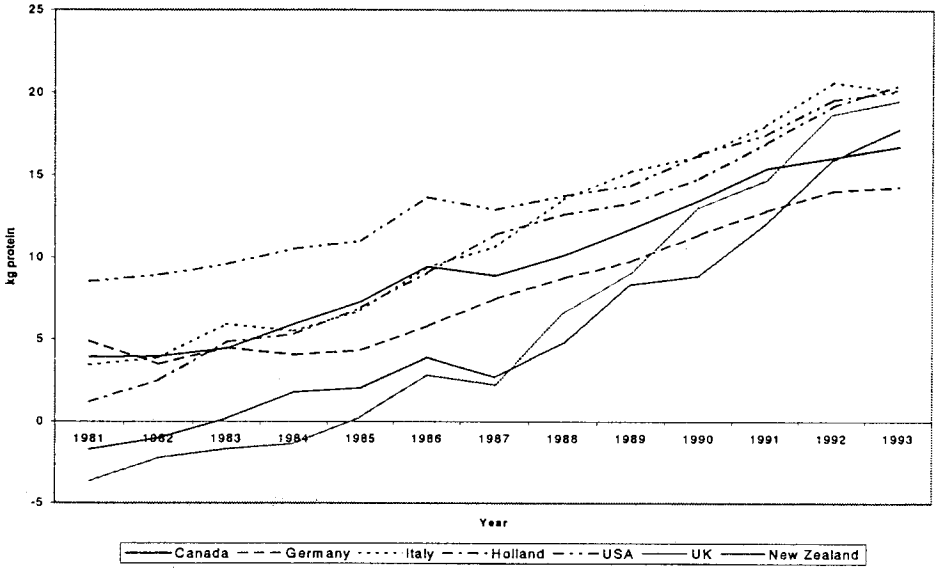


Figure 2. The genetic trend for protein yield for some countries from Interbull evaluations on UK base

The consequence of selection on effective population size

One of the consequences of directional selection is a reduction in the genetic variance and increase in the rate of inbreeding as families superior for the selected trait contribute more offspring to the next generation. Any increase in the rate of inbreeding implies a reduction in Ne, as both are inversely related; $\Delta F = 1/(2Ne)$. In the past few decades, most dairy cattle populations have been subjected to intensive selection for yields of milk, fat and protein. However, in the past decade or so, as a result of several changes such as the availability of more accurate genetic evaluations, global sire rankings, increasing international trade in semen and embryos and the use of new reproductive techniques, even higher intensities of selection, especially in the BB and CB pathways, have become possible. These factors are expected to increase the rate of inbreeding and hence reduce Ne. For instance, the use of BLUP animal model evaluations to select sires and bull dams increases the probability of co-selection of close relations (Belonsky and Kennedy 1988) and hence and increase in inbreeding. Increased international trade in stock, semen and embryos may imply the use of fewer top sires world wide. In a world wide study Leitch (1994) showed that the 10 most popular sires and sons for Holstein Friesians represented 62% of the young bulls sampled in each country, and the 10 most popular maternal grandsons represented 63% of sampled sons. Moreover, the use of reproductive techniques such as artificial insemination and embryo transfer, imply that superior individuals can be exploited more intensively in the breeding programme

In this section, current trends in effective population size and genetic variance in various dairy cattle populations are examined and the implications discussed.

Trends in effective population size and the influence of various factors

In a theoretical study, over a wide range of numbers of bulls tested, Goddard (1992) reported an optimum number of about 6 bulls to be selected to maximize net present value (future genetic gains discounted to an equivalent present value) in the Holstein global population. To understand current trends in effective population size in the Holstein population, the number of sires of sons used each year in several countries was examined based on Interbull results. Figure 3 shows the percentage of bulls with progeny test results in various countries which are sons of the 5 most used sires over the period 1981 to 1991. In general, this percentage is increasing in a number of countries implying that the number of sires of sons used each year is decreasing.

Considering the same trend in 18 countries, Wickham and Banos (1998) indicated that the 5 most heavily used sires accounted for 50% of all bulls born in 1990 which have been progeny tested compared with 14% in 1970.

The percentage of bulls by the 5 sires with most sons each year, summed over five countries (Canada, Denmark, USA, New Zealand and Australia) for the Jersey breed is shown in Figure 4. A similar trend for the Holstein summed over the seven countries represented in Figure 3 is included for comparison. While the percentage is increasing slightly with time, the rate of increase is lower in the Jersey breed than in the Holstein, over the same time period.

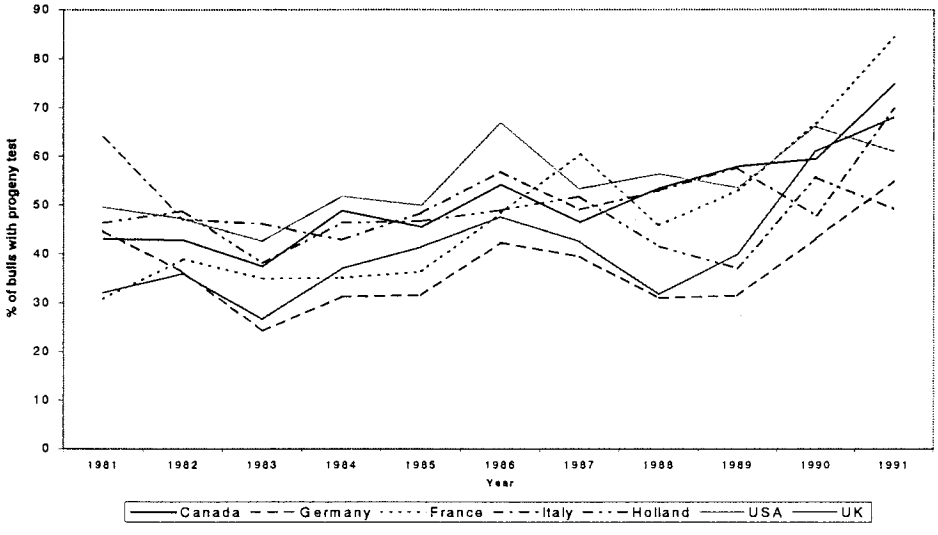


Figure 3. The percentage of bulls with progeny test results which are sons of the 5 most used sires in some countries

Figure 4. Percentage of bulls with progeny test which are sons of the 5 top bulls with most sons

Not only has the percentage of progeny tested sons of the top 5 sires increased with time, the percentage of these top 5 sires in common use across several countries has also increased. Tables 3 and 4 show the percentages of the top 5 most used sires common among several countries over two five year time periods (1981-1985 and 1988-1992). The percentages for the top 10 most used sires are also given in these Tables. The results indicate a higher proportion of the same bulls are being used in the bull breeding section of the progeny testing programmes of several countries between 1988 – 1992 when compared with 1981 – 1985. For instance of the top 5 bulls with most sons between 1988 –1992, three of the bulls were the same in France and Holland and 2 were the same in the USA, Germany, the UK and Canada. Considering the top 10 sires over the 2 time periods indicates an even higher number of bulls were in common use in several countries. A similar analysis for the Jersey population based on 5 countries in Figure 4, also indicated an increase in the proportion of bulls in common use. However, these proportions were not as high as in the Holstein, with at most 2 bulls in common use between any two of these countries when the top 5 sires with most sons were considered. Again the proportion of bulls in common use were higher when the top 10 sires with most sons were considered, with Canada, USA and Australia using the same 5 sires.

Table 3. Percentage of the 5 bulls with most sons in the period 1981-1985 which are common among several countries

	CANADA	GERMANY	FRANCE	ITALY	HOLLAND	USA	UK
CANADA	100						
GERMANY	20 (10)	100					
FRANCE	0 (0)	0 (0)	100				
ITALY	40 (30)	20 (10)	20 (30)	100			
HOLLAND	0 (0)	0 (0)	20 (20)	0 (10)	100		
USA	20 (30)	20 (10)	0 (0)	40 (20)	0 (10)	100	
UK	0 (0)	20 (10)	0 (0)	0 (0)	0 (0)	0 (0)	100

Percentages in brackets are those when the 10 bulls with the most sons are considered.

Table 4. The percentage of the 5 bulls with most sons in the period 1988-1992 which are common among several countries

	CANADA	GERMANY	FRANCE	ITALY	HOLLAND	USA	UK
CANADA	100						
GERMANY	0 (20)	100					
FRANCE	40 (30)	0 (10)	100				
ITALY	0 (30)	0 (10)	20 (50)	100			
HOLLAND	20 (30)	0 (0)	60 (50)	20 (50)	100		
USA	20 (30)	40 (20)	20 (30)	0 (30)	20 (30)	100	
BRITAIN	40 (50)	20 (20)	40 (40)	40 (40)	40 (40)	40 (30)	100

Percentages in brackets are those when the 10 bulls with the most sons are considered.

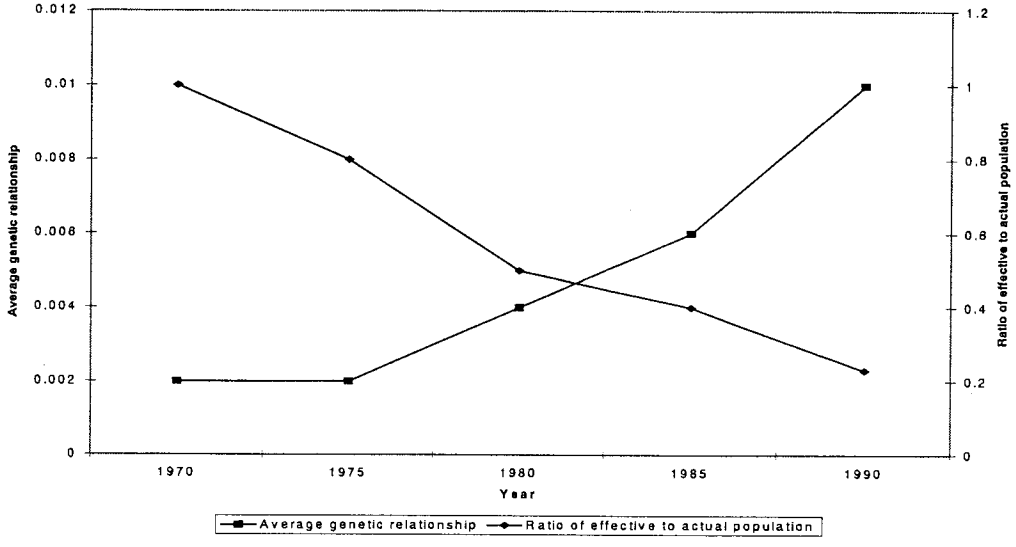


Figure 5. Trends in genetic relationship among sires for 18 Holstein Friesian populations (Wickham and Banos, 1998)

These results attest to a declining Ne of both the world wide B&W and Jersey populations. Wickham and Banos (1998) examined the actual trend in Ne and average additive genetic relationship among bulls in 18 countries in the Interbull evaluations and their results are presented in Figure 5. The effective population size of bulls each year was calculated as a function (harmonic mean) of the number of sires of sons and dams of sons. Similar results for UK B&W data only is given in Figure 6.

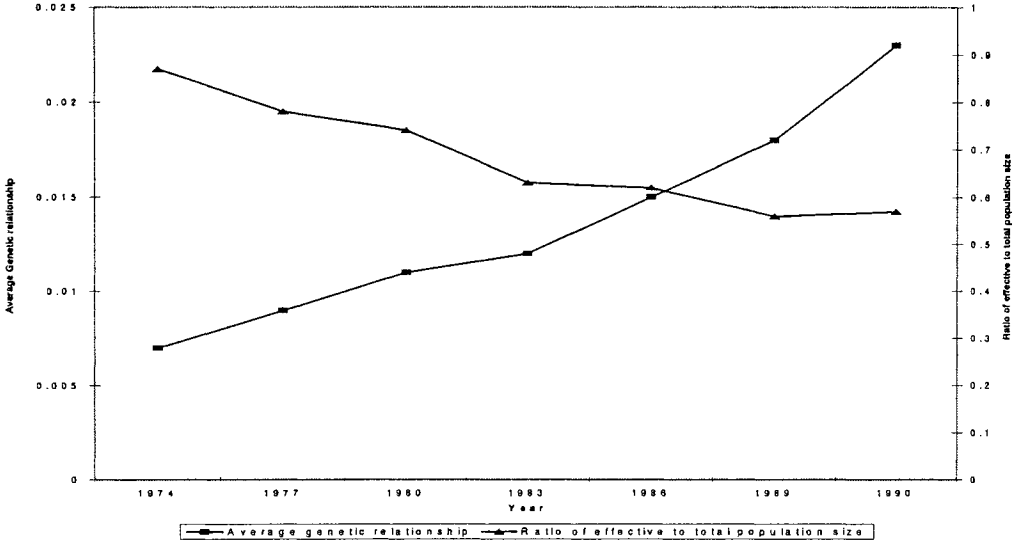


Figure 6. Trends in genetic relationship among sires for UK Holstein Friesians

Considering bulls born in 1990 over the 18 countries (Figure 5), Ne was only 22% of the actual population size compared with 100% for bulls born in 1970. At the same time the average genetic relationship of sires of sons born in 1990 was six times that of sons born in 1970. The results from the UK data followed the same trend with the average relationship of bulls in 1990, three time those born in 1974. A similar result was obtained for the Jersey breed in the UK. These results confirm a decreasing trend in Ne with time in both the B&W and Jersey populations.

Estimate of total sires to breed bulls

Considering the top 5 or 10 sires with the most sons in the seven countries represented in Figure 2, the total number of different bulls used to breed bulls in the B&W breed across these countries is given in Table 5. Similar results for the Jersey breed are also represented in the table. The 10 bulls with most sons in these countries accounted for 65% to 93% of all bulls with progeny test results in the Interbull test analysis, and therefore gives a good approximation of the total of different bulls used to breed sons in these countries. In general the total number of bulls decreased with time in both breeds with 28 sires used to breed bulls in 1991 in each breed in these countries. This number is substantially higher than the optimum recommended by Goddard (1992) and Goddard and Smith (1990) for the B&W population. Thus while total Ne is generally decreasing with time, it seems the current level is higher than the recommended optimum. However, this ignores the average relationship and therefore levels of inbreeding associated with the sires used in the several countries.

Table 5. Total number of Holstein Friesian sires used to breed bulls in seven countries* based on the top 5 or 10 sires with the most sons

	Total number of sires used to breed bulls			
	Holstein Friesian Breed		Jersey Breed	
Year	Based on the top 5 bulls	Based on the top 10 bulls	Based on the top 5 bulls	Based on the top 10 bulls
1981	22	41	27	45
1983	19	37	22	43
1985	22	42	21	39
1987	15	36	23	42
1989	12	31	23	37
1991	10	28	16	28
1992	15	29	18	26

*Countries included were Canada, Germany, France, Italy, Holland, USA and UK for the Holstein Friesian breed and Canada, Denmark, USA, New Zealand and Australia for the Jersey breed.

Influence of genetic correlations between countries on effective population size

The genetic correlation between countries arising from differences in selection objectives and genotype by environment interaction (G×E) for individual milk traits determine the extent to which they select the same sires of sons and hence affect global Ne. In a simulation study Goddard (1992) examined the influence of r_g on a number of bulls selected if 2000 bulls were tested in more than two countries under two situations: In the first situation, there were differences in the selection objectives of the countries with a correlation of 0.9 between each pair of countries and each country was assumed to have selected the best two bulls in terms of their selection objective as bull sires. In this situation, the total number of bulls selected continued to increase as the number of countries increased to at least 10, with a total of about 8 bulls selected. In the second situation, selection objectives were the same but there was G×E in individual traits and again with each country selecting the best two local bulls as bull sires. The total number of bulls selected increased as the number of countries increased but the number of bulls selected reached a maximum of 5.

In general accounting for r_g ($r_g < 1$) due to G×E and/or differences in selection objectives between countries increased the total number of bulls to be selected and hence increased global Ne. In practice, with respect to the first situation, Leitch (1994) carried out a comparison of selection indices for dairy cattle in ten countries. The largest differences observed between countries were in terms of the relative emphasis on production and type and relative emphasis on milk yield (positive or negative). Also 5 of the 10 countries included non-productive traits such as mastitis resistance, somatic cell count, milking speed and temperament in their indices. Leitch also reported rank correlations of 0.90 – 0.92 between Canadian LPI and ILQM (Italy), ISU (France) and PIN II (now ITEM, UK) and 1.00 with US TPI. However, considering only the production component of LPI, she observed rank correlations of 0.88 to 0.92 with the production index in the UK, Netherlands, Italy and France, but a correlation of unity with several countries: USA, New Zealand, Germany and Israel. In spite of these high correlations between the various selection criteria she observed a substantial variation in the sires and maternal grandsires of sons progeny tested in 1993 in the countries she studied. Thus while it seems that the present indices ensure variation in bulls sampled by different countries, differences between selection indices in different countries will become more obvious as countries include other traits relating to profitability in their indices rather than concentrate on milk traits. Meuwissen (1998) indicated that genetic

correlations of below 0.7 – 0.8 between different selection objectives are needed, if these objectives are to give rise to breeding schemes that are substantially different.

Considering the second situation of Goddard (1992), Interbull Centre estimates genetic correlations among countries using sire proofs, which are then utilized by the Centre in multi-trait across country evaluation (MACE). These correlations are a reflection of G×E among countries as well as differences in methods for national evaluations. Presented in Table 6 are the genetic correlations among seven countries from protein yield from Interbull results.

Table 6. Interbull Genetic Correlations* among countries for protein yield

	CANADA	GERMANY	FRANCE	ITALY	HOLLAND	USA	UK
CANADA	1.00						
GERMANY	0.91	1.00					
FRANCE	0.95	0.92	1.00				
ITALY	0.93	0.90	0.95	1.00			
HOLLAND	0.93	0.93	0.93	0.91	1.00		
USA	0.96	0.89	0.95	0.95	0.93	1.00	
UK	0.93	0.90	0.93	0.91	0.93	0.91	1.00

*From Interbull run in February 1998

These correlations vary from 0.89 between USA and Germany to 0.96 between USA and Canada. These less than unity correlations imply bull rankings in each country are different and should increase the total number of bulls available for selection and hence increase Ne. The effect of these correlations on the total number of bulls available for selection and the percentage average relationship among these bulls considering the top bulls on protein yield in each of the seven countries in the previous section, is presented in Table 7. Thus if selection for global protein yield is based on the top 5 bulls on protein yield in all seven countries, a total of 11 bulls will be available for selection, that is, 120% more than selecting only from the top 5 of the national list of an exporter. If the top 50 bulls are considered, 93 bulls will be available for selection, 86% more than selecting from the top 50 of any national list. Considering 18 countries in the Interbull evaluation, Wickham and Banos (1998) reported that, selection for protein yield based on a "top 50" list of all these countries would lead to a total of 128 bulls being available for selection, which is 150% more than selection based on the top 50 of any national list.

Table 7. Total number of bulls available for selection based on the top 5 bulls for protein yield in seven countries*

No. of Top Bulls	No. of bulls available for selection	% Average genetic relationship
5	11	3.3
10	24	3.4
20	46	4.4
30	63	4.4
40	82	4.5
50	93	4.9

*Countries included were Canada, Germany, France, Italy, Holland, USA and UK

Although the percentage average relationship among the available bulls for selection increases as more top bulls are considered (Table 7), in all cases it was less than 6.25%, the additive genetic relationship between a sire and his progeny in the fourth generation. It seems that international evaluation by accounting for genetic correlations among countries results in more bulls being available for selection on a global scale. This should slow down the rate of decrease in global diversity.

Influence of inbreeding depression

In a simulation study Goddard (1992) concluded that inbreeding depression (D) largely determines the optimal Ne for the global Holstein population. Inbreeding depression results in a reduction of milk production. Table 8 presents estimates of inbreeding depression for milk traits in some dairy breeds in the UK (Mrode and Swanson, 1998), USA (Wiggans *et al.,* 1995) and Canada (Miglior *et al.,* 1992). The estimates for the Holstein breed for milk yield vary from –13.1kg in the UK to -29.6kg in the USA per 1% of inbreeding. Expressed as a percentage of the mean, these represent a reduction of 0.25% to 0.31% of the mean milk production of cows born in 1990. The estimates for the Jersey breed are similar to the Holstein when expressed in percentage terms.

Table 8. Estimates of inbreeding depression for milk traits by per 1% inbreeding in Holstein Friesians and Jerseys

	Holstein Friesian		Jersey		
Trait	UK	USA	UK	USA	CANADA
Milk (kg)	-13.1	-29.6	-6.9	-21.3	-9.8
Fat (kg)	-0.54	-1.08	-0.64	-1.03	-0.55
Protein (kg)	-0.42	-0.97	-0.28	-0.88	-

However, inbreeding depression not only reduces milk production, it also greatly affects fertility, survival and general hardiness. Studies evaluating the influence of D on Ne have mostly accounted for its effect through reduction of production and may therefore have under-estimated its effects. Over a wide range of bulls tested (50-

1000) Goddard and Smith (1990) observed that the optimum number of bulls to maximize net response ($\Delta G - D\Delta F$) was about 4-7 bulls assuming 0.25% depression in economic merit per % inbreeding. With 10 sires used per generation (5 years), they estimated the level of inbreeding to give the optimum response to be 0.25% per year. In general they concluded that 8-12 bulls per generation gave more than 95% of the maximum response except in extreme cases where very low numbers of bulls were tested.

How do current estimates of inbreeding compare with the levels for the optimum scheme of Goddard and Smith (1990)? Table 9 shows current rates of inbreeding for several Holstein Friesian populations.

Table 9. Estimates of rates of inbreeding in some Holstein Friesian populations

Country	Rate of Inbreeding (%/year)	Reference
Canada	0.21 (bulls only)	Miglior and Burnside (1995)
USA	0.20	Wiggans et al (1995)
France	0.20*	Maignel et al (1996)
Netherlands	0.10	Te Braake et al (1994)
Italy	0.10 (bulls only)	Miglior (personal communication)
UK	0.02	Mrode and Swanson (1998)

*Assuming a generation interval of 5.5 years

Most of the estimates are between 0.1% to 0.2% per year and are within the levels for the maximum response considering 10 bulls per generation. However, these studies are retrospective and future rates of inbreeding are expected to increase. For instance, considering all the dairy breeds in the USA, Wiggans et al (1995) observed that the rate of change in annual increase in inbreeding was the greatest in the Holstein in spite of the large population size. In the UK, the proportion of bulls and cows with inbreeding levels greater than zero increased from 45% and 66% respectively for animals born in the period 1971 to 1975 to 81% and 90% for animals born in the period

1986-1990. Similarly in the Netherlands the percentage of cows which were inbred rose from 33% for cows born in 1980 to 98% for cows born in 1996 (Gerben DeJong, personal communication). These imply that most subsequent matings are likely to increase the rate of inbreeding. This current increasing trend in the percentage of animals which are inbred is partly due to selection of sires and bull dams on the basis of estimated breeding value. Thus the same bulls are being used in the BB and CB pathways as explained earlier.

The rates of inbreeding observed in the less numerical breeds such as Jerseys and Ayrshires tend to be slightly higher than those for the Holstein Friesian breed (Wiggans *et al.,* 1995, Mrode and Swanson, 1998). In the UK data (Mrode and Swanson, 1998) higher levels of inbreeding in these breeds was due mainly to an increase in the proportion of animals with higher levels of inbreeding rather than a rapid increase in total number of animals inbred. The increase in the proportion of animals with higher levels of inbreeding is due to small population size and fewer numbers of bulls tested each year in these breeds.

In addition, further increases in annual rates of inbreeding are expected in the near future as a result of advances in reproductive physiology. Cloning and in vitro embryo production (IVEP) could result in breeding programmes with intense female selection and shorter generation intervals. For instance, with IVEP it could be possible that large numbers of embryos could be sold into the commercial population, creating a large dispersed nucleus in which both adult and juvenile MOET schemes can be exploited. In this situation, Goddard (1992) indicated that breeding programmes based on maximizing net genetic gain would lead to annual rates of inbreeding of 1% per year, with a resultant drastic reduction in Ne.

Trend in genetic variance

One of the major consequences of selection is a reduction in genetic variance. In spite of the rapid rate of genetic progress in milk traits observed in the Holstein Friesian population, most recent estimates of heritabilities have risen to 0.40 in the last decade from the traditional values of 0.25 to 0.30 in several populations (Visscher and Thompson, 1992). In European countries such an increase in genetic variance could partly be attributed to the substantial immigration of Holsteins. However, heritability estimates in the USA Holsteins have either not changed or show a slight increase (Van Vleck *et al.,* 1988; Dimov *et al.,* 1995) but definitely not a decline in genetic variance

(Va). Thus the increases in estimates of Va have not been consistent with expectation as a result of selection and reduced population size. This could possibly be attributed to better management, improved recording, better statistical methods for data analysis and the effects of mutation.

New Va attributable to mutation (Vm) increases at a rate of the order of 10^{-3} times the environmental variance (Ve) per generation in a variety of characters (Falconer and Mackay, 1996). As Hill (1989), indicated, if values of $Vm = 10^{-3}Ve$ are realistic, then mutation can provide useful variation in long term directionally selected populations, provided a proportion of mutants increase the trait.

Discussion

The results reveal a decreasing trend in Ne in the global Holstein population and also in the less numerical breeds such as the Jersey. Not only is the number of bull sires used decreasing with time, the percentage in common use among several countries is on the increase. Indications are that the current downward trend in Ne will continue given the very high percentage of inbred cows in many Holstein populations. The consequence of this is obviously a continual decline in genetic diversity in the Holstein population, therefore ways of arresting this decline in Ne and hence conserving the available genetic diversity must be examined.

Firstly, as Wickham and Banos (1998) remarked, the very high usage of popular sires of sons in several countries, identifies an opportunity for co-operation between bull studs to test sons from a wider range of sires and hence maintain a large effective population size to protect long term genetic gain.

Selection on the basis of pedigree information and BLUP has become the standard practice in most countries. This will inevitably push up the rate of inbreeding and hence reduce Ne as it increases the probability of co-selection of close relatives. Several simulation studies have examined the use of selection indices constructed from BLUP evaluations in which the emphasis given to ancestral information is reduced (Verrier *et al.*, 1993; Grundy *et al.*, 1998). One of the simple indices studied by Grundy et al (1998) involved the use of artificially inflated heritability by a factor of 2 or 3 in BLUP evaluations to restrict family information. They reported that after 25 generations of selection, cumulative response was about 3% less than standard BLUP but the final level of inbreeding was 23% to 35% lower than from BLUP. In general, they observed that the magnitude in reduction in the rate of inbreeding compared to

BLUP selection was inversely proportional to the amount of family information included in the modified selection indices. The possibility of utilizing such modified indices as national selection indices should be examined in practise in order to conserve the available genetic diversity. The substantial influence of pedigree information on International proofs by the Interbull Centre, especially in countries where foreign bulls have little or no information, has been demonstrated by Mrode and Swanson (1997). The adoption of indices with restrictions on family information would therefore also be beneficial even with international evaluations.

One of the factors that influences Ne is the genetic correlation among countries, resulting from differences in selection objectives and G×E. This determines to what extent different countries select the same sires of sons. To ensure different bulls are sampled by different countries, selection objectives should be defined to include other traits that relate to profit such as health and reproduction rather than concentrate on milk traits. More comprehensive indices may result in lower genetic correlations among the selection objectives of different countries than the present estimates of 0.88 – 0.92. Genetic correlations of lower than 0.7 – 0.8 are needed if these different selection objectives are to result in breeding schemes that are substantially different (Meuwissen, 1998). Also the inclusion of more functional traits in selection objectives should ensure genetic variability in any one trait is not significantly reduced.

In addition, the results have demonstrated the substantial increase in the number of bulls available for selection following the inclusion of genetic correlations among countries in International evaluations by the Interbull Centre. For the less numerical breeds with few bulls tested each year, the effects of international evaluations are even more pronounced, offering greater opportunities to select from a wider range of bulls and also increase selection differential. This should avoid polarized usage of a few "trendy" bulls and hence conserve global genetic diversity. The effect of this will be maximized as all countries adopt Interbull evaluations and more appropriate data and methodology are utilized to improve on the current estimate of these genetic correlations.

Improvements in current advances in reproductive technology are probable with the implication of higher intensity of selection in females and shorter generation interval. With IVEP, the application of MOET type schemes could be quite pronounced in the commercial populations. Under such a situation rates of inbreeding, as Goddard (1992) indicated, could be as high as 1% per year and the adoption of appropriate

breeding schemes such as the factorial mating design (Woolliams, 1989), equalizing family sizes, constraining the rate of inbreeding (Meuwissen, 1998) and deliberately choosing less closely related animals, would become necessary to substantially reduce the rate of inbreeding with little effect on genetic gain.

Conclusion

The results have indicated substantial rates of progress of about 2% in milk traits in most B&W populations in the last 5 years. This however has been accompanied by a decreasing trend in Ne. The number of bull sires used is decreasing with time while the percentage in common usage is increasing. The average relationship among bulls, rates of inbreeding and percentage of inbred cows all show an increase with years.

Firstly, the very high usage of popular sires of sons in several countries identifies the opportunity for co-operation between bull studs to test sons from a wider range of sires and hence maintain a large Ne. International evaluations by the Interbull Centre account for G×E and therefore offers opportunities to select from a wide range of top bulls. This, in addition to comprehensive national selection objectives which include other traits which relate to profit in addition to milk traits, should ensure different countries sample different bulls when progeny testing and hence increase Ne. The effectiveness of selection using indices with restrictions on family information in maintaining the present rate of genetic gain while reducing the level of inbreeding have been demonstrated. The possibility of utilizing such modified indices in national selection schemes should be exploited. Current rates of inbreeding could increase substantially due to advances in reproductive technology. Special breeding schemes such as factorial mating designs and equalizing family sizes may have to be adopted to reduce rates of inbreeding to conserve available genetic diversity.

References

Belonsky, G. M. and Kennedy, B. W. 1988. Selection on individual phenotype and best linear unbiased predictor of breeding value in a closed swine herd. Journal of Animal Science 66: 1124 –1131.

Dekkers, J. C. M. 1992. Structure of breeding programs to capitalize on reproductive technology for genetic improvement. Journal of Dairy Science 75: 2880-2891.

Dimov, G., Albuquerque, L. G., Keown, J. F., VanVleck, L. D. and Norman, H. D. 1995. Variance of interaction effects of sire and herd for yield traits of Holsteins in California, New York and

Pennsylvania with an animal model. Journal of Dairy Science 78: 939-946.

Falconer, D. S. and Mackay, T. F. C. 1996. Quantitative genetics. Longman, Essex, England.

Goddard, M. G. 1992. Optimal effective population size for the global population of black and white dairy cattle. Journal of Dairy Science 75: 2902-2911.

Goddard, M. G. and Smith, C. 1990. Optimum number of bull sires in dairy cattle breeding. Journal of Dairy Science 73: 1113-1122.

Grundy, B., Luo, Z. W., Villanueva, B. and Woolliams, J.A. 1998. The use of Mendelian indices to reduce the rate of inbreeding in selection programmes. Journal of Animal Breeding and Genetics 115: 39-51.

Hill, W. G. 1989. Mutation and maintenance of quantitative genetic variation. Evolution and Animal Breeding, edited by W. G. Hill and T. F. C. Mackay. CAB International, Wallingford, Oxon, UK.

Leitch, H. 1994. Comparison of international selection indices for dairy cattle breeding. Proceedings of the open session of the Interbull annual meeting, Ottawa Canada. Bulletin No 10 International Bull Evaluation Service, Uppsala, Sweden.

Maignel, L., Boichard, D. and Verrier, E. 1996. Genetic variability of French dairy breeds estimated from pedigree information. Proceedings of the open session of Interbull meeting, Veldhoven, the Netherlands. Bulletin No 14: 49-54.

Meuwissen, T. H. E. 1998. Risk management and the definition of breeding objectives. Proceedings of the 6th World Congress on Genetics applied to Livestock Production, Armidale, Australia. 25: 347-354.

Miglior, F., Szkotnicki, B., and Burnside, E. B. 1992. Analysis of levels of inbreeding and inbreeding depression in Jersey cattle. Journal of Dairy Science 75: 1112-1118.

Miglior, F. and Burnside, E. B. 1995. Inbreeding of Canadian Holstein cattle. Journal of Dairy Science 78: 1163-1167.

Mrode, R. A. and Swanson, G. J. T. 1997. Deriving simple equations for evaluations of bulls in the Interbull MACE system. Proceedings of the 1997 Interbull meeting, Vienna, Austria, Bulletin No 16: 8-12.

Mrode, R. and Swanson, G. J. T. 1998. Calculation of inbreeding coefficients and accounting for inbreeding in genetic coefficients. Paper in preparation.

Nicholas, F. W. and Smith, C. 1983. Increased rates of genetic change in dairy cattle by embryo transfer and splitting. Animal Production 36: 341-353.

Rendel, J. M. and Robertson, A. 1950. The use of progeny testing with artificial insemination in dairy cattle. Journal of Genetics 50: 21-31.

Schmidt, G. H. and VanVleck, L. D. 1974. Principles of Dairy Science. San Francisco, W. H. Freeman and Co.

Te Braake, M .F. H., Groen, A. F. and Van Der Lugt, A. W. 1994. Trends in inbreeding in Dutch black and white dairy cattle. Journal of Animal Breeding and Genetics 111: 356-366.

VanVleck, L. D. 1981. Potential genetic impact of artificial insemination, sex selection, embryo

transfer, cloning and selfing in dairy cattle. New Technologies in Animal Breeding, edited by Brackett, B. G., Seidel, G. E. and Seidel, S. M. Academic Press, London, England.

VanVleck, L. D., Dong, M. C. and Wiggans, G. R. 1988. Genetic (co)variances for milk and fat yields in California, New York and Wisconsin by an animal model by restricted maximum likelihood. Journal of Dairy Science 71: 3053-3060.

Verrier, E., Colleau, J. J. and Foulley, J. L. 1993. Long term effects of selection based on animal model BLUP in a finite population. Theoretical and Applied Genetics 87: 446-454.

Visscher, P. M. and Thompson, R. 1992. Univariate and multivariate parameter estimates for milk production traits using an animal model. 1. Description and results of REML analysis, Genetics, Selection, Evolution 24: 415-430.

Wickham, B. W. and Banos, G. 1998. Impact of international evaluations on dairy cattle breeding. Proceedings of the 6th World Congress on Genetics applied to Livestock Production, Armidale, Australia, 23: 315-320.

Wiggans, G. R., Van Raden, R. M. and Zuurbier, J. 1995. Calculation and use of inbreeding coefficients for genetic evaluation of United Sates Dairy Cattle. Journal of Dairy Science 78: 1584-1590.

Woolliam, J. A. 1989. Modifications to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. Animal Production 49: 1-14.

Genetic Improvement and Inbreeding Problem in Swine

HIROSHI TAKAHASHI
Global Pig Farms, Inc. Gunma 377-0052, Japan

Abstract

The effects of inbreeding on age in days at 105 kg (D105) of body weight and average backfat thickness (ABF) were estimated using a single trait animal model which included a partial linear regression coefficient of D105 or ABF on the inbreeding coefficient (F) of the pigs. Performance data for D105 and ABF of 3,403 Duroc pigs, born between July 1985 and June 1987, were collected at Akachi Genetics, Inc., Gunma, Japan. A mixed animal model used for the analysis included the fixed effects of birth-year-month (BYM), sex of the pig, parity of the dam (gilt vs. sow, MAT) and a partial linear regression on F of the pig. Random effects were litter within BYM and MAT and the additive genetic effect (breeding value) of pigs within litter. After forming the relationship matrix of the 376 parents, the Fs of the 3,403 pigs were calculated. The average F value was 0.013 $\pm$ 0.033 with the range of zero and 0.250. Slightly more than 71% were not inbred. The number of pigs having less than 0.100 F was 3,302, which represented approximately 97% of the total pigs. The estimates of the effects of inbreeding per 10% increase in F were + 1.47 d for D105 and - 0.17 mm for ABF.

Introduction

Genetic improvement can be accomplished only if genetically superior individuals are selected and then mated to reproduce the next generation. However, the true genetic merit of individuals are never known but can be estimated from the available data. A mixed animal model (Henderson, 1984) is the model preferred by animal breeders to obtain estimated breeding values. The estimators and predictors have desirable properties and the model is general, so that any known effects peculiar to the data set can be accounted for.

Most swine breeders today, either consciously or unconsciously, practice some types of inbreeding. As Lush (1933) noted, linebreeding, though mild, is a form of inbreeding. Inbreeding can also be used to fix a desirable type. Progeny produced by crossing two inbred parents of different lines can be expected to be more uniform. This increased uniformity in phenotype can be considered desirable for the mass production systems typical of many commercial pork production enterprises.

Swine breeders are now attempting to limit the introduction of new genetic

components as live pigs, because of the higher risk of introducing new disease organisms. Hence, almost all the replacement gilts are reproduced on their own farms and only when necessary are a minimum number of boars or semen purchased from outside sources. Unless replacement herd boars are always purchased from different outside sources, some level of inbreeding is likely to build up within a herd.

In theory, two consequences of inbreeding are well established: (1) changes of mean value and (2) changes of variance (Falconer, 1981). The most striking observed consequence of inbreeding is the reduction of the mean phenotypic value which is known as inbreeding depression.

When the average inbreeding coefficient of a population is $\bar{F}$, the within-line additive genetic variance is reduced to $(1 - \bar{F})\ \sigma^2_a$ while the between line additive genetic variance increases to $2\bar{F}\ \sigma^2_a$. The additive genetic variance in the whole population is the sum, and thus is equal to $(1 + \bar{F})\ \sigma^2_a$.

Dickerson *et al.* (1954) clarified two methods to estimate the effects of inbreeding: (1) the comparison of linecrosses with parental lines of known inbreeding and relationship within the same season of farrowing, and (2) the linear regression of performance on F among the unselected progeny within the same line and season. Therefore, to estimate the effects of inbreeding on the phenotypic value, the effects of selection and environment must be accounted for if the linear regression coefficient of phenotypic value on F is used.

Selection is the process of discrimination by which individuals of divergent phenotypes are made to differ in rate of reproduction (Dickerson *et al.*, 1954). Response to selection can be estimated by using mixed model methodology, even when records from a control group are not available. Results from a Monte Carlo simulation (Sorensen and Kennedy, 1984) and from a selection experiment (Blair and Pollak, 1984) revealed that a mixed model analysis could be used without a control population to partition adequately the phenotypic trend into its genetic and environmental components. Both studies strongly suggested that for the mixed model approach to succeed in separating genetic and environmental trends, strong genetic ties between contemporary groups are required. Swine data collected within a herd will generally meet this requirement because of the strong genetic ties among females which are mostly produced within the herd.

A systematic procedure for constructing a numerator relationship matrix among animals is well established (Emik and Terrill, 1949; Plum, 1954). The

off-diagonal elements of the relationship matrix, when they are multiplied by the additive genetic variance, represent the additive genetic covariances between animals. The k^{th} diagonal element of the matrix is $(1 + F_k)$, where F_k is the inbreeding coefficient of the k^{th} animal. Therefore, when the complete relationship matrix of all the animals is used, not only the known additive genetic covariances between the animals but also the increased additive genetic variances due to inbreeding are accounted for.

The purpose of this study was to estimate the effects of inbreeding on age in days and average backfat thickness at 105 kg of body weight from a mixed animal model. After solving the corresponding mixed model equations and for each trait separately, estimates were obtained for the linear regression coefficients of records on F of individual animals.

Materials and Methods

Description of data

Performance data for D105 and ABF from 3,403 Duroc pigs were collected at Akachi Genetics, Inc., Seta-Gun, Japan. The pigs were born between July 1985 and June 1987 and were from 497 litters. Very little, if any, castration and preselection prior to performance testing was practiced, because all boars were candidates for either sale or herd replacement. Hence, unless pigs had severe illness or were injured, all Duroc pigs born within the period were measured for both traits. Evidence of little preselection of boars for testing is demonstrated by the fact that the average number of pigs tested per litter was slightly over 6.8 pigs. Sexes of the pigs were either intact males or females.

Birth information was recorded on each litter. Individual pigs were weighed off-test and ultrasonically measured for backfat thickness using a Lean-Meater® at approximately 105 kg. Measurements of backfat were taken 5 cm off the midline at three points; above the point of the elbow, the last rib, and the stifle joint; and averaged for ABF. Both D105 and ABF were adjusted to a 105 kg basis using the formulas recommended by the National Swine Improvement Federation (1987) with the desired weight of 105 kg.

The number of male and female pigs and the total in each birth-year-month (BYM) subgroup are shown (Table 1).

Table 1. Distribution of male and female pigs and the total number of pigs, sires and litters in each birth-year-month (BYM)

BYM	No.litters	No.sires	No.pigs	No.males	No.females
8507	18	5	103	49	54
8508	21	6	134	57	77
8509	25	7	176	88	88
8510	21	7	137	72	65
8511	24	7	146	79	67
8512	21	7	150	72	78
8601	27	7	196	102	94
8602	27	7	230	134	96
8603	20	9	137	76	61
8604	20	10	124	70	54
8605	21	10	176	81	95
8606	21	8	136	75	61
8607	21	9	145	74	71
8608	23	11	152	85	67
8609	20	9	125	67	58
8610	21	13	134	76	58
8611	17	12	127	69	58
8612	23	10	162	88	74
8701	27	11	175	97	78
8702	11	9	77	40	37
8703	14	8	110	53	57
8704	26	11	191	97	94
8705	22	9	135	87	48
8706	6	4	25	19	6
Total	497		3,403	1,807	1,596

The number of pigs within the BYM were distributed rather uniformly across the BYM, except for the last BYM, June 1987. This happened because only a small portion of the pigs born in this BYM had completed the test at the time the data were obtained. The 25 pigs representing that particular BYM could have been either the fastest growing pigs and hence were able to complete the test, pigs born at the beginning of the BYM, or the mixture of both. The most likely reason would be the latter and hence they were included in the analysis. The distribution of the number of males and females within each BYM was quite even, and, hence, also provided evidence for little or no preselection of boars for testing. The numbers of sires and litters were likewise distributed uniformly across the BYM, except again for the last BYM.

In addition to the pigs that were tested, 44 of their sires and 226 dams whose individual performance had not been recorded were added to the data set to increase genetic ties among the pigs, making a total of 3,673 animals in the analysis. There were 21 sires and 48 dams with both parents unknown and they comprised the base population. The remaining 3,604 pigs had both parents identified and were traced back to 69 ancestors in the base population.

Calculation of coefficient of inbreeding

Emik and Terrill (1949) and Plum (1954) developed a systematic procedure for calculating the coefficient of inbreeding (F) by constructing a numerator relationship matrix. One important aspect of the matrix is that the off-diagonal elements are Wright's coefficient of numerator relationship between animals. Hence, a relationship matrix comprised of only the parents is required to calculate F of all the animals as:

$$F_i = a_{sd}/2 \qquad [1]$$

where F_i is the coefficient of inbreeding of the i^{th} animal and a_{sd} is the sire by dam off-diagonal element of the relationship matrix.

To calculate F of the 3,673 pigs, the relationship matrix of their 376 parents was constructed. Then, F of all the pigs was calculated by using the respective sire by dam off-diagonal elements as in [1].

Models

The data were analyzed separately for D105 and ABF with the same model:

$$y_{ijknq} = m_i + s_j + p_k + f_{ikn}b + l_{ikn} + a_{ijknq} + e_{ijknq} \qquad [2]$$

where y_{ijknq} is an adjusted record of either D105 or ABF of the q^{th} pig of the j^{th} sex in the n^{th} litter farrowed in the i^{th} birth-year-month (BYM) from a dam of the k^{th} maturity group (MAT), m_i is the fixed effect of the i^{th} BYM ($i = 1,...,24$), s_j is the fixed effect of the j^{th} sex ($j = 1,2$), p_k is the fixed effect of the k^{th} MAT of the dam ($k = 1,2$), f_{ikn} is the covariate term of F , l_{ikn} is the random effect of the n^{th} litter ($n = 1,...,497$), a_{ijknq} is the random effect of the additive genetic value of the q^{th} pig ($q = 1,...,3,403$), e_{ijknq} is the residual error peculiar to the record and b is the partial linear regression coefficient of records on F.

The model in [2] can be expressed in matrix notation as:

$$\mathbf{y} = \mathbf{X_1 m} + \mathbf{X_2 s} + \mathbf{X_3 p} + \mathbf{f b} + \mathbf{Z l} + [\,\mathbf{0}, \mathbf{I}\,] \begin{bmatrix} \mathbf{a_0} \\ \mathbf{a_1} \end{bmatrix} + \mathbf{e}$$

$$= \mathbf{X\beta} + \mathbf{f b} + \mathbf{Z l} + [\,\mathbf{0}, \mathbf{I}\,]\,\mathbf{a} + \mathbf{e} \qquad [3]$$

where **y**, **m**, **s**, **p**, **f**, **l**, $\mathbf{a}_0$, $\mathbf{a}_1$, and **e** are vectors of records (3,403 × 1), BYM effects (24 × 1), sex effects (2 × 1), MAT effects (2 × 1), F as a covariate (3,403 × 1), random litter effects (497 × 1), random additive genetic effects of the 270 parents without their own records (270 × 1), random additive genetic effects of pigs (3,403 × 1) and residual errors (3,403 × 1), respectively. Design matrices for **m**, **s**, **p**, and **l** are $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$ and **Z**, respectively. Also , $\beta' = [\mathbf{m}', \mathbf{s}', \mathbf{p}']$, $\mathbf{a}' = [\mathbf{a}_0', \mathbf{a}_1']$ and $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3]$. The additive genetic effects of the 270 parents without their own records ($\mathbf{a}_0$) are augmented to the breeding vales of the 3,403 pigs ($\mathbf{a}_1$) to provide extra genetic ties and thus the dimensions of vector **a** is 3,673 × 1. The design matrix corresponding to vector **a** has dimensions of 3,403 × 3,673; zero matrix of order 3,403 × 270 augmented with a unit matrix of order 3,403. The partial linear regression coefficient of records on F is denoted by b in [3]. With the assumptions that the first and the second moments are:

$$E\begin{bmatrix}\mathbf{y}\\ \mathbf{l}\\ \mathbf{a}\\ \mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{X}\beta + \mathbf{f}b\\ \mathbf{0}\\ \mathbf{0}\\ \mathbf{0}\end{bmatrix} \text{ and } Var\begin{bmatrix}l\\ a\\ e\end{bmatrix} = \begin{bmatrix}\mathbf{I}\sigma_l^2 & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{A}\sigma_a^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{I}\sigma_e^2\end{bmatrix}$$

where **A** is the numerator relationship matrix of 3,673 pigs and σ^2_l, σ^2_a, and σ^2_e are variances of litters, additive genetic effects and residual errors, the mixed model equations (MME) for model [3] are:

$$\begin{bmatrix}\mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{f} & \mathbf{X}'\mathbf{Z} & 0 & \mathbf{X}'\\ \mathbf{f}'\mathbf{X} & \mathbf{f}'\mathbf{f} & \mathbf{f}'\mathbf{Z} & 0 & \mathbf{f}'\\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{f} & \mathbf{Z}'\mathbf{Z} + \mathbf{I}\alpha_l & 0 & \mathbf{Z}'\\ 0 & 0 & 0 & \begin{bmatrix}0 & 0\\ 0 & I\end{bmatrix} + A^{-1}\alpha_a & \\ \mathbf{X} & \mathbf{f} & \mathbf{z} & & \end{bmatrix}\begin{bmatrix}\hat{\beta}\\ \hat{b}\\ \hat{l}\\ \hat{a}_0\\ \hat{a}_1\end{bmatrix} = \begin{bmatrix}\mathbf{X}'\mathbf{y}\\ \mathbf{f}'\mathbf{y}\\ \mathbf{Z}'\mathbf{y}\\ 0\\ \mathbf{y}\end{bmatrix}$$

where $\alpha_l = \sigma^2_e / \sigma^2_l$ and $\alpha_a = \sigma^2_e / \sigma^2_a$. Because $\mathbf{A}\,\sigma^2_a$ accounts for all the additive genetic variances and covariances of the pigs, σ^2_l accounts for the covariance between littermates after $\sigma^2_a / 2$ is removed. Or equivalently, the components of σ^2_l are:

$$\sigma^2_l = \text{cov (LM)} - \sigma^2_a / 2$$
$$= \sigma^2_d / 2 + \sigma^2_c$$

where cov (LM) represents the covariance between littermates and σ^2_d and σ^2_c are variances of dominance genetic effects and common environmental effects, respectively.

Takahashi *et al.* (1990) reported the estimates of α_l and α_a as 8.716 and 2.295 for D105, respectively, and 6.979 and 2.996 for ABF, respectively. Those calculated α_l's and α_a's were used in MME to estimate the linear regression coefficient of the records on F of the pigs.

Results and Discussion

The average F values of the 64 sires and the 312 dams were 0.012 ± 0.040 and 0.006 ± 0.023, respectively (Table 2). The minimum F values of both the sires and dams were zero, while the maximum F values of sires and dams were 0.250 and 0.125, respectively. In Table 3, the average parental F values of sex-status subgroups are shown. Status 1 represents a group of parents without their own performance records and status 2 a group with records. There were 44 sires of Status 1, all of which had F of zero. The remaining 20 sires had performance records and their average F was 0.038 ± 0.064. There were 226 dams of Status 1. Their F averaged 0.004 ± 0.022. The remaining 86 dams had their own performance records and the averaged 0.011 ± 0.025 in F.

Table 2. Average F values of the 64 sires and the 312 dams

Sex	No.	Ave.F	S.D.	Min.F	Max.F
Sires	64	0.012	0.040	0.000	0.250
Dams	312	0.006	0.023	0.000	0.125

Table 3. Average parental F values of sex by status subgroups

Sex	Status[a]	No.	Ave.F	S.D.	Min.F	Max.F
Sires	1	44	0.000	0.000	0.000	0.000
Sires	2	20	0.038	0.064	0.000	0.250
Dams	1	226	0.004	0.022	0.000	0.125
Dams	2	86	0.011	0.025	0.000	0.125

[a]Status 1 represents a group of parents without performance records and 2 of those with their own records.

Table 4. Average F values by sex of the 3,403 pigs

Sex	No.	Ave.F	S.D.	Min.F	Max.F
Males	1,807	0.012	0.031	0.000	0.250
Females	1,596	0.015	0.035	0.000	0.250

Table 5. Distribution of pigs by level of F

F	No.
0.000	2,426
0.004	17
0.008	77
0.016	219
0.020	2
0.031	320
0.035	6
0.039	6
0.047	34
0.063	159
0.078	9
0.094	27
0.125	53
0.133	9
0.141	9
0.143	4
0.250	26

After constructing the relationship matrix for the 376 parents, Fs of the 3,403 pigs, which included those of 20 sires and 86 dams that had individual records, were calculated. The average F was 0.012 ± 0.031 for the males and 0.015 ± 0.035 for gilts (Table 4). Both sexes had minimum and maximum values of zero and 0.250, respectively. In Table 5, the distribution of the pigs for each level of F are shown. Slightly more than 71% of the pigs were not inbred. The number of pigs having less

than 0.100 for F was 3,302, which represented approximately 97% of the total pigs.

The partial linear regression coefficient (b) of records on F of the pigs in [3] was estimable, because in the current data set there were two boars born in the first BYM that were from dams of the second MAT, and had differing F values (zero and 0.037). The estimates of b were 14.700 and -1.659, respectively, for D105 and for ABF. Since the estimates of b for both traits were expressed per unit of F, the effects of inbreeding per 10% increase of F were to increase D105 by 1.47 d and to reduce ABF by 0.17 mm.

By using the method of least squares (LS) and the current data set, the effects of inbreeding on D105 and ABF were also estimated. The statistical models used for LS contained the same fixed effects, including a partial linear regression coefficient (b_{LS}) of records on F of the pigs, as was used with the mixed model in [2]. But all the random effects in [2] were not fitted and hence fell into residual error in the LS model thus permitting the estimate of b_{LS} to be estimable. The estimates of the effects of inbreeding on the phenotypic values of D105 and ABF were 1.02 d and -0.14 mm, respectively, per 10% increase of F. These results, compared to the results from MME, indicated that the estimates obtained from LS were regressed more toward zero.

The estimates of the effects of inbreeding on D105 and ABF obtained in this analysis with MME were similar to those reported by Bereskin and Hetzer (1981, 1984). They used data obtained from an unselected control population of more than 1,000 pigs and collected over the period of 1955 to 1967. The F values of the pigs ranged from zero to approximately 0.40. By using the least squares method and a model that included the linear regression coefficients of records on F of the pigs, they obtained estimates of the effects of inbreeding on age in days to 100 kg (D100) and ABF of 0.89 d and -0.17 mm, respectively, per 10% increase of F.

Using data obtained from a large crossbreeding experiment, Schneider *et al.* (1982) calculated estimates of the effects of inbreeding on D100 and ABF. The average F value of the pigs was approximately 0.05. The estimates for D100 and ABF were 6.00 d and -1.37 mm, respectively, per 10% increase of F. Although the magnitude of the estimates reported by Schneider *et al.* (1982) was larger than those obtained in this analysis, they were in the same direction.

Use of the dominance relationship matrix (D) in MME was studied by Maki-Tanila and Kennedy (1986). Data were simulated for three levels of frequency of the favorable gene (0.1, 0.5 and 0.9) and several levels of dominance within the range

of -1 and 1 with a limited number of loci involved. The data were generated for three generations with selection. In the third generation, the average F value was 0.140 ± 0.003. Use of D in MME, however, failed to adjust for the effect of inbreeding. After reviewing the study, Kennedy *et al.* (1988) concluded that with inbreeding, inclusion of D did not account for inbreeding depression and the effects of inbreeding depression might be accounted for by including the inbreeding coefficient as a covariate in the model.

Conclusions

For an accurate genetic evaluation of animals, known sources of effects, both genetic and environmental, should be incorporated into a model. With an animal model, the additive animal effect accounts for the change in additive genetic values of animals that is due to selection. Use of the complete **A** accounts for the known additive genetic variance-covariance structure as well as the increase of the additive genetic variance due to inbreeding of the whole population. Inbreeding depression affects phenotypic values and, hence, use of F as a covariate in an animal model can account for the change in phenotypic values resulting from inbreeding. When the major interest in a swine breeding herd is the evaluation of sires by using progeny performance data, inbreeding depression should not be ignored. Because progeny of sires raised on the farm will tend to be inbred while progeny of purchased sires from the outside will more likely be non-inbred, accounting for inbreeding depression, especially for D105, provides an opportunity for fair comparisons among both sires raised on the farm and those purchased from the outside.

References

Bereskin, B. and H. O. Hetzer. 1981. Reciprocal recurrent selection in swine: postweaning growth. J. Anim. Sci. 53: 1446.

Bereskin, B. and H. O. Hetzer. 1984. Reciprocal recurrent selection in swine: carcass traits. J. Anim. Sci. 58: 821.

Blair, H. T. and E. J. Pollak. 1984. Estimation of genetic trend in a selected population with and without the use of a control population. J. Anim. Sci. 58: 878.

Dickerson, G. E., C. T. Blunn, A. B. Chapman, R. M. Kottman, J. L. Krider, E. J. Warwick and J. A. Whatley, Jr. 1954. Evaluation of selection in developing inbred lines of swine. Mo. Agric. Exp. Stn. Res. Bull. 551. (Also No. Central Reg. Publ. 38).

Emik, L. O. and C. E. Terrill. 1949. Systematic procedures for calculating inbreeding coefficients. J.

Heredity 40: 51.

Falconer, D. S. 1981. Introduction to Quantitative genetics. Longman, Inc., New York, NY.

Henderson,C. R. 1984. Applications of linear models in animal breeding. Univ. Guelph, Guelph, Canada.

Kennedy, B. W., L. R. Schaeffer and D. A. Sorensen. 1988. Genetic properties of animal models. J. Dairy Sci. 71 (Supple. 2): 17.

Lush, J. L. 1933. Linebreeding. Iowa Agric. and Home Econ. Exp. Stn. Bull. No. 301.

Maki-Tanila, A. and B. W. Kennedy. 1986. Mixed model methodology under genetic models with a small number of additive and nonadditive loci. Proc. 3rd World Congr. Genet. Appl. Livest. Prod. 12: 443.

National Swine Improvement Federation. 1987. Guidelines for uniform swine improvement programs (Rev. Ed.). Program Aid 1157. U.S.D.A., Washington, DC.

Plum, M. 1954. Computation of inbreeding and relationship coefficients. J. Heredity 45: 92.

Schneider, J. F., L. L. Christian and D. L. Kuhlers. 1982. Crossbreeding in swine: genetic effects on pig growth and carcass merit. J. Anim. Sci. 54: 747.

Sorensen, D. A. and B. W. Kennedy. 1984. Estimation of response to selection using least-squares and mixed model methodology. J. Anim. Sci. 58: 1097.

Takahashi, H., L. L. Christian, T. Sugimoto and A. Nibe. 1990. Estimates of heritabilities and common environmental effects of littermates for age at 105 kg of body weight and backfat thickness of Duroc pigs. Jpn. J. Zootech. Sci. 61: 255.

Mapping Quantitative Trait Loci and Marker Assisted Selection in Small Domesticated Populations

SHIZHONG XU
Department of Botany and Plant Sciences,
University of California, Riverside, CA 92521, USA

Abstract

The general concept and theory of quantitative genetics of finite loci are described. A method of partitioning the genetic variance into variance components contributed by individual loci are developed. The key to the partitioning of genetic variance components is the variation in the identical-by-descent (IBD) value shared by relatives. The IBD value of a putative quantitative trait locus (QTL) is inferred from multiple markers via a multipoint method. QTL variances are estimated and tested in the general framework of maximum likelihood. The simplex and EM algorithms for solving the maximum likelihood solutions are described. Marker assisted selection (MAS) via the best linear unbiased prediction (BLUP) is discussed. A new concept, called marker assisted mating (MAM), is proposed. MAM can help breeders to design the best mating strategy to avoid inbreeding depression and maintain genetic variance of small populations. The new concept and methodology of QTL mapping and MAS are fully consistent with the classical framework for the genetic analysis of quantitative characters.

Introduction

Estimating genetic variances and covariances is the first step towards understanding the genetic mechanisms of quantitative traits. It is also the basis for developing efficient selection procedures because selection acts by consuming existing genetic variances. In classical quantitative genetics, genetic variances are estimated by the phenotypic resemblance between relatives. Molecular techniques allow the new area of molecular quantitative genetics to further partition the overall genetic variance into components corresponding to chromosomal segments of the genome (Goldgar 1990). Each chromosomal segment may contain one or more loci that control the traits of interest. These loci are referred to as quantitative trait loci (QTLs). In the terminology of molecular quantitative genetics, locating QTLs and estimating QTL effects are called mapping QTLs.

Data collected from well designed crossing experiments are usually analyzed using the fixed model approach. Usually a single line cross is analyzed. Linkage disequilibrium is the key to QTL mapping using the fixed model approach. Under the fixed model, data (phenotypic values) are described by a

linear model with a variable that indicates the QTL genotype as the independent variable and the effect of allelic substitution of the QTL as the regression coefficient (Lander & Botstein 1989). The QTL genotype itself, however, is not observable. The usual practice is to use the genotypes of linked markers to guess the missing genotype of the QTL.

The random model approach of QTL mapping is more appropriate than the fixed model when data are sampled from outbred populations (Xu & Atchley 1995). Under the random model approach, variance of QTL segregation (the second moment) rather than the effect of allelic substitution (the first moment) is estimated (Haseman & Elston 1972; Schork 1993; Fulker & Cardon 1994). Sometimes, the random model approach is referred to as the robust method (Amos 1994; Olson 1995) because knowledge of the actual genetic mechanism, such as the number of alleles, the allelic frequencies or the linkage phases, is not absolutely required. Many populations of natural organisms may be in approximate linkage equilibrium, whereby marker genotypes provide no information about the genotype of a QTL. This situation invalidates the applicability of fixed model approach to outbred populations. The random model methodology, however, is fully consistent with the classical framework for the genetic analysis of quantitative characters. The primary innovation is the partitioning of the genetic variance into variance components associated with particular markers and those associated with the polygenic background. All of the existing models and types of gene action that have been elaborately studied by biometrical geneticists, including dominance and epistatic effects and fixed environmental factors are readily incorporated into the framework.

Theory and Methods

1. Genetic variance in small populations

Assume that a small population of size N_e was isolated from a base population t generations ago. In the absence of mutation, the additive genetic variance of a quantitative trait at the current generation (t) decreases monotonically compared with that of the base population (generation 0). The relationship is

$$\sigma_A^2(t) = \sigma_A^2(0)(1 - F_t) \approx \sigma_A^2(0)\exp\left(-\frac{t}{2N_e}\right) \qquad (1)$$

where $\sigma_A^2(t)$ and $\sigma_A^2(0)$ are the additive genetic variances at generation t and 0, respectively, and F_t is the average inbreeding coefficient of the population in the current generation (Falconer & Mackay 1996). To estimate the genetic variance of the population, N individuals are sampled from the current generation (t) and their phenotypic values are analyzed. If the pedigree relationships of the N individuals are known exactly back to the original generation (0) and the relationships are fully taken into account, the estimate will be $\hat{\sigma}_A^2(0)$. According to equation (1), in the absence of mutation, this variance is greater than the variances in all subsequent generations. This perhaps is the reason why one always tries to incorporate all pedigree information in estimating σ_A^2, hoping to obtain a maximum estimate of σ_A^2. However, less attention has been paid to the utilities of $\hat{\sigma}_A^2(0)$. From evolutionary standpoint, $\hat{\sigma}_A^2(0)$ may help us to understand the genetic history of the small population. From breeding point of view, however, $\hat{\sigma}_A^2(t)$ may be more useful than $\hat{\sigma}_A^2(0)$, because it is $\sigma_A^2(t)$ that determines the actual response to selection occurred in the current generation.

Directly estimating $\sigma_A^2(t)$ is impossible using individuals in the current generation, but $\sigma_A^2(t-1)$ can be estimated. In fact, it is much easier to estimate $\sigma_A^2(t-1)$ than to estimate $\sigma_A^2(0)$. When the population size is not too small, $\hat{\sigma}_A^2(t) \approx \hat{\sigma}_A^2(t-1)$. To estimate $\sigma_A^2(t-k)$ for $k \le t$ in general, one simply uses the pedigree information back to k generations and treats all individuals in that generation as completely unrelated. For instance, if one uses pedigree information only back to one generation, only full-sib or half-sib families are considered because all parents are treated as independent. Then the estimated genetic variance would be $\hat{\sigma}_A^2(t-1)$. Sib analyses ($k = 1$), are more convenient, and perhaps more appropriate, than complex pedigree analyses ($k > 1$). Of course, one can arbitrarily increases k to obtain a higher estimate of the variance. In subsequent discussions, only sib analyses are considered. Thus, the estimated genetic variance is $\hat{\sigma}_A^2(t-1)$.

2. Genetic covariance between relatives under finite loci

Consider N individuals with known family origins sampled from a finite population in generation t (the current generation). Assume that we have collected both phenotypic and molecular data for all N individuals and their parents. Our purpose here is to estimate the total genetic variance as well as variances contributed by different chromosome regions in the parental population (generation t - 1).

The phenotypic value of a quantitative trait for the j-th individual, y_j, is expressed by the following mixed linear model:

$$y_j = x_j b + \sum_{k=1}^{n} (a_{jk} + d_{jk}) + e_j \qquad (2)$$

where x_j is an $1 \times m$ design vector for a set of fixed effects, b is an $m \times 1$ vector of fixed effects, n is the total number of loci contributing to the trait, a_{jk} and d_{jk} are the additive and dominance effects, respectively, for the k-th locus, and e_j is the environmental error with a $N(0, \sigma_e^2)$ distribution. Epistatic effects are assumed to be absent. All the genetic effects are random variables with normal distributions, i.e., $a_{jk} \sim N[0, (1 + f_{jk})\sigma_{a_k}^2]$ and $d_{jk} \sim N(0, \sigma_{d_k}^2)$, where f_{jk} indicates the inbreeding event of individual j at locus k. Because all parents are assumed to be unrelated, f_{jk} can be non-zero only if there are selfings. In domesticated animals which do not allow selfing, $f_{jk} = 0$. The total additive and dominance genetic variances are defined by $\sigma_A^2 = \sum_{k=1}^{n} \sigma_{a_k}^2$ and $\sigma_D^2 = \sum_{k=1}^{n} \sigma_{d_k}^2$. The phenotypic variance of the trait is, therefore, $\sigma_P^2 = \sigma_A^2 + \sigma_D^2 + \sigma_e^2$.

The expectation of y_j is

$$E(y_j) = x_j b. \qquad (3)$$

The variance is

$$Var(y_j) = \sum_{k=1}^{n} (1 + f_{jk})\sigma_{a_k}^2 + \sum_{k=1}^{n} \sigma_{d_k}^2 + \sigma_e^2 \qquad (4)$$

Note that f_{jk} is not the usual inbreeding coefficient defined in the context of probability; it indicates the status of identical-by-descent (IBD) between the

paternal and maternal alleles of individual j at locus k. The covariance between y_i and y_j is

$$\mathrm{Cov}(y_i, y_j) = \sum_{k=1}^{n} \pi_{ij}^{k} \sigma_{a_k}^2 + \sum_{k=1}^{n} \theta_{ij}^{k} \sigma_{d_k}^2 \qquad (5)$$

where π_{ij}^{k} is the IBD value shared by i and j at the k-th locus, and θ_{ij}^{k} indicates the event that i and j share IBD genotype (both alleles are IBD). These IBD variables, π_{ij}^{k} and θ_{ij}^{k}, can be non-zero if i and j are sibs, otherwise, they are zero because members of different families are assumed to be independent.

Denote $\Delta_{a_k} = \sigma_{a_k}^2 - \frac{1}{n}\sigma_A^2$ as the deviation of the locus specific variance from the average variance across loci so that $\sigma_{a_k}^2 = \Delta_{a_k} + \frac{1}{n}\sigma_A^2$. Similarly, denote $\Delta_{d_k} = \sigma_{d_k}^2 - \frac{1}{n}\sigma_D^2$ so that $\sigma_{d_k}^2 = \Delta_{d_k} + \frac{1}{n}\sigma_D^2$. The variance and covariance can be expressed by

$$\mathrm{Var}(y_j) = (1 + \frac{1}{n}\sum_{k=1}^{n} f_{jk})\sigma_A^2 + \sum_{k=1}^{n}(1 + f_{jk})\Delta_{a_k} + \sigma_D^2 + \sigma_e^2 \qquad (6)$$

and

$$\mathrm{Cov}(y_i, y_j) = \frac{1}{n}\sum_{k=1}^{n} \pi_{ij}^{k}\sigma_A^2 + \sum_{k=1}^{n} \pi_{ij}^{k}\Delta_{a_k} + \frac{1}{n}\sum_{k=1}^{n} \theta_{ij}^{k}\sigma_D^2 + \sum_{k=1}^{n} \theta_{ij}^{k}\Delta_{d_k} \qquad (7)$$

It can be easily shown that the finite locus model approaches to the classical polygenic model when the assumptions of polygenic model hold. With equal effects across loci,

$\Delta_{a_k} = \Delta_{d_k} = 0$ for k = 1, ..., n. For infinite number of unlinked loci,

$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f_{jk} \to F_j$, $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \pi_{ij}^{k} \to A_{ij}$ and $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \theta_{ij}^{k} \to D_{ij}$, where F_j is the usual inbreeding coefficient, A_{ij} is the additive relationship between i and j, and D_{ij} is the coefficient of fraternity (Lynch and Walsh 1998).

In matrix notation, the model can be expressed as

$$y = Xb + \sum_{k=1}^{n} a_k + \sum_{k=1}^{n} d_k + e \qquad (8)$$

where $y = \begin{bmatrix} y_1 & \cdots & y_N \end{bmatrix}^T$, $X = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{N1} & \cdots & x_{Nm} \end{bmatrix}$, $b = \begin{bmatrix} b_1 & \cdots & b_m \end{bmatrix}^T$,

$a_k = \begin{bmatrix} a_{1k} & \cdots & a_{Nk} \end{bmatrix}^T$, $d_k = \begin{bmatrix} d_{1k} & \cdots & d_{Nk} \end{bmatrix}^T$ and $e = \begin{bmatrix} e_1 & \cdots & e_N \end{bmatrix}^T$.

The expectation and variance-covariance matrix are

$$E(y) = Xb \tag{9}$$

and

$$Var(y) = V = \sum_{k=1}^{n} \Pi^k \sigma^2_{a_k} + \sum_{k=1}^{n} \Theta^k \sigma^2_{d_k} + I\sigma^2_e, \tag{10}$$

respectively, where $\Pi^k = \begin{bmatrix} \pi^k_{11} & \cdots & \pi^k_{1N} \\ \vdots & \ddots & \vdots \\ \pi^k_{1N} & \cdots & \pi^k_{NN} \end{bmatrix}$, $\Theta^k = \begin{bmatrix} \theta^k_{11} & \cdots & \theta^k_{1N} \\ \vdots & \ddots & \vdots \\ \pi^k_{1N} & \cdots & \theta^k_{NN} \end{bmatrix}$,

$\pi^k_{jj} = 1 + f_{jk}$ and $\theta^k_{jj} = 1$.

Inbreeding coefficient, additive relationship and coefficient of fraternity can be considered as average measurements of various kinds of IBD variables across different loci. Their values depend on the pedigree relationships of individuals concerned. The corresponding IBD measurements at the locus level, f_{jk}, π^k_{ij} and θ^k_{ij}, however, depend not only on the pedigree relationships but also on the actual genotype of each individual at the locus of interest. Conditional on pedigrees, these locus specific IBDs are still variables. The variance of the IBD measurements across loci provides the key to the random model method of QTL mapping. In all subsequent discussion, f_{jk}, π^k_{ij} and θ^k_{ij} are referred to as IBD variables. Consider the genotypic configurations of progenies from mating type $A_1A_2 \times A_3A_4$. There are four possible types of progeny, each with an equal frequency. The four possible genotypes are A_1A_3, A_1A_4, A_2A_3 and A_2A_4. If two sibs are sampled from this family, ignoring the order of sampling, there are 10 possible sib-pairs. Suppose we observe a pair of sibs with genotypic configuration A_1A_3-A_1A_3, we know immediately that they have received exactly the same alleles

from their parents and thus $\pi_{ij}^{k} = 1$. The two sibs behave like identical twins for this locus. If A_1A_3-A_2A_4 is observed, then we know that they do not share any IBD alleles, and thus behave like two unrelated individuals. If we happen to know the genotypes of two sibs at a particular locus, the covariance between them at this locus can be different from what is expected. For example,

$\mathrm{Cov}(a_{ik}, a_{jk}) = 1\sigma_{a_k}^2$ for a pair of sibs with genotypic configuration A_1A_3-A_1A_3 and $\mathrm{Cov}(a_{ik}, a_{jk}) = 0\sigma_{a_k}^2$ with A_1A_3-A_2A_4. It is incorrect to say $\mathrm{Cov}(a_{ik}, a_{jk}) = \frac{1}{2}\sigma_{a_k}^2$ if one already knows that i and j share no IBD allele at the locus of interest.

Full-sib families, half-sib families and sibs resulting from a selfing parent are the three types of families considered in this paper. There has been no scientific term for sibs derived from a single parent (selfing). For simplicity of description, I will call them selfed-sibs in subsequent discussion. In domesticated animals, it is common to have a hierarchical family structure of mixed sibships, i.e., a half-sib family consists of several full-sib families. The distributions of various IBD variables are given in Table 1 for full-sibs, Table 2 for half-sibs and Table 3 for selfed-sibs. The expectations and variances of these IBD variables are summarized in Table 4. Selfed-sibs have the largest IBD variances, and thus have the highest power of separating QTL variance from the polygenic variance. Half-sibs have the lowest power due to the smallest IBD variance. In addition, half-sib analysis does not provide an estimate of the dominance variance.

3. Mapping QTLs under the random model

Given the expectation and the variance-covariance matrices, under the assumption that y has a multivariate normal distribution, a likelihood function can be constructed. The maximum likelihood method requires that the number of QTLs and the IBD matrices are known *a priori*, whereas, this information is not known. In fact, one of the main purposes of QTL mapping is to find the number of QTLs. In addition, it may be difficult for the maximum likelihood method to handle too many unknown parameters. The usual treatment is to decompose the total genetic effect into the effect of the q-th QTL and the sum of the effects of the remaining QTLs. The model is written as

$$y = Xb + a_q + d_q + a_A + d_D + e \qquad (11)$$

Table 1. Distribution of the identical-by-descent variables shared by full-sibs resulting from matting of A_1A_2 with A_3A_4, where A_1 and A_3 are the paternal alleles of the parents, and A_2 and A_4 are the maternal alleles of the parents. The 16 possible sib-pairs are classified into four states, (a) state 11: share both alleles, (b) state 10: share paternal alleles, (c) state 01: share maternal alleles, and (d) state 00: share no allele.

Sib pair	Probability	π_{ij}^k	θ_{ij}^k	State
A_1A_3-A_1A_3	1/16	1	1	11
A_1A_3-A_1A_4	1/16	1/2	0	10
A_1A_3-A_2A_3	1/16	1/2	0	01
A_1A_3-A_2A_4	1/16	0	0	00
A_1A_4-A_1A_3	1/16	1/2	0	10
A_1A_4-A_1A_4	1/16	1	1	11
A_1A_4-A_2A_3	1/16	0	0	00
A_1A_4-A_2A_4	1/16	1/2	0	01
A_2A_3-A_1A_3	1/16	1/2	0	01
A_2A_3-A_1A_4	1/16	0	0	00
A_2A_3-A_2A_3	1/16	1	1	11
A_2A_3-A_2A_4	1/16	1/2	0	10
A_2A_4-A_1A_3	1/16	0	0	00
A_2A_4-A_1A_4	1/16	1/2	0	01
A_2A_4-A_2A_3	1/16	1/2	0	10
A_2A_4-A_2A_4	1/16	1	1	11

Table 2. Distribution of the identical-by-descent variables shared by half-sibs resulting from matings of a father A_1A_2 with randomly selected females.

Sib pair[a]	Probability	π_{ij}^k
A_1A_X-A_1A_y	1/4	1/2
A_1A_X-A_2A_y	1/4	0
A_2A_X-A_1A_y	1/4	0
A_2A_X-A_2A_y	1/4	1/2

[a]: A_X or A_y represents a random allele from the population.

Table 3. Distribution of the identical-by-descent variables shared by selfed sibs of a single parent A_1A_2.

Sib pair	Probability	π_{ij}^k	θ_{ij}^k
A_1A_1-A_1A_1	1/16	2	1
A_2A_2-A_2A_2	1/16	2	1
A_1A_1-A_1A_2	1/8	1	0
A_2A_2-A_1A_2	1/8	1	0
A_1A_2-A_1A_1	1/8	1	0
A_1A_2-A_2A_2	1/8	1	0
A_1A_2-A_1A_2	1/4	1	1
A_2A_2-A_1A_1	1/16	0	0
A_1A_1-A_2A_2	1/16	0	0

Table 4. Expectations and variances of the identical-by-descent variables of a single locus in full-sibs, half-sibs and selfed-sibs.

	f_{jk}		π_{ij}^k		θ_{ij}^k	
	$E(f_{jk})$	$Var(f_{jk})$	$E(\pi_{ij}^k)$	$Var(\pi_{ij}^k)$	$E(\theta_{ij}^k)$	$Var(\theta_{ij}^k)$
Full-sib	0	0	1/2	1/8	1/4	3/16
Half-sib	0	0	1/4	1/16	0	0
Selfed-sib	1/2	1/4	1	1/4	3/8	15/64

where $a_A = \sum_{k \neq q}^{n} a_k$ is the sum of the additive effects of the remaining QTLs (additive polygenic effect) and $d_D = \sum_{k \neq q}^{n} d_k$ is the sum of the dominance effects of the remaining QTLs (dominance polygenic effect). Accordingly, the total additive variance is partitioned into variance explained by the q-th QTL and the variance contributed collectively by all other loci. The latter is called the polygenic variance and still denoted by $\sigma_A^2 = \sum_{k \neq q}^{n} \sigma_{a_k}^2$ for notational convenience but with $\sigma_{a_q}^2$ excluded. Similarly, the dominance variance can be partitioned into $\sigma_{d_q}^2$ and $\sigma_D^2 = \sum_{k \neq q}^{n} \sigma_{d_k}^2$. The variance-covariance matrix of y is now rewritten as

$$\text{Var}(y) = V = \Pi^q \sigma_{a_q}^2 + \Theta^q \sigma_{d_q}^2 + A\sigma_A^2 + D\sigma_D^2 + I\sigma_e^2, \qquad (12)$$

where $A = \{A_{ij}\}_{N \times N}$ and $D = \{D_{ij}\}_{N \times N}$. With this model, the list of unknown parameters becomes much shorter: $\theta = \left\{ b \ \sigma_{a_q}^2 \ \sigma_{d_q}^2 \ \sigma_A^2 \ \sigma_D^2 \ \sigma_e^2 \right\}$. The likelihood function is

$$L(y|\theta) \propto |V|^{-\frac{1}{2}} \text{Exp}\left[-\frac{1}{2}(y - Xb)^T V^{-1} (y - Xb) \right]. \qquad (13)$$

In QTL mapping, the parameters of interest are $\left\{ \sigma_{a_q}^2 \quad \sigma_{d_q}^2 \right\}$, the remaining elements in θ are called nuisance parameters. These two parameters of interest can be tested for significance by fitting the model with all parameters and obtaining the natural log of the likelihood of the data, $\ln(L_1)$, and refitting without those two parameters, obtaining $\ln(L_0)$, the log of the likelihood under the null hypothesis. The test statistic,

$$\lambda = -2[\ln(L_0) - \ln(L_1)] \qquad (14)$$

is asymptotically distributed as a χ^2 with degrees of freedom equal to the number of parameters tested. If there is a QTL at a particular chromosome position, the IBD matrices at that position will match Π^q and Θ^q and the test statistic will

show a peak at that position. The chromosome position is allowed to vary so that the whole chromosome is searched from one end to the other for the evidence of QTLs.

The IBD matrices of a QTL cannot be observed and must be inferred from markers linked to the putative QTL. There will be some uncertainty associated with the statistical inference of IBD. What we can obtain is the conditional distributions of the IBD variables given marker information. When QTL mapping is conducted, the conditional expectations of the IBD variables are used in place of the true IBDs.

For full-sibs, the conditional expectation of π_{ij}^q given marker information is calculated by

$$E(\pi_{ij}^q | I_M) = \hat{\pi}_{ij}^q = p_{11}^q + \tfrac{1}{2}(p_{10}^q + p_{01}^q), \tag{15}$$

where p_{11}^q, p_{10}^q, p_{01}^q and p_{00}^q represent the conditional probabilities that i and j share two IBD alleles, one IBD allele (paternal), one IBD allele (maternal) and no IBD allele, respectively. These conditional probabilities are calculated using a multipoint method which will be discussed in the next section. The conditional expectation of θ_{ij}^q is

$$E(\theta_{ij}^q | I_M) = \hat{\theta}_{ij}^q = p_{11}^q . \tag{16}$$

For half-sibs, the conditional expectation of π_{ij}^q is

$$E(\pi_{ij}^q | I_M) = \hat{\pi}_{ij}^q = \tfrac{1}{2} p_1^q + 0 p_0^q = \tfrac{1}{2} p_1^q , \tag{17}$$

where p_1^q is the conditional probability that i and j share one IBD allele (paternal) and p_0^q is the probability that they share no IBD allele ($p_0^q + p_1^q = 1$).

For selfed-sibs, the conditional expectations of π_{ij}^q and θ_{ij}^q are

$$E(\pi_{ij}^q | I_M) = \hat{\pi}_{ij}^q = 2p_5^q + 1p_4^q + 1p_3^q + 1p_2^q + 0p_1^q = 2p_5^q + (p_4^q + p_3^q + p_2^q) \tag{18}$$

and

$$E(\theta_{ij}^q | I_M) = \hat{\theta}_{ij}^q = p_5^q + p_2^q \tag{19}$$

respectively, where p_5^q is the conditional probability that the two sibs are identical homozygotes, i.e., $A_1A_1 - A_1A_1$ or $A_2A_2 - A_2A_2$, p_4^q is the conditional probability that the first sib is homozygote while the second one is heterozygote, i.e., $A_1A_1 - A_1A_2$ or $A_2A_2 - A_1A_2$, p_3^q is the conditional

probability that the first sib is heterozygote while the second one is homozygote, i.e., $A_1A_2 - A_1A_1$ or $A_1A_2 - A_2A_2$, p_2^q is the conditional probability that both sibs are heterozygotes, i.e., $A_1A_2 - A_1A_2$, and p_1^q is the conditional probability that the two sibs are different homozygotes, i.e., $A_1A_1 - A_2A_2$ or $A_2A_2 - A_1A_1$. These conditional probabilities are calculated using a multipoint method. In selfed families, a sib has a chance to be inbred. The IBD of an individual with itself is $\pi_{jj}^q = 1 + f_{jq}$. The conditional expectation of π_{jj}^q is

$$E(\pi_{jj}^q | I_M) = 1 + E(f_{jq} | I_M) = 1 + \hat{f}_{jq} = 1 + p_q \qquad (20)$$

where p_q is the conditional probability that the two alleles of j are IBD at the QTL, i.e., j is either A_1A_1 or A_2A_2.

4. Multipoint inference of IBD states

In outbred populations considered here, markers may be partially informative. In such situations, two flanking markers may not extract the maximum amount of information about the segregation of a putative QTL, and because of this, markers outside the interval can provide additional information. In these situations, a more desirable procedure than interval mapping is to use all markers simultaneously, a procedure called multipoint mapping (Fulker *et al.* 1995; Kruglyak & Lander 1995).

Consider M ordered markers on the chromosome of interest. If a marker is fully informative, the IBD states of each marker shared by sibs are observed (known). Otherwise, the probabilities of IBD states can be inferred based on the observed marker genotypes. I will first demonstrate the multipoint method in full-sibs. Denote the probabilities of IBD states of marker k by p_{11}^k, p_{10}^k, p_{01}^k and p_{00}^k, respectively, for the four states. Assume that the QTL is located between marker k and k+1 for M-1 ≥ k ≥ 1. What we want is to calculate $\Pr(\pi_{ij}^q | \pi_{ij}^1 \cdots \pi_{ij}^M) = \left[p_{11}^q \quad p_{10}^q \quad p_{01}^q \quad p_{00}^q \right]^T$. When the marker IBD states are not directly observed, $\Pr(\pi_{ij}^q | \pi_{ij}^1 \cdots \pi_{ij}^M)$ should be denoted by $\Pr(\pi_{ij}^q | I_M)$, where I_M means marker information. Under Haldane map function, the sequence $\left\{ \pi_{ij}^1 \cdots \pi_{ij}^k \pi_{ij}^q \pi_{ij}^{k+1} \cdots \pi_{ij}^M \right\}$ forms a reversible Markov chain with the following

transition probability matrix between π_{ij}^{k} and π_{ij}^{l}:

$$T_{kq} = \begin{bmatrix} \Psi^2 & \Psi(1-\Psi) & (1-\Psi)\Psi & (1-\Psi)^2 \\ \Psi(1-\Psi) & \Psi^2 & (1-\Psi)^2 & (1-\Psi)\Psi \\ (1-\Psi)\Psi & (1-\Psi)^2 & \Psi^2 & \Psi(1-\Psi) \\ (1-\Psi)^2 & (1-\Psi)\Psi & \Psi(1-\Psi) & \Psi^2 \end{bmatrix}$$

where $\Psi = r_{kl}^2 + (1 - r_{kl})^2$ and r_{kl} is the recombination fraction between loci k and l. With the multipoint method, we will first calculate the probabilities of observing the marker genotypes given each of the four IBD states at the QTL. These four conditional probabilities are calculated as follows:

$$\Pr(I_M|\pi_{ij}^{q} = \tfrac{1+1}{2}) = P_{11} = 1^T D_1 T_{12} D_2 \cdots T_{kq} D_{(11)} T_{qk+1} \cdots D_{M-1} T_{M-1M} D_M 1,$$

$$\Pr(I_M|\pi_{ij}^{q} = \tfrac{1+0}{2}) = P_{10} = 1^T D_1 T_{12} D_2 \cdots T_{kq} D_{(10)} T_{qk+1} \cdots D_{M-1} T_{M-1M} D_M 1,$$

$$\Pr(I_M|\pi_{ij}^{q} = \tfrac{0+1}{2}) = P_{01} = 1^T D_1 T_{12} D_2 \cdots T_{kq} D_{(01)} T_{qk+1} \cdots D_{M-1} T_{M-1M} D_M 1, \text{ and}$$

$$\Pr(I_M|\pi_{ij}^{q} = \tfrac{0+0}{2}) = P_{00} = 1^T D_1 T_{12} D_2 \cdots T_{kq} D_{(00)} T_{qk+1} \cdots D_{M-1} T_{M-1M} D_M 1.$$

where $D_k = \mathrm{diag}[p_{11}^{k} \quad p_{10}^{k} \quad p_{01}^{k} \quad p_{00}^{k}]$, $D_{(11)} = \mathrm{diag}[1 \quad 0 \quad 0 \quad 0]$, $D_{(10)} = \mathrm{diag}[0 \quad 1 \quad 0 \quad 0]$, $D_{(01)} = \mathrm{diag}[0 \quad 0 \quad 1 \quad 0]$ and $D_{(01)} = \mathrm{diag}[0 \quad 0 \quad 1 \quad 0]$.

According to the Bayesian theorem, the probability of IBD state of QTL given marker information is

$$\Pr(\pi_{ij}^{q}|I_M) = \frac{\Pr(\pi_{ij}^{q})\Pr(I_M|\pi_{ij}^{q})}{\Sigma \Pr(\pi_{ij}^{q})\Pr(I_M|\pi_{ij}^{q})} \qquad (21)$$

where $\Pr(\pi_{ij}^{q})$ is the prior distribution of the IBD state. Recall that $\Pr(\pi_{ij}^{q} = \tfrac{1+1}{2}|I_M) = p_{11}^{q}$, $\Pr(\pi_{ij}^{q} = \tfrac{1+0}{2}|I_M) = p_{10}^{q}$, $\Pr(\pi_{ij}^{q} = \tfrac{0+1}{2}|I_M) = p_{01}^{q}$ and $\Pr(\pi_{ij}^{q} = \tfrac{0+0}{2}|I_M) = p_{00}^{q}$. The prior probability is the unconditional distribution which is 1/4 for each of the four possible IBD states. Therefore, the explicit forms of equation (21) are $p_{11}^{q} = c^{-1}P_{11}$, $p_{10}^{q} = c^{-1}P_{01}$, $p_{01}^{q} = c^{-1}P_{01}$ and $p_{00}^{q} = c^{-1}P_{00}$, where $c = P_{11} + P_{10} + P_{01} + P_{00}$.

The multipoint method for half-sibs are similarly defined. There are two IBD states: $\pi_{ij}^q = \frac{1+0}{2} = 1/2$ and $\pi_{ij}^q = \frac{0+0}{2} = 0$. Therefore, the dimension of the Markov chain is 2×2, rather than 4×4. The transition matrix between the IBD states of loci k and l is

$$T_{kl} = \begin{bmatrix} \Psi & 1-\Psi \\ 1-\Psi & \Psi \end{bmatrix}$$

The conditional probabilities of observing marker genotypes given the shared IBD state of the QTL are

$$\Pr(I_M | \pi_{ij}^q = \tfrac{1+0}{2}) = P_1 = 1^T D_1 T_{12} D_2 \cdots T_{kq} D_{(1)} T_{qk+1} \cdots D_{M-1} T_{M-1M} D_M 1$$

and

$$\Pr(I_M | \pi_{ij}^q = \tfrac{0+0}{2}) = P_0 = 1^T D_1 T_{12} D_2 \cdots T_{kq} D_{(0)} T_{qk+1} \cdots D_{M-1} T_{M-1M} D_M 1,$$

where

$$D_k = \begin{bmatrix} p_1^k & 0 \\ 0 & p_0^k \end{bmatrix}, \; D_{(1)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } D_{(0)} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

The prior probability of the IBD state of QTL is 1/2 for each state, leading to $p_1^q = c^{-1} P_1$ and $p_0^q = c^{-1} P_0$, where $c = P_1 + P_0$.

In the multipoint method of selfed-sibs, I first define the conditional probabilities of marker IBD states, P_t, for t=1,...,5, given each of the five IBD states of QTL using the Markov chain method. The transition probability matrix of the five IBD states between loci k and l is

$$T_{kl} = \begin{bmatrix} \Psi^2 - \frac{1}{2}(1-\Psi)^2 & \Psi(1-\Psi) & \Psi(1-\Psi) & (1-\Psi)^2 & \frac{1}{2}(1-\Psi)^2 \\ \frac{1}{2}\Psi(1-\Psi) & \Psi^2 & (1-\Psi)^2 & \Psi(1-\Psi) & \frac{1}{2}\Psi(1-\Psi) \\ \frac{1}{2}\Psi(1-\Psi) & (1-\Psi)^2 & \Psi^2 & \Psi(1-\Psi) & \frac{1}{2}\Psi(1-\Psi) \\ \frac{1}{2}(1-\Psi)^2 & \Psi(1-\Psi) & \Psi(1-\Psi) & \Psi^2 & \frac{1}{2}(1-\Psi)^2 \\ \frac{1}{2}(1-\Psi)^2 & \Psi(1-\Psi) & \Psi(1-\Psi) & (1-\Psi)^2 & \Psi^2 - \frac{1}{2}(1-\Psi)^2 \end{bmatrix}$$

The prior probabilities of the five IBD states of QTL are $\left[\frac{1}{8} \quad \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{8}\right]$. The conditional probabilities of the IBD states of QTL given marker information are,

$p_1^q = \frac{1}{8} c^{-1} P_1$, $p_2^q = \frac{1}{4} c^{-1} P_2$, $p_3^q = \frac{1}{4} c^{-1} P_3$, $p_4^q = \frac{1}{4} c^{-1} P_4$ and $p_5^q = \frac{1}{8} c^{-1} P_5$

where $c = \frac{1}{8} P_5 + \frac{1}{4} P_4 + \frac{1}{4} P_3 + \frac{1}{4} P_2 + \frac{1}{8} P_1$.

The multipoint calculation of IBD probability of an individual with itself is $p_q = c^{-1}P_q$, where $c = P_q + Q_q$, P_q is the probability of observing the marker genotypes given the individual being homozygote at the QTL, and Q_q is the probability of observing the marker genotypes given the individual being heterozygote at the QTL. P_q and Q_q are calculated using the Markov chain approach:

$$P_q = 1^T D_1 T_{12} D_2 \cdots T_{kq} D_{(1)} T_{qk+1} \cdots D_{M-1} T_{M-1M} D_M 1$$

and

$$Q_q = 1^T D_1 T_{12} D_2 \cdots T_{kq} D_{(0)} T_{qk+1} \cdots D_{M-1} T_{M-1M} D_M 1,$$

where the transition matrix is

$$T_{kl} = \begin{bmatrix} \Psi & 1-\Psi \\ 1-\Psi & \Psi \end{bmatrix}$$

and the D matrices are defined as

$$D_k = \begin{bmatrix} p_k & 0 \\ 0 & q_k \end{bmatrix}, \; D_{(1)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } D_{(0)} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Note that p_k is the conditional probability that the two alleles of marker k are IBD given the observed genotype of marker k and $q_k = 1 - p_k$.

The hidden Markov model method provides a general machinery for the multipoint calculation of the conditional probabilities of IBD states of QTLs. The additional work left to the practitioners is to calculate the probabilities of IBD states of an individual marker given the observed genotypes of this marker. Computing marker IBD probabilities is much easier than computing IBD probabilities of QTLs. For example, in a selfed family, if the parent is of genotype A_1A_1 at a marker locus, all the sibs will be of the same genotype, and there is no information about the IBD states in the progenies. The probabilities of the IBD states will take their prior probabilities. If the parent is A_1A_2 and the sib-pair is of type $A_1A_1 - A_1A_2$, then the IBD state is type 4, which leads to $p_4^k = 1$ and $p_i^k = 0$ for $i = 1,2,3,5$. For the same parent, if the sib-pair is of type $A_1A_1 - A_?A_?$, i.e., the second sib has a missing genotype, then the sib-pair can

be of type 5 with a probability 1/4 or type 4 with a probability 1/2 or type 1 with a probability 1/4. Therefore, $p_5^k = 1/4, p_4^k = 1/2, p_3^k = 0, p_2^k = 0$ and $p_1^k = 1/4$. A general method of calculating marker IBD probabilities can be derived following the Bayesian method suggested by Haseman & Elston (1972). However, description of the method is very tedious and one has to resort to a computer program.

5. Algorithms of searching for ML solutions

The likelihood is a complicated function of the genetic variance components. Analytical solutions are hard to derive and a numerical approach must be resorted. There are numerous ways to search for the ML solutions, but only two algorithms are commonly used in this context. The first algorithm is the simplex method and the second one is the EM (expectation-maximization) algorithm.

The simplex method: This method was developed by Nelder & Mead (1965). It is a general searching algorithm for any complicated objective function. Let θ be a $k \times 1$ vector of the unknown parameters and $f(\theta)$ is the objective function which needs to be minimized. Note that maximizing a likelihood function is equivalent to minimizing the negative of the likelihood function. A simplex is defined by a set of $k+1$ points in R^k joined by lines. In two dimensions, a simplex is a triangle. In three dimensions it is a tetrahedron. The basic idea of the simplex method is that iteration process generates a new simplex by reflecting one vertex in the hyperplane spanned by the other vertices of the simplex. Suppose the function to be optimized has a maximum at a given vertex; then we expect that the function will have a minimum at the reflected vertex. Let $\theta^{(m)}$ (a $k \times 1$ vector) be vertex such that

$$f(\theta^{(m)}) = \max_i f(\theta^{(i)}), \quad i = 1, \cdots, k+1$$

Let $\theta^{(e)}$ be the vertex such that

$$f(\theta^{(e)}) = \min_i f(\theta^{(i)}), \quad i = 1, \cdots, k+1.$$

Let

$$\theta^{(0)} = \frac{1}{k} \sum_{i \neq m}^{k+1} \theta^{(i)}$$

The reflection of the point $\theta^{(m)}$ is given by

$$\theta^{(r)} = (1+\alpha)\theta^{(0)} - \alpha\theta^{(m)}, \tag{22}$$

where α is a reflection factor. Contraction is defined by

$$\theta^{(c)} = \beta\theta^{(m)} + (1-\beta)\theta^{(0)} \tag{23}$$

where β is a contraction factor. Similarly, the expansion is defined by

$$\theta^{(e)} = \gamma\theta^{(r)} + (1-\gamma)\theta^{(0)} \tag{24}$$

where γ is an expansion factor. The values of the three factors are chosen based on personal experience. The choice that $\alpha = 1, \beta = \frac{1}{2}$ and $\gamma = 2$ given by Nelder & Mead (1965) seems to perform well. The iteration process of the simplex method consists of the following steps.

(I) Determine $\theta^{(0)}, \theta^{(e)}, \theta^{(m)}, \theta^{(r)}$.

(II) Try reflection, and if $f(\theta^{(r)}) \geq f(\theta^{(e)})$, replace $\theta^{(m)}$ by $\theta^{(r)}$ and restart with the new simplex.

(III) If $f(\theta^{(r)}) < f(\theta^{(e)})$, we proceed with expansion. The expansion succeeds if $f(\theta^{(m)}) > f(\theta^{(e)})$, and in this case $\theta^{(r)}$ replaces $\theta^{(m)}$. In case expansion fails, we replace $\theta^{(m)}$ also by $\theta^{(r)}$, and the process starts with the new simplex.

(IV) If the reflection is such that $f(\theta^{(m)}) > f(\theta^{(r)})$, we replace $\theta^{(m)}$ by $\theta^{(r)}$ and make a contraction move. Now if $f(\theta^{(m)}) > f(\theta^{(c)})$, the contraction is successful, $\theta^{(m)}$ is replaced by $\theta^{(c)}$, and a new simplex results. In case of failure, the simplex is shrunk about the lowest function value $\theta^{(e)}$ by $\theta^{(i)}$, and $\theta^{(i)}$ is replaced by $\frac{1}{2}(\theta^{(i)} + \theta^{(e)})$.

The stopping criterion generally uses a pre-assigned number $\varepsilon > 0$ and stop if

$$\left\{\frac{1}{k}\sum_{i=1}^{k+1}[f(\theta^{(i)}) - f(\theta^{(0)})]^2\right\}^{\frac{1}{2}} < \varepsilon. \tag{25}$$

The optimal value of θ must be searched within the space of R^k. The three actions in the simplex method (reflection, contraction and expansion), however, do not necessarily guarantee that all vertices of θ are within R^k. For example, the variance components cannot be negative, but a reflection of θ may contain a negative variance component. The non-negativity of variance components can be achieved by reparameterization. Instead of directly searching for the value of a variance component, we search for ξ for $-\infty < \xi < +\infty$ and

assign e^{ξ} to the variance component. This will automatically restrict the variance component within $(0,+\infty)$.

The EM algorithm: This algorithm was developed by Dempster *et al.* (1977) for ML problem involving missing observations. The likelihood function in this study can be formulated as incomplete data problem. The missing data are the genetic effects. I now demonstrate the EM algorithm using a simple additive effect genetic model:

$$y = Xb + a_q + a_A + e \tag{26}$$

where a_q is an $N \times 1$ vectors for the effects of QTL and $a_A = \sum_{k \neq q}^{n} a_k$ is the additive polygenic effect. The expectation and variance of y are $E(y) = Xb$ and $Var(y) = V = \Pi^q \sigma^2_{a_q} + A\sigma^2_A + I\sigma^2_e$. The likelihood of the complete data is

$$L(y, a_q, a_A|\theta) = (2\pi\sigma^2_e)^{-N/2} Exp\left\{-\frac{1}{2\sigma^2_e}(y - Xb - a_q - a_A)^T(y - Xb - a_q - a_A)\right\} \times f(a_q|\theta) f(a_A|\theta) \tag{27}$$

where

$$f(a_q|\theta) = (2\pi\sigma^2_{a_q})^{-N/2} \left|\Pi^q\right|^{-\frac{1}{2}} Exp\left\{-\frac{1}{2\sigma^2_{a_q}} a_q^T (\Pi^q)^{-1} a_q\right\} \tag{28}$$

and

$$f(a_A|\theta) = (2\pi\sigma^2_A)^{-N/2} |A|^{-\frac{1}{2}} Exp\left\{-\frac{1}{2\sigma^2_A} a_A^T A^{-1} a_A\right\}. \tag{29}$$

The actual likelihood is obtained by integrating out the missing data:

$$L(y|\theta) = \int_{a_q, a_A \in \Omega} L(y, a_q, a_A|\theta) da_q da_A . \tag{30}$$

Instead of maximizing $l(y|\theta) = \ln[L(y|\theta)]$ directly, the EM algorithm proceeds by using an initial estimate $\theta^{(0)}$ and maximizing

$$E\left[\ln L(y, a_q, a_A|\theta)\right], \tag{31}$$

with respect to θ, where the expectation is taken with respect to a_q and a_A

conditional on y and $\theta^{(0)}$. The solution after the maximization is denoted by $\theta^{(1)}$. The initial value $\theta^{(0)}$ is then replaced by $\theta^{(1)}$ and the iteration is continued until a convergence criterion is satisfied. The explicit process of the EM steps is described below.

In the E-step, we need to calculate the conditional expectations and conditional variances of a_q, a_A and e given y and $\theta^{(t)}$. First, we establish the following expectation and variance-covariance matrices:

$$E\begin{bmatrix} y \\ a_q \\ a_A \\ e \end{bmatrix} = \begin{bmatrix} Xb \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ and } Var\begin{bmatrix} y \\ a_q \\ a_A \\ e \end{bmatrix} = \begin{bmatrix} V & \Pi^q\sigma^2_{a_q} & A\sigma^2_A & I\sigma^2_e \\ \Pi^q\sigma^2_{a_q} & \Pi^q\sigma^2_{a_q} & 0 & 0 \\ A\sigma^2_A & 0 & A\sigma^2_A & 0 \\ I\sigma^2_e & 0 & 0 & I\sigma^2_e \end{bmatrix}.$$

According to the theorem of conditional distribution (Graybill 1976, p 106), the conditional expectations are

$$E(a_q|y) = \sigma^2_{a_q}\Pi^q V^{-1}(y - Xb), \tag{32}$$

$$E(a_A|y) = \sigma^2_A A V^{-1}(y - Xb) \tag{33}$$

and

$$E(e|y) = \sigma^2_e V^{-1}(y - Xb). \tag{34}$$

The conditional variances are

$$Var(a_q|y) = \sigma^2_{a_q}\Pi^q(I - \sigma^2_{a_q} V^{-1}\Pi^q), \tag{35}$$

$$Var(a_A|y) = \sigma^2_A A(I - \sigma^2_A V^{-1} A) \tag{36}$$

and

$$Var(e|y) = \sigma^2_e(I - V^{-1}\sigma^2_e). \tag{37}$$

These conditional expectations and variances are then used to calculate the conditional expectations of the following quadratic terms:

$$E[a_q^T(\Pi^q)^{-1}a_q|y] = E(a_q^T|y)(\Pi^q)^{-1}E(a_q|y) + Tr[Var(a_q|y)(\Pi^q)^{-1}], \tag{38}$$

$$E(a_A^T A^{-1} a_A|y) = E(a_A^T|y)A^{-1}E(a_A|y) + Tr[Var(a_A|y)A^{-1}] \tag{39}$$

and

$$E(e^T e|y) = E(e^T|y)E(e|y) + Tr[Var(e|y)]. \quad (40)$$

The M-step is quite simple. It only involves replacing the ML estimates in the complete data case by the conditional expectations, as given below,

$$\sigma^2_{a_q} = \frac{1}{N} E[a_q^T (\Pi^q)^{-1} a_q | y], \quad (41)$$

$$\sigma^2_A = \frac{1}{N} E(a_A^T A^{-1} a_A | y), \quad (42)$$

$$\sigma^2_e = \frac{1}{N} E(e^T e|y), \quad (43)$$

and

$$b = (X^T V^{-1} X)^{-1} X^T V^{-1} Y. \quad (44)$$

The estimate of θ in the M-step is denoted by $\theta^{(t+1)}$ which is used to initiate the E-step in the next cycle. The iteration continues until a convergence criterion is satisfied, e.g.,

$$\left\| \theta^{(t+1)} - \theta^{(t)} \right\| < \varepsilon. \quad (45)$$

Although the EM algorithm appears to involve $(\Pi^q)^{-1}$ which does not always exist, directly computing $(\Pi^q)^{-1}$ is not needed. This can be shown by

$$\begin{aligned} E[a_q^T (\Pi^q)^{-1} a_q | y] &= E(a_q^T|y)(\Pi^q)^{-1} E(a_q|y) + Tr[(\Pi^q)^{-1} Var(a_q|y)] \\ &= (y - Xb)^T V^{-1} \sigma^4_{a_q} \Pi^q V^{-1} (y - Xb) + Tr[\sigma^2_{a_q} (I - \sigma^2_{a_q} V^{-1} \Pi^q)]. \end{aligned} \quad (46)$$

I have described two algorithms, the simplex method and the EM algorithm. Although neither method is the most efficient one in terms of computational speed, both are well behaved in the sense that convergence to a local maximum is usually guaranteed.

6. Marker assisted selection

In animal production, the ultimate goal of QTL mapping is to use the mapped QTLs to facilitate selection. Theory and methods of marker assisted selection have been well developed for line crossing experiments (Lande & Thompson 1990). Because the allelic effects are treated as fixed, a multiple

regression approach is usually applied. In outbred populations, however, because the allelic effects of QTLs are treated as random, the best linear unbiased prediction (BLUP) technique (Henderson 1975) must be applied to marker assisted selection. The BLUP marker assisted selection was first introduced by Fernando & Grossman (1989), and subsequently extended by Goddard (1992).

Assume that there are n QTLs each mapped on a particular position of a chromosome. The variances of these QTLs and the IBD matrices are assumed to be known or estimated with sufficient accuracy. The expectation step of the EM algorithm automatically provides the BLUP formula of the QTL effects, i.e.,

$$\hat{a}_q = E(a_q | y) = \sigma^2_{a_q} \Pi^q V^{-1}(y - X\hat{b}). \quad (47)$$

Therefore, the estimated breeding value (EBV) is

$$EBV = \sum_{k=1}^{n} \hat{a}_k = (\sum_{k=1}^{n} \sigma^2_{a_k} \Pi^k) V^{-1}(y - X\hat{b}). \quad (48)$$

When the number of QTLs is large and each QTL has a small effect,

$\sum_{k=1}^{n} \sigma^2_{a_k} \Pi^k \approx A\sigma^2_A$, resulting in

$$EBV \approx \sigma^2_A A V^{-1}(y - X\hat{b}), \quad (49)$$

which is the estimated breeding value under the polygenic model. Therefore, the efficiency of marker assisted selection relative pedigree selection depends on the deviation of $\sum_{k=1}^{n} \sigma^2_{a_k} \Pi^k$ from $A\sigma^2_A$. The amount of deviation is expected to be large when there are a few QTLs with uneven contributions to the trait.

In small breeding populations, mating strategy is important in maintaining genetic variance and diminishing inbreeding depression. In addition to marker-assisted selection, results of QTL mapping can be used to facilitate mating design, called marker-assisted mating. As described earlier, the reduction in additive genetic variance of a population is proportional to the average inbreeding coefficient of the population, i.e., $\sigma^2_A = \sigma^2_A(0)(1 - F)$. After individual QTLs are identified, the above equation becomes

$$\sigma_A^2 = \sum_{k=1}^{n} \sigma_{a_k}^2 (0)(1 - F_k) \qquad (50)$$

where F_k is the average inbreeding coefficient at locus k (the k-th QTL). With molecular marker information, we can manipulate matings of the selected individuals so that their offspring have minimum chance of inbreeding at more important QTLs. This will reduce the average inbreeding coefficients at these QTLs, as illustrated in the following example. Assume that a major QTL has been identified and the population size is so small that full-sib mating cannot be avoided, we can avoid inbreeding for this locus while the mating is still between full-sibs. For example, a highly polymorphic marker cosegregates with the major QTL so that the QTL genotype of each individual is actually observed. Let Q_1Q_3, Q_1Q_4, Q_2Q_3 and Q_2Q_4 be the four possible QTL genotypes of full-sibs resulted from mating parents $Q_1Q_2 \times Q_3Q_4$. Assume that the whole population consists of this single full-sib family. All matings will occur between full-sibs, but a mating between $Q_1Q_3 \times Q_2Q_4$ will generate an offspring with absolutely no chance of inbreeding at this locus. On the other hand, a mating between $Q_1Q_3 \times Q_1Q_3$ will produce an offspring with 50% of chance to be inbred. The distribution of the inbreeding event in the offspring resulted from full-sib mating can be deduced from Table 1. If a pair of full-sibs share no IBD alleles, their mating (Type I mating) will generate absolutely no inbred offspring. If a pair of full-sibs share one IBD allele, the mating (Type II mating) will produce offspring with 1/4 of chance to be inbred. A mating between full-sibs sharing two IBD alleles (Type III mating) will generate offspring with 1/2 of chance to be inbred. The probabilities of the three types of matings are 1/4, 1/2 and 1/4, respectively. With marker information, we may be able to choose all Type I full-sib mating so that their offspring are all non-inbred for the locus in consideration. If a quantitative trait is controlled by several QTLs, we should first order these QTLs by their contributions to the total phenotypic variance. Avoidance of inbreeding should be considered with high priority for larger QTLs than for smaller ones. High priority of avoidance of inbreeding should also be placed to loci that show a high level of inbreeding depression.

Discussion

Development of the random model QTL mapping is based on the animal model approach in which the allelic effects of all individuals are included as model effects. Information of allelic inheritance from parents to offspring is not required in the method, rather, only the final IBD sharings between sibs are used. Because of this, the method does not require information of parental marker linkage phases. However, if that information is available, a reduced animal model may be used (Cantet & Smith 1991). In the reduced animal model, the allelic effects of a progeny are expressed as linear functions of the allelic effects of the parents. Because a full-sib family only share a maximum of four parental alleles, the number of effects in the model can be reduced significantly when the family size is large. The functional relationships of allelic effects between parents and progenies must be derived based on the information of allelic inheritance from parents to progenies. This will need the parental marker linkage phases when multipoint method is used to infer the allelic inheritance of a linked QTL. The reduced animal model is computationally more efficient than the animal model. It may also provide better estimates of the genetic variances due to the use of linkage phase information. The animal model presented in this study, however, cannot use linkage phase information, even if it is available. To incorporate linkage phase information into the animal model, one needs to modify the multipoint method. For example, in the full-sib analysis, instead of lumping the 16 sib-pair genotypic configurations of a marker (see Table 1) into four IBD states, one should use all the 16 possible sib-pair configurations of all markers to infer the probabilities of all 16 possible configurations of a putative QTL. The 16 configurations of the QTL are then lumped into the four IBD states which are eventually applied to the animal model. Further investigation is necessary to quantify the improvement of using linkage phase information relative to the situation when linkage phases are not known.

For a given variance explained by a QTL, the IBD-based linkage analysis usually shows a lower statistical power in QTL detection than the interval mapping using line crossing experiments. This appears to be caused by the difference between the statistical methods applied, because a fixed model

approach has been used in interval mapping of line crosses. In a previous study (Xu 1998), however, I have shown that the difference is not due to statistical methods, rather, it is due to the difference in genetic samplings. A simple cross involves only two alleles. Traditional interval mapping utilizes only a single cross or family, and thus deals with a few alleles. The IBD-based linkage analysis, however, usually deals with a large number of small families and thus a large number of alleles. When the overall sample sizes and the allelic variances are the same for the two sampling strategies, the strategy with smaller number of alleles is favored and thus has a higher power. When the same sampling strategy is applied to both the fixed model and the random model approaches, I found that the two approaches have virtually identical power.

Statistical power of QTL detection highly depends on the sampling strategy. Most power studies tend to ignore the effect of sampling strategy on the statistical power. These studies usually assume that the parents sampled are heterozygotes and the variance of the allelic effects among the parents are given. The random process only occurs in the sampling of these parental alleles to form the progenies. In fact, there is another level of random process which occurs in the sampling of the parents. This type of random sampling is called genetic drift. The variance of the allelic effects of the sampled parents is a representative of the genetic variance of the population in which these parents are sampled. The overall genetic variance is a parameter, but the variance among the sampled parents is a variable because it can vary depending on the sample size of the parents and the particular set of parents selected. The sampled genetic variance has a scaled chi-square distribution with degrees of freedom determined by the number of parents sampled. Consider a single locus, q, and define $\alpha_j = v_j^s - v_j^d$ as the effect of allelic substitution of the two alleles carried by parent j, for $j = 1, \cdots, p$, where v_j^s and v_j^d are the paternal and maternal allelic effects, respectively. If the allelic effects are normally distributed, α_j will be distributed as $N(0, \sigma_{a_q}^2)$, where $\sigma_{a_q}^2$ is the additive variance of locus q in the population in which these parents are sampled. The genetic variance among the sampled parents is actually

$$s_{a_q}^2 = \frac{1}{p}\sum_{j=1}^{p}\alpha_j^2 . \qquad (51)$$

Because $\alpha_j^2 / \sigma_{a_q}^2$ has a chi-square distribution with one degree of freedom,

$$s_{a_q}^2 = \frac{\sigma_{a_q}^2}{p} \sum_{j=1}^{p} \frac{\alpha_j^2}{\sigma_{a_q}^2} = \frac{\sigma_{a_q}^2}{p} \chi_p^2 \qquad (52)$$

has a scaled chi-square distribution with p degrees of freedom. Therefore, the expectation and variance of $s_{a_q}^2$ are

$$E(s_{a_q}^2) = \frac{\sigma_{a_q}^2}{p} E(\chi_p^2) = \frac{\sigma_{a_q}^2}{p} p = \sigma_{a_q}^2 \qquad (53)$$

and

$$Var(s_{a_q}^2) = \left(\frac{\sigma_{a_q}^2}{p}\right)^2 Var(\chi_p^2) = \left(\frac{\sigma_{a_q}^2}{p}\right)^2 (2p) = 2\frac{\sigma_{a_q}^4}{p}. \qquad (54)$$

This variance is called drift variance which can be diminished by increasing p. Genetic drift can cause significant deviation of the estimated from the true genetic variance, even though the sample sizes (numbers of progenies within families) are large. From the standpoint of estimating genetic parameter, one should sample as many parents as possible to reduce the bias of estimation. However, when the total number of progenies is fixed, increasing the number of parents will decrease the numbers of progenies within families. The latter will eventually increase the error of the estimated genetic variance. There is an optimal sampling strategy in which the estimation is the best, i.e., estimation with the smallest joint effect of the bias and estimation error. From the standpoint of detecting QTL, genetic drift is not necessarily a bad thing because, by chance, $s_{a_q}^2$ may be larger than $\sigma_{a_q}^2$. If $s_{a_q}^2 > \sigma_{a_q}^2$, genetic drift has a positive influence on the power. Unfortunately, by chance, $s_{a_q}^2$ can also be smaller than $\sigma_{a_q}^2$. If so, genetic drift has a negative impact on the power. Overall, genetic drift will decrease the power because $Pr(s_{a_q}^2 < \sigma_{a_q}^2) > Pr(s_{a_q}^2 > \sigma_{a_q}^2)$ due to the asymmetrical nature of the distribution of $s_{a_q}^2$. For instance, $Pr(s_{a_q}^2 < \sigma_{a_q}^2) = 0.68$ for p = 1 and $Pr(s_{a_q}^2 < \sigma_{a_q}^2) = 0.52$ for p = 100. Numerical evaluations conducted by Muranty (1996) show that at least a few parents should be sampled to warrant a reasonable power in QTL detection. Again, when the total number of progenies is fixed, reducing drift error

by increasing the number of parents may cause a loss in power due to the decreased numbers of progenies within families. There is an optimal number of sampled parents; any deviation from the optimal number will cause a loss in power (Xu 1998).

The IBD-based variance component method is particularly designed for linkage study in outbred populations. Computer simulations show that when the mapping population consists of many small families, the method is inefficient in detecting small QTLs. Association study via the candidate gene approach may be the choice for mapping QTLs with small effects (Risch & Merikangas 1996). With the candidate gene approach, one directly observes genotypes of QTLs. The allelic effects have a global nature instead of being nested within families. The whole population is now treated as a "single big family". Because genotypes of the candidate gene are observed and segregation of the QTL is assured, genetic drift has been prevented. The number of alleles of a candidate gene is usually small, resulting in small estimation errors of the allelic effects. The consequence of no drift and small estimation errors is an increased statistical power.

Acknowledgments

This research was supported by the National Institutes of Health Grant GM55321-01 and the USDA National Research Initiative Competitive Grants Program 97-35205-5075.

References

Amos, C.I. 1994. Robust variance-components approach for assessing genetic linkage in pedigrees. Am. J. Hum. Genet. 54: 535-543.

Cantet, R. J. C. & C. Smith. 1991. Reduced animal model for marker assisted selection using best linear unbiased prediction. Genet. Res. 23: 221-233.

Dempster, A.P., N.M. Laird & D.B. Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm. J. Roy. Stat. Soc. B 39: 1-38.

Falconer, D. S. & T. F. C. Mackay. 1996. Introduction to quantitative genetics, ed 4. Longman, NY.

Fernando, R. L. & M. Grossman. 1989. Marker assisted selection using best linear unbiased prediction. Genet. Sel. Evol. 21: 467-477.

Fulker, D. W. & L. R. Cardon. 1994. A sib-pair approach to interval mapping of quantitative trait loci. Am. J. Hum. Genet. 54: 1092-1103.

Fulker, D.W., S.S. Cherny & L.R. Cardon. 1995. Multipoint interval mapping of quantitative trait loci using sib pairs. Am. J. Hum. Genet. 56: 1224-1233.

Goddard, M. E. 1992. A mixed model for analyses of data on multiple genetic markers. Theor. Appl. Genet. 83: 878-886.

Goldgar, D. E. 1990. Multipoint analysis of human quantitative genetic variation. Am. J. Hum. Genet. 47: 957-967.

Graybill, F. A. 1976. Theory and Application of the Linear Model. Duxbury Press. North Scituate, Massachusetts.

Haseman, J. K. & R. C. Elston. 1972. The investigation of linkage between a quantitative trait and a marker locus. Behave. Genet. 2: 3-19.

Henderson, C.R. 1975. Best linear unbiased estimation and prediction under a selection model. Biometrics. 31: 423-447.

Kruglyak, L., & E.S. Lander. 1995. Complete multipoint sib-pair analysis of qualitative and quantitative traits. Am. J. Hum. Genet. 57: 439-454.

Lande, R. & R. Thompson. 1990. Efficiency of marker assisted selection in the improvement of quantitative traits. Genetics 124: 743-756.

Lander, E. S. & D. Botstein. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121: 185-199.

Lynch, M. & B. Walsh. 1998. Genetics and Analysis of Quantitative Traits. Sinaure Associate, Sunderland, Massachusetts.

Muranty, H. 1996. Power of tests for quantitative trait loci detection using full-sib families in different schemes. Heredity 76: 156-165.

Nelder, J. A. & R. Mead. 1965. A simplex method for function minimization. The Computer Journal 7: 308-313.

Olson, J.M. 1995. Robust multipoint linkage analysis: an extension of the Haseman-Elston method. Genet. Epidemiol. 12: 177-193.

Risch, N. & K. Merikangas 1996. The future of genetic studies of complex human diseases. Science 273: 1516-1517.

Schork, N. J. 1993. Extended multipoint identity-by-descent analysis of human quantitative traits: efficiency, power, and modeling considerations. Am. J. Hum. Genet. 53: 1306-1393.

Xu, S. & W.R. Atchley. 1995. A random model approach to interval mapping of quantitative trait loci. Genetics 141: 1189-1197.

Xu, S. 1998. Mapping quantitative trait loci using multiple families of line crosses. Genetics 148: 517-524.

Conservation of Rare Breeds in Spanish Rural Agro-economy

DÉA DE LIMA VIDAL
Agricultural Economics Unit, University of Zaragoza, Spain
(medea.p@mailcity.com)

Introduction

The animal production systems in the Mediterranean Basin, dominated by herds of small ruminants have diverse characteristics in time and space, according to the objectives and importance of the physical, social, economic and environmental conditions. All this has lead to different organization of the herds, markets, agrobiological resource use, the reproduction calendar, etc , in accordance with management adopted. These systems are very flexible with regard to the capacity of the territory use. Within the Mediterranean Basin there are many contrasts that reflect differences in the living standard of the various peoples (Daza-Andrada, 1997).

The forage system has a well documented effect on the configuration, management and economic results of the households. The nutritional, forage and pastoral expenditures, vary according to the implemented production system (Vidal, 1997). Among the households of specialized zones, the expenditures which influence forage production can reach 70% to 74% of the total expenditure. Thus, forage yield represents the main influence on total income (Henot, 1983).

With regard to autochthonous agricultural breeds, in recent decades governmental abandoment and extended neglect have been contrasted by expansion of the territorial range and increasing economic and social importance of these breeds. In addition, some breeds or races have experienced decline and even extinction. During agricultural development, the autochthonous breeds appear to have gained equal status with commercial breeds (Sañudo, 1994).

The objectives of the present work are, to characterize and differentiate distinct livestock households of the semi-arid area of *Aragón* (Spain) with regard to specific aspects of structure and economy in relation to the forage system. In addition, the ovine racial diversity in relation to this household groups modelling, with special attention to the autochthonous breeds, is discussed.

Methodology

The data was obtained by direct survey of 101 representative farm households raising sheep or sheep and cereals, located in the semi-arid zone of *Aragón* Community (Fig.1; Vidal, 1997). The survey provided information on structural, technical and economic aspects of the households. The information covered the annual production cycle: October of 1994 - October 1995. A total of fifteen variables for the management of the forage systems was classified in three types: seven ratios of structure; four economic ratios of specific expenditures; three of global economic results; and a productive orientation indicator were considered (Table 1). These indicators, were processed statistically by Multiple Correspondence Factorial Analysis. Subsequently, the first three factors (32% of the absorbed inertia), were used in Cluster Analysis to model the households according to their implemented system.

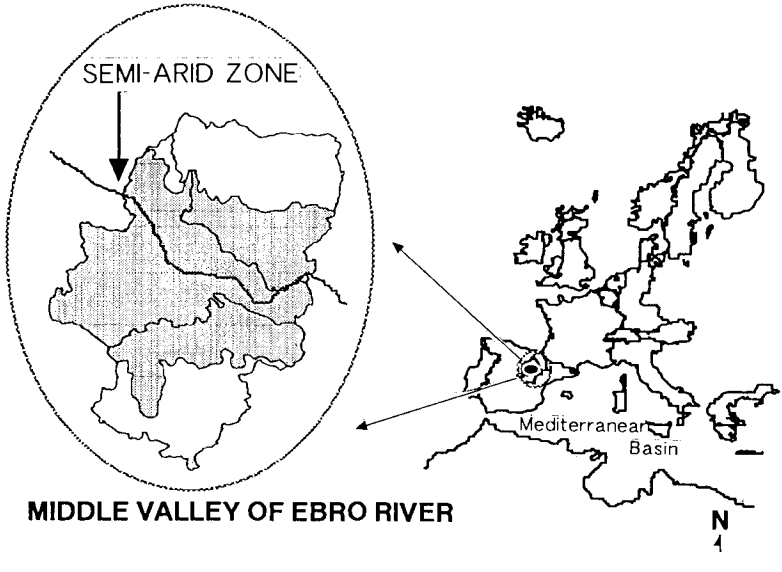


Fig.1 *Aragón* semi-arid zone delimitation according to two climatologic classifications (Spain) (Vidal, 1997)

The types of households could be differentiated by their characteristic economic and structural elements. The characteristics and determinant factors of the forage systems implemented, and their relationships to economic structure, labour and herd productivities are analyzed and discussed.

In order to analyze the racial diversity in relation to household modelling, the ovine racial classification of the Official Inventory of the Spanish Ministry of Agriculture has been used (MAPA, 1997).

Table 1. Variables used and theirs definitions

STRUCTURAL VARIABLES
1 - UAA : hectares Useful Agrarian Area.
2 - % FA/UAA: percentage of Forage Area with respect to the UAA
3 - % Irrig.SFC/FA: percentage of specific forage cultivation for sheep and goat in irrigable area with respect to the FA
4 - %Dry SFC/FA: percentage of specific forage cultivation for sheep and goat in dry area with respect to the FA
5 - SGLU/ha SFC*FA : stocking rate in units of sheep and goat with respect to the hectares of FA
6 - SGLU/ ha FL: stocking rate in sheep and goat livestock units with respect to the surfaces in hectares that are grazed not in the principal cultivation, but in by-products : Fallow Lands (fallows, stubbles andcrops)
7 - Forage Chain: expresses the degree of autonomy in the forage compilation (complete, partial, no forage mechanization, no forage area). (Variable of qualitative type).

SPECIFIC EXPENDITURES
1 - Conc.exp./SGLU: expenditures in concentrate feed for the sheep and goat with respect to SGLU
2 - Forage exp./SGLU: expenditures in voluminous feed for the sheep and goat with respect to SGLU
3 - Fix.salar./SGLU: expenditures in fixed salary in relationship to SGLU
4 - Ev. salar./SGLU: expenditures in eventual salary in relationship to SGLU

ECONOMIC RESULTS
1 - FGVA : Farm Gross Value Added
2 - FGVA /ALU: the relatIonship between the farm gross value added and the units of available total annual labour
3 - ov&capFGVA/head (pts/head): the specific ovine and caprine FGVA in relationship to the ewe or goat

PRODUCTIVE ORIENTATION
1 - % OCFP/AFP: the percentage of the ovine and caprine final production in relationship to agrarian final production.

Results and Discussion

The diversity of the forage systems studied has been grouped into five types (Fig.2). The average characteristics of these groups with reference to the variables used are shown (Table 2).

The main structural characteristic that differentiates the five types of forage systems is its character of dryness or irrigation. The *two irrigable types are referred as G1 and G2*, and differ by: (1) the size of the Useful Agrarian Area; (2) by the composition of the Forage Area (FA); and (3) availability and type of labour. Of these the FA and watered area have similar importance. On the other hand, the difference between the *three types of dry systems (G3,G4 and G5)* was based mainly on: a) the herd size; b) the composition of the FA; and c) the availability of rented grass surface by Ovine and Caprine Livestock Units. The foregoing factors (a-c) are reflected in the stocking rate related to the FA.

The economic differences in the five defined types were determined globally by the productive orientation and the proportional incidence of Ovine and Caprine Final Production within the Agrarian Final Production. The value unitary production varied with the dimension of the herd in all the defined groups. It can be said that there are economies of scale as well as influences by the different reproductive management methods in the economy of the groups of dryland systems.

In the groups of irrigable systems, as well as in those of prevailing dryland, the value unitary production was increased in relation to the importance of the Specific Forage Cultivation in the composition of the Forage Area.

There is a relationship between labour productivity and the stocking rates of the household (Hamrouni, 1993). In the groups of prevailing dryness, the increase in productivity of the labour would depend on the increase in stocking rate on the Fallow Lands, and has a positive repercussion in the specific livestock productivity. The same trend was not observed between the specific stocking rate and labour productivity.

In contrast, in irrigable systems group, the relationship shows an opposite trend: the increase in labour productivity does not correspond to an increase in the stocking rate, either by specific or by Fallow Lands. In the case of these, both, stocking rate and specific livestock productivity, decreased as the labour productivity increased. This could be caused by the productive orientation of these

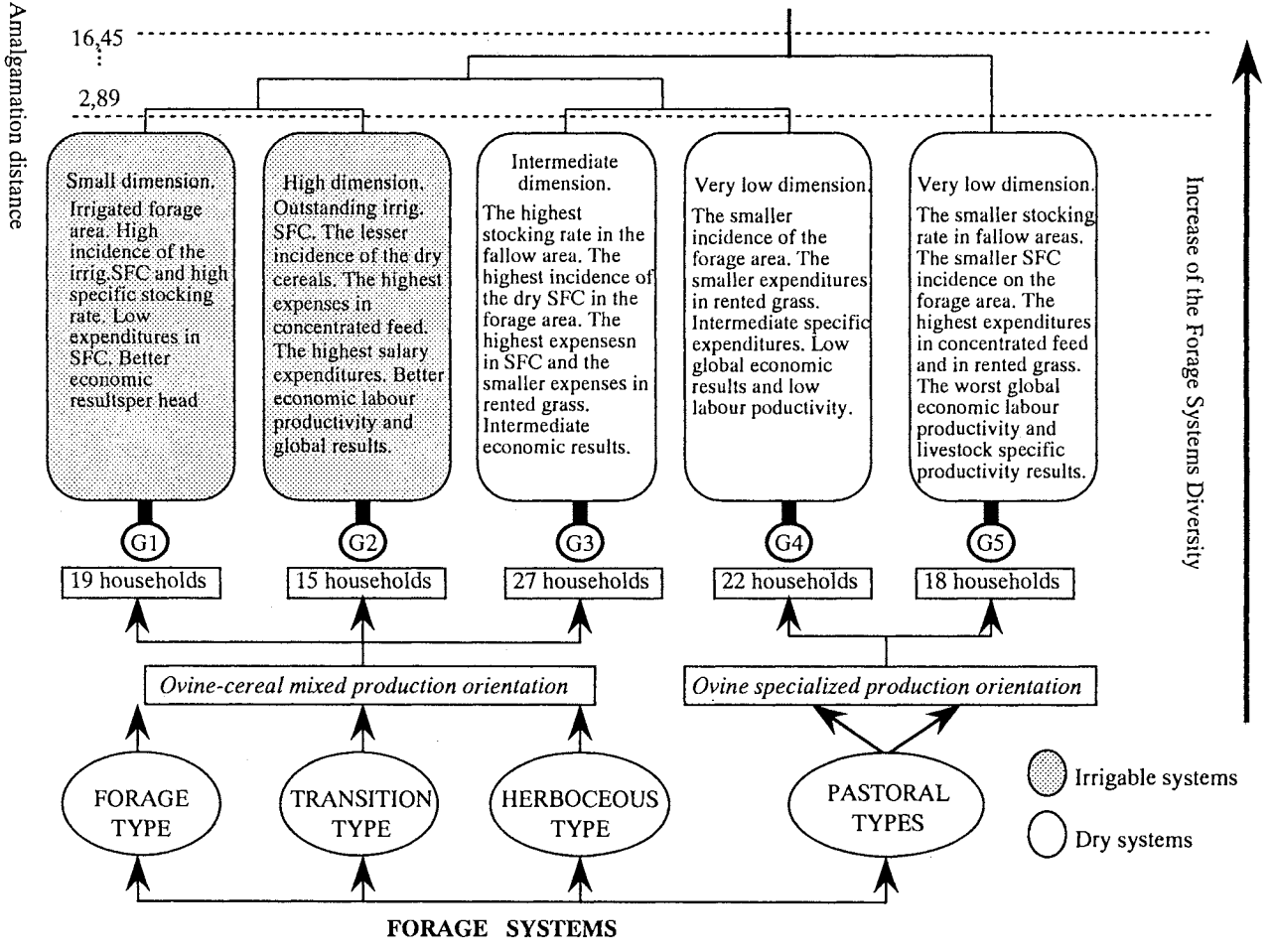


Fig.2 Typological diversity of forage systems

Table 2. Means and variation coefficient (VC) of the forage systems characteristics

	GROUP 1		GROUP 2		GROUP 3		GROUP 4		GROUP 5	
	Mean	VC	Mean	VC	Mean	VC	Mean	VC	Mean	VC
Structural Variables										
Hectares Useful Agrarian Area (UAA)	46.98	0.72	356.32	1.64	186.5	0.96	35.03	1.25	35.03	1.92
% FA/UAA	40.69	0.83	34.17	0.90	25.44	1.09	15.05	1.91	27.35	1.51
% Irrig.SFC/FA	22.78	1.37	15.90	1.41	21.14	1.51	19.80	1.80	13.89	2.34
% Dry SFC/FA	28.87	1.22	17.44	1.98	36.35	1.06	14.56	2.19	13.10	2.43
Specific livestock load (SGLU/ ha SFC*FA)	3.25	1.75	0.83	2.90	0.83	2.05	1.61	3.32	0.80	1.81
Fallow Lands livestock load (SGLU/ha FL)	3.37	1.21	2.05	1.26	10.47	3.70	6.85	1.98	2.42	2.74
Forage chain	complete or partial		partial		complete or partial		partial or don't have forage chain		partial or don't have forage chain	
Specific Expenses										
Expenses in sheep & goat concentrate feed (pts/SGLU)	8.602	0.60	10.974	0.61	8.682	0.41	9.713	0.44	11.562	0.53
Expenses in sheep & goat voluminous feed (pts/SGLU)	1.818	2.29	11.348	0.91	5.552	0.92	9.318	0.68	10.882	0.58
Expenses in sheep & goat eventual salaried (pts/SGLU)	170	2.31	2.356	1.19	0	0	0	0.00	0	0
Expenses in sheep & goat fixed salaried (pts/SGLU)	205	4.24	7.227	1.03	6.813	1.17	2.36	3.26	4.971	1.71
Productive Orientation										
% Ovine & caprine final production/agrarian final production	76.03	0.33	51.89	0.61	71.28	0.33	90.36	0.18	93.69	0.19
Economic Results										
Farm Gross Value Added (FGVA) pts	7,549,081	0.50	2,7105,991	2.04	13,476,892	0.59	4,368,434	0.43	3,468,987	0.60
FGVA/ALU (pts/ALU)	3,372,664	0.43	10,046;358	1.81	5,659,739	0.44	2,469,554	0.50	1,784,485	0.61
Ov & cap FGVA/head (pts/head)	11.064	0.25	7.032	0.38	10.547	0.17	9.003	0.22	6.521	0.34

groups, that is a mix between ovine-cereal.

The Forage Area is present in all of the types of forage systems defined. In the groups of greater availability of FA (G1 and G2), Specific Forage Cultivation prevailed, by and large in those of smaller FA surface (G4 and G5) comprised the natural grassland and scrub.

Group 1 type is "*Forage*", since it has an important Specific Forage Cultivation irrigable base, near a high Fallow Land with emphasis on alfalfa. Group 3 type is "*Herbaceous*" or "*Grain Producer*", because it represents privileged resource derivatives from the diversity in the Specific Forage Cultivation (dryness and irrigable) and the important percentage of dry cereals. On the other hand, according to Guerin *et al.* (1994), Groups 4 and 5 both possess forage area type "*Pastoral*". This types represent at the end of spontaneous vegetation or of ancient installation involving direct utilization by the animals such as: permanent meadows, mediocre lawns, scrubs, grassland and fallows. Group 2 consist of a form of FA in transition, since it possesses a combination of mechanized SFC, with an significant percentage of scrubs and grassland. This type would be the type "*Transition*" or "*Forage-Pastoral*".

The diversity of the spanish autochthonous sheep breeds used in these systems is shown (Table 3).

The *Manchega* Spanish ovine autochthonous race, whose cache is normal, appear mainly in the Forage Irrigable Systems (G1, Forage type and G2, Transition type) with ovine-cereal mixed productive orientation and better global economic results and labour productivity. This race is currently being promoted in Spain (Abecia, 1998; Sañudo, 1994), and is predominant among others autochthonous races in these irrigables systems.

In Forage Dry Systems G3, G4 and G5 (defined above) the *Ojinegra* and *Roya Bilbilitana* spanish ovine autochthonous races predominate. The *Ojinegra* race is being promoted like the *Manchega* race, while the *Roya Bilbilitana* race is being studed in Spain as a race to officially conserved by the European Union, since its cache is reduced and in danger of extinction (MAPA, 1997; FAO, 1992).

Conclusions

The *Rasa Aragonesa* race, whose cash is normal, appears in all the Forage Systems in high proportions. Nevertheless, the *Roya Bilbilitana* race, which

Table 3. The diversity of the Spanish auchthonous sheep breeds in the forage systems group

	Rasa Aragonesa Spanish Sheep Race	**Incorporated and Others Autochothonous Spanish Sheep Races***			**Non-Authocthonous Spanish Sheep Races*** **
	(households % total group)	(households % total group)	local name	(households % intra-group)	(households % total group)
Group 1	42.1	31.54	Manchega Ojinegra Roya Bilbilitana I. or W.** races	15.6 10.6 5.3 0	26.36
Group 2	26.7	46.67	Manchega Ojinegra Roya Bilbilitana Ripollesa I. or W.** races	20 6.7 6.7 6.7 6.7	26.63
Group 3	59.25	25.92	Ojinegra Roya Bilbilitana Segurena Montanna Pirenaica I. or W.** races	14.8 3.7 3.7 3.7 3.73	14.83
Group 4	59	31.72	Ojinegra Roya Bilbilitana Castellana I. or W.** races	18.18 9.1 4.5 0	9.28
Group 5	44.44	44.44	Manchega Ojinegra Roya Bilbilitana "Mask" I. or W.** races	16.6 5.6 16.6 5.6 0	11.12

* Not excluding cases with the presence of foreign races in the Group 3
** I. : incorporated; W. : without definited race
*** Not excluding cases with presence of W. in the Group 4

is rare, also shows viability in all the modelling systems. This demonstrates the high capacity for adaptation of these races to the great diversity of structural and economics conditions of the studied area, with emphasis on the dryland systems.

The analysis verifies that an "in situ" conservation strategy is appropriate to farm households in semi-arid conditions. Specific economic to help these nucleus herds will need to be developed.

References

Abecia, A. 1998. Personal communication. Faculty of Veterinaria, Zaragoza University, Spain.

Daza-Andrada A. 1997. Reproducción y sistemas de explotación del ganado ovino. Ed. Mundi-Prensa, Madrid, Spain. 384 pp.

FAO, 1992. Animal Genetic Resources Information. FAO Animal Production paper, 10: 81 pp.

Guerin, G., Leger, F. and Pflimlin, A. 1994. "Stratégie d'alimentation: méthodologie d'analyse et de diagnostic de l'utilisation et de la gestion des surfaces fourragères et pastorales". INRA, Department Alimentation et Systèmes Fourragers et Départment Systèmes d'Exploitation d'Elevage. 36 pp.

Hamrouni, S. 1993. Tipificación estructural, técnica y económica y nivel de intensificación de las explotaciones ovinas del Valle Medio del Ebro en el marco y dinámica de las agriculturas desarrolladas. Ph.D. Thesis, Unidad de Economía Agraria de la Universidad de Zaragoza, España. 400 pp + annexes.

Henot, A.Y. 1983. Incidence des systèmes fourragers sur le revenu d'une explotaition latière en Bretagne. Etude par simulation. Fourrages 96: 333.

MAPA (Ministerio de Agricultura, Pesca y Alimentación, España), 1997. Evaluación de recursos agrarios. Dirección General de la Producción Agraria. Madrid, Spain.

Sañudo, C. 1994. Las razas ganaderas tradicionales (II). Su importancia y recuperación. Mundo ganadero 4: 22-29.

Vidal, D. de L. 1997. The incidence of the Forage System in the Ovine and Ovine-Cereal houlseholds in Semi-Arid zones of the Ebro River Middle Valley (Spain). Ph.D. Thesis, Unidad de Economía Agraria de la Universidad de Zaragoza, España. 517 pp + annexes.

Genetic Distance and Classification of Domestic Animals Using Genetic Markers

YOSHITAKA NAGAMINE, HIDEAKI TAKAHASHI* and HISATO TAKEDA*
Tohoku National Agricultural Experiment Station
Shimo-kuriyagawa, Morioka, 020-0198, Japan
*National Institute of Agrobiological Resources
Kannon-dai, Tsukuba, 308-8602, Japan

Abstract

Genetic distance is used to classify and understand evolutionary relationships between populations, such as species, which have long divergent time. The properties of genetic distance measures in classification have not been investigated. We compared the accuracy of 6 genetic distance measures to classify populations which have recently diverged in simulation studies. Effect of varied population sizes from generation to generation were also considered. Nei's standard distance and minimum distance generally showed the best results under various conditions. However, the difference in accuracy among distance measures was not significant. The number of applied loci had a larger effect on accuracy than the type of distance measures. The four distance measures had a linear relationship with generation at a constant population size. None of distance measures, however, had a linear relationship when population size varied. The effect of varied population size on the average distance was large over the first few generations but became less when number of generations was increased.

Introduction

Various methods to measure the genetic distances between populations have been described by Nei (1987). Properties of these genetic distance measures using genetic markers, such as blood and protein polymorphism and micro satellite DNA, have been discussed mainly from an evolutionary perspective. Genetic distances help understand evolutionary relationships between populations.

Conservation of traditional livestock breeds is receiving increasing international attention (Hall and Bradley 1995). Traditional livestock breeds may carry useful genes for agriculture. In addition, indigenous breeds are thought worthy to conserve for their cultural merit. Measuring the genetic distance between breeds in a species provides information on their genetic diversity and aids classification of populations. The main characteristics of livestock breeds compared to other populations

in evolutionary perspective are number of generations from the point of divergence and role of mutation. Since livestock are domesticated and improved by man, the divergent time between breeds is short from an evolutionary perspective. The role of mutation of marker genes might be very small to make the genetic difference between two breeds. Consequently it is necessary to determine if the properties of genetic distance measures are appropriate for understand the relationships among livestock breeds. The objectives of this paper were to investigate the accuracy of classification using various genetic distance measures with small number of generations in simulation study. The effect of varied population sizes from generation to generation are considered to simulate the realistic divergence between breeds.

Method of computing simulation

Genetic distance measures: The property of eight genetic distance measures were compared. Mathematical symbols used mainly followed the paper by Takezaki and Nei (1996). Here, X_{ij} and Y_{ij} are the frequencies of the i th allele at the j th locus in population x and y, respectively. m_j is the number alleles at the j th locus and r is the number of loci.

1) Sanghvi's (1953) distance

$$X^2=(1/r)\sum_j^r \sum_i^{mj} (x_{ij}-y_{ij})^2 / (x_{ij}+y_{ij})$$

The Sanghvi's distance was multiplied by 2 originally but we ignored the multiplier 2 because it had no effect in this simulation.

2) Roger's (1972) distance

$$D_R=(1/r)\sum_j^r \sqrt{\sum_i^{mj} (x_{ij}-y_{ij})^2 / 2}$$

3) Nei's (1973) minimum genetic distance

$$D_m = (J_X + J_Y) / 2 - J_{XY}$$

4) Nei's (1972) standard genetic distance

$$D_S = -\ln(J_{XY} / \sqrt{J_X J_Y})$$

Here, $$J_x = (1/r)\sum_j^r \sum_i^{mj} X_{ij}^2$$

$$J_y = (1/r)\sum_j^r \sum_i^{mj} y_{ij}^2$$

$$J_{xy} = (1/r)\sum_j^r \sum_i^{mj} x_{ij} y_{ij}$$

5) Nei's (1983) D_A

$$D_A = (1/r)\sum_j^r (1 - \sum_i^{mj} \sqrt{x_{ij} y_{ij}})$$

6) Cavalli-Sforza and Edwards' (1967) distance

$$Dc = (2/r\pi)\sum_j^r \sqrt{2(1 - \sum_i^{mj} \sqrt{x_{ij} y_{ij}})}$$

Model tree and population: A simple model tree was used to test the accuracy of the 6 genetic distance measures (Figure 1). Twenty populations were split from the base population and they were named the first split populations. They were monoecious diploid populations and individuals were mated randomly in each population through 20 generations. Each population was split into two subpopulations to generate the 21st generation. Then 40 subpopulations, the second split populations, were generated. Individuals in each second split population were mated randomly for 5 to 20 generations. Genetic distances between second split populations at the 5th, 10th, 15th and 20th generation after generating subpopulation, i.e. 25th, 30th, 35th and 40th generations from the base population, were examined. We did 3 types of simulation with different conditions regarding population sizes: (1) All populations, 20 of the first and 40 of the second split population, had size 100 constantly, (2) each population chose a size, 90, 100 or 110 randomly, or (3) 80,100 or 120 randomly from the first to the final generation.

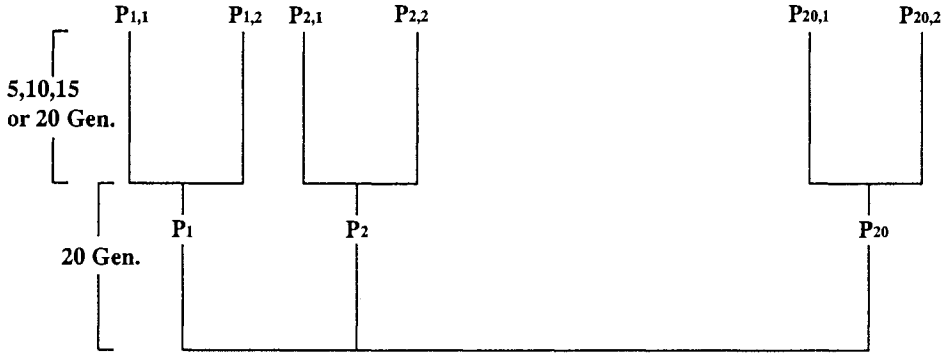


Fig.1 Model tree used in simulation. The first split populations, P_1, P_2,.... and P_{20} are mated randomly through 20 generations and the second split populations, $P_{1,1}$, $P_{1,2}$,... and $P_{20,2}$ are mated randomly through 5, 10, 15 or 20 generations.

Loci and alleles: Ten and twenty loci were generated and 2 allelic genes were assumed at each locus. Since one out of two genes were chosen randomly with equal probability to generate the base population, expected heterozygosity at each locus was 0.5.

Accuracy of distance measures: Since there were 40 subpopulations, a particular subpopulation could have 39 combinations of distance with the other subpopulations. The expected distance between a pair of subpopulations, which are from the same first split population, should be minimum. Subpopulations, i.e. the second split subpopulations, from the first split population P_1 were designated to be $P_{1,1}$ and $P_{1,2}$. $P_{1,1}$ might have 39 combinations with all other subpopulations including the combination with its partner subpopulations $P_{1,2}$. $P_{1,2}$ might also have 38 combinations if the combination with $P_{1,1}$ was excluded. The genetic distances between a pair of subpopulations $P_{1,1}$ and $P_{1,2}$ was expected to be minimum from 77 (= 39 + 38) combinations. When the distance between a pair of subpopulations had the minimum distance from 77 combinations, the genetic distance measure was judged as correct or accurate.

Results and Discussion

Number of correct pairs: The average number of correct pairs which showed the

minimum distances between paired subpopulations from 20 pairs after 200 replications is shown (Table 1 and 2). When a larger number of generations and varied population sizes were applied, the number of correct pairs was reduced. The number of correct pairs for Sanghvi's distance, for example, were 16.08 at the 5th generation with a constant population size, 100, and the number was reduced to 4.36 at the 20th generation. Varied population sizes resulted a reduction in number of correct pairs. Number of correct pairs was reduced from 16.08 to 11.99 with varied population sizes, between 90 and 110, for Sanghvi's distance. When population sizes varied between 80 and 120, it was reduced to 7.40 at the 5th generation. The largest number of correct pairs were observed from Nei's standard genetic distance (Table 1). The smallest number of correct pairs were generally observed from Cavalli-Sforza and Edwards' distance. Nei's standard distance showed the largest number and Nei's minimum distance showed the second largest numbers in all generations at any varied population size (Table 1 and 2). Generally, Cavalli-Sforza and Edwards' distance showed the smallest number and Roger's distance showed the second smallest number of correct pairs. However, among the 6 measures of genetic distance, no significant differences were observed under the same conditions for generation and population size. Increasing the number of loci from 10 to 20, increased the number of correct pairs significantly (< 1%) for the all distance measures. Under the same conditions for population size and generation, the minimum number of correct pairs from all distance measures with 20 loci was always larger than the maximum number of correct pairs from all measures with 10 loci. When 10 loci was used, for example, with constant population size at the 20th generation, the maximum number was 4.91, which was obtained by Nei's standard distance, from all distance measures. Though this value was significantly (<1%) smaller than the minimum value of 8.65, from Cavalli-Sforza and Edwards' distance with 20 loci. When a large number of generations was applied, the effect of loci number was very clear. Number of correct pairs from Nei's standard distance measure at 20th generation, for example, increased from 4.91 with 10 loci to 10.66 with 20 loci.

Relationship between distance and generation: At generation 0 of the subpopulations, which is the 20th generation from the base population, expected genetic distance is zero. A distance measures, except Roger's distance and Cavalli-Sforza and Edwards' distance, showed the liner relationship with generation. None of distance measures, however, had a linear relationship when population size varied. From the 0 to the 5th generation, distances increased rapidly in all measures with population sizes varying

Table 1. Number of correct pairs from 20 pairs using 10 loci

Generation and Pop. size	Distance measures					
	Sanghvi	Roger	Nei Min	Nei St	Nei DA	Cavalli
Gen.5						
100	16.08 (2.08)	15.51 (2.18)	16.20 (2.04)	16.34 (1.99)	16.06 (2.07)	15.34 (2.17)
90-110	11.99 (2.56)	11.42 (2.51)	12.18 (2.51)	12.36 (2.50)	11.98 (2.54)	11.28 (2.52)
80-120	7.40 (2.31)	6.93 (2.29)	7.55 (2.27)	7.63 (2.30)	7.43 (2.31)	6.94 (2.30)
Gen.10						
100	9.87 (2.58)	9.50 (2.34)	10.07 (2.50)	10.36 (2.48)	9.85 (2.57)	9.19 (2.44)
90-110	6.90 (2.27)	6.61 (2.34)	7.10 (2.28)	7.29 (2.28)	6.88 (2.30)	6.40 (2.37)
80-120	4.38 (1.84)	4.11 (1.86)	4.52 (1.92)	4.66 (1.95)	4.36 (1.83)	3.98 (1.87)
Gen.15						
100	6.28 (2.21)	6.12 (2.30)	6.67 (2.44)	6.94 (2.45)	6.18 (2.20)	5.64 (2.32)
90-110	4.56 (1.89)	4.45 (2.01)	4.65 (1.95)	4.89 (2.02)	4.49 (1.89)	4.15 (1.96)
80-120	3.05 (1.85)	2.93 (1.74)	3.17 (1.83)	3.30 (1.88)	3.01 (1.83)	2.73 (1.76)
Gen.20						
100	4.36 (1.90)	4.40 (1.95)	4.60 (1.96)	4.91 (2.00)	4.28 (1.89)	3.95 (1.89)
90-110	3.20 (1.62)	3.14 (1.70)	3.40 (1.78)	3.58 (1.80)	3.12 (1.62)	2.89 (1.61)
80-120	2.27 (1.47)	2.29 (1.42)	2.36 (1.51)	2.49 (1.53)	2.25 (1.49)	2.14 (1.36)

Nei Min: Nei's minimum distance, Nei St: Nei's standard distances.
Cavalli: Cavalli-Sforza and Edwards' distance.
Generation: Gen.5, Gen10, Gen.15 and Gen.20 are 5th,10th,15th and 20th generation, respectively, after the divergence.
Pop.size: Population sizes are (1) constant at 100 (2) varied 90,100 or 110 or (3) varied 80,100 or 120.

Table 2. Number of correct pairs from 20 pairs using 20 loci

Generation and Pop. size	Distance measures					
	Sanghvi	Roger	Nei Min	Nei St	Nei DA	Cavalli
Gen.5						
100	19.82 (0.47)	19.77 (0.53)	19.85 (0.45)	19.87 (0.42)	19.82 (0.47)	19.73 (0.58)
90-110	17.36 (1.65)	16.77 (1.81)	17.41 (1.69)	17.52 (1.60)	17.39 (1.61)	16.72 (1.82)
80-120	11.36 (2.29)	10.65 (2.24)	11.36 (2.36)	11.35 (2.37)	11.38 (2.30)	10.62 (2.23)
Gen.10						
100	17.32 (1.78)	16.66 (2.04)	17.52 (1.66)	17.68 (1.66)	17.27 (1.81)	16.25 (2.18)
90-110	13.16 (2.32)	12.32 (2.53)	13.47 (2.26)	13.71 (2.30)	13.10 (2.35)	12.14 (2.56)
80-120	8.03 (2.25)	7.48 (2.26)	8.13 (2.32)	8.22 (2.31)	8.01 (2.23)	7.35 (2.26)
Gen.15						
100	13.00 (2.37)	12.26 (2.45)	13.35 (2.34)	13.66 (2.36)	12.87 (2.31)	11.65 (2.41)
90-110	9.14 (2.69)	8.73 (2.56)	9.60 (2.56)	9.81 (2.55)	9.07 (2.64)	8.41 (2.63)
80-120	6.30 (2.39)	5.94 (2.36)	6.42 (2.39)	6.57 (2.38)	6.25 (2.39)	5.72 (2.37)
Gen.20						
100	9.91 (2.50)	9.38 (2.46)	10.20 (2.47)	10.66 (2.41)	9.69 (2.56)	8.65 (2.41)
90-110	6.61 (2.15)	6.25 (2.15)	6.99 (2.17)	7.31 (2.19)	6.50 (2.16)	5.85 (2.15)
80-120	4.37 (1.74)	4.22 (1.71)	4.58 (1.78)	4.78 (1.86)	4.37 (1.78)	3.98 (1.65)

Nei Min: Nei's minimum distance. Nei St: Nei's standard distances.
Cavalli: Cavalli-Sforza and Edwards' distance.
Generation: Gen.5, Gen10, Gen.15 and Gen.20 are 5th, 10th, 15th and 20th generation, respectively, after the divergence.
Pop. size: Population sizes are (1) constant at 100, (2) varied 90, 100 or 110 or (3) varied 80,100 or 120.

between 80 and 120. When population sizes varied between 90 and 110, average distances also increased but they rose slightly from the 0 to 5th generation. The effect of varied population size was clear over the first few generations. Though with later generations, the effect of varied population size was less.

Nei's standard distance and minimum distance generally showed the largest number of correct pairs in various conditions. Numbers of correct pairs among the genetic distance measures, however, were small and not significantly different. Nei (1978) and Gorman and Renzi (1979) recognized the number of loci was important to estimate unbiased genetic distance. In our simulation, number of loci had much larger effect on the accuracy than the type of distance measures. In electrophoretic surveys, the number of loci studied is generally few (Nei 1987). Though the number of loci detected using microsatellite marker is much larger and easy to increase (Takahashi *et al.,* 1998). Microsatellite markers can improve the accuracy in classification of domestic animals. The difference in the property of linearity with divergent time among the distance measures was found. Roger's distance and Cavalli-Sforza and Edwards' distance showed the non-linearity even with a constant population size. All distance measures, however, showed rapid increase of average distance from 0 to 5th generation with varied population size. At the 20th generation the average distance between populations with a constant populations size and varied population size was very small. This suggests that the effect of varied population size is smaller with increasing number of generations.

Acknowledgments

We thank Dr. D. A. Vaughan for his helpful comments on our manuscript.

References

Cavalli-Sforza, L.L. and Edwards, A.W.F., 1967. Phylogenetic analysis: models and estimation procedures. Amer. J. Hum. Genet. 19: 233-257.

Gorman, G.C. and Reinz, J., 1979. Genetic Distance and heterozygosity estimates in electrophoretic studies. Copeia: 242-249.

Hall, S.J.G. and Bradley, D.G.,1995. Conserving livestock breed biodiversity. Trends in Ecology and Evolution 10: 267-270.

Nei, M., 1972. Genetic distance between populations. Amer. Nat. 106: 283-291.

Nei, M., 1973. The theory and estimation of genetic distance: In: Genetic Structure of Populations, (Morton N, ed). Honolulu: University of Hawaii: 45-54.

Nei, M., Tajima, F. and Tateno, Y., 1983. Accuracy of estimated phylogenetic trees from molecuar data.

J. Mol. Evol. 19: 153-170.

Nei, M., 1987. Molecular Evolutionary Genetics. New York; Columbia University Press.

Roger, J.S.,1972. Measures of genetic similarity and genetic distance: In: Studies in Genetics VII. University of Texas Publication 7213, Austin: University of Texas: 145-153.

Sanghvi, L.D.,1953. Comparison of genetical and morphological methods for a study of biological differences. Amer. J. Phys. Anthropol. 11: 385-404.

Takahashi, H., Nirasawa, K., Nagamine, Y., Tsudzuki, M. and Yamamoto, Y., 1998. Genetic relationship among Japanese native breeds of chicken based on microsatellite DNA polymorphisms. J. Hered. 89: 543-546.

Takezaki, N. and Nei, M.,1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. Genetics 144: 389-399.

Session 2. Current Status and Perspectives of Effective Use of Sperm

The Role of Sertoli Cells in Differentiation and Exclusion of Spermatogenic Cells

AKIKO SHIRATSUCHI, TOMOJI MAEDA and YOSHINOBU NAKANISHI
Faculty of Pharmaceutical Sciences, Kanazawa University
Kanazawa, Ishikawa 920-0934, Japan

Abstract

In order to clarify the molecular basis of the functional interaction between spermatogenic and Sertoli cells, we identified molecules responsible for the specific adhesion of spermatogenic cells to Sertoli cells and for the phagocytosis of apoptotic spermatogenic cells by Sertoli cells. Adopting an expression cloning procedure, we isolated a cDNA coding region for a spermatogenic cell protein whose expression gave a cultured cell line the ability to bind to Sertoli cells. The protein, 243 amino acids with the N-terminal signal peptide and a C-terminal Cys-rich region, turned out to be the rat homologue of a testicular protein called Tpx-1 whose function was not determined. A polyclonal anti-body raised against bacterially expressed Tpx-1 significantly inhibited the binding of spermatogenic cells to Sertoli cells. These results indicated that Tpx-1 is a testicular cell adhesion molecule responsible for the specific interaction between spermatogenic and Sertoli cells. We established a quantitative phagocytosis assay in which apoptotic spermatogenic cells were selectively engulfed by Sertoli cells. Cell surface exposure of phosphatidylserine, which is observed in many apoptotic cells, occurred with apoptotic spermatogenic cells, and their phagocytosis by Sertoli cells was specifically inhibited in the presence of PS-containing liposomes. These results indicated that PS is exposed on apoptotic spermatogenic cells and serves as a marker for subsequent phagocytosis by Sertoli cells.

Introduction

The mammalian spermatogenic pathway consists of a complex series of events, such as proliferation and differentiation of spermatogonial stem cells, meiotic division of differentiating spermatogenic cells, and morphogenic maturation of spermatids. Throughout the spermatogenic pathway, spermatogenic cells remain in close contact with somatic Sertoli cells in the seminiferous tubules. It is thus presumed that a variety of materials and information are transferred between the two cell types. On the other hand, more than half of the differentiating spermatogenic cells undergo apoptotic cell death and are excluded through phagocytosis possibly by Sertoli cells. We developed a primary culture of rat testicular cells, in which spermatogenesis proceeds in terms of the occurrence of testis-specific gene expression (Tamaru *et al*, 1990). In our previous

study (Mizuno *et al.*, 1996), spermatogenic differentiation was abrogated and dead spermatogenic cells accumulated when spermatogenic cells and Sertoli cells were placed on the opposite side of a permeable membrane. These results indicated that both differentiation and exclusion of spermatogenic cells occur in primary culture depending on their direct association with Sertoli cells. We here analyzed the molecular basis for this functional association between the two cell types.

Materials and Methods

Primary culture of rat testicular cells---Dispersed testicular cells were prepared from testes dissected from 20-day-old Donryu rats and primary cultured at 32.5 °C in a mixture of F12 and L15 media (1:1) with 10% fetal calf serum (FCS) and norepinephrine (1 μg/ml) as described previously (Nagao, 1989).

Cell adhesion assay---Testicular cells were primary cultured on coverslips (ϕ=3 cm) for 3 days. Cultured spermatogenic cells consisting mostly of spermatocytes were detached from Sertoli cells by pipetting and spun down, and the cells (5×10^5) were added to a monolayer of the Sertoli cells (1.5×10^6). After a three-hour culture, the coverslips were taken off with forceps and washed by shaking them exactly five times in the culture medium. The number of spermatogenic cells remaining attached to the Sertoli cells was determined under a phase-contrast microscope. Ten randomly-chosen microscopic fields were examined, and the cell numbers are presented relative to that of initially loaded spermatogenic cells taken as 100; i. e., the cell adhesion index. To determine the function of cloned cDNA, the binding of Jurkat Tag cells expressing the DNA to Sertoli cells was examined as described above.

Cloning of Tpx-1 cDNA---Poly(A)-containing RNA (5 μg) prepared from cultured spermatogenic cells was used to synthesize cDNA using a commercial kit (Great Lengths cDNA Synthesis Kit; Clontech), and the resulting cDNA was ligated with the pcDNAI/Amp vector (Invitrogen). The cDNA library consisted of 2×10^6 independent clones with insert sizes ranging from 0.5 to 3 kbp. The library (200 μg DNA) was introduced into Jurkat Tag cells (5×10^7) by electroporation using Gene Pulser (Bio-Rad Laboratories) at 0.25 kV with a capacitance of 960 μF. The cells were cultured in RPMI1640 medium with 10% FCS for 48 h and then subjected to screening in the cell adhesion assay with Sertoli cells. The cells remaining attached to the Sertoli cells were detached by pipetting and selected again in the same adhesion assay. Plasmid DNA was then recovered from the selected cells and amplified in *E. coli* TOP10F'. The DNA was introduced again into Jurkat Tag cells (1×10^7), and the cells were subjected to another

round of screening. After four rounds of screening, plasmid DNA was extracted from the selected cells and about 50 cDNA clones were sequenced. We obtained three clones that contained reasonably long open reading frames. The inserts of these clones were ligated with the pHook-2 vector (Invitrogen), and the resulting DNA (40 μg) was introduced into Jurkat Tag cells (1×10^7) by electroporation. After a 48-h culture, the cells expressing the cDNA (1×10^5) were collected using phOx-coated magnetic beads (Capture-Tec pHook-2 kit; Invitrogen) and subjected to the final cell adhesion assay with Sertoli cells. About 80% of the cells selected by the above procedure expressed an introduced DNA when examined with a control pHook-2-*lacZ* DNA (data not shown). Eventually, one clone coded for a protein possessing the desired activity was obtained.

Preparation of anti-Tpx-1 antibody---The region between nucleotide positions 345 and 807 (amino acids 89-243) of clone #97 was inserted into a pET-15b (Novagen)-derived His-tag vector, and the resulting DNA was introduced into *E. coli* BL21. Expression of the His-tagged Tpx-1 was induced with 1 mM isopropyl β-D-thiogalactoside. The His-tagged Tpx-1 recovered as inclusion bodies were separated on a 15% polyacrylamide gel containing SDS. A portion of the gel containing the fusion protein was excised and the protein was electroeluted using Maxyield-NP (Atto). Tpx-1 fused with glutathione S-transferase (GST) was similarly prepared by inserting the region between nucleotide positions 345 and 807 into a vector (pGX-KG; Amersham Pharmacia Biotech), and *E. coli* BL21 was transformed with the resulting DNA. The GST-Tpx-1 fusion protein was purified from *E. coli*, as was the His-tagged protein. GST was purified using glutathione Sepharose (Amersham Pharmacia Biotech) under standard procedures. Rabbits were immunized with about 0.4 mg of the purified His-tagged Tpx-1, and the antibody titer was determined by Western blots of GST-Tpx-1. The total IgG fraction was obtained from sera by successive fractionation with ammonium sulfate precipitation and DEAE-cellulose chromatography. IgG specific to Tpx-1 was further purified by affinity chromatography with His-tagged Tpx-1-conjugated Ni-NTA agarose (Qiagen).

Western blots---Proteins were separated on a 12% polyacrylamide gel containing SDS and electrophoretically transferred onto a polyvinylidene difluoride membrane (Millipore). The membrane was blocked with 0.2% skim dry milk and incubated with a primary antibody in a buffer consisting of 10 mM Tris-HCl (pH 8), 0.15 M NaCl, and 0.5% Tween 20. The membrane was washed and reacted with a secondary antibody conjugated with alkaline phosphatase. The signals were detected using the Immun-Star

system (Bio-Rad Laboratories).

Immunohistochemistry---Spermatogenic cells prepared from 20-day-old rats were fixed with 4% paraformaldehyde and 2% sucrose for 15 min at room temperature, then with methanol for 1 min at room temperature. The fixed cells were treated first with affinity-purified anti-Tpx-1 IgG, then with a FITC-conjugated secondary antibody, and examined under a confocal laser microscope.

Phagocytosis assay---Spermatogenic cells were recovered by gentle pipetting from testicular cells co-cultured in collagen-coated multiwell plates for two days. On the other hand, Sertoli cells were obtained by removing the spermatogenic cells from a co-culture of testicular cells maintained on Chamber Slides (Nunc). Most of the recovered spermatogenic cells were spermatocytes, and the Sertoli cell culture was about 90% pure. The recovered spermatogenic cells were maintained with no added cells for two days, labeled with biotin (NHS-LS-Biotin; Pierce), and added back to the Sertoli cell culture maintained in Chamber Slides. Spermatogenic cells (about 2.5×10^5) were mixed with Sertoli cells (about 3×10^4) in 0.15 ml of medium, and the phagocytosis reaction carried out at 32.5 °C for 2 h, except in the time course experiment. Phosphate-buffered saline was then added and unreacted spermatogenic cells washed out by pipetting 15 strokes two times. The mixture was further treated with trypsin (0.5 mg/ml) for 3 min at room temperature, after which those cells detached from the culture slides were removed. The remaining cells were fixed with 2% paraformaldehyde/0.1% glutaraldehyde/0.05% Triton X-100/phosphate-buffered saline. The fixed cells were supplemented with Fluorescein Avidin D (Vector) and kept for 20 min at room temperature. The biotinylated spermatogenic cells were detected under a fluorescence/phase-contrast microscope. The ratio of the number of positively stained Sertoli cells to total Sertoli cells (100-150) was determined in each microscopic field. Eight to ten fields from different culture wells were examined in each experiment, and the results statistically treated. The mean and standard deviations of a typical example from at least three independent experiments were presented as the phagocytic index. Under these conditions, we routinely obtained a phagocytic index of 13-20.

Liposome preparation---Dried lipid films containing various phospholipids (2 mmoles) were swollen in 10 mM Tris-HCl, pH 7.4/0.15 M NaCl and sonicated for 10 min on ice. The liposomes were composed of either phosphatidylcholine (PC) only, or a combination of PC and another phospholipid at a molar ratio of 7:3. Fluorescence-labeled liposomes were prepared as above in the presence of L-α-phosphatidylethanolamine-N-(lissamine rhodamine B sulfonyl) (Avanti Polar-Lipids)

at 1% of total phospholipids. The engulfment of fluorescent liposomes by Sertoli cells was analyzed using a confocal laser microscope.

Annexin V-binding assay---Spermatogenic cells, which had been maintained in single-culture for various periods, were supplemented with FITC-conjugated annexin V (Bender MedSystems) and propidium iodide, and the mixture left on ice for 15 min. The fluorescence from FITC and propidium iodide was simultaneously determined with the cultured cells (10^4) in a flowcytometer.

Results

Molecular cloning of Tpx-1 responsible for association of spermatogenic and Sertoli cells (Maeda *et al.*, 1998)

Expression cloning of a cell adhesion molecule of spermatogenic cells---When rat testicular cells were primary cultured, spermatogenic cells were maintained in association with Sertoli cells that grew as a monolayer. In order to identify the molecule(s) responsible for the specific interaction between the two cell types, we first established a quantitative cell adhesion assay. Cultured spermatogenic cells were isolated and loaded on the Sertoli cell culture, and the mixture was left for 3 h. The culture was then washed carefully and cell adhesion indices were determined as described in the Materials and Methods section. In this assay, spermatogenic cells were shown to bind to Sertoli cells much more efficiently than did a cultured T-cell line, Jurkat Tag.

We then aimed at cloning the molecule(s) responsible for this interaction, using an expression cloning procedure. We chose the human T-cell line Jurkat Tag as parent cells for the following reasons; the cells grow in suspension, do not bind to Sertoli cells, and express SV40 large T-antigen that allows a plasmid containing the SV40 replication origin to replicate extrachromosomally. A cDNA library prepared from the mRNA of cultured spermatogenic cells was introduced into Jurkat Tag cells, and the cells that acquired the ability to bind to Sertoli cells were selected as described in the Materials and Methods section. After four rounds of screening, plasmids containing the cDNA were recovered from the selected cells and sequenced. We obtained three cDNA clones that contained the complete coding sequences for reasonably large peptides. To determine whether these cDNA coded for proteins with the desired activity, the final cell adhesion assay was conducted. Two of the clones were inserted into the pHook-2 vector, and the resultant DNA was introduced into Jurkat Tag cells. We were unable to ligate the insert of the third clone with pHook-2 for an unknown reason. Cells

expressing the cDNA were selected using magnetic beads, and their ability to bind to Sertoli cells was examined. The cells expressing one of the clones bound to Sertoli cells as efficiently as did spermatogenic cells, but those expressing the other or the vector alone did not. These results showed that a cDNA clone that coded for a protein with the desired activity was obtained.

Structure of the cell adhesion protein---The cDNA clone encoded a protein with 243 amino acids including an N-terminal hydrophobic sequence and a Cys-rich region at the C-terminal half. Characteristically, the Cys residues repeated five times with an interval of eight amino acids near the C-terminus. The entire amino acid sequence of this protein showed a significant similarity to that of a testicular protein called Tpx-1 or autoantigen 1 (AA1); 85, 65, and 68% identity with mouse, guinea pig, and human proteins, respectively. We thus concluded that the protein is the rat homologue of Tpx-1. Tpx-1 is a member of the CRISP (standing for Cys-rich secretory proteins) family of proteins, which contain the N-terminal signal peptide and are rich in Cys residues at the C-terminal half.

Role of Tpx-1 in spermatogenic and Sertoli cell adhesion---In order to assess the role of Tpx-1 in the specific adhesion of spermatogenic cells to Sertoli cells, a polyclonal antibody was raised against a bacterially-expressed His-tagged Tpx-1. The specificity of the antibody was first examined by Western blots. *E. coli* proteins containing a GST-Tpx-1 fusion protein were analyzed with anti-Tpx-1 and anti-GST antibodies in the presence and absence of corresponding antigen proteins. The fusion protein of about 44 kDa was detectable with either antibody. When the reaction with primary antibodies was conducted in the presence of GST-Tpx-1 or GST, the signal with anti-Tpx-1 disappeared only in the presence of the fusion protein, while that with anti-GST was abolished by the addition of either protein. These results showed that the antibody raised against bacterially-expressed Tpx-1 specifically recognizes the Tpx-1 protein in Western blots. The same antibody was then examined to determine whether it reacts with native Tpx-1. When spermatogenic cells of 20-day-old rats were immunohistochemically analyzed with the affinity-purified anti-Tpx-1 antibody, most of the cells showed signals. We then assessed the effect of anti-Tpx-1 on the adhesion between spermatogenic and Sertoli cells. The adhesion was inhibited in a manner responding to the amount of added antibody, and the maximum inhibition was about 50%. The control normal rabbit IgG showed little effect. These results indicated that Tpx-1 is responsible, at least in part, for the adhesion between spermatogenic and Sertoli cells.

Phagocytosis of apoptotic spermatogenic cells by Sertoli cells (Shiratsuchi *et al.*, 1997)

Establishment of a quantitative phagocytosis assay---In our previous phagocytosis experiments (Mizuno *et al.*, 1996), spermatogenic cells that adhered to Sertoli cells were not rigorously distinguished from those engulfed. In the present experiment, we treated the phagocytosis reaction with trypsin to eliminate spermatogenic cells that were attached to the surface of Sertoli cells. Even after extensive washing, a significant number of spermatogenic cells remained associated with Sertoli cells. These cells were distinguishable from phagocytosed cells when examined carefully under fluorescence/phase-contrast microscopy. We speculate that the tight association of spermatogenic cells with Sertoli cells is an important step toward subsequent phagocytosis. Stained particles of various sizes were observed within the Sertoli cells, and we regarded these as phagocytosed spermatogenic cells.

In order to analyze the phagocytosis reaction in a quantitative manner, we defined the phagocytic index as follows: the number of Sertoli cells positive for phagocytosis was determined as a percentage relative to the total number of Sertoli cells present in each microscopic field. We first determined a time course for the phagocytosis reaction. Spermatogenic cells that had been single-cultured for about 40 h were subjected to the phagocytosis reaction. The reaction seemed to continue during the first 2 h and reached a plateau at about index 20. Our previous experiments suggested that Sertoli cells selectively phagocytose degenerating spermatogenic cells in culture (Mizuno *et al.*, 1996). To further examine this possibility, spermatogenic cells were single-cultured for one and three days and subjected to a phagocytosis assay. The phagocytic index increased as the culture continued, while viability of spermatogenic cells, as assessed by trypan blue exclusion, decreased. These results support the above hypothesis that dying or dead spermatogenic cells are preferable targets for phagocytosis by Sertoli cells.

Inhibition of phagocytosis by liposomes containing acidic phospholipids---For identifying the molecule(s) that participates in the cell-to-cell recognition between degenerating spermatogenic cells and Sertoli cells, we examined the effect of several compounds that are known to be recognized by the putative phagocytosis receptors present on the surface of macrophages. The synthetic RGDS peptide, which contains an amino acid sequence bound by members of the integrin superfamily, was first tested, but the phagocytic index remained unchanged in the presence of this peptide. We then examined whether phospholipids are involved in cell-to-cell recognition, by adding

liposomes that consist of various phospholipids to the phagocytosis reaction. Liposomes containing phosphatidylserine (PS), an acidic phospholipid, inhibited phagocytosis in a dose-dependent manner, whereas those containing neutral phospholipids, PC and phosphatidylethanolamine, had little effect. The inhibitory effect was not specific for PS but seemed to be common to the acidic phospholipid; the addition of liposomes containing either phosphatidylinositol or cardiolipin, caused a significant reduction in the phagocytic index. Although the efficacy of inhibition somewhat varied among those anionic liposomes, the significance of this difference is not certain at the present time. These results suggest the involvement of acidic phospholipids in the phagocytosis of spermatogenic cells by Sertoli cells. Phosphoester compounds related to PS were examined next, to determine whether they affected spermatogenic cell phagocytosis. The addition of glycerophosphoryl-L-serine led to a significant reduction of phagocytosis. Phospho-L-serine showed marginal inhibition at higher concentrations whereas phagocytosis was unaffected in the presence of its optical isomer, phospho-D-serine. These results suggest that the inhibitory effect of PS-liposomes was executed through not only the serine residue, but also through the more complicated structure of the phospholipid.

To examine the way these liposomes inhibit phagocytosis, fluorescence-labeled liposomes were used instead of spermatogenic cells as targets for Sertoli cells. Liposomes containing acidic phospholipids were more efficiently engulfed than were liposomes consisting of neutral phospholipids. This coincided well with the above results, indicating that anionic liposomes compete with spermatogenic cells to be phagocytosed by Sertoli cells.

Exposure of PS to the surface of dying spermatogenic cells---Since the above results suggested that PS was recognized by Sertoli cells, we examined whether degenerating spermatogenic cells expose PS, which is normally restricted in the inner leaflet of the plasma membrane, to the cell surface. Spermatogenic cells were single-cultured for various periods and subjected to flowcytometric analysis with FITC-labeled annexin V, which specifically binds to PS. Since it was necessary to detect PS on the surface of spermatogenic cells, we analyzed only those cells whose plasma membrane remained intact. For that purpose, the cells were simultaneously treated with propidium iodide, which binds to DNA and stains the nucleus of cells whose plasma membranes are damaged and permealized. We observed two distinct populations of spermatogenic cells in terms of their propidium iodide positivity; it was presumed that the cells with less staining possessed intact plasma membranes, while the membranes of cells that were

more intensely stained with the reagent were damaged. The ratio of propidium iodide-positive cells increased as the culture period was prolonged. This indicated that spermatogenic cells degenerate during single-culture.

When the binding of annexin V to cells that presumably possessed intact plasma membranes was analyzed, two peaks were found, which probably represented annexin V-negative and -positive cell populations. On the other hand, most of the propidium iodide-positive cells appeared to be also positive with annexin V. It is likely that annexin V bound to PS present in both the outer and inner leaflets of the plasma membrane of damaged cells. As the culture continued, the relative number of cells which were propidium iodide-negative and FITC-positive, gradually increased. These results suggest that spermatogenic cells that are dying, but which still retain membrane integrity, are bound by annexin V. From the above results, we presume that during the early stages of apoptotic death, PS translocates from the inner leaflet to the outer leaflet of the spermatogenic cell membrane and serves as a signal for phagocytosis by Sertoli cells.

Discussion

Tpx-1 as a testicular cell adhesion molecule

A number of cell adhesion proteins such as cadherins, Ca^{2+}-independent cell adhesion molecules and integrins exist in the testis, seemingly on the surface of both spermatogenic and Sertoli cells. Although such proteins are likely candidates responsible for the specific contact between these two cell types, no direct evidence of this has been presented. We attempted in this study to isolate the cDNA for the molecule(s) that defines the binding of rat spermatogenic cells to Sertoli cells, employing an expression cloning method. One such cDNA clone, obtained after repeated screening in a cell adhesion assay with Sertoli cells, coded for a protein distinct from authentic cell adhesion molecules. The protein turned out to be the rat homologue of a known testicular protein called Tpx-1 or AA1. The synthesis of the mRNA and protein of AA1 (guinea pig Tpx-1) becomes detectable at the pachytene spermatocyte stage, reaching a peak in round spermatids, and it is thereafter stably present in spermatogenic cells. Tpx-1 may not be solely responsible for the adhesion of spermatogenic cells to Sertoli cells, since an anti-Tpx-1 polyclonal antibody inhibited it by only 50%. Other testicular proteins are likely to also be involved in the specific interaction between the two cell types.

Tpx-1 was first identified as a protein encoded by one of the randomly-cloned

genomic fragments from mouse chromosome 17. It was eventually revealed to be the mouse homologue of a guinea pig protein named AA1, a component of the acrosome. Tpx-1 consists of about 240 amino acids with a presumed N-terminal signal peptide and a Cys-rich C-terminal half, with no sequences identified that are considered to be transmembrane regions or N-linked glycosylation sites. Within the Cys-rich region, five Cys residues appear every nine amino acids near the C-terminus. These characteristics of Tpx-1 indicate that it is a member of the CRISP family of proteins. Another member of this family is an epididymal protein called acidic epididymal glycoprotein (AEG) that is synthesized in the epididymal epithelium and exists attached to the plasma membrane covering the sperm head. AEG has been suggested to be involved in sperm-egg fusion. Since more than 55% of the amino acid sequence of Tpx-1 is identical to that of AEG, these two proteins could possess similar biological functions. We here showed that Tpx-1 functions as a cell adhesion protein not for the sperm-egg interaction but for the association between spermatogenic and Sertoli cells.

PS-dependent phagocytosis of apoptotic spermatogenic cells by Sertoli cells

By adopting the phagocytic index, it became feasible to quantitatively analyze spermatogenic cell phagocytosis by Sertoli cells. We showed that Sertoli cells are responsible, at least in culture, for heterophagic elimination of degenerating spermatocytes. As discussed below, Sertoli cells appeared to recognize and phagocytose apoptotic cells in a manner similar to that of macrophages, a professional phagocyte. Macrophages attack a variety of cell types by infiltrating various areas. Sertoli cells however, are presumed to phagocytose only testicular cells since they are localized within the seminiferous tubules in the testis.

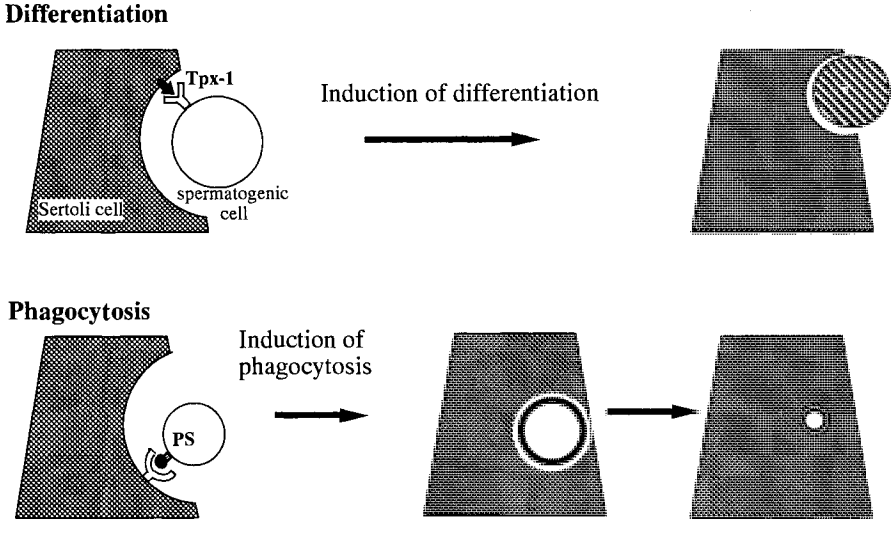
We tried several approaches in order to clarify the mechanism by which Sertoli cells selectively recognize and phagocytose apoptotic spermatogenic cells. $\alpha_v\beta_3$ integrin, which mediates phagocytosis of apoptotic neutrophils by macrophages, was seemingly not involved in spermatogenic cell phagocytosis, since the addition of the RGDS peptide, which binds to members of the integrin superfamily and inhibits their interaction with specific ligands, did not influence spermatogenic cell uptake by Sertoli cells. In contrast, acidic phospholipids, PS in particular, most likely act as a phagocytosis signal for spermatogenic cells. Liposomes containing acidic phospholipids, when present in the phagocytosis reaction, brought about a great reduction in the phagocytic index, whereas the addition of liposomes containing neutral phospholipids had little effect. As degeneration of the cultured spermatogenic cells progressed, we observed a significant increase in the number of cells that exposed PS

to the cell surface. These results suggest that the amount of cell surface PS, and probably other acidic phospholipids as well, increased during the apoptosis of spermatogenic cells and served as a signal for phagocytosis by Sertoli cells.

Membrane phospholipids of normal cells are localized asymmetrically with regard to the two leaflets of the membrane bilayer; that is, PC and sphingomyelin are mostly present in the outer leaflet, while other phospholipids including anionic PS, phosphatidylinositol and phosphatidic acid, are confined to the inner leaflet. It has been suggested that such asymmetry is lost upon induction of apoptosis, and that all phospholipids are then redistributed evenly in the membrane bilayer. This would result in the exposure of phospholipids that normally exist in the cytoplasmic side to the cell surface, and these phospholipids could help in the distinguishing of apoptotic cells from normal cells. Among those phospholipids, PS has been proposed to be a phagocytosis signal. Translocation of PS from the inner to the outer plasma membrane leaflet has been reported with a variety of apoptotic cells, including thymocytes, vascular smooth muscle cells, and cultured cell lines, as well as with aged red blood cells. We showed here that this is also the case with apoptotic spermatogenic cells.

Cell surface PS was shown to be responsible, at least in part, for thymocyte phagocytosis by macrophages and homophagic elimination of vascular smooth muscle cells, since the addition of PS-containing liposomes significantly inhibited phagocytosis. PS thus most likely serves as a common phagocytosis signal for several different phagocytes, such as macrophages, vascular smooth muscle cells, and Sertoli cells, but the modes of recognition of target cells by these phagocytes, are presumably somewhat different. We speculate that phagocytosis receptors present on the surface of macrophages and vascular smooth muscle cells are related to, but distinct from, those of Sertoli cells. It has been proposed that scavenger receptors are responsible for recognition of cell surface PS. Among members of the scavenger receptor family, CD36, SR-BI, and macrosialin bind to membrane acidic phospholipids including PS. Class B scavenger receptors, CD36 and SR-BI, were also shown to be involved in recognition and phagocytosis of apoptotic cells. However, only a limited number of examples have been reported, and the identity of the phagocyte receptor for PS is still a matter of conjecture. Whether Sertoli cells express any particular type of scavenger receptor, is under investigation.

In conclusion, our findings should contribute to a deeper understanding as to the functional interaction between spermatogenic cells and Sertoli cells (see the figure below).



References

Maeda, T., Sakashita, M., Ohba, Y. and Nakanishi, Y. 1998. Molecular cloning of the Tpx-1 responsible for the interaction between spermatogenic and Sertoli cells. Biochem. Biophys. Res. Commun. 248: 140-146.

Mizuno, K., Shiratsuchi, A., Masamune, Y. and Nakanishi, Y. 1996. The role of Sertoli cells in the differentiation and exclusion of rat testicular germ cells in primary culture. Cell Death Differ. 3: 119-123.

Nagao, Y. 1989. Viability of meiotic prophase spermatocytes of rats is facilitated in primary culture of dispersed testicular cells on collagen gel by supplementing epinephrine or norepinephrine: evidence that meiotic prophase spermatocytes complete meiotic divisions in vitro. In Vitro Cell. Dev. Biol. 25: 1088-1098.

Shiratsuchi, A., Umeda, M., Ohba, Y. and Nakanishi, Y. 1997. Recognition of phosphatidylserine on the surface of apoptotic spermatogenic cells and subsequent phagocytosis by Sertoli cells of the rat. J. Biol. Chem. 272: 2354-2358.

Tamaru, M., Nagao, Y., Taira, M., Tatibana, M., Masamune, Y. and Nakanishi, Y. 1990. Selective activation of testis-specific genes in cultured rat spermatogenic cells. Biochim. Biophys. Acta 1049: 331-338.

Role of Epididymal Secretory Proteins in Sperm Maturation

NAOMICHI OKAMURA
College of Medical Technology & Nursing and Tsukuba Advanced Research Alliance
(TARA researcher for Sakabe-Shoun Project),
University of Tsukuba, Tsukuba, Ibaraki 305-8577, Japan

Abstract

In mammals, testicular sperm must undergo the process of maturation in the male genital tract and capacitation in the female genital tract before being able to fertilize *in vivo*. During transit through the epididymis, immature sperm are affected by the continuously changing microenvironment which is formed by the secretory and absorptive activities of the epithelial cells lining the ductus epididymidis. Many studies have been performed to identify the epididymis specific proteins which relate to the development of sperm motility and fertilizing ability. We have purified several region-specific secretory proteins from porcine epididymal fluid, cloned their cDNAs and determined their roles in the sperm maturation. Among them, two proteins are reviewed in the present paper. A 16 kDa protein (HE1 homologue) which is specifically secreted from the middle and distal caput and corpus epididymis has been found to bind free cholesterol and regulate the lipid composition of the sperm plasma membranes during the maturation in the epididymis. In addition, the 23 kDa epididymis-specific glutathione peroxidase-like protein (EGLP) which is specifically secreted from the proximal caput epididymis. This protein has been found to be enzymatically quiescent in the epididymal fluid but thought to protect sperm from the premature acrosome reaction and maintain sperm fertilizing ability in the epididymis.

Introduction

In mammals, structurally and functionally immature sperm leaving the testis change into sperm possessed with fertilizing ability during the transit through the epididymis. This process is called "post testicular sperm maturation". The epithelial cells of the ductus epididymidis provide the microenvironment for this process by their secretory and absorptive activities. The interaction between sperm and the epididymal secretions is essential for the development of the fertilizing capacity in sperm. Moreover, the exact molecular mechanism involved in sperm maturation remains to be elucidated.

The ductus epididymidis is differentiated morphologically and functionally from the initial segment to cauda epididymis. Different genes are expressed in the different epididymal regions, resulting in the changes in the protein contents of the luminal fluids among various epididymal regions (Fig.1). So, the immature sperm move through the duct and have sequential contact with different proteins. It is anticipated that region specific secretory proteins in the epididymis especially play key roles in sperm maturation. Isolation of the epididymis-specific gene transcripts from the epididymal cDNA library has been actively attempted by the differential screening

method (Kirchoff *et al.*, 1990; 1993; Osterhoff *et al.*, 1997). In addition, purification of epididymis specific proteins, direct determination of their biological activities and cloning their cDNAs have also been tried extensively (Okamura *et al.*, 1995a; 1995b; 1997; Syntin *et al.*, 1996; Jin *et al.*, 1997). Among these proteins, two epididymis-specific proteins which were purified and whose physiological functions were determined (Kiuchi *et al.*, 1997; Okamura *et al.*, 1997) are reviewed in the present paper with respect to the interaction between these proteins and the epididymal sperm in their maturation process.

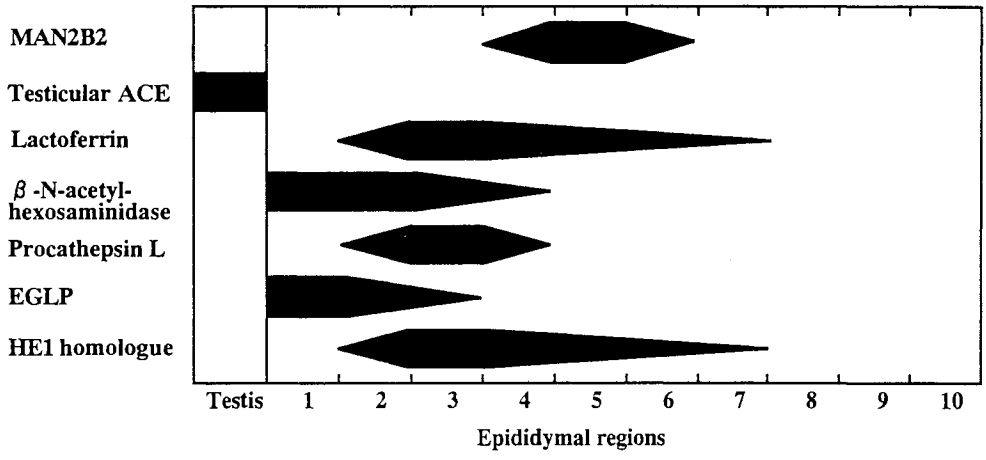


Fig.1 Diagram of the region specific gene expression in the porcine epididymis. Porcine epididymis was divided into ten regions. Regions 1, 2, 3 and 4 are proximal, middle, anterior half of distal and posterior half of distal caput epididymis, respectively. Regions 5, 6 and 7 are proximal, middle and distal corpus epididymis, respectively. Regions 8, 9 and 10 are proximal, middle and distal cauda epididymis, respectively. The distribution of mRNA was determined by Northern blot analysis. MAN2B2: 135 kDa α-D-mannosidase. ACE: angiotensin I-converting enzyme.

16 kDa choresterol-binding protein

The change in the lipid composition of the sperm plasma membranes is one of the important features of epididymal sperm maturation. It has been shown that the total cellular contents of lipids decrease while sperm pass through the epididymis (Davis, 1980; Poulos *et al.*, 1973; 1975; Adams and Johnson, 1977; Rana *et al.*, 1991). Phospholipids loss during the sperm maturation is highly selective; whereas choline phospholipids are rather stable, the amounts of phosphatidylethanolamine, phosphatidylserine and phosphatidylinosotol decline greatly (Jones, 1998). The

cholesterol content in sperm membranes have also been shown to decrease in boar (Nikolopoulou, Soucek and Vary, 1985), hamster (Awano, Kawaguchi and Mohri, 1993), rat (Hall, Hadley and Doman, 1991), stallion (Lopez and Souza, 1991), and ram (Parks and Hammerstedt, 1985), whereas they increase in goat (Adams and Johnson, 1977) during maturation. The mechanisms that mediate these changes in lipid content during the sperm maturation have not been fully elucidated yet. It is possible that some lipid-exchange proteins in epididymal fluid would mediate the transfer of phospholipids and/or cholesterol from the sperm plasma membrane to other acceptors in the epididymal fluid.

We purified a 16 kDa protein from the porcine cauda epididymal fluid and identified it as a cholesterol-binding protein. The 16 kDa protein is a major cholesterol-binding protein in epididymal fluid (Fig.2). The 16 kDa protein amounted to about 1% of total proteins in the porcine cauda epididymal fluid. This protein was also found to be expressed specifically in the distal caput and corpus epididymis and to be absent in testis, rete testis, and the proximal caput and cauda epididymis. It is first secreted into the luminal fluid as a 19 kDa glycoprotein, whose sugar chain(s) are processed to form 16 kDa protein during passage through the epididymis. An immunocytochemical examination, using a polyclonal antibody raised against the purified 16 kDa protein, showed that the 16 kDa protein does not bind to the porcine epididymal sperm.

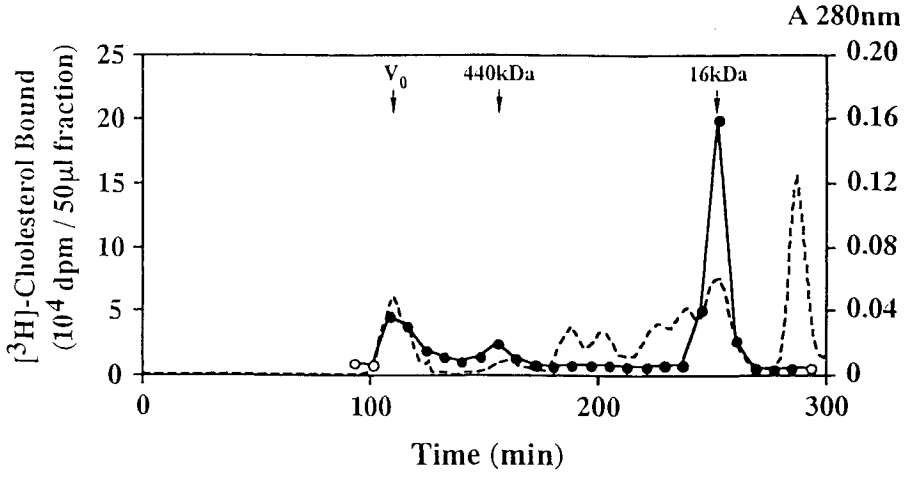


Fig. 2. Separation of the [^{3}H]-cholesterol-binding proteins in the porcine cauda epididymal fluid by Superdex 200 FPLC. The cauda epididymal fluid was incubated with 0.5 μM [^{3}H]-cholesterol at 0 °C for 30 min and then was fractionated by a Superdex 200 FPLC column. Radioactivity in 50 μl aliquots of each fraction was counted and shown by open circle. Broken line shows absorbance at 280 nm.

Analysis of its cDNA revealed that the 16 kDa protein is a porcine homologue of the major secretory protein of the human epididymis, HE1 (Fig.3; Kirchoff *et al.*, 1990). mRNA of the HE1 homologue was found in the epididymis of all species studied so far (Ellerbrock *et al.*, 1994; Perry *et al.*, 1995; Kirchoff *et al.*, 1996; Frohlich and Young, 1996; Kiuchi *et al.*, 1997). This suggests that it plays some fundamental roles in the epididymis. Baker *et al.* (1993) reported that some chromatographic fractions of ram cauda epididymal fluid and seminal plasma, which have an activity to transfer free cholesterol from liposome to cauda epididymal sperm, contain proteins with the similar N-terminal amino acid sequences to that of HE1.

```
          -19                  -1 +1                                  30
Porcine    MHFLAAAFL LLTLSASALA EPVHFRDCGS GVGVIKEVNV NPCPTQPCQL
Dog        MRLLVAAFL LLALGDLGPG GAVHFKDCGS AVGVIKELNV NPCP-QPCKL
M.fas.     MRFLAATFL LLALSTAAQA EPVQFKDCGS VDGVIKEVNV SPCF-QPCQL
Human      MRFLAATFL LLALSTAAQA EPVQFKDCGS VDGVIKEVNV SPCF-QPCQL
Chimp.     MRFLAATFL LLALSTAAQA EPVQFKDCGS VDGVIKEVNV SPCF-QPCQL
                                                                 80
Porcine   HKGQSYSVNV TFTSNTQSKG SKAVVHGIVM GVPIPFPIPD PDGCKSGINC
Dog       HKGQSYSVNV TFTSNIPSQS SKAVVHGIVL GVAVPFPIPE ADGCKSGINC
M.fas.    SKGQSYSVNV TFTSNIQSKS SKAVVHGILM GVPVPFPIPE PDGCKSGINC
Human     SKGQSYSVNV TFTSNIQSKS SKAVVHGILM GVPVPFPIPE PDGCKSGINC
Chimp.    SKGQSYSVNV TFTSNIQSKS SKAVVHGILM GVPVPFPIPE PDGCKSGINC
                                                                130
Porcine   PIQKDQTYSY LNKLPVKAEY PSIKLVVEWK LQDDNDQCLF CWQIPVQIES --
Dog       ---KDKTYSY LNKLPVKNEY PSIKLVVQWM LLGDNNQHLF CWEIPVQIEG --
M.fas.    ---KDKTYSY LNKLPVKSEY PSIKLVVEWQ LQDDKNQSLF CWEIPVQIVS HL
Human     ---KDKTYSY LNKLPVKSEY PSIKLVVEWQ LQDDKNQSLF CWEIPVQIVS HL
Chimp.    ---KDKTYSY LNKLPVKSEY PSIKLVVEWQ LQDDKNQSLF CWEIPVQIVS HL
```

Fig. 3. Alignment of the amino acid sequences of HE1 homologues. The deduced porcine sequence is aligned with those of dog (Ellerbrock *et al.*, 1994), macaque (Perry, Jones and Hall, 1995), human (Kirchoff, Osterhoff and Young, 1996), and chimpanzee (Frohlich and Young, 1996). -1 and +1 show the suggested signal sequence cleavage site and E numbered with a +1 is N-terminus of the purified 16 kDa protein.

We have recently found that the purified 16 kDa protein specifically binds free cholesterol in a time and concentration dependent manner and the amounts of cholesterol-binding reach saturation. The dissociation constant of the cholesterol-16 kDa protein complex was 2.3 μM, which is similar to that of the recombinant fatty acid binding protein (L-FABP) of rat liver: 1.53 μM (Nemecz and Schroeder, 1991). The maximal binding capacity of the purified 16 kDa protein for cholesterol was determined to be 58.8 pmol/μg protein. Binding stoichiometry was calculated to be 0.94 mol/mol in this study, which is also similar to the values reported for other cholesterol-binding proteins such as L-FABP: 0.83 (Nemecz and Schroeder, 1991) and VIP21/caveolin: about 1.0 (Murata *et al.*, 1995). This suggests that one mole of cholesterol binds to one mole of the purified 16 kDa protein.

Analysis of amino acid sequence homology revealed that the porcine HE1 homologue showed no significant homology with the other cholesterol-binding proteins such as heart-FABP (X98558; Gerbens *et al.*, 1997), liver-FABP (P49924; Perozzi *et al.*, 1993), nor the human caveolin (Z18951, S49856; Glenney, 1992). It is concluded that the porcine HE1 homologue is a novel type of the cholesterol-binding protein which is specifically expressed in the restricted part of the epididymis.

Recently, it has been reported that bovine milk contains a protein named EPV20 which displays 79% sequence similarity to HE1 (Larsen *et al.*, 1997). EPV20 has six cysteine residues which are disulfide-linked in 1-6, 2-3 and 4-5 pattern. Those six cysteins are well conserved among HE1 homologues including the porcine 16 kDa protein. So, the fact that the 16 kDa protein is a cholesterol-binding protein might provide a clue to find out the biological activity of the HE1 homologues and other proteins containing the same domains with three disulfide bonds such as Derf2, the major mite allergen (Ichikawa *et al.*, 1998) and esr16, a protein found in moth trachea (Meszaros and Morton, 1996).

It was also found that the 16 kDa protein greatly decreased the binding of cholesterol to sperm by competitively binding cholesterol. Furthermore, the 16 kDa protein was found to reduce the amount of bound cholesterol in sperm (Fig.4). This was not merely due to the fact that 16 kDa protein binds free cholesterol and decreases its concentration in the medium. The 16 kDa protein seems to extract the bound cholesterol directly from sperm, because in the absence of the 16 kDa protein, the bound cholesterol remained on the sperm even after free cholesterol was washed out.

The change in the cholesterol contents in the sperm membranes should affect membrane fluidity, membrane stability against the cold shock, protein behavior in the membrane, and the motile and fertile abilities of mature sperm. Although little is known about the molecular mechanism involved in the alteration of the membrane

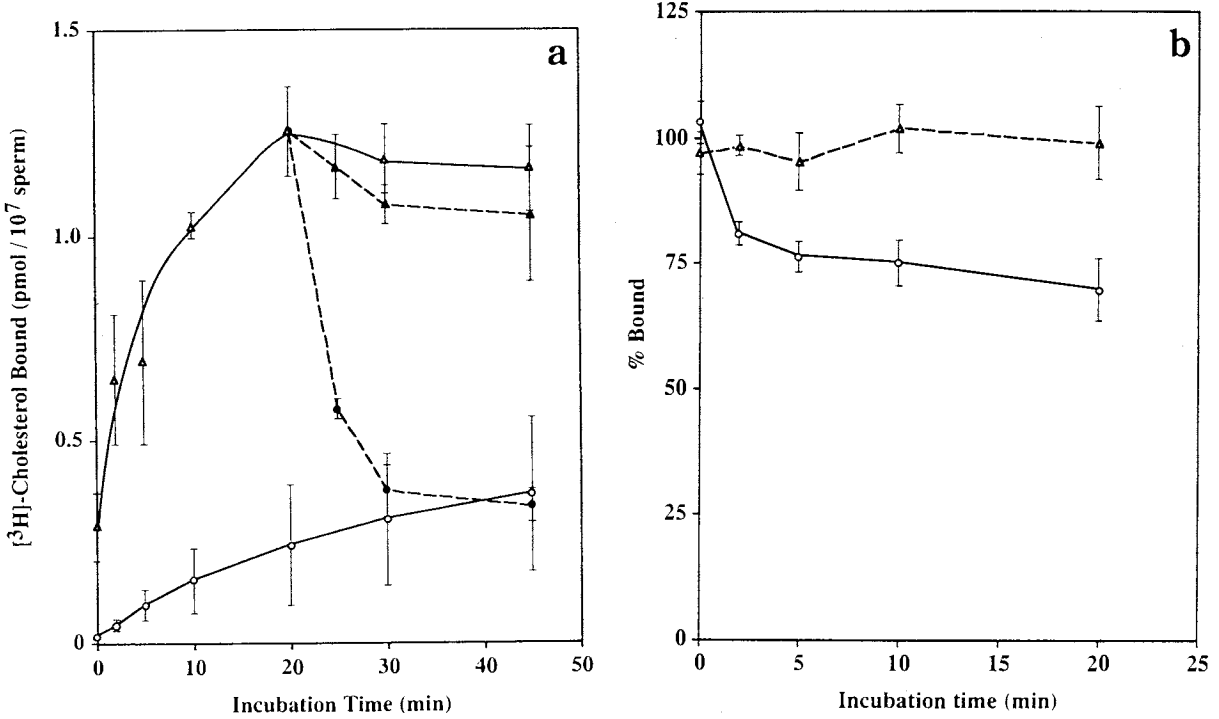


Fig. 4. **a**: Time course of [^{3}H]-cholesterol-binding to the porcine cauda epididymal sperm. 0.05 μM of [^{3}H]-cholesterol was incubated with the porcine cauda epididymal sperm (2×10^8 cells/ml) at 32 °C for various intervals of time in the absence (Δ) or the presence (O) of the 16 kDa protein (200 μg/ml). At 20 min, 200 μg/ml of the 16 kDa protein (●) or phosphate buffered saline (▲) was added to the incubation mixture which did not contained the 16 kDa protein (Δ). **b**: Release of bound [^{3}H]-cholesterol from the porcine cauda epididymal sperm. [^{3}H]-cholesterol-bound sperm was incubated at 32 °C for appropriate intervals of time in the absence (Δ) or the presence (O) of the purified 16 kDa protein (200 μg/ml). Amounts of the [^{3}H]-cholesterol bound to sperm were then determined.

lipids during sperm maturation, the 16 kDa protein, a HE1 homologue, is indicated to be one of the important regulators of the cholesterol content of sperm membranes.

The capacitation and subsequent acrosome reaction are preceded by a decrease in the cholesterol content of the sperm plasma membrane (Hammerstedt *et al.*, 1979; Davis, 1980). Recently, we have found that the 16 kDa protein mRNA is also expressed in female reproductive organs such as the ovary, oviduct and uterus. Furthermore, the purified 16 kDa protein has been found to stimulate capacitation and acrosome reaction induced *in vitro* in the presence and absence of caffeine, high concentration of Ca^{2+} and fatty acid free BSA. These results suggest that the 16 kDa cholesterol-binding protein, HE1 homologue, partly mediates the post testicular acquisition of sperm fertility by binding and extracting cholesterol from sperm membranes both in the male and female genital tracts. But, in connection with the function of cholesterol in sperm maturation and capacitation, further study must be done to know the exact roles of the HE1 homologue in epididymal sperm maturation.

23 kDa epididymis-specific glutathione peroxidase-like protein

The epididymis has important roles in protecting sperm from the oxidative damage. Mammalian sperms are rich in unsaturated fatty acids in membranes, and oxygen free radicals react with them to form organic hydroperoxides. Although lipid peroxidation is required to increase the capacity of sperm to bind to zona pellucida (Aitken, Clarkson and Fishel, 1989), it has been correlated with the male infertility because of the loss of motility or morphological abnormality of sperm (Aitken and Clarkson, 1987; Alvalez and Storey, 1989; Rao *et al.*, 1989). The peroxidative process is thought to damage the fluidity and integrity of the sperm plasma membrane, resulting in a loss of sperm function (Aitken and Fisher, 1994). Lipid peroxidation enhances the destruction of phospholipids in sperm plasma membranes, which induces the premature acrosome reaction in the epididymis (Ungemach *et al.*, 1985; Bennet *et al.*, 1987).

In addition to superoxide dismutase (SOD), it has been suggested that glutathione peroxidase (GPX) protects sperm from peroxidative damage in the epididymis. It has also been reported that mouse sperm and epididymal fluid contain high concentrations of glutathione and selenium-dependent GPX activity (Alvalez and Storey, 1984; 1989). Recently, cDNAs encoding epididymis-specific GPX-like proteins (EGLP) have been cloned from mouse, rat and monkey, and they were found to be distinct from either cytosolic or placental secreted GPXs (Ghyselinck, Jimenez and Dufore, 1991; Perry *et al.*, 1992). Cytosolic and placental secreted GPXs contain a selenocysteine residue at the catalytic site, and their mRNAs contain the selenocysteine

codon (UGA). But in EGLP mRNA, it was found to be replaced by the cysteine codon (UGY). Although there is no direct evidence that EGLP is enzymatically active in the epididymis, it has been suggested that it protects sperm from the deleterious effects of oxygen free radicals by preventing the accumulation of organic hydroperoxides.

In order to see whether EGLP exerts a physiologically significant amount of activity in the epididymal fluid, we purified EGLP from the porcine cauda epididymal fluid and its activities were quantified in comparison with selenocysteine containing erythrocyte GPX. The purified EGLP was found to consist of four identical 23 kDa subunits. EGLP showed only an extremely low activity when hydrogen peroxide was used as substrate (Table 1). The affinity of EGLP for hydrogen peroxide was less than 0.1% of the value for erythrocyte GPX. The Vmax of EGLP was also much lower (1/150) than that of the erythrocyte enzyme. The same tendency was observed when organic hydroperoxides were used as substrate; the activity of EGLP was as much as 1000-fold lower than the erythrocyte GPX (Table 2).

Table 1. Comparison of the kinetic parameters between the purified EGLP and bovine erythrocyte GPX

Enzyme	Km (mM)	Vmax (μmol/min/mg protein)	Turnoer rate (Min^{-1}/$subunit^{-1}$)
EGLP	0.31 ± 0.01	3.7 ± 0.1	85
Bovine erythrocyte GPX	(0.24 ± 0.02)× 10^{-3}	580 ± 20	12000

The GPX activities were determined at 37 °C in the presence of 4 mM glutathione and various concentrations of hydrogen peroxide.

Table 2. Comparison of the activities with different substrates between the purified EGLP and the bovine erythrocyte GPX

	GPX activity (nmol/min/mg protein)	
Substrate	EGLP	Bovine erythrocyte GPX
Hydrogen peroxide	1.69 ± 0.26	3200 ± 95
13 - Hydroperoxy - (9,11)- octadecadienoic acid	4.16 ± 0.55	4050 ± 133
15 - Hydroperoxy - (5,8,11,13) - eicosatetraenoic acid	3.54 ± 0.33	3000 ± 160

The GPX activities were determined at 37 °C in the presence of 4 mM glutathione and 20 μM of various substrates.

Furthermore, it was found that the concentrations of glutathione in the porcine epididymal fluids were 14.2~24.6 μM which is much lower than the optimal concentration of glutathione for the GPX activity of EGLP (around 5 mM). All these data strongly suggest that the GPX activity of EGLP is almost negligible in porcine epididymal fluid, though the possibility can not be dismissed completely that there may be unknown mechanisms for activation of the protein. It is also possible that EGLP has a substrate specificity different from that of the cytosolic GPX. In porcine epididymal fluid, the purified EGLP is the only source of GPX. So EGLP is unlikely to protect porcine epididymal sperm from the oxidative damages at least by its GPX activity.

It has been proposed that some proteins secreted from the epididymal epithelial cells bind to sperm and/or affect sperm functions. EGLP is specifically secreted from the caput epididymis and is a major protein component in the caput epididymal fluid. Using polyclonal antibodies, the localization of EGLP on sperm was analyzed. Consistent with the results reported by Jimenez *et al.* (1992), EGLP is localized in the acrosomal region of the epididymal sperm (Fig.5).

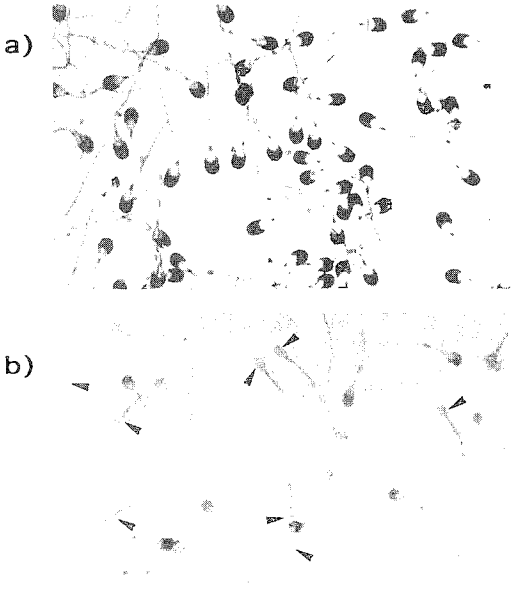


Fig. 5. Immunocytochemical localization of EGLP on sperm. Porcine cauda epididymal sperm before (a) and after (b) the induction of acrosome reaction were incubated with antisera raised against the purified 23 kDa EGLP subunit and the localization of EGLP was observed. The stained samples were observed under the light microscope at ×200. The acrosome reacted sperm are indicated by arrowheads.

It was also found that EGLP disappeared from the acrosomal region during the acrosome reaction induced by 1.8 mM $CaCl_2$, 10 mM caffeine and 0.1% fatty acid free BSA. EGLP was detected only on the acrosome intact sperm. These results indicate that EGLP in sperm can be used as a marker to distinguish acrosome-intact sperm from the acrosome-reacted ones. Furthermore, it was found that EGLP inhibited the acrosome reaction induced by high calcium, caffeine and fatty acid free BSA. Namely, the pretreatment of the sperm with the purified EGLP significantly reduced the percentage of the acrosome-reacted sperm.

It has been reported that the lipid hydroperoxides accumulate on the membranes of maturating sperm in the epididymis. Lipid peroxides are known to induce acrosome reaction by enhancing the activity of sperm phospholipase A_2 (Ungemach *et al.*, 1985; Bennet *et al.*, 1987; Goldman *et al.*, 1992). Namely, lipid hydroperoxides stimulates phospholipase A_2 activity, which results in the formation of lysophospholipids in the sperm plasma membrane. This destabilizes the structure of the sperm plasma membrane and induces the membrane fusion, leading to the acrosome reaction (Parrish *et al.*, 1988).

As the acrosome reaction must not occur in the epididymis, there should be some mechanisms to restrain the lipid peroxides from inducing the acrosome reaction. As the epididymis expresses relatively low catalase mRNA (Zini and Schlegel, 1996), it has been indicated that GPX and SOD is responsible for the degradation of the lipid peroxides accumulated in sperm. SOD has been reported to be expressed and secreted in the epididymis (Jow *et al.*, 1993; Perry *et al.*, 1993). The conventional, selenium-containing plasma glutathione peroxidase, GPX3 was also found to be expressed in the mouse epididymis, but it was not detected in sperm or in the epididymal fluid (Schwaab *et al.*, 1995; 1998), suggesting that GPX3 does not directly participate in the degradation of the lipid peroxides in sperm. On the other hand, it has recently been reported that the testis-expressed phospholipid hydroperoxide glutathione peroxidase (PHGPX) is present in the rat epididymal sperm (Godeas *et al.*, 1997). In addition, Jin *et al.* (1997) have reported that lactoferrin is secreted from the epididymis and binds to sperm. It is suggested that lactoferrin binds free iron and prevents it from catalyzing lipid peroxidation chain reactions in the epididymal sperm.

At least in boar, as mentioned above, EGLP is unlikely to exert a protective effect on sperm by its enzymatic activity. But EGLP has been found to protect transfected mammalian culture cells from the oxidative stress (Vernet *et al.*, 1996). A hypothesis is that EGLP binds to the lipid peroxides and prevents them from interacting with phospholipase A_2, which results in the inhibition of the acrosome reaction. In this connection, it is interesting that, in sperm-egg recognition, enzymes

such as galactosyltransferase and proacrosin, produce their effects not by enzymatic activities but by specific interactions between the catalytic site of the enzyme and its substrate. Although the acceptor in the sperm for EGLP is not identified yet, EGLP seems to support sperm maturation in the epididymis by maintaining the acrosome integrity.

Acknowledgments

This work was supported in part by a grant from the Ministry of Education (05044121, 10680682) to N.O. and partly by Tsukuba Advanced Research Alliance (TARA Sakabe-Shoun project).

References

Adams, C.S., and Johnson, A.D. 1977. Comp. Biochem. Physiol., **58**, 409-411.

Aitken,R.J. and Clarkson,J.S. 1987. J. Reprod. Fertil., **81**, 459-469.

Aitken,R.J., Clarkson,J.S., and Fishel,S. 1989. Biol. Reprod., **40**, 183-197.

Aitken, R.J. and Fisher, H. 1994. Bioessays, **16**, 259-267.

Alvarez,J.G. and Storey,B.T. 1984. Biol. Reprod., **30**, 833-841.

Alvalez,J.G. and Storey,B.T. 1989. Gamete Res., **23**, 77-90.

Awano, M., Kawaguchi, A., and Mohri, H. 1993. J. Reprod. Fertil., **99**, 375-383.

Baker, C.S., Magargee, S.F., and Hammerstedt, R.H. 1993. Biol. Reprod., **48** (suppl 1), 86.

Bennet,P.J., Moatti,J.P., Mansat,A., Ribbes,H., Cayrac,J.C., Pontonnier,F., Chap,H., and Doust-Blazy,C. 1987. Biochim. Biophys. Acta, **919**, 255-265.

Davis, B.K. 1980. Arch. Androl., **5**, 249-254.

Davis, B.K. 1981. Proc. Natl. Acad. Sci., USA, **78**, 7560-7564.

Ellerbrock, K., Pera, I., Hartung, S., and Ivell, R. 1994. Int. J. Androl., **17**, 314-323.

Frohlich, O., and Young, L.G. 1996. Biol. Reprod., **54**, 857-864.

Gerbens, F.N.A., Rettenberger, G., Lenstra, J.A., Veerkamp, J.H., and tePas, M.F.W. 1997. Mamm. Genome, **8**, 328-332.

Ghyselinck,N.B., Jimenez,C., and Dufore,J.P. 1991. J. Reprod. Fertil., **93**, 461-466.

Glenney, J.R. 1992. FEBS Lett., **314**, 45-48.

Godeas, C., Tramer, F., Micali, F., Soranzo, M., Sandri, G. and Panfili, E. 1997. Biol. Reprod., **57**, 1502-1508.

Goldman, R., Ferber, E. and Zort, U. 1992. FEBS Letters, **309**, 190-192.

Hammerstedt, R.H., Keith, A.D., Hay, S., Deluca, N., and Amann, R.P. 1979. Arch. Biochem. Biophys., **196**, 7-12.

Hall, J.C., Hadley, J., and Doman, T. 1991. J. Androl., **12**, 76-87.

Ichikawa, S., Hatanaka, H., Yuuki, T., Iwamoto, N., Kojima, S., Nishiyama, C., Ogura, K., Okumura, Y., and Inagaki, F. 1998. J. Biol. Chem., **273**, 356-360.

Jimenez,C., Lefrancois,A.M., Ghyselinck,N.B., and Dufaure,J.P. 1992. J. Endocrinol., **133**, 197-203.

Jin, Y.Z., Bannai, S., Dacheux, F., Dacheux, J.L. and Okamura, N. 1997. Mol. Reprod. Develop., **47**, 490-496.

Jones, R. 1998. J. Reprod. Fertil., **Suppl. 53**, 73-84.

Jow, W.W., Schlegel, P.N., Cichon, Z., Phillips, D., Goldstein, M. and Bardin, C.W. 1993. J. Androl., **14**, 439-447

Kirchhoff, C., Krull, N., Pera, I., and Ivell, R. 1993. Mol. Reprod. Develop., **34**, 8-15.

Kirchhoff, C., Osterhoff, C., Habben, I., and Ivell, R. 1990. Int. J. Androl., **13**, 155-167.

Kirchhoff, C., Osterhoff, C., and Young, L. 1996. Biol. Reprod., **54,** 847-856.

Kiuchi, S., Jin, Y.Z., Tamba, M., Hiramoto, S., Baba, T., Dacheux, F., Dacheux, J.L., and Okamura, N. 1997. Int. J. Androl., **20** (suppl 1), 110.

Larsen, L.B., Ravn, P., Boisen, A., Berglund, L., and Petersen, T.E. 1997. Eur. J. Biochem., **243**, 437-441.

Lopez, M.L., and de Souza, W. 1991. Mol. Reprod. Develop., **28,** 158-168.

Meszaros, M., and Morton, D.B. 1996. Insect Biochem. Mol. Biol., **26**, 7-11.

Murata, M., Peranen, J., Schreiner, R., Wieland, F., Kurzchalia, T. V., and Simons, K. 1995. Proc. Natl. Acad. Sci. USA, **92,** 10339-10343.

Nemecz, G., and Schroeder, F. 1991. J. Biol. Chem., **266,** 17180-17186.

Nikolopoulou, M., Soucek, D.A., and Vary, J.C. 1985. Biochim. Biophys. Acta, **815,** 486-498.

Okamura, N., Iwaki, Y., Hiramoto, S., Tamba, M., Bannai, S., Sugita, Y., Syntin, P., Dacheux, F., and Dacheux, J.L. 1997. Biochim. Biophys. Acta, **1336**, 99-109.

Okamura, N., Tamba, M., Liao, H.J., Onoe, S., Sugita, Y., Dacheux, F., and Dacheux, J.L. 1995a. Mol. Reprod. Develop., **42,** 141-148.

Okamura, N., Tamba, M., Uchiyama, Y., Sugita, Y., Dacheux, F., Syntin, P., and Dacheux, J.L. 1995b. Biochim. Biophys. Acta, **1245,** 221-226.

Osterhoff, C., Ivell, R., and Kirchhoff, C. 1997. DNA and Cell Biol., **16,** 379-389.

Parks, J.E., and Hammerstedt, R.H. 1985. Biol. Reprod. **32,** 653-668.

Parrish, J.J., Susko-Parrish, J.L., Winer,M.A. and First, N.L. 1988. Biol. Reprod., **38**, 1171-1180.

Perozzi, G., Baril, D., Murgia, C., Kelly, D., Begbie, R., and King, T.P. 1993. J. Nutr. Biochem., **4**, 699-705

Perry, A.C., Jones, R., and Hall, L. 1993. Biochem. J., **293**, 21-25.

Perry, A.C., Jones, R., and Hall, L. 1995. Gene, **153,** 291-292.

Perry,A.C.F., Jones,R., Niang,L.S.P., Jackson,R.M., and Hall,L. 1992. Biochem. J., **285**, 863-870.

Poulos, A., Brown Woodman, P.D., White, I.G., and Cox, R.I. 1975. Biochim. Biophys. Acta, **388,** 12-18.

Poulos, A., Voglmayr, J.K., and White, I.G. 1973. Biochim. Biophys. Acta, **306,** 194-202.

Rana, A.P., Majumder, G.C., Misra, S., and Ghosh, A. 1991. Biochim. Biophys. Acta, **1061,** 185-196.

Rao,B., Soufir,J.C., Martin,M., and David,G. 1989. Gamete Res., **24**, 127-134.

Schwaab, V., Baud, E., Ghyselinck, N., Mattei, M.G., Dufaure, J.P. and Drevet, J.R. 1995. Gene, **167**, 25-31.

Schwaab, V., Lareyre, J.J., Vernet, P., Pons, E., Faure, J., Dufaure, J.P. and Drevet, J.R. 1998. J. Reprod. Fertil., **Suppl. 53**, 157-162.

Syntin, P., Dacheux, F., Druart, X., Gatti, J.L., Okamura, N., and Dacheux, J.L. 1996. Biol. Reprod., **55**, 956-974.

Ungemach,F.R. 1985. in " Free radicals in liver injury. " Eds. Poli,G., Cheeseman,K.H., Dianzani,M.U., and Slater,T.F. pp.127-134, IRL Press,Washington.

Vernet, P., Rigaudiere, N., Ghyselinck, N., Dufaure, J.P. Drevet, J.R. 1996. Biochem. Cell Biol., **74**, 125-131.

Zini, A. and Schlegel, P.N. 1996. J. Androl., **17**, 473-480.

Sperm Evaluation: What Should We Be Testing?

R. A. P. HARRISON
Laboratory of Gamete Signalling, The Babraham Institute
Cambridge CB2 4AT, UK

Abstract

In conserving animal genetic resources, it is likely that semen evaluation will need to be carried out largely to examine the impact of new technology on sperm function, during development of more effective semen preservation procedures. At the present time, semen evaluation methods are not yet sufficiently discerning, and there is an urgent need for improvement. In seeking these improvements, it is the author's contention that (a) the natural heterogeneity of sperm populations should be more widely recognised and (b) clear distinction should be made between "preservation environments" and the "activating" environment of the female tract: most tests should examine sperm population response to a similar activating environment. Recent studies have shown that the fertilizing spermatozoon exerts several profound post-fertilization effects on the development of the resultant zygote, and these will also need to be taken into account; notably, methodological opportunities now exist for sensitive assessment of DNA integrity, which have so far been applied very largely only to human sperm. There is currently much concern as to the damage that can be caused by reactive oxygen species. Although human sperm seem to be particularly sensitive, knowledge of other sperm species' vulnerability is limited. With respect to potential oxidative damage, it is often forgotten that during its natural life *in vivo* the spermatozoon never encounters oxygen levels as high as those of the outside atmosphere. It would therefore seem prudent to consider preserving and processing semen in an oxygen-depleted environment.

Amazing advances in reproductive technology are currently being achieved (e.g. successful fertilization via intracytoplasmic injection of oocytes with freeze-dried sperm; ability to transplant spermatogenic cells between individuals and even species - see later articles in this volume). However, it will almost certainly be some time before these advances emerge from the specialised research laboratory, and they may never become applicable to the maintenance or rescue of endangered wild animal species. At least for the time being, therefore, the conservation of animal genetic resources must depend very largely upon the use of AI and conventional semen preservation (whether in liquid or frozen state). It is clear that the success of semen preservation using current methods is not universal, varying between species, breed and even individual. Obviously, it is highly undesirable simply to abandon semen types that respond poorly to preservation, because probably those semen samples in themselves represent a specific genetic resource! Thus improvement of preservation methods remains a major goal. This

improvement will not be achievable without evaluating the impact of the methodological modifications on sperm fertility. At present, evaluation methods themselves are inadequate and give imprecise results. In this paper I aim to provide a framework within which choices and development of more focused evaluation methods may be made.

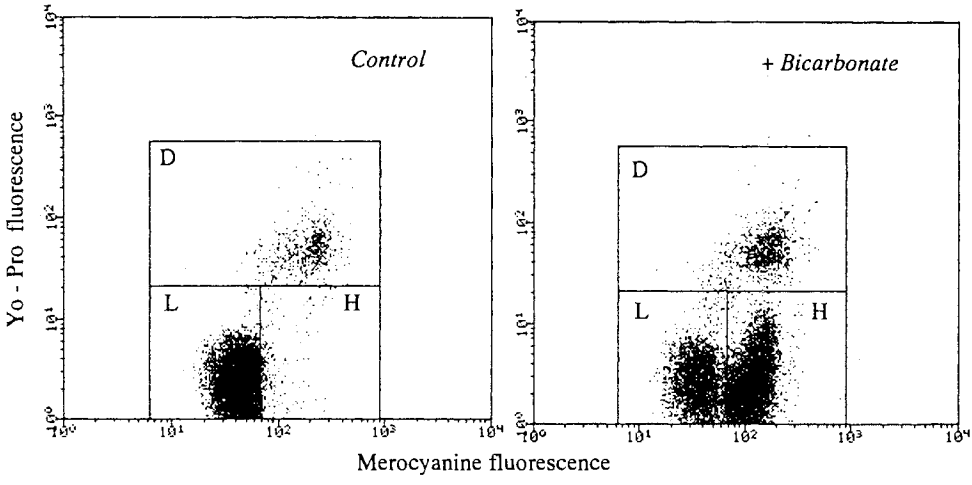
Ideally, fertility should be measured as the sperm's ability to bring about the eventual live birth of normal offspring. In the past, field trials have been used to evaluate such overall ability. However, not only are field trials slow and expensive but they are likely to be completely impractical where rare breeds or species are involved. Great efforts therefore continue to be made to find sperm evaluation procedures that are rapid, simple and sensitive. The problem is that the spermatozoon must be capable of many functions in order to play its normal physiological role (see Table 1). Some of these functions are obvious; for example, the sperm cell must be viable (i.e. capable of maintaining its normal intracellular homeostatic processes within an intact plasma membrane). Other functions are much less obvious (e.g. role in supporting embryo development). We need to find key sperm characteristics that are crucial to fertility. Many traditional evaluation parameters such as viability are so fundamental to any physiological capability as to be only of very general value (e.g. a dead sperm is obviously incapable of fertilizing; on the other hand, a live sperm may also prove to be infertile).

Table 1. Levels of sperm functional capability

Cellular integrity: maintain basal intracellular environment (Na^+, K^+, Ca^{2+}, pH, ATP etc)
Cellular potential: express motility, undergo capacitational changes in response to suitable extracellular environment
Zona penetration: undergo zona-induced acrosome reaction and express hyperactivated motility
Fertilization: fuse with oocyte and initiate early development
Embryo development: form normal implanting blastocyst
Live young: support full foetal development

In order to achieve actual fertilization, it is clear that various sperm functions must be specifically activated (e.g. ability to undergo an acrosome reaction as a result of contact with the zona pellucida). This activation, termed globally "capacitation" (see Harrison, 1996), apparently takes place via intracellular signalling pathways that are set in train by changes in the sperm's extracellular milieu (*in vivo*, by the environments of the female reproductive tract). It would therefore seem logical to test the sperm's ability to respond to such "fertilizing" conditions. In my own laboratory, we have been focusing on the boar sperm's response to bicarbonate/CO_2 because bicarbonate has been identified as an essential milieu component for successful IVF in pig as in several other animal species (see Suzuki *et al.*, 1994 and references therein). Boar sperm are particularly good models because they are unusually quiescent as they emerge from the epididymis (see Tajima *et al.*, 1987); they therefore offer excellent possibilities for studying the signalling mechanisms that "switch-on" sperm. We have observed several cellular changes which are specifically induced by bicarbonate, including changes in membrane lipid architecture (Harrison *et al.*, 1996a) and enhancement of motility (Tajima *et al.*, 1987; Abaigar *et al.*, 1999). These latter changes are of particular interest since they take place in most cells within a few minutes after the start of exposure to bicarbonate; we therefore think they may constitute very early stages in the overall capacitation process (which takes some 2-3 hours to complete in the majority of boar sperm: Hunter and Dzuik, 1968; Yoshida, 1987; Yoshida *et al.*, 1993).

Figure 1 shows the change in plasma membrane lipid architecture caused by bicarbonate, while Figure 2 shows the stimulation of motility by bicarbonate; Figure 3 shows differences between individual boars with respect to their plasma membrane lipid response to bicarbonate. A key feature of our strategy has been that response to bicarbonate is measured in relation to the effect of parallel incubation in a control medium which is essentially identical in all respects to the test medium save for its lack of bicarbonate/CO_2. Figures 1-3 illustrate in particular two of our relevant findings: namely that individual cells within the sperm samples show considerable differences in their response, and that proportions of responding cells vary greatly between samples from different boars. It may be emphasized that examination of the extended sperm samples themselves (i.e. under 'preservation' conditions) do not reveal such differences. Figure 3 also shows that the membrane



(Modified from Harrison, 1996)

Figure 1. Bicarbonate induction of changes in plasma membrane lipid architecture in boar spermatozoa, detected by flow cytometry. "Control": incubation for 10 min in a bicarbonate-free Hepes-buffered Tyrode's-based medium; "+Bicarbonate": incubation for 10 min in the Tyrode's-based medium supplemented with 15 mM bicarbonate/5% CO_2. Washed spermatozoa were incubated at 38 °C in the presence of Yo-Pro (fluorescent probe for viability), after which aliquants were mixed with merocyanine and subjected to flow cytometry 90 sec later. The rectangles within the plots delineate: D, dead cells (stained with Yo-Pro); L, live cells showing low merocyanine staining; H, live cells showing high merocyanine staining. Fluorescence is expressed as arbitrary units. Further experimental details are provided in Harrison *et al.* (1996).

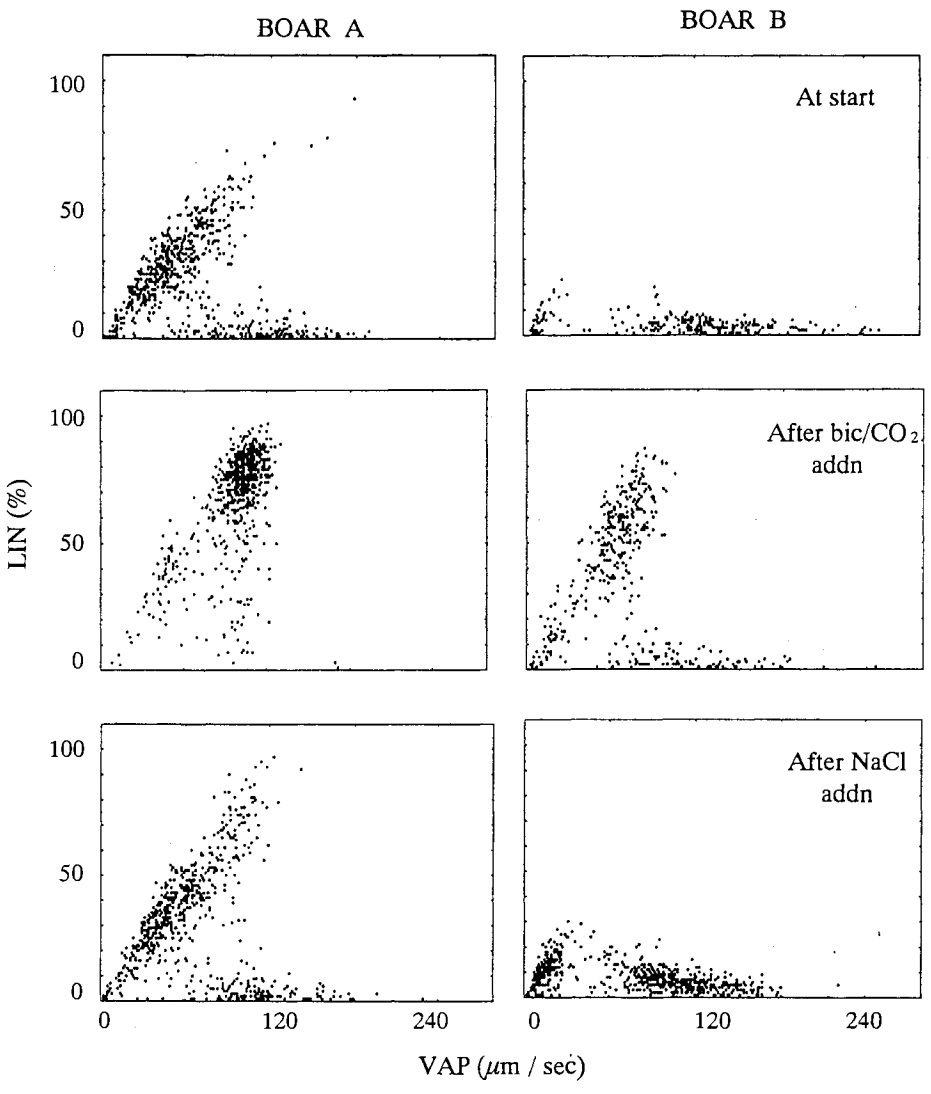
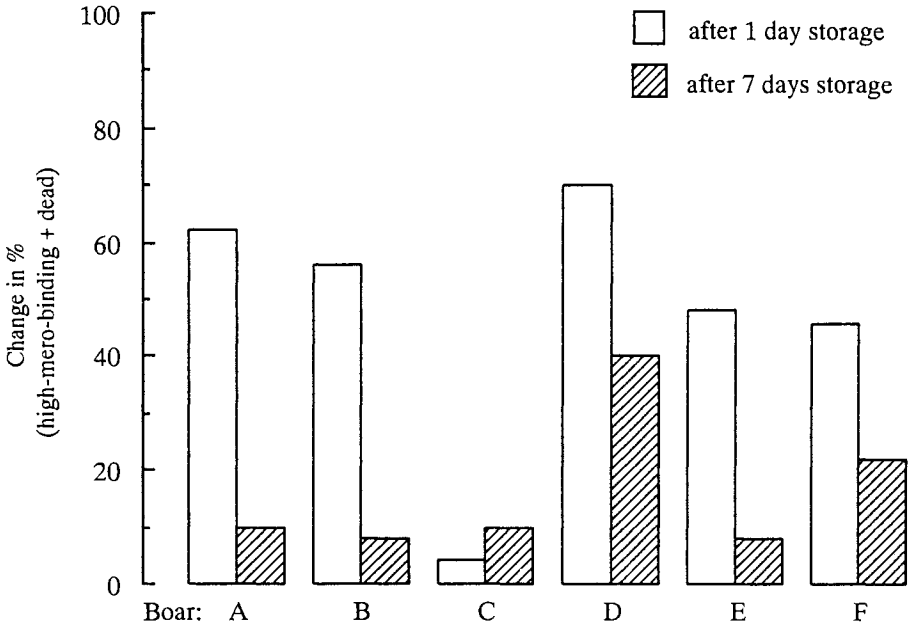


Figure 2. Bicarbonate induction of enhanced motility in boar spermatozoa. Washed spermatozoa (prepared separately from semen samples from two different boars) were incubated at 38 °C for 10 min in a bicarbonate-free Hepes-buffered Tyrode's-based medium; to half the suspension, bicarbonate/CO_2 was added to a final concentration of 15 mM/5%, while 15 mM NaCl was added to the other half (as control treatment). Samples of the suspensions were observed and recorded by video microscopy, whence motion descriptors for individual sperm were measured from the recordings using a Hobson Sperm Tracker (W.V.Holt and R.A.P.Harrison: unpublished data; further experimental details are provided in Holt *et al.*, 1996; Abaigar *et al.*, 1999). Path linearity (LIN = mean straight line velocity/mean path velocity) has been plotted against mean path velocity (VAP); LIN is a reflection of progressive motility, VAP a reflection of tail beating. "At start": motility just before addition of bicarbonate; "After bic/CO_2 addn": motility 12 min after addition of bicarbonate; "After NaCl addn": motility 18 min after addition of NaCl.



(Modified from Harrison *et al.*, 1996b)

Figure 3. Variation among boars with respect to sperm plasma membrane response to bicarbonate. Sperm-rich fractions from the ejaculates of six different boars were diluted and stored at ambient temperature in Beltsville Thawing Solution (BTS; Johnson *et al.*, 1988) for either 1 or 7 days. Spermatozoa were isolated from the diluted samples by washing through Percoll after which they were incubated in bicarbonate-containing or bicarbonate-free Hepes-buffered Tyrode's-based medium and analysed for merocyanine staining, as described in Fig.1. Results from a single set of samples (tested in parallel) are presented as the combined bicarbonate-inducible changes in high merocyanine-binding live cells or dead cells (i.e. those in boxes H and D respectively in Fig.1). Standard deviation of the assay = ±5.3 cells per 100 (estimated from a series of several similar experiments in which duplicate assays were performed).

lipid response to bicarbonate declines during liquid preservation of boar semen in a widely used commercial extender, though how closely this decline is correlated with the well-established loss of fertility during liquid storage (Waberski *et al.*, 1994) remains to be demonstrated.

One of the problems in developing laboratory tests of sperm function is the difficulty in relating cellular processes to actual fertilizing ability (see Amann and Hammerstedt, 1993; Hammerstedt, 1996). In most species, *in vivo* testing is precluded because of the impossibility of obtaining statistically meaningful data from single sperm samples, while the use of *in vitro* fertilization (where possible) is time-consuming and very demanding on resources and skill. A simple but sensitive laboratory bioassay is therefore needed. Of the sperm's several tasks at fertilization itself, the ability to pass through the zona pellucida appears to be the most demanding (see Liu and Baker, 1994; Yanagimachi, 1994; Johnson *et al.*, 1995). Recently, we developed a procedure for testing this in boar sperm, using cryopreserved zona-intact pig eggs (Lynham and Harrison, 1998). We were able to show (Figure 4) that the sperm developed zona penetrating ability only under fertilizing conditions, and that the time-base of development of zona penetrating ability was closely similar to the development of overall fertilizing ability. Recently several groups have tried to relate sperm ability to bind to the zona pellucida to fertility (Fazeli *et al.*, 1993, 1995a, 1995b; Coddington *et al.*, 1994; Codde and Berger, 1995; Meyers *et al.*, 1996), with varying conclusions. In our assay of boar sperm function, in which we specifically differentiated between sperm that had only bound to the zona surface and sperm that had penetrated into the zona matrix, there was little or no relationship between zona binding and zona penetration: considerable numbers of sperm bound to the zona pellucida under non-fertilizing conditions (see Figure 4). In general, assays of sperm-zona interaction have not distinguished between binding and penetration; our results suggest strongly that estimation of zona penetrating ability would be a sensitive and meaningful quantitative parameter of sperm fertilizing capability.

An intriguing if worrying conundrum in current sperm evaluation is the apparent discrepancy between the numbers of sperm that after incubation under fertilizing conditions show cell characteristics considered as possible indicators of capacitation and those that actually show an ability to fertilize. The original (and still valid) definition of the capacitated state in sperm is the ability to initiate

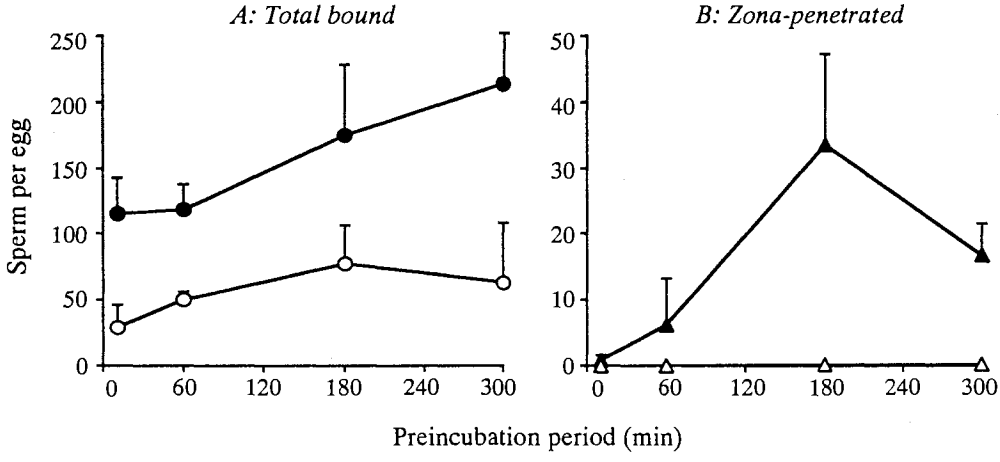


Figure 4. Development of boar sperm ability to penetrate the pig zona pellucida. Solid symbols: "Capacitating medium" (Hepes-buffered Tyrode's-based medium containing 15 mM bicarbonate/5% CO_2 and 2 mM caffeine); open symbols: "Non-capacitating medium" (Hepes-buffered Tyrode's-based medium without supplementation). Washed spermatozoa were preincubated for various periods of time in one or other of these two media, and then coincubated for 10 min with groups of cumulus-free eggs in the same medium (The eggs were members of a large cryopreserved pool previously prepared from ovaries taken from immature females). The groups of sperm-egg complexes were then transferred to fresh medium, after which some of each group were assessed immediately for total numbers of bound sperm while the rest of the group was incubated for a further 60 min before being assessed for numbers of zona-penetrated sperm. Results of three independent experiments are presented (means + SD). Further details are given in Lynham and Harrison (1998).

fertilization very soon after the start of gamete coincubation (see Harrison, 1996). However, in recent times, more specific, "molecular" changes have begun to be taken as direct parameters of capacitation (e.g. changes in chlortetracycline binding pattern: Fraser *et al.* 1995; ability to undergo the acrosome reaction in response to zona pellucida components: Shi and Roldan, 1995). Assays of such in-vitro capacitation characteristics detect rather high percentages of altered sperm: depending on the parameter used, figures of between 10 and 50% are the norm (examples for boar sperm: Mattioli *et al.*, 1996; Harkema. *et al.*, 1998). However, calculations suggest that far fewer may be capable of fertilizing. Table 2 shows data derived from published reports of pig IVF. Relating numbers of penetrated sperm with those potentially able to make contact with eggs in an in-vitro fertilization system is very imprecise, since it must be based on very many, as-yet-untested assumptions. However, in our studies (Lynham and Harrison, 1998), of the sperm that bound to the surface of stored pig eggs during brief gamete coincubation, only some 15% showed any degree of zona penetrating ability during a subsequent 3-hr incubation. These low numbers were certainly more in accord with the fertilization calculations in Table 2. Data from other species similarly suggests very low rates of sperm fertilizing success *in vitro* (rat: Nishimura *et al.*, 1982; mouse: Siddiquey and Cohen, 1982; rabbit: Lambert *et al.*, 1978; hamster: Stewart-Savage and Bavister, 1988; human: Van der Ven *et al.*, 1989).

One probable reason for the discrepancy between numbers of sperm deemed 'capacitated' and those actually capable of zona penetration is that the former measures a single sperm parameter/function whereas the latter depends on the expression of several functions (Hammitt *et al.*, 1989; Amann and Hammerstedt, 1993; Amann *et al.*, 1993). For example, the acrosome reaction is clearly a necessary prerequisite for zona penetration (see Yanagimachi, 1994), and therefore a number of studies have evaluated sperm ability to undergo the acrosome reaction (see Troup *et al.*, 1994; Fénichel, 1995; Mortimer, 1995); on the other hand, other investigations have indicated that motility also plays a crucial role in zona penetration (Stauss *et al.*, 1995; Olds-Clarke, 1996). Obviously, for a given sperm to fertilize, it must express several functions in a carefully controlled temporal sequence (see Amann and Hammerstedt, 1993; Amann *et al.*, 1993). Fulfilment of this criterion is inevitably of low likelihood since the different functions are by no means necessarily 'coupled' or related. For accuracy of semen

Table 2. Calculation of proportions of fully competent sperm in boar semen from *in vitro* fertilization data

Authors	Fertilization Medium	Volume per dish (ml)	Sperm per dish	Oocytes per dish	Coincubation period (h)	Fertilization rate (%)	Penetrated sperm per oocyte	Competent sperm per thousand*
Coy *et al.*, 1993	TCM-199 + 12% foetal calf serum + 2 mM caffeine at pH 7.4	2	6×10^5	15	4	55.3	1.44	0.036
Mori *et al.*, 1996	Bracket & Oliphant's + 1% BSA +5 mM caffeine at pH 7.6	0.1	1×10^4	10	2.5	72.1	2.2	2.2
		0.1	1×10^4	10	4	97.3	8.8	8.8
Xu *et al.*, 1996	TCM-199 + 10% foetal calf serum + 10 mM caffeine at pH 7.4	1	3.125×10^4	10	6	91.5	1.4	0.448

*: calculated by multiplying the number of penetrated sperm per oocyte by the number of oocytes, dividing by the total number of sperm in the fertilization volume, and multiplying by 1000.

evaluation, it may well be necessary to monitor several dynamic characteristics simultaneously in the individual cells, a technical ability which is as yet largely beyond our capabilities.

The future of sperm evaluation obviously resides in the continued development of "*in vitro*" assays. However, such development risks overlooking the very large overall differences between in-vitro and in-vivo fertilization (Harrison, 1996). A particularly important issue is the need for sperm to survive in the female tract for long periods of time (see Drobnis and Overstreet, 1992; Hunter, 1995). Natural or artificial insemination usually takes place following oestrus detection, which may be several (or many) hours in advance of actual ovulation; once deposited in the female tract, the sperm must await the eggs' arrival. Strong circumstantial evidence suggests that ability to survive this waiting period is a key contributory factor in the overall in-vivo fertility of the sperm sample (Waberski *et al.*, 1994; Rota *et al.*, 1997; Troedsson *et al.*, 1998). In-vitro fertilizing tests, on the other hand, traditionally demand a rapid sperm response. Thus it seems pertinent to consider including 'survival' tests in the repertoire of sperm evaluation procedures (e.g. Holt *et al.*, 1997; Rota *et al.*, 1997). Recent studies have indicated that *in vivo* those sperm that will eventually fertilize the egg(s) populate a localised region, the oviductal isthmus, where they form close relationships with the epithelial cells (Drobnis and Overstreet, 1992; Hunter, 1995). Considerable evidence has now been gathered to show that oviductal epithelial cells prolong sperm survival in a specific manner, at the same time reducing sperm activity (Chian and Sirard, 1995; Smith and Nothnick, 1997; Smith, 1998; Suarez, 1998). The exact mechanism of this process is as yet unknown. However, the attention currently being paid to this aspect of sperm physiology is likely to provide novel ways of testing sperm fitness. Tests of survival may be particularly important with respect to the development of new preservation technology, because considerable number of recent studies have reported that after cooling or cryopreservation sperm exhibit behaviour and/or surface characteristics that appear to mimic capacitation (see Watson, 1995; Maxwell and Watson, 1996; also Cormier *et al.*, 1997; Fuller and Whittingham, 1997). Capacitation is considered a 'destabilizing' process which will eventually lead to more rapid cell death (see Harrison, 1996), thus capacitation-like changes induced by cryopreservation are clearly counterproductive in view of the need for sufficient numbers of sperm to survive in the female tract until the time of

ovulation. A further interesting consideration is that capacitated cells appear to bind less strongly to oviductal epithelial cells (Smith, 1998), thus one can hypothesise that prematurely "capacitated" sperm may be less able to profit from the specific protecting power of the isthmic oviduct.

At this point, it may be worth discussing briefly the concept that the female tract in some way 'selects' sperm, with the result that a high proportion of sperm that encounter the egg are capable of fertilization (Drobnis and Overstreet, 1992; Hunter, 1995). At first sight, this might seem to render unnecessary any concern that only a very low proportion of sperm in the ejaculate seem capable of fertilizing eggs. However, it should be remembered that there is much evidence to indicate that the likelihood of fertilization depends on the numbers of 'fit' sperm inseminated (see Amann and Hammerstedt, 1993). Thus semen evaluation must be carried out on unselected sperm samples because it is important to know what proportion of sperm are 'fit'. All other things being equal, use of AI will involve the insemination of approximately constant numbers of (live) unselected sperm, and it will be the sample which contains the largest proportion of 'fit' sperm that will stand the greatest chance of fertilizing successfully.

The approaches to semen evaluation described above have all focused on fertilizing ability. However, the sperm play further roles in the development of live young (Chaykin and Watson, 1983; Setchell *et al.*, 1988; Janny and Menezo, 1994), the subtleties of which are still being uncovered. Full consideration should be given to finding means of testing relevant sperm capabilities in these roles. Four general areas of sperm activity have so far been identified. In order of involvement in embryonic development, they are:- the contribution of the sperm centrosomes to syngamy and subsequent mitosis (Navarra *et al.*, 1995); the sperm's initiation of correct Ca^{2+} signalling processes within the zygote (Swann and Ozil, 1994); the contribution of the paternally derived genome to correct differentiation of embryonic tissue (Norris *et al.*, 1990); and the sperm's contribution of a full haploid complement of potentially functional genes to the diploidy of the embryo. The importance of the first area is still to be defined (see Navarra *et al.*, 1995), but nonetheless its potential should not be ignored. The second area relates to sperm-borne egg-activating factor(s) known as oscillogens; as these have not yet been clearly identified (see Swann *et al.*, 1998; Tesarik, 1998), nothing can be proposed with respect to testing their presence within the sperm cell or their potential activity.

The third area relates to the concept of genetic imprinting, in which identical genes passing through the male or female parent acquire differing phenotypic potentials within foetal development; the most obvious expression of imprinting is the global role played by the male-derived genome in the correct differentiation of embryonic and extra-embryonic tissues (Norris *et al.*, 1990), though more specific examples of imprinting effects are continually being uncovered (Reik, 1996). Again, the molecular identification of imprinting has not yet progressed to a point at which any advance testing of the sperm genome for its imprinting adequacy could be made. The fourth area, the general contribution of the sperm's haploid genomic complement to the embryo/foetus, is less easy to define; one aspect would obviously be the development of male and female offspring following fertilization by Y- and X-bearing sperm respectively. A preliminary approach to this would be to employ more widely the newly developed micromethods for examining the state of sperm genetic material. There is considerable evidence that abnormalities of chromatin structure are associated with poor reproductive performance (see among others Evenson *et al.*, 1980; Rodriguez *et al.*, 1985; Engh *et al.*, 1993; Evenson *et al.*, 1994; Sailer *et al.*, 1995; Sakkas *et al.*, 1996), though there is rather little evidence to suggest that cryopreservation procedures cause such abnormalities (Royere *et al.*, 1988; Evenson *et al.*, 1994; Sailer *et al.*, 1995).

Chromatin packing (Ward and Coffey, 1991), and the degree of interaction between sperm protamines and DNA, can be assessed fluorimetrically using either acridine orange (Evenson *et al.*, 1980; Sailer *et al.* 1995) or chromomycin A_3 (Bianchi *et al.*, 1993; Sakkas *et al.*, 1996). Damage to DNA itself, in the form of strand breakage, can be assessed either by TUNEL (Terminal deoxynucleotidyl transferase-mediated dUTP-biotin 'nick' end labelling) assay or by 'comet' assay. In the TUNEL assay, DNA strand ends are labelled by attaching biotinylated deoxyuridine nucleotide, and detecting the presence of the biotinylated nucleotide by secondary labelling with streptavidin conjugated with either alkaline phosphatase or fluorescein. The amount of labelling, quantified either by image analysis of sperm smears (Bianchi *et al.*, 1993; Sakkas *et al.*, 1996; Twigg *et al.*, 1998a) or by flow cytometry (Sun *et al.*, 1997), correlates with the degree of strand breakage. In the 'comet' assay (Hughes *et al.*, 1996), sperm are immobilized in agarose (as a layer on a slide), their chromatin decondensed and their DNA uncoiled, after which the layer is placed in an electric field; electrophoresis of the

DNA proceeds, when fragmented DNA migrates more rapidly than intact DNA. Visualized by ethidium bromide staining, the 'intact' DNA appears as a compact grouping with the DNA fragments spread to one side (hence the adjective 'comet'), and the degree of fragmentation can be estimated by image analysis. Up till now, TUNEL and "comet" assays have been almost exclusively applied to human sperm, related to the concern that the rapid proliferation of ICSI (intracytoplasmic sperm injection) as a means of overcoming human male-factor infertility might involve the injection of sperm whose DNA is severely compromised; even if successful births were thereby achieved, the individuals might have enhanced predispositions to subsequent medical problems.

While it might be argued that ICSI is unlikely to become a widely used animal breeding technology, one reason for the recent intensified attention to human sperm DNA damage has been the demonstration that such damage can be caused by reactive oxygen species (ROS), to which human sperm seem particularly vulnerable (Aitken, 1997; Twigg *et al.*, 1998a;b). Cellular targets other than DNA (e.g. plasma membrane phospholipids) have also been shown to be sensitive to oxidation (Hammerstedt, 1993; Storey, 1997; Twigg *et al.*, 1998a), but overall, there is uncertainty as to which targets (or potential functions) are most at risk (Aitken *et al.*, 1998; Ellington *et al.*, 1998). The actual source and/or conditions of generation of the ROS will clearly have an important influence. For example, ROS that are water-soluble or that are generated extracellularly will tend to affect different cellular targets from those that are lipid-soluble or generated intracellularly, and the protective ability of particular antioxidants added to the sperm's environment will vary similarly (see for example Twigg *et al.*, 1998a). Although rather little evidence has been obtained so far to show that ROS species are responsible for decline in sperm function during semen preservation, this lack of evidence may simply be due to a failure to pose the right questions in the correct way. For example, although the sperm may be in some way protected from or insensitive to ROS during preservation or in the preservation medium, the preservation process may render the sperm more vulnerable to ROS once they are in a more "physiological" environment. Even in freshly obtained spermatozoa, ROS-based damage appears to be enhanced by incubation in "fertilization" media (c.f. Estop *et al.*, 1993; Gomez and Aitken, 1996; Twigg *et al.*, 1998b), in environments depleted of the wide range of antioxidants that abound in body fluids

(Halliwell and Gutteridge, 1990; Rice-Evans and Gopinathan, 1995). An aspect which appears to have received rather little attention is the fact that sperm never encounter *in vivo* the levels of oxygen to which they are usually exposed *in vitro* - i.e. atmospheric: 20% (Zhil'tsov, 1970; Free *et al.*, 1976; Fischer and Bavister, 1993). Preservation and processing of sperm under lowered oxygen conditions might therefore be of great benefit (see Hammerstedt, 1993; Vishwanath and Shannon, 1997; Griveau and Le Lannou, 1998).

Concluding remarks

Semen evaluation has been a subject of great attention for more than 50 years. Despite (or maybe because of!) this long history, advances have not in my opinion kept up with the forefront of our knowledge of sperm physiology, and current methods remain rather basic and therefore perhaps inadequate. In this paper, rather than cataloguing the many semen evaluation procedures that have been proposed, I have tried to present lines of thinking that may encourage the development of new and more focused procedures. While the main aim during such development must be to keep the details of sperm physiology firmly in the forefront, there is an important concomitant need for students to keep abreast of new ideas and techniques in other fields of cell biology. Although it is a specialised cell with its own peculiarities, the spermatozoon is none the less just another mammalian cell type, with features and mechanisms in common with other cells. It is a well-established fact in practical science that breakthroughs in one field often depend on advances in other, apparently unrelated fields!

Acknowledgement

The author's research is supported by the UK Biotechnology and Biological Sciences Research Council. Grateful thanks are also due to Dr. W.V. Holt for permission to publish some of our unpublished collaborative research and to Prof. P.F. Watson for reading the manuscript and commenting helpfully on it.

References

Abaigar, T., Holt, W.V., Harrison, R.A.P., del Barrio, G. 1999. Sperm subpopulations in boar (*Sus scrofa*) and gazelle (*Gazella dama mhorr*) semen as revealed by pattern analysis of computer-assisted motility assessments. Biology of Reproduction 60: 32-41.

Aitken, R.J. 1997. Molecular mechanisms regulating human sperm function. Molecular Human Reproduction 3: 169-173.

Aitken, R.J., Gordon, E., Harkiss, D., Twigg, J.P., Milne, P., Jennings, Z., Irvine, D.S. 1998. Relative impact of oxidative stress on the functional competence and genomic integrity of human

spermatozoa. Biology of Reproduction 59: 1037-1046.

Amann, R.P., Hammerstedt, R.H. 1993. *In vitro* evaluation of sperm quality: an opinion. Journal of Andrology 14: 397-406.

Amann, R.P., Hammerstedt, R.H., Veeramachaneni, D.N.R. 1993. The epididymis and sperm maturation: a perspective. Reproduction, Fertility and Development 5: 361-381.

Bianchi, P.G., Manicardi, G.C., Bizzaro, D., Bianchi, U., Sakkas, D. 1993. Effect of deoxyribonucleic protamination on fluorochrome staining and in situ nick-translation of murine and human mature spermatozoa. Biology of Reproduction 49: 1083-1088.

Chaykin, S., Watson, J.G. 1983. Reproduction in mice: spermatozoa as factors in the development and implantation of embryos. Gamete Research 7: 63-73.

Chian, R.-C., Sirard, M.-A. 1995. Fertilizing ability of bovine spermatozoa cocultured with oviduct epithelial cells. Biology of Reproduction 5: 156-162.

Codde, J.M., Berger, T. 1995. In vivo fertility of rams in relation to sperm-zona pellucida binding and sperm-zona pellucida penetration of ovine oocytes. Theriogenology 44: 901-906.

Coddington, C.C., Oehninger, S.C., Olive, D.L., Franken, D.R., Kruger, T.F., Hodgen, G.D. 1994. Hemizona index (HZI) demonstrates excellent predictability when evaluating sperm fertilizing capacity in *in vitro* fertilization patients. Journal of Andrology 15: 250-254.

Cormier, N., Sirard, M.-A., Bailey, J.L. 1997. Premature capacitation of bovine spermatozoa is initiated by cryopreservation. Journal of Andrology 18: 461-468.

Coy, P., Martítez, E., Ruiz, S., Vázquez, J.M., Roca, J., Matas, C. 1993. Sperm concentration influences fertilization and male pronuclear formation in vitro in pigs. Theriogenology 40: 539-546.

Drobnis, E.Z., Overstreet, J.W. 1992. Natural history of mammalian spermatozoa in the female reproductive tract. Oxford Reviews of Reproductive Biology 14: 1-45.

Ellington, J.E., Evenson, D.P., Fleming, J.E., Brisbois, R.S., Hiss, G.A., Broder, S.J., Wright, R.W. 1998. Coculture of human sperm with bovine oviduct epithelial cells decreases sperm chromatin structural changes seen during culture in media alone. Fertility and Sterility 69: 643-649.

Engh, E., Schølberg, A., Clausen, O.P.F., Purvis, K. 1993. DNA flow cytometry of sperm from normozoospermic men in barren couples and men of proven fertility. International Journal of Fertility 38: 305-310.

Estop, A.M., Munne, S., Jost, L.K., Evenson, D.P. 1993. Studies on sperm chromatin structure alterations and cytogenetic damage of mouse sperm following *in vitro* incubation. Studies on *in vitro*-incubated mouse sperm. Journal of Andrology 14: 282-288.

Evenson, D.P., Darzynkiewicz, Z., Melamed, M.R. 1980. Relation of mammalian sperm chromatin heterogeneity to fertility. Science 210: 1131-1133.

Evenson, D.P., Thompson, L., Jost, L. 1994. Flow cytometric evaluation of boar semen by the sperm chromatin structure assay as related to cryopreservation and fertility. Theriogenology 41: 637-651.

Fazeli, A.R., Steenweg, W., Bevers, M.M., de Loos, F.A.M., van den Broek J., Colenbrander, B. 1993. Development of a sperm zona pellucida binding assay for bull semen. Veterinary Record 132: 14-16.

Fazeli, A.R., Holt, C., Steenweg, W., Bevers, M.M., Holt, W.V., Colenbrander, B. 1995a. Development of a sperm hemizona binding assay for boar semen. Theriogenology 44: 17-27.

Fazeli, A.R., Steenweg, W., Bevers, M.M., van den Broek, J., Bracher, V., Parlevliet, J.,

Colenbrander, B. 1995b. Relation between stallion sperm binding to homologous hemizonae and fertility. Theriogenology 44: 751-760.

Fénichel, P. 1995. Acrosomal function and sperm fertilizing ability. In: Human Sperm Acrosome Reaction. Eds P. Fénichel., J. Parinaud. Colloque INSERM/John Libbey Eurotext Ltd, Montrouge, France. pp. 315-325.

Fischer, B., Bavister, B.D. 1993. Oxygen tension in the oviduct and uterus of rhesus monkeys, hamsters and rabbits. Journal of Reproduction and Fertility 99: 673-679.

Fraser, L.R., Abeydeera, L.R., Niwa, K. 1995. Ca^{2+}-regulating mechanisms that modulate bull sperm capacitation and acrosomal exocytosis as determined by chlortetracycline analysis. Molecular Reproduction and Development 40: 233-241.

Free, M.J., Schluntz, G.A., Jaffe, R.A. 1976. Respiratory gas tensions in tissues and fluids of the male rat reproductive tract. Biology of Reproduction 14: 481-488.

Fuller, S.J., Whittingham, D.G. 1997. Capacitation-like changes occur in mouse spermatozoa cooled to low temperatures. Molecular Reproduction and Development 46: 318-324.

Gomez, E., Aitken, J. 1996. Impact of in vitro fertilization culture media on peroxidative damage to human spermatozoa. Fertility and Sterility 65: 880-882.

Griveau, J.F., Grizard, G., Boucher, D., Le Lannou, D. 1998. Influence of oxygen tension on function of isolated spermatozoa from ejaculates of oligozoospermic patients and normozoospermic fertile donors. Human Reproduction 13: 3108-3113.

Halliwell, B., Gutteridge, J.M.C. 1990. The antioxidants of human extracellular fluids. Archives of Biochemistry and Biophysics 280: 1-8.

Hammerstedt, R.H. 1993. Maintenance of bioenergetic balance in sperm and prevention of lipid peroxidation: a review of the effect on design of storage preservation systems. Reproduction, Fertility and Development: 675-690.

Hammerstedt, R.H. 1996. Evaluation of sperm quality: identification of the subfertile male and courses of action. Animal Reproduction Science 42: 77-87.

Hammitt, D.G., Martin, P.A., Callanan, T. 1989. Correlations between heterospermic fertility and assays of porcine semen quality before and after cryopreservation. Theriogenology 32: 385-399.

Harkema, W., Harrison, R.A.P., Miller, N.G.A., Topper, E.K., Woelders, H. 1998. Enhanced binding of zona pellucida proteins to the acrosomal region of intact boar spermatozoa in response to fertilizing conditions: a flow cytometric study. Biology of Reproduction 58: 421-430.

Harrison, R.A.P. 1996. Capacitation mechanisms, and the role of capacitation as seen in eutherian mammals. Reproduction, Fertility and Development 8: 581-594.

Harrison, R.A.P., Ashworth, P.J.C., Miller, N.G.A. 1996a. Bicarbonate/CO_2, an effector of capacitation, induces a rapid and reversible change in the lipid architecture of boar sperm plasma membranes. Molecular Reproduction and Development 45: 378-391.

Harrison, R.A.P., Ashworth, P.J.C., Miller, N.G.A. 1996b. Assessment of sperm function under fertilizing conditions. Reproduction in Domestic Animals 31: 25-30.

Holt, C., Holt, W.V., Moore H.D.M. 1996. Choice of operating conditions to minimize sperm subpopulation sampling bias in the assessment of boar semen by computer-assisted semen analysis. Journal of Andrology 17: 587-596.

Holt, C., Holt, W.V., Moore, H.D.M., Reed, H.C.B., Curnock, R.M. 1997. Objectively measured boar sperm motility parameters correlate with the outcomes of on-farm inseminations: results of two fertility trials. Journal of Andrology 18: 312-323.

Hughes, C.M., Lewis, S.E.M., McKelvey-Martin, V.J., Thompson, W. 1996. A comparison of

baseline and induced DNA damage in human spermatozoa from fertile and infertile men, using a modified comet assay. Molecular Human Reproduction 2: 613-619.

Hunter, R.H.F. 1995. Ovarian endocrine control of sperm progression in the Fallopian tubes. Oxford Reviews of Reproductive Biology 17: 85-124.

Hunter, R.H.F., Dziuk, P.J. 1968. Sperm penetration of pig eggs in relation to the timing of ovulation and insemination. Journal of Reproduction and Fertility 15: 199-208.

Janny, L., Menezo, Y.J.R. 1994. Evidence for a strong paternal effect on human preimplantation embryo development and blastocyst formation. Molecular Reproduction and Development 38: 36-42.

Johnson, L.A., Aalbers, J.G., Grooten, H.J.G. 1988. Artificial insemination of swine: fecundity of boar semen stored in Beltsville TS (BTS), Modified Modena (MM), or MR-A and inseminated on one, three and four days after collection. Zuchthygiene 23: 49-55.

Johnson, L.R., Pilder, S.H., Bailey, J.L., Olds-Clarke, P. 1995. Sperm from mice carrying one or two *t* haplotypes are deficient in investment and oocyte penetration. Developmental Biology 168: 138-149.

Lambert, R.D., Blanchette, L.G., Tourigny, M. 1978. The relationship between sperm concentration and fertilization in vitro of rabbit ova. Animal Reproduction Science 1: 43-47.

Liu, D.Y., Baker, H.W.G. 1994. A new test for the assessment of sperm-zona pellucida penetration: relationship with results of other sperm tests and fertilization *in vitro*. Human Reproduction 9: 489-496.

Lynham, J.A., Harrison, R.A.P. 1998. Use of stored pig eggs to assess boar sperm fertilizing functions in vitro. Biology of Reproduction 58: 539-550.

Mattioli, M., Barboni, B., Lucidi, P., Seren, E. 1996. Identification of capacitation in boar spermatozoa by chlortetracycline staining. Theriogenology 45: 373-381.

Maxwell, W.M.C., Watson, P.F. 1996. Recent progress in the preservation of ram semen. Animal Reproduction Science 42: 55-65.

Meyers, S.A., Liu, I.K.M., Overstreet, J.W., Vadas, S., Drobnis, E.Z. 1996. Zona pellucida binding and zona-induced acrosome reactions in horse spermatozoa: comparisons between fertile and subfertile stallions. Theriogenology 46: 1277-1288.

Mori, T., Hirayama, M., Suzuki, K., Shimizu, H., Nagai, T. 1996. Effect of casein phospho peptides and Ca^{2+} on penetration of boar spermatozoa into pig oocytes matured in vitro. Biology of Reproduction 55: 364-369.

Mortimer, D. 1995. *In vitro* induction of the human sperm acrosome reaction. In: Human Sperm Acrosome Reaction. Eds P. Fénichel., J. Parinaud. Colloque INSERM/John Libbey Eurotext Ltd, Montrouge, France. pp. 295-313.

Navarra, C.S., Simerly, C., Zoran, S., Schatten, G. 1995. The sperm centrosome during fertilization in mammals: implications for fertility and reproduction. Reproduction, Fertility and Development 7: 747-754.

Nishimura, H., Niwa, K., Miyake, M., Iritani, A. 1982. The effect of sperm concentration during preincubation and at insemination on fertilization of rat eggs in vitro. Journal of Experimental Zoology 223: 75-78.

Norris, M.L., Barton, S.C., Surani, M.A.H. 1990. The differential roles of parental genomes in mammalian development. Oxford Reviews of Reproductive Biology 12: 225-244.

Olds-Clarke, P. 1996. How does poor motility affect sperm fertilizing ability? Journal of Andrology 17: 183-186.

Reik,W. 1996. Genetic imprinting: the battle of the sexes rages on. Experimental Physiology 81: 161-

172.

Rice-Evans, C.A., Gopinathan, V. 1995. Oxygen toxicity, free radicals and antioxidants in human disease: biochemical implications in atherosclerosis and the problems of premature neonates. Essays in Biochem 29: 39-63.

Rodriguez, H., Ohanian, C., Bustos-Obregon, E. 1985. Nuclear chromatin decondensation of spermatozoa in vitro: a method for evaluating the fertilizing ability of ovine semen. International Journal of Andrology 8: 147-158.

Rota, A., Ström, B., Linde-Forsberg, C., Rodriguez-Martinez, H. 1997. Effects of Equex STM paste on viability of frozen-thawed dog spermatozoa during in vitro incubation at 38℃. Theriogenology 47: 1093-1101.

Royere, D., Hamamah, S., Nicolle, J.C., Barthelemy, C., Lansac, J. 1988. Freezing and thawing alter chromatin stability of ejaculated human spermatozoa: fluorescence acridine orange staining and Feulgen-DNA cytophotometric studies. Gamete Research 21: 51-57.

Sailer, B.L., Jost, L.K., Evenson, D.P. 1995. Mammalian sperm DNA susceptibility to *in situ* denaturation associated with the presence of DNA strand breaks as measured by the terminal deoxynucleotidyl transferase assay. Journal of Andrology 16: 80-87.

Sakkas, D., Urner, F., Bianchi, P.G., Bizzaro, D., Wagner, I., Jaquenoud, N., Manicardi, G., Campagna, A. 1996. Sperm chromatin anomalies can influence decondensation after intracytoplasmic sperm injection. Human Reproduction 11: 837-843.

Setchell, B.P., D'Occhio, M.J., Hall, M.J., Laurie, M.S., Tucker, M.J., Zupp, J.L. 1988. Is embryonic mortality increased in normal female rats mated to subfertile males? Journal of Reproduction and Fertility 82: 567-574.

Shi, Q-X., Roldan, E.R.S. 1995. Bicarbonate/CO_2 is not required for zona pellucida- or progesterone-induced acrosomal exocytosis of mouse spermatozoa but is essential for capacitation. Biology of Reproduction 52: 540-546.

Siddiquey, A.K.S., Cohen, J. 1982. In-vitro fertilization in the mouse and the relevance of different sperm/egg concentrations and volumes. Journal of Reproduction and Fertility 66: 237-242.

Smith, T.T. 1998. The modulation of sperm function by the oviductal epithelium. Biology of Reproduction 58: 1102-1104.

Smith, T.T., Nothnick, W.B. 1997. Role of direct contact between spermatozoa and oviductal epithelial cells in maintaining rabbit sperm viability. Biology of Reproduction 56: 83-89.

Stauss, C.R., Votta, T.J., Suarez, S.S. 1995. Sperm motility hyperactivation facilitates penetration of the hamster zona pellucida. Biology of Reproduction 53: 1280-1285.

Stewart-Savage, J., Bavister, B.D. 1988. Success of fertilization in golden hamsters is a function of the relative gamete ratio. Gamete Research 21: 1-10.

Storey, B.T. 1997. Biochemistry of the induction and prevention of lipoperoxidative damage in human spermatozoa. Molecular Human Reproduction 3: 203-213.

Suarez, S.S. 1998. The oviductal sperm reservoir in mammals: mechanisms of formation. Biology of Reproduction 58: 1105-1107.

Sun, J-G., Juriskova, A., Casper, R.F. 1997. Detection of deoxyribonucleic acid fragmentation in human sperm: correlation with fertilization in vitro. Biology of Reproduction 56: 602-607.

Suzuki, K., Ebihara, M., Nagai, T., Clarke, N.G.E., Harrison, R.A.P. 1994. Importance of bicarbonate/CO_2 for fertilization of pig oocytes *in vitro*, and synergism with caffeine. Reproduction, Fertility and Development 6: 221-7.

Swann, K., Ozil, J-P. 1994. Dynamics of the calcium signal that triggers mammalian egg activation. International Review of Cytology 152: 183-222.

Swann, K., Parrington, J., Jones, K.T. 1998. On the search for the sperm oscillogen. Molecular Human Reproduction 4: 1010-1012.

Tajima, Y., Okamura, N., Sugita, Y. 1987. The activating effects of bicarbonate on sperm motility and respiration at ejaculation. Biochimica Biophysica Acta 924: 519-529.

Tesarik, J. 1998. Sperm-induced calcium oscillations. Oscillin - reopening the hunting season. Molecular Human Reproduction 4: 1007-1009.

Troedsson, M.H.T., Liu, I.K.M., Crabo, B.G. 1998. Sperm transport and survival in the mare: a review. Theriogenology 50: 807-818.

Troup, S.A., Lieberman, B.A., Matson, P.L. 1994. The acrosome reaction to ionophore challenge test: assay reproducibility, effect of sexual abstinence and results for fertile men. Human Reproduction 9: 2079-2083.

Twigg, J., Fulton, N., Gomez, E., Irvine, D.S., Aitken R.J. 1998a. Analysis of the impact of intracellular reactive oxygen species generation on the structural and functional integrity of human spermatozoa: lipid peroxidation, DNA fragmentation and effectiveness of antioxidants. Human Reproduction 13: 1429-1436.

Twigg, J., Irvine, D.S., Houston, P., Fulton, N., Michael, L., Aitken, R.J. 1998b. Iatrogenic DNA damage induced in human spermatozoa during sperm preincubation: protective significance of seminal plasma. Molecular Human Reproduction 4: 439-445.

Van Der Ven, H.H., Hoebbel, K., Al-Hasani, S., Diedrich, K., Krebs, D. 1989. Fertilization of human oocytes in capillary tubes with very small numbers of spermatozoa. Human Reproduction 4: 72-76.

Vishwanath, R., Shannon, P. 1997. Do sperm cells age? A review of the physiological changes in sperm during storage at ambient temperature. Reproduction, Fertility and Development 9: 321-331.

Waberski, D., Weitze, K.F, Lietmann, C., Lubbert zur Lage, W., Bortolozzo, F.P, Willmen, T. Petzoldt, R. 1994. The initial fertilizing capacity of longterm-stored liquid boar semen following pre- and postovulatory insemination. Theriogenology 41: 1367-1377.

Ward, W.S., Coffey, D.S. 1991. DNA packaging and organization in mammalian spermatozoa: comparison with somatic cells. Biology of Reproduction 44: 569-574.

Watson, P.F. 1995. Recent developments and concepts in the cryopreservation of spermatozoa and the assessment of their post-thawing function. Reproduction, Fertility and Development 7: 871-891.

Xu, X., Seth, P.C., Harbison, D.S., Cheung, A.P., Foxcroft, G.R. 1996. Semen dilution for assessment of boar ejaculate quality in pig IVM and IVF systems. Theriogenology 46: 1325-1337.

Yanagimachi, R. 1994. Mammalian fertilization. In: The Physiology of Reproduction, Second Edition. Vol. I. Eds E. Knobil, J.D. Neill. New York: Raven Press Ltd. pp 189-317.

Yoshida, M. 1987. In vitro fertilization of pig oocytes matured in vivo. Japanese Journal of Veterinary Science 49: 711-718.

Yoshida, M., Cran, D.G., Pursel, V.G. 1993. Confocal and fluorescence microscopic study using lectins of the distribution of cortical granules during the maturation and fertilization of pig oocytes. Molecular Reproduction and Development 36: 462-468.

Zhil'tsov, N.Z. 1970. Oxygen concentration in the epididymis of rabbits. Doklady-vsesoiuznaia Akademiia sel'skokhoziaistvennykh Nauk im V I Lenina 5: 34-35. (Russian).

Using Coordinated Computer-Assisted Motility Analysis (CASA) and Western Phosphoamino Acid Detection to Elucidate Significant Signal Transduction Pathways in Live Sperm

JOSEPH S. TASH, Ph.D.
Department of Molecular and Integrative Physiology, and the Center for Reproductive Sciences, University of Kansas Medical Center, Kansas City, KS 66160, USA

In animal reproduction, assessment of sperm motility is a fundamental method for screening breeders of potential value in animal husbandry programs and for detection of readily identifiable infertility due to dismotility (Holt *et al.*, 1997; Garner, 1997; Working, 1988). In humans, infertility affects 15-19% of the general population and almost 50% of infertility cases in couples are attributable to the male (Pontonnier and Bujan, 1993). Of the male factor patients in whom a single abnormal semen parameter is presented (37% of all male factor patients), 24 out of the 37% are characterized by asthenospermia (i.e. reduced motility) that is not due to agglutination, accessory fiber abnormalities, infection, cold shock or prolonged exposure to seminal plasma (Sigman *et al.*, 1991; McConnell, 1991).

A classic example of a known cause of congenital infertility that affects male sperm function is immotile cilia syndrome (ICS) and an associated subset of these with Kartagener's syndrome (Kartagener, 1933). In ICS, the dynein arms are missing or abnormal (Eliasson *et al.*, 1977) resulting in immotile sperm and infertility (Baccetti *et al.*, 1979). It has been suggested that ICS is caused by at least two different genes (somatic and testicular ICS), since some male patients with respiratory tract ICS produce normal sperm and have fathered children without Assisted Reproductive Technologies (ART: Phillips *et al.*, 1995; Conraads *et al.*, 1992). It should be noted that the incidence of patients with ICS affecting sperm is rather rare at about 1:20,000. Thus when technical and analysis artifacts have been ruled out, the overwhelming proportion of male factor patients has asthenospermia of unproven cause. Some of the proposed origins for these kinds of asthenospermia include epididymal dysfunction, varicocele, inhibitory factors in seminal plasma, deficiencies in seminal plasma and

defects in the axoneme, ATP generation, Ca^{2+} regulation, and cAMP signaling (McConnell, 1991).

Signal transduction in sperm motility

The incidence of male-factor cases due to reduced sperm motility of unknown cause supports the need to understand the signal transduction pathways regulating sperm motility. Once specific targets in the signal transduction cascade are identified, it can be determined whether these targets represent lesions in the expression of motility in cases of asthenospermia. Conversely, these targets could be employed as potential targets in the development of sperm-directed fertility management. An important role for cAMP in the initiation and maintenance of sperm motility has been known for a long time (Tash and Mann, 1973; Garbers *et al.*, 1973), reviewed by Tash (1989) and Tash and Bracho (1994). Protein kinase activation and enhanced phosphorylation of endogenous proteins are concomitant events associated with elevated cAMP (Tash *et al.*, 1986; Tash and Means, 1982).

Identification of important targets of regulation

Using reactivated models, targets of protein phosphorylation that have been proposed to be involved in regulation of sperm motility have been identified in a number of diverse species. In trout, for example, a 15 kDa protein was identified whose phosphorylation was increased very rapidly during activation of motility (Hayashi and Morisawa, 1988; Hayashi *et al.*, 1987) in concert with an increase in cAMP. However, analysis of the phosphoamino acid residue in this protein revealed phosphotyrosine (pY) rather than phosphoserine (pS) or phosphothreonine (pT). A subsequent paper (Jin *et al.*, 1994b) identified a three-protein complex that was isolated from trout testis that contained a 15 kDa protein that could be phosphorylated *in vitro* in a cAMP-dependent manner. Monoclonal antibodies against the complex recognized a larger protein at 38 kDa and not the 15 kDa protein (Jin *et al.*, 1994a). In a note added in proof in that paper, the authors stated that the antibodies that were developed to characterize the 38 kDa protein, in fact, recognized a larger protein whose size was not given. The major pY-containing proteins of this size in sea urchin sperm appear to be associated with the sperm head and do not change during activation of motility in the intact sperm (Bracho *et al.*, 1998). In this connection, a slightly larger cAMP-dependent phosphoprotein (18-20 kDa) as well as a soluble 55 kDa

phosphoprotein were identified in reactivated Ciona sperm correlated with motility (Dey and Brokaw, 1991). That study demonstrated that a synthetic tyrosine kinase inhibitor blocked motility and ^{32}P-incorporation in the reactivated model, but had no effect on pY levels when analyzed by anti-pY antibodies. Several other studies have identified phosphoprotein targets in sperm that are either dynein subunits or tightly bound to flagella (Bracho *et al.*, 1998; Gingras *et al.*, 1996; Si and Okuno, 1995; Ashizawa *et al.*, 1995; Stephens and Prior, 1995; Ashizawa *et al.*, 1994). In bovine sperm, a 55 kDa pY-containing soluble protein was identified whose phosphorylation state is correlated with motility (Vijayaraghavan *et al.*, 1997b). This protein proved not to be RII of PKA (cAMP-dependent protein kinase), another soluble protein that we and others have identified in sperm to be correlated with sperm motility (Tash, 1989; Paupard *et al.*, 1988; Noland *et al.*, 1987). The phosphorylation of this protein was stimulated by cAMP suggesting a role for crosstalk between PKA and PKY signaling cascades. In mouse sperm, a flagellar 65 kDa cAMP-dependent phosphoprotein was identified whose microtubule-sliding activity was stimulated in live sperm by bicarbonate ions (Si and Okuno, 1995). When considering all of the above data that has been obtained in reactivated models, no clear consistent picture can be made of what are the key proteins involved in regulation of sperm flagellar motility because each system varies, giving different values for the size of phosphoproteins as well as their localization. In contrast, using live intact sperm to examine phosphoproteins coupled to sperm motility, we have identified an axonemal 130 kDa phosphoprotein (FP130) that is the major flagellar phosphoprotein in sea urchin sperm. This protein is phosphorylated on pT and not pS or pY. Most importantly, a functionally similar pT-containing protein of nearly identical size is the major pT protein present in mouse sperm (Tash and Bracho, 1998). In addition, a major pT-modified protein identical in size to the mouse motility-coupled protein is present in live human sperm. The striking similarity of the pT pattern in live human and sea urchin sperm combined with the fact that the field of likely critical regulatory targets has been significantly narrowed, strengthens the utility of such models to produce meaningful data.

In terms of the significance to sperm fertility, dynein is the primary flagellar component that must be triggered after sperm exit the testis in order for motility to be expressed. This is true for echinoderms to humans. Intact testicular sperm do not show progressive motility when placed in a buffer in which mature sperm show normal motility. However under conditions that stimulate cAMP or in reactivated models in which phosphorylation is increased, the testicular sperm can display forward

progression motility (Jaiswal and Majumder, 1996; Voglmayr and White, 1979; Cascieri *et al.*, 1976). In mammals, sperm must undergo additional maturation during transit through the epididymis that allows them to be become ready to be fully motile by the time the cauda epididymis is reached. These maturation changes must allow dynein ATPase activity to be expressed (Mohri and Yanagimachi, 1980). A link between the acquisition of the ability of the sperm to display vigorous sperm motility during epididymal transit and successful fertilization of eggs has been known for a long time (Brackett *et al.*, 1978; Hoppe, 1975). A positive correlation between stimulated cAMP-dependent pathways and stimulated fertilization has been demonstrated in a number of systems including mouse, monkey, and humans (Fenichel *et al.*, 1996; Aitken *et al.*, 1986; Monks *et al.*, 1986; Boatman and Bavister, 1984; Fraser, 1981; Fraser, 1979). We have recently demonstrated that FP130 phosphorylation is altered in response to the egg chemotactic peptide, speract (Tash and Bracho, 1999). One of the discernable responses to speract is a change in swimming pattern from circular to straight trajectories. Thus, by altering FP130 phosphorylation, speract may contribute to the change in flagellar mechanochemical dynamics that changes circular to straight swimming patterns.

In studies that have focused on treatments that stimulate sperm cAMP and motility, positive outcomes in the treatment of infertility were achieved (Zheng and Zhang, 1997; Sharma *et al.*, 1996; Wolf *et al.*, 1989). In such studies however, it is not possible to distinguish cAMP effects on motility from the known effects of cAMP on the capacitation process (Galantino Homer *et al.*, 1997; Leclerc *et al.*, 1996a; Fraser, 1990). Thus, in order to identify important sperm components that couple motility and signal transduction to successful fertilization, it is useful to also examine species in which sperm motility pathways are not occurring in a background that includes capacitation-associated signal transduction. Sea urchin sperm have been a long standing model for this type of analysis (for example, see Shapiro *et al.*, 1990; Christen *et al.*, 1986; Trimmer and Vacquier, 1986).

Reactivated models

Our laboratory was the first to utilize [γ-^{32}P]ATP to reactivate detergent permeabilized sperm to identify endogenous phosphoproteins (pPrs) associated with activation of motility. In all sperm that have been examined, elevated cAMP results in increased phosphorylation of multiple proteins (reviewed in Tash, 1989). A number of

studies suggest that soluble factors are involved in this process (Leclerc and Gagnon, 1996b; Tash *et al.*, 1986; Tash *et al.*, 1984; Murofushi *et al.*, 1982; Brandt and Hoskins, 1980). However, the degree to which soluble factors are required for initiation of motility remains unresolved since all models examined thus far start with sperm that have already undergone some degree of motility initiation prior to preparation of the immotile model (Leclerc and Gagnon, 1996b; San Agustin and Witman, 1994; Dey and Brokaw, 1991; Tash *et al.*, 1986; Ishiguro *et al.*, 1982). In the process of characterizing the mechanism of action of axokinin (RII), we discovered that the presence of detergent alone had a major stimulatory effect on reactivation of permeabilized dog sperm (Tash, 1989). This brought us to question whether detergent treatment might allow interaction between signal transduction pathway enzymes that, *in vivo*, might not normally occur. However, in order to be able to add exogenous ^{32}P-ATP for sperm activation and subsequent analysis of phosphorylation, membrane compartmentalization is destroyed. Furthermore, if protein phosphorylation and sperm motility are coupled, then any motility present in the sperm prior to stimulation will be expressed as a background phosphorylation. This, combined with the destruction of compartmentalization makes more difficult the identification of primary regulatory components. The problem is further magnified by diversity of samples between individuals and collection procedures that allow variable levels of motility expression prior to experimental activation (Jones and Murdoch, 1996; Brokaw, 1984; Morton *et al.*, 1978). Another possibility is that some regulatory proteins may undergo rapid phosphorylation and that dephosphorylation may be extremely slow or non-existent. In this case, subsequent attempts to tag the targets with ^{32}P in reactivation experiments would either fail or give highly variable levels of incorporation, depending on the extent of 'background' motility and phosphorylation prior to preparation of the model.

Development of improved methods for collection and storage of live sperm

Recent experiments that we conducted in microgravity on the Space Shuttle necessitated the development of improved methods to collect and store live sperm in an immotile state (Tash and Bracho, 1999). To solve the problem of motility expression prior to activation, we sought to minimize sperm exposure to normal seawater during spawning and activation during storage. This was achieved by modification of the conditions outlined by Christen *et al.* (1986) into MES sperm storage buffer, MSSB (Table 1). This buffer contained higher potassium (50 mM),

lower pH (6.0) and 5 mM MES to buffer the lower pH (Bracho *et al.*, 1997). Using such buffers, sperm could be stored for up to 96 hr at 5 °C with no significant reduction in the motility that was achieved by activating the sperm by dilution into sea water when compared to the level of motility obtained in the freshly collected sperm (Bracho *et al.*, 1997). Using the same principles, a similar collection method has been developed for mouse sperm (Tash and Bracho, 1998). Further testing is currently underway for a similar method to apply to turkey sperm. Success in our endeavors to develop a method for turkey sperm will greatly help productivity in the turkey industry since current technologies yield sperm that are optimally fertile for only 3-6 hr (Donoghue *et al.*, 1998; Brillard, 1993).

Table 1. Composition of Buffers for Collection, Storage and Activation of Immotile Sperm. Note: All concentrations are given in mM.

	ASW	HSW	MSSB
PH	8.0	8.3	6.0
KCl	10	0	50
MES	—	—	5
TAPS	5	5	5
NaCl	425	425	425
$MgCl_2$	27	27	27
$MgSO_4$	29	29	29
$CaCl_2$	10	10	10
$NaHCO_3$	2.4	2.4	2.4

Signal transduction in live sperm

Using the new sperm collection method, one could then ask whether the elimination of 'background' motility would result in lower background phosphorylation during initiation of motility. Furthermore, the experiment was performed with live sperm rather than permeabilized sperm to eliminate the problem of destroyed compartmentalization, which might create fortuitous kinase-substrate interactions. Since ^{32}P-ATP could not be used with live sperm to examine internal phosphorylation, we used phosphoamino acid antibodies to examine the phosphorylation changes (Galantino Homer *et al.*, 1997; Berruti and Martegani, 1989). Results showed that in immotile sperm, as expected, the levels of pS, pT and pY were

extremely low and limited predominantly to a few low M_r proteins in the sperm head (Bracho *et al.*, 1998). These results were in sharp contrast to reactivated sperm experiments which showed numerous phosphoproteins in which the identification of critically important proteins becomes uncertain (Ahmad *et al.*, 1995; Tash *et al.*, 1986; Takahashi *et al.*, 1985; Leclerc and Gagnon, 1996b; Dey and Brokaw, 1991; Tash and Means, 1983). Initiation of motility by dilution of immotile sperm into seawater produced the phosphorylation of two dominant sets of proteins within 20 sec. One set of proteins (32-29 kDa) contained predominantly pS, appeared to be associated with flagella and was extracted under conditions that extract outer arm dynein (OAD). The other major motility-related phosphoprotein contained predominantly pT, had a molecular weight of 130 kDa (named FP130), and appeared to remain bound to the axoneme after OAD extraction. In these sperm, no phosphorylation was detected at 56 kDa that would correspond to RII, either in the whole sperm, soluble material nor any of the flagellar fractions (Bracho *et al.*, 1998). Another important new observation yielded by the live sperm model is that the phosphoproteins revealed at 20 sec of activation are the same after 20 min of activation. This supports the hypothesis that in the intact sperm the signal transduction kinases and phosphatases are directed to a very limited number of targets.

Role of signal transduction anchoring proteins

In this regard, a family of proteins, AKAP, (A kinase anchoring protein) that anchor cAMP-dependent protein kinase to specific compartments of cells by binding to RII, has been identified (reviewed by Coghlin *et al.*,1993). In mammalian sperm, the major fibrous sheath protein has been shown to a member of the AKAP family (Carrera *et al.*, 1994). A role for AKAP in the regulation of sperm motility has been suggested by recent work of Vijayaraghavan *et al.* (1997a). This group showed that a permeable peptide that represents the amphipathic helical RII binding motif of AKAP caused inhibition of bovine sperm motility. A control peptide with identical sequence, except for a single proline substitution that renders the peptide inactive, did not inhibit sperm motility. They also demonstrated that the reversible inhibition of motility by the active AKAP peptide was dependent on the presence of external Ca^{2+}. In addition, they found that inhibitors of PKA did not inhibit basal motility or stimulation of motility via activators of the PKA pathway. This latter result is in contrast to work demonstrating an inhibitory effect of permeable inhibitors of PKA on sperm motility (Leclerc and

Gagnon, 1996b; San Agustin and Witman, 1994; Brokaw, 1987; Tash *et al.*, 1986; Ishiguro *et al.*, 1982). It has been suggested that the ability of inhibitors of PKA to affect sperm motility depends on the timing of inhibitor addition relative to activation of motility, as well as the ability of endogenous protein phosphatases to reverse activation-dependent phosphorylations (Tash and Bracho, 1994; San Agustin and Witman, 1994). In experiments described in the preliminary results, we have demonstrated that sea urchin sperm motility is stimulated by the same AKAP peptide that inhibit bovine sperm motility. In addition, in studies performed in collaboration with Dr. Stuart Moss (U. Pennsylvania) we have identified an 82 kDa AKAP in outer-arm dynein containing high salt extracts of the flagella as well as in the salt-extracted axonemes. It should be noted that sea urchin sperm do not contain a fibrous sheath. These new results emphasize the need to study the role of AKAPs in signal transduction in sperm from different species. The fact that AKAP inhibitors showed opposing effects in mammalian versus sea urchin sperm raise several questions: 1) Do mammalian sperm also have AKAPs bound to the axoneme, and 2) do the axonemal AKAPs help to target stimulatory pathways for motility, whereas the fibrous sheath AKAPs target inhibitory pathways for motility? In interpreting the results using these AKAP inhibitors, it must be noted that the protein sequence used to synthesize these inhibitors is homologous to adrenal AKAPs rather than sperm AKAPs. Whether this difference in sequence homology is of functional significance also remains to be determined. Finally, it should also be mentioned that the AKAP inhibitors mentioned above had no effect on mouse sperm motility (Vijayaraghavan, personal communication).

One possible explanation for the divergent results could be that the fibrous sheath anchors predominantly enzymes that inhibit motility such as protein phosphatases. On the other hand, the axonemal AKAPs may anchor stimulatory enzymes such as cAMP-dependent protein kinase. AKAPs have been shown to bind both kinases as well as phosphatases (Vo *et al.*, 1998; Scott, 1997; Lester and Scott, 1997; Klauck *et al.*, 1996; Coghlan *et al.*, 1995). The response to AKAP disruption would be determined by the prevailing enzyme released from its docking site. For example, since sea urchin sperm do not contain a fibrous sheath, the predominant enzyme would be stimulatory to motility. In sperm with a fibrous sheath, the amount of phosphatase released may overwhelm the stimulatory enzymes. This could also explain why the stimulatory effect of AKAP inhibitors is very rapid, whereas the inhibition of motility of mammalian sperm by these peptides takes longer. In order to explore

whether such mechanisms indeed occur, it will be necessary to compare the molecular structure of the axonemal and fibrous sheath AKAPs and directly determine whether these AKAPs differ with respect to the signal transduction enzymes that they tether.

Evolutionary significance of important regulators of sperm motility

Recent studies in *Chlamydomonas* have identified a 138 kDa subunit of inner arm dynein 1 (IAD1) whose phosphorylation regulates flagellar beating (King and Dutcher, 1997; Habermacher and Sale, 1997). Furthermore, the function of this subunit in the regulation of IAD function appears to be regulated by interplay between cAMP, Ca^{2+}, and PP1 (Habermacher and Sale, 1996). A major aspect of this interplay that must be mentioned is that cAMP inhibits motility in *Chlamydomonas* and Ca^{2+} stimulates motility. Thus, *Chlamydomonas* is not the best model to study regulation of flagellar motility that may also occur in sperm. We identified a pT-containing phosphoprotein in sea urchin sperm flagella whose phosphorylation was rapidly increased during initiation of motility (Bracho *et al.*, 1998). This protein remained tightly bound to the sperm axoneme after flagellar extraction with 0.6 M NaCl. Nonetheless, whether the 130 kDa sea urchin protein is related to the *Chlamydomonas* 138 kDa IAD subunit or not is a question that merits investigation. Our recent studies suggest that human sperm have a similar protein (Tash and Bracho, 1998). Thus, it appears that FP130 represents a family of axonemal proteins of significance to the regulation of flagellar motility. In order to be able to answer the ultimate question regarding the physiological importance of FP130 to sperm function, we have utilized a sperm cell for which much has already been published regarding the mechanisms involved in expression of flagellar motility. In addition, many mechanochemical and regulatory components of sea urchin sperm have been cloned and found to be relatives to components in mammalian sperm or spermatogenic cells (Neesen *et al.*, 1997; Milisav *et al.*, 1996; Yuan *et al.*, 1995). The sea urchin sperm also offers the ability to provide investigators with large quantities of homogenous cells, which are an excellent starting point for the purification of FP130. Once purified, the protein may then be used to generate peptide sequence and immunochemical probes that can be used to generate reagents for obtaining molecular information not only in sea urchins, but can be used to generate immunochemical and molecular information in other species as well, including the mouse and human.

Conclusion

In summary, lesions in sperm motility remain a major component of male factor infertility. Although many soluble and insoluble sperm phosphoproteins have been identified that are correlated with changes in motility, identification of flagellar components that are the primary link between signal transduction pathways and the flagellar axoneme remained elusive. The development of new models for investigation of the initiation and modulation of motility in live sperm that have never become motile prior to experimental activation will be a valuable tool to identify phosphoproteins that are part of the signal transduction pathway of initiation of axonemal function. Such new collection and storage methods will prove useful in developing similar technologies for species, such as the turkey, for which such methods are not currently available. A key and powerful aspect of the new model is that phosphorylation can be examined in live sperm where compartmentalization is intact. Our recent studies have demonstrated that initiation of motility in sea urchin sperm is associated with very rapid phosphorylation of pS and pT on very few proteins, and that these proteins are flagellar, some of which are tightly bound to the axoneme. In particular, the evidence points to the axonemal protein, FP130, as an important component of the signal transduction pathways that couple protein phosphorylation to regulation of sperm motility. In addition, evidence suggests that human sperm also contain a FP130 homologue. The second messenger pathways that regulate FP130, the localization of FP130 within the axoneme, the presence and functional identification of FP130-related proteins in mouse and human sperm, how FP130 regulates flagellar function, its developmental significance with regard to normal spermatogenesis and ultimately, the expression of motility in the mature sperm, and how the phosphorylation state and integrity of FP130 are altered in sperm with reduced motility and fertility, are all questions that can now be asked with a degree of certainty than has been heretofore not possible

References

Ahmad, K., Bracho, G.E., Wolf, D.P., and Tash, J.S. (1995) Regulation of human sperm motility and hyperactivation components by calcium, calmodulin, and protein phosphatases. *Arch. Androl.* **35**, 187-208.

Aitken, R.J., Irvine, S., and Kelly, R.W. (1986) Significance of intracellular calcium and cyclic adenosine 3',5'-monophosphate in the mechanisms by which prostaglandins influence human sperm function. *J. Reprod. Fertil.* 77, 451-462.

Ashizawa, K., Kamiya, T., Tamura, I., and Tsuzuki, Y. (1994) Inhibition of flagellar motility of fowl spermatozoa by L-carnitine: its relationship with respiration and phosphorylation of axonemal proteins. *Mol. Reprod. Dev.* **38**, 318-325.

Ashizawa, K., Wishart, G.J., Hashimoto, K., and Tsuzuki, Y. (1995) Dephosphorylation of a 30-kDa protein of fowl spermatozoa by the addition of myosin light chain kinase substrate peptide inhibits the flagellar motility. *Biochem. Biophys. Res. Commun.* **215**, 706-712.

Baccetti, B., Burrini, A.G., Maver, A., Pallini, V., and Renieri, T. (1979) "9 + 0" immotile spermatozoa in an infertile man. *Andrologia.* **11**, 437-443.

Berruti, G. and Martegani, E. (1989) Identification of proteins cross-reactive to phosphotyrosine antibodies and of a tyrosine kinase activity in boar spermatozoa. *J. Cell Sci.* **93**, 667-674.

Boatman, D.E. and Bavister, B.D. (1984) Stimulation of rhesus monkey sperm capacitation by cyclic nucleotide mediators. *J. Reprod. Fertil.* **71**, 357-366.

Bracho, G.E., Fritch, J.J., and Tash, J.S. (1997) A method for preparation, storage and activation of large populations of immotile sea urchin sperm. *Biochem. Biophys. Res. Commun.* **237**, 59-62.

Bracho, G.E., Fritch, J.J., and Tash, J.S. (1998) Identification of flagellar proteins that initiate the activation of sperm motility in vivo. *Biochem. Biophys. Res. Commun.* **242**, 231-237.

Brackett, B.G., Hall, J.L., and Oh, Y.K. (1978) In vitro fertilizing ability of testicular, epididymal, and ejaculated rabbit spermatozoa. *Fertil. Steril.* **29**, 571-582.

Brandt, H. and Hoskins, D.D. (1980) A cAMP-dependent phosphorylated motility protein in bovine epididymal sperm. *J. Biol. Chem.* **255**, 982-987.

Brillard, J.P. (1993) Sperm storage and transport following natural mating and artificial insemination. *Poult. Sci.* **72**, 923-928.

Brokaw, C.J. (1984) Cyclic AMP-dependent activation of sea urchin and tunicate sperm motility. *Ann. N. Y. Acad. Sci.* **438**, 132-141.

Brokaw, C.J. (1987) Regulation of sperm flagellar motility by calcium and cAMP-dependent phosphorylation. *J. Cell Biochem.* **35**, 175-184.

Carrera, A., Gerton, G.L., and Moss, S.B. (1994) The major fibrous sheath polypeptide of mouse sperm: structural and functional similarities to the A-kinase anchoring proteins. *Dev. Biol.* **165**, 272-284.

Cascieri, M., Amann, R.P., and Hammerstedt, R.H. (1976) Adenine nucleotide changes at initiation of bull sperm motility. *J. Biol. Chem.* **251**, 787-793.

Christen, R., Schackmann, R.W., and Shapiro, B.M. (1986) Ionic regulation of sea urchin sperm motility, metabolism and fertilizing capacity. *J. Physiol. Lond.* **379**, 347-365.

Coghlan, V.M., Bergeson, S.E., Langeberg, L., Nilaver, G., and Scott, J.D. (1993) A-kinase anchoring proteins: a key to selective activation of cAMP-responsive events? *Mol. Cell Biochem.* **127-128**, 309-319.

Coghlan, V.M., Perrino, B.A., Howard, M., Langeberg, L.K., Hicks, J.B., Gallatin, W.M., and Scott, J.D. (1995) Association of protein kinase A and protein phosphatase 2B with a common anchoring protein. *Science* **267**, 108-111.

Conraads, V.M., Galdermans, D.I., Kockx, M.M., Jacob, W.A., Van, S.C., and Coolen, D. (1992) Ultrastructurally normal and motile spermatozoa in a fertile man with Kartagener's syndrome. *Chest* **102**, 1616-1618.

Dey, C.S. and Brokaw, C.J. (1991) Activation of Ciona sperm motility: phosphorylation of dynein polypeptides and effects of a tyrosine kinase inhibitor. *J. Cell Sci.* **100**, 815-824.

Donoghue, A.M., Holsberger, D.R., Evenson, D.P., and Froman, D.P. (1998) Semen donor selection by in vitro sperm mobility increases fertility and semen storage in the turkey hen. *J. Androl.* **19**, 295-301.

Eliasson, R., Mossberg, B., Camner, P., and Afzelius, B.A. (1977) The immotile-cilia syndrome. A congenital ciliary abnormality as an etiologic factor in chronic airway infections and male sterility. *N. Engl. J. Med.* **297**, 1-6.

Fenichel, P., Gharib, A., Emiliozzi, C., Donzeau, M., and Menezo, Y. (1996) Stimulation of human sperm during capacitation in vitro by an adenosine agonist with specificity for A2 receptors. *Biol. Reprod.* **54**, 1405-1411.

Fraser, L.R. (1979) Accelerated mouse sperm penetration in vitro in the presence of caffeine. *J. Reprod. Fertil.* **57**, 377-384.

Fraser, L.R. (1981) Dibutyryl cyclic AMP decreases capacitation time in vitro in mouse spermatozoa. *J. Reprod. Fertil.* **62**, 63-72.

Fraser, L.R. (1990) Adenosine and its analogues, possibly acting at A2 receptors, stimulate mouse sperm fertilizing ability during early stages of capacitation. *J. Reprod. Fertil.* **89**, 467-476.

Galantino Homer, H.L., Visconti, P.E., and Kopf, G.S. (1997) Regulation of protein tyrosine phosphorylation during bovine sperm capacitation by a cyclic adenosine 3'5'-monophosphate-dependent pathway. *Biol. Reprod.* **56**, 707-719.

Garbers, D., First, N.L., and Lardy, H.A. (1973) The stimulation of bovine epididymal sperm metabolism by cyclic nucleotide phosphodiesterase inhibitors. *Biol. Reprod.* **8**, 589-598.

Garner, D.L. (1997) Ancillary tests of bull semen quality. *Vet. Clin. North Am. Food Anim. Pract.* **13**, 313-330.

Gingras, D., White, D., Garin, J., Multigner, L., Job, D., Cosson, J., Huitorel, P., Zingg, H., Dumas, F., and Gagnon, C. (1996) Purification, cloning, and sequence analysis of a Mr = 30,000 protein from sea urchin axonemes that is important for sperm motility. Relationship of the protein to a dynein light chain. *J. Biol. Chem.* **271**, 12807-12813.

Habermacher, G. and Sale, W.S. (1996) Regulation of flagellar dynein by an axonemal type-1 phosphatase in Chlamydomonas. *J. Cell Sci.* **109**, 1899-1907.

Habermacher, G. and Sale, W.S. (1997) Regulation of flagellar dynein by phosphorylation of a 138-kD inner arm dynein intermediate chain. *J. Cell Biol.* **136**, 167-176.

Hayashi, H. and Morisawa, M. (1988) Phosphorylation of a protein and the initiation of flagellar motility in rainbow trout spermatozoa. *Adv. Exp. Med. Biol.* **231**, 459-466.

Hayashi, H., Yamamoto, K., Yonekawa, H., and Morisawa, M. (1987) Involvement of tyrosine protein kinase in the initiation of flagellar movement in rainbow trout spermatozoa. *J. Biol. Chem.* **262**, 16692-16698.

Holt, C., Holt, W.V., Moore, H.D., Reed, H.C., and Curnock, R.M. (1997) Objectively measured boar sperm motility parameters correlate with the outcomes of on-farm inseminations: results of two fertility trials. *J. Androl.* **18**, 312-323.

Hoppe, P.C. (1975) Fertilizing ability of mouse sperm from different epididymal regions and after washing and centrifugation. *J. Exp. Zool.* **192**, 219-222.

Ishiguro, K., Murofushi, H., and Sakai, H. (1982) Evidence that cAMP-dependent protein kinase and a protein factor are involved in reactivation of triton X-100 models of sea urchin and starfish spermatozoa. *J. Cell Biol.* **92**, 777-782.

Jaiswal, B.S. and Majumder, G.C. (1996) In-vitro initiation of forward motility in testicular spermatozoa. *Int. J. Androl.* **19**, 97-102.

Jin, Z.X., Inaba, K., Manaka, K., Morisawa, M., and Hayashi, H. (1994a) Monoclonal antibodies against the protein complex that contains the flagellar movement-initiating phosphoprotein of Oncorhynchus keta. *J. Biochem. Tokyo.* **115**, 885-890.

Jin, Z.X., Nakajima, T., Morisawa, M., and Hayashi, H. (1994b) Isolation and properties of a protein complex containing flagellar movement-initiating phosphoprotein from testes of a white salmon. *J. Biochem. Tokyo.* **115**, 552-556.

Jones, R.C. and Murdoch, R.N. (1996) Regulation of the motility and metabolism of spermatozoa for storage in the epididymis of eutherian and marsupial mammals. *Reprod. Fertil. Dev.* **8**, 553-568.

Kartagener, M. (1933) Pathogenese der bronchiektasien. I. Bronctiektasien bei situs viscerum inversus. *Beitr. Klin. Tuberk.* **83**, 489-501.

King, S.J. and Dutcher, S.K. (1997) Phosphoregulation of an inner dynein arm complex in Chlamydomonas reinhardtii is altered in phototactic mutant strains. *J. Cell Biol.* **136**, 177-191.

Klauck, T.M., Faux, M.C., Labudda, K., Langeberg, L.K., Jaken, S., and Scott, J.D. (1996) Coordination of three signaling enzymes by AKAP79, a mammalian scaffold protein. *Science* **271**, 1589-1592.

Leclerc, P., de Lamirande, E., and Gagnon, C. (1996a) Cyclic adenosine 3',5'monophosphate-dependent regulation of protein tyrosine phosphorylation in relation to human sperm capacitation and motility. *Biol. Reprod.* **55**, 684-692.

Leclerc, P. and Gagnon, C. (1996b) Phosphorylation of Triton X-100 soluble and insoluble protein substrates in a demembranated/reactivated human sperm model. *Mol. Reprod. Dev.* **44**, 200-211.

Lester, L.B. and Scott, J.D. (1997) Anchoring and scaffold proteins for kinases and phosphatases. *Recent. Prog. Horm. Res.* **52**, 409-429.

McConnell, J.D. (1991) Abnormalities in sperm motility: Techniques of evaluation and treatment. In: *Infertility in the male* (Lipshultz, L.I. and Howards, S.S., Eds.) pp. 254-276, Mosby Year Book, St. Louis, MO.

Milisav, I., Jones, M.H., and Affara, N.A. (1996) Characterization of a novel human dynein-related gene that is specifically expressed in testis. *Mamm. Genome* **7**, 667-672.

Mohri, H. and Yanagimachi, R. (1980) Characteristics of motor apparatus in testicular, epididymal and ejaculated spermatozoa. A study using demembranated sperm models. *Exp. Cell Res.* **127**, 191-196.

Monks, N.J., Stein, D.M., and Fraser, L.R. (1986) Adenylate cyclase activity of mouse sperm during capacitation in vitro: effect of calcium and a GTP analogue. *Int. J. Androl.* **9**, 67-76.

Morton, B.E., Sagadraca, R., and Fraser, C. (1978) Sperm motility within the mammalian epididymis: species variation and correlation with free calcium levels in epididymal plasma. *Fertil. Steril.* **29**, 695-698.

Murofushi, H., Ishiguro, K., and Sakai, H. (1982) Involvement of cyclic AMP-dependent protein kinase and a protein factor in the regulation of the motility of sea urchin and starfish spermatozoa. In: *Biological functions of microtubules and related structures* (Sakai, H., Mohri, H., and Borisy, G.G., Eds.) pp. 163-176, Academic Press, New York, NY.

Neesen, J., Koehler, M.R., Kirschner, R., Steinlein, C., Kreutzberger, J., Engel, W., and Schmid, M. (1997) Identification of dynein heavy chain genes expressed in human and mouse testis: chromosomal localization of an axonemal dynein gene. *Gene* **200**, 193-202.

Noland, T.D., Abumrad, N.A., Beth, A.H., and Garbers, D.L. (1987) Protein phosphorylation in intact bovine epididymal spermatozoa: identification of the type II regulatory subunit of cyclic adeno-

sine 3',5'-monophosphate-dependent protein kinase as an endogenous phosphoprotein. *Biol. Reprod.* **37**, 171-180.

Paupard, M.C., Macleod, J., Wasco, W., and Orr, G.A. (1988) Major 56,000-dalton, soluble phosphoprotein present in bovine sperm is the regulatory subunit of a type II cAMP-dependent protein kinase. *J. Cell Biochem.* **37**, 161-175.

Phillips, D.M., Jow, W.W., and Goldstein, M. (1995) Testis factors that may regulate gene expression: evidence from a patient with Kartagener's syndrome. *J. Androl.* **16**, 158-162.

Pontonnier, F. and Bujan, L. (1993) [How to recognize and classify male infertility]. *Rev. Prat.* **43**, 941-947.

San Agustin, J.T. and Witman, G.B. (1994) Role of cAMP in the reactivation of demembranated ram spermatozoa. *Cell Motil. Cytoskeleton* **27**, 206-218.

Scott, J.D. (1997) Dissection of protein kinase and phosphatase targeting interactions. *Soc. Gen. Physiol. Ser.* **52**, 227-239.

Shapiro, B.M., Cook, S., Quest, A.F., Oberdorf, J., and Wothe, D. (1990) Molecular mechanisms of sea-urchin sperm activation before fertilization. *J. Reprod. Fertil. Suppl.* **42**, 3-8.

Sharma, R.K., Tolentino-MV, J., Thomas-AJ, J., and Agarwal, A. (1996) Optimal dose and duration of exposure to artificial stimulants in cryopreserved human spermatozoa. *J. Urol.* **155**, 568-573.

Si, Y. and Okuno, M. (1995) Activation of mammalian sperm motility by regulation of microtubule sliding via cyclic adenosine 5'-monophosphate-dependent phosphorylation. *Biol. Reprod.* **53**, 1081-1087.

Sigman, M., Lipshultz, L.I., and Howards, S.S. (1991) Evaluation of the subfertile male. In: *Infertility in the male* (Lipshultz, L.I. and Howards, S.S., Eds.) pp. 179-210, Mosby Year Book, St. Louis, MO.

Stephens, R.E. and Prior, G. (1995) Dynein inner arm heavy chain identification in cAMP-activated flagella using class-specific polyclonal antibodies. *Cell Motil. Cytoskeleton* **30**, 261-271.

Takahashi, D., Murofushi, H., Ishiguro, K., Ikeda, J., and Sakai, H. (1985) Phosphoprotein phosphatase inhibits flagellar movement of Triton models of sea urchin spermatozoa. *Cell Struct. Funct.* **10**, 327-337.

Tash, J.S. (1989) Protein phosphorylation: the second messenger signal transducer of flagellar motility. *Cell Motil. Cytoskeleton* **14**, 332-339.

Tash, J.S. and Bracho, G.E. (1994) Regulation of sperm motility: emerging evidence for a major role for protein phosphatases. *J. Androl.* **15**, 505-509.

Tash, J.S. and Bracho, G.E. (1998) Identification of phosphoproteins coupled to initiation of motility in live epididymal mouse sperm. *Biochem. Biophys. Res. Commun.* **in press**,

Tash, J.S. and Bracho, G.E. (1999) Microgravity alters protein phosphorylation changes during initiation of sea urchin sperm motility. *FASEB J.* **13**, in press.

Tash, J.S., Hidaka, H., and Means, A.R. (1986) Axokinin phosphorylation by cAMP-dependent protein kinase is sufficient for activation of sperm flagellar motility. *J. Cell Biol.* **103** , 649-655.

Tash, J.S., Kakar, S.S., and Means, A.R. (1984) Flagellar motility requires the cAMP-dependent phosphorylation of a heat-stable NP-40-soluble 56 kd protein, axokinin. *Cell* **38**, 551-559.

Tash, J.S. and Mann, T. (1973) Adenosine 3':5'-cyclic monophosphate in relation to motility and senescence of spermatozoa. *Proc. R. Soc. Lond. B. Biol. Sci.* **184**, 109-114.

Tash, J.S. and Means, A.R. (1982) Regulation of protein phosphorylation and motility of sperm by cyclic adenosine monophosphate and calcium. *Biol. Reprod.* **26**, 745-763.

Tash, J.S. and Means, A.R. (1983) Cyclic adenosine 3',5' monophosphate, calcium and protein phosphorylation in flagellar motility. *Biol. Reprod.* **28**, 75-104.

Trimmer, J.S. and Vacquier, V.D. (1986) Activation of sea urchin gametes. *Annu. Rev. Cell Biol.* **2**, 1-26.

Vijayaraghavan, S., Goueli, S.A., Davey, M.P., and Carr, D.W. (1997a) Protein kinase A-anchoring inhibitor peptides arrest mammalian sperm motility. *J. Biol. Chem.* **272**, 4747-4752.

Vijayaraghavan, S., Trautman, K.D., Goueli, S.A., and Carr, D.W. (1997b) A tyrosine-phosphorylated 55-kilodalton motility-associated bovine sperm protein is regulated by cyclic adenosine 3',5'-monophosphates and calcium. *Biol. Reprod.* **56**, 1450-1457.

Vo, N.K., Gettemy, J.M., and Coghlan, V.M. (1998) Identification of cGMP-dependent protein kinase anchoring proteins (GKAPs). *Biochem. Biophys. Res. Commun.* **246**, 831-835.

Voglmayr, J.K. and White, I.G. (1979) Effects of rete testis and epididymal fluid on the metabolism and motility of testicular and post-testicular spermatozoa of the ram. *Biol. Reprod.* **20**, 288-293.

Wolf, D.P., VandeVoort, C.A., Meyer, H.G., Zelinski, W.M., Hess, D.L., Baughman, W.L., and Stouffer, R.L. (1989) In vitro fertilization and embryo transfer in the rhesus monkey. *Biol. Reprod.* **41**, 335-346.

Working, P.K. (1988) Male reproductive toxicology: comparison of the human to animal models. *Environ. Health Perspect.* **77**, 37-44.

Yuan, L., Liu, J.G., and Hoog, C. (1995) Rapid cDNA sequencing in combination with RNA expression studies in mice identifies a large number of male germ cell-specific sequence tags. *Biol. Reprod.* **52**, 131-138.

Zheng, R.L. and Zhang, H. (1997) Effects of ferulic acid on fertile and asthenozoospermic infertile human sperm motility, viability, lipid peroxidation, and cyclic nucleotides. *Free Radic. Biol. Med.* **22**, 581-586.

Spermatogonial Transplantation

TAKEHIKO OGAWA
Department of Urology, Yokohama City University School of Medicine
Yokohama 236-0004, Japan

Abstract

The spermatogonial stem cell is the foundation of spermatogenesis and it continues to divide throughout the lifetime of a male. It divides to renew itself as well as to produce daughter cells which sequentially differentiate into spermatozoa. Transplantation of spermatogonial stem cells from a fertile mouse to the testis of sterile mouse has resulted in the development of donor cell derived spermatogenesis. In successful cases, the recipient mice have become fertile. Furthermore, spermatogenesis of the rat and hamster has occurred in the immunodeficient mouse testis following spermatogonial transplantation. It has also been shown that spermatogonial stem cells can be cryopreserved for later transplantation. These new techniques will be useful for basic spermatogenesis research and for future application to reproductive technology.

Introduction

Spermatogenesis is a complex process including several distinct but continuous steps of cell proliferation and differentiation which ultimately lead to the formation of spermatozoa. It can be divided into three phases. First, expansion of spermatogonia occurs on the basement membrane of the seminiferous tubule. The spermatogonial stem cells commit to differentiation division followed by sequential divisions of progeny cells, which give rise to numerous differentiated spermatogonia in the basal compartment of the seminiferous tubules. The second phase consists of meiotic division of spermatocytes which move into the adluminal compartment of the seminiferous tubules. During the meiotic division, homologous recombination of chromosomes takes place. The third phase, spermiogenesis, is characterized by dramatic morphological changes of spermatids into spermatozoa. All these processes are supported by surrounding somatic cells, especially Sertoli cells. Due to this complexity, it has not been possible to reproduce this process in vitro even though a number of trials have been done.

The foundation of spermatogensis resides in the spermatogonial stem cells that are located on the basement membrane of seminiferous tubules as a subset of spermatogonia. They have an enormous potential to keep renewing and to give rise to differentiated daughter cells throughout the life of a male. One spermatogonial stem

cell theoretically can produce 4,096 spermatozoa (Russell *et al.*, 1990), although more than half of them will disappear mainly through apoptotic cell death (Tegelenbosch and de Rooij, 1993; Barratt, 1995). Nonetheless, the number of spermatozoa produced from a single stem cell is considerable. While a single spermatozoon transmits only a defined set of genetic information to the offspring through fertilization of an oocyte, the spermatogonial stem cell carries the entire genetic code of the male. Through meiotic recombination, resultant haploid cells acquire genomic diversity. Therefore, any two spermatozoa from a single animal could be different in terms of their genetic composition.

A spermatogonial transplantation technique has been developed in rodents by Brinster and colleagues (Brinster and Zimmermann, 1994; Brinster and Avarbock, 1994). Through this technique, spermatogonial stem cells can be collected, maintained in vitro, then transplanted to recipient testes to establish spermatogenesis. The recipient mouse became fertile and produced offspring carrying the donor genotype.

Development of Spermatogonial Transplantation

The spermatogonia of a fertile mouse can be transplanted to the testis of a sterile mouse (Brinster and Zimmerman, 1994). The testis cells of a donor mouse were collected using a two-step enzymatic digestion protocol to make a cell suspension. The cells were then introduced into the seminiferous tubules of a recipient mouse testis. The donor cells migrated to the basement membrane of the recipient seminiferous tubules and formed colonies that developed spermatogenesis. Recipient mice were treated with busulfan to eliminate endogenous germ cells prior to the transplantation. Mice carrying the white spotting mutation (*W*) were also available as recipients because they lack germ cells due to a mutation in the c-kit proto-oncogene receptor which is essential for germ cell development during embryogenesis. In successful cases of spermatogonial transplantation, the recipient mice became fertile and had offspring through mating with females. Some offspring showed the donor haplotype which indicated that donor derived spermatogenesis was complete to form functional sperm in the recipient testis (Brinster and Avarbock, 1994).

Procedure of Spermatogonial Transplantation

Transgenic mice or rats that carry the *lacZ* gene as a marker were used as

donor animals in most cases to help distinguish the donor cell derived spermatogenesis from residual endogenous spermatogenesis of a recipient. In an experiment where the hamster was used as donor animal, histological examination was performed to distinguish morhpological differences between hamster spermatogenesis and spermatogenesis of the recipient mouse (Ogawa *et al.*, 1999) because transgenic hamsters carrying a *lacZ* marker gene are not available.

The testis cells of donor animals were collected by a two-step enzymatic digestion (Bellve *et al.*, 1977). Briefly, decapsulated testis tissue was treated with collagenase to remove interstitial tissues, followed by trypsin and EDTA digestion for making single cell suspensions from seminiferous tubules. The cells were suspended in DMEM+10% FBS at concentrations between 20 - 200×10^6 /ml. This cell suspension includes different stages of germ cells and some somatic cells.

One of the key preparations for spermatogonial transplantation is to render recipient mice devoid of germ cells. Busulfan treatment proved to be one of the best methods to eliminate germ cells including stem cells. Busulfan at a dose of 40 mg/kg or more was given intraperitoneally to the recipient mouse 40 to 60 days prior to the transplantation. *W* mutant mice are also available as recipient mice because they lack germ cells due to a mutation in the c-kit receptor (Brinster and Zimmermann, 1994).

The procedure of donor cell injection into the seminiferous tubules of a recipient mouse can be performed through different routes: direct injection into seminiferous tubules, injection into the rete cavity, or injection through the efferent duct (Ogawa *et al.*, 1997).

In case of xenotranplantation, immunodeficient mice, nude or SCID, were used as recipients to avoid immunological rejection. Several months after transplantation, the recipient mouse testes were taken out for *lacZ* staining. The seminiferous tubules containing donor spermatogenesis expressing the *lacZ* transgene will stain blue which makes it possible to unequivocally identify donor-derived spermatogenesis.

Donor testes cells can also be frozen in the same way as cultured somatic cells for later use in transplantation. Briefly, collected testis cells were suspended in DMEM +10% FBS followed by drop-wise addition of freezing medium (FBS, DMEM+10% FBS, DMSO in a ratio of 1:3:1) in a volume equal to the original cell suspension. Cells were aliquoted 1.0 ml per freezing vial, placed in an insulated container at -70 °C for more than 12 hours and then moved into liquid nitrogen (-196 °C). The thawed cells

were resuspended in DMEM+10% FBS and introduced into recipient testes in the same way as freshly collected cells.

Xenogeneic Spermatogonial Transplantation and Cryopreservation of Spermatogonial Stem Cells.

In 1996, Brinster and colleagues reported that rat spermatogenesis had proceeded in the immunodeficient mouse testis (Clouthier *et al.*, 1996). The success of xenogeneic spermatogonial transplantation has expanded the possibility for applications of this technique to animals other than mice. In addition, they also reported that spermatogonial stem cells could be cryopreserved in the same manner as somatic cells (Avarbock *et al.*, 1996). Considering the species specific protocols necessary for sperm cryopreservation, the comparatively easy preservation of stem cells will be beneficial for animals for which reliable procedures of sperm cryopreservation have not been developed.

Spermatogenesis of each species has a unique defined duration from the onset of stem cell differentiation to the completion of spermatozoa formation. It is also known that the length of each step of spermatogenesis is rigidly determined. Based on numerous observations and experimental results on the relationship between germ cells and Sertoli cells, it has been supposed that Sertoli cells are regulating almost every detail of spermatogenesis including the length of the spermatogenic processes. Combining the classic [^{3}H]-thymidine incorporation method and the new xenogeneic spermatogonial transplantation technique, it has been shown that rat spermatogenesis in the mouse seminiferous tubules is supported by mouse Sertoli cells and proceeded with the timing typical for rat spermatogenesis. This result demonstrated that the cell differentiation process of spermatogenesis is regulated exclusively by the germ cells. Sertoli cells are indispensable but they do not regulate the duration of spermatogenesis (Franca *et al.*, 1998).

Very recently, Brinster and colleagues have successfully transplanted hamster testis cells that resulted in spermatogenesis in the recipient mouse testis (Ogawa *et al.*, 1999). Considering the phylogenetic distance between mouse and hamster, which 1.6 times longer than that of mouse and rat, the results might suggest that xenogeneic spermatogonial transplantation is also possible in other animals for production of donor spermatozoa. In addition, frozen hamster testis cells transplanted to mouse testes

also resulted in spermatogenesis. Taken together, it might be possible to freeze the testis cells of a particular species, and at a later time the cells can be thawed and transplanted into the testis of not only the same but also a different species.

Conclusion

The technique of spermatogonial transplantation both in the same species and inter-species has tremendous potential for development of future applications to reproductive technologies. I expect that cryopreservation of diploid spermatogonial stem cells can be achieved in a broad, if not universal, range of species. Therefore, the spermatogonial transplantation technique together with cryopreservation of spermatogonial stem cells would be a powerful tool in conservation of genetic resources, such as endangered species and valuable agricultural animals. Now we can preserve spermatogonial stem cells from those animals for future transplantation when the recipient animal systems become available to support donor spermatogenesis and subsequent transmission of genetic material.

One of the next important steps is the development of a culture system for spermatogonia which will allow expansion of spermatogonia, leading to their genetic modification. Then it will be possible to make genetically modified animals through spermatogonial transplantation in combination with other assisted reproductive technologies.

Acknowledgements

I would like to express my gratitude to Prof. Ralph L. Brinster for his heartful encouragement along with precious suggestions and instructions through all experiments included in this review. I also acknowledge helpful discussions and suggestions from my colleagues, especially, Ina Dobrinski and Makoto Nagano.

References

Avorbock MR, Brinster CJ, Brinster RL. 1996. Reconstitution of spermatogenesis from frozen spermatogonial stem cells. Nature Med 2: 693-696.

Barratt CLR. 1995. Spermatogenesis, In Gametes - The spermatozoon (Eds. JG Grudzinskas and JL Yovich) Cambridge University Press, New York, pp. 250-267.

Bellve AR, Cavicchia JC, Millette CF, O'Brien DA, Bhatnagar YM, Dym M. 1977. Spermatogenic cells of the prepubertal mouse. Isolation and morphological characterization. J Cell Biol 74: 68-85.

Brinster RL and Zimmermann JW, 1994. Spermatogenesis following male germ-cell transplantation. Proc Natl Acad Sci USA, 91:11298-11302.

Brinster RL and Avorbock M, 1994. Germ line transmission of donor haplotype following spermatogonial transplantation. Proc Natl Acad Sci USA, 91:11303-11307.

Clouthier DE, Avarbock MR, Shanna DM, Hammer RE, Brinster RL, 1996. Rat spermatogenesis in mouse testis. Nature 381: 418-421.

Franca LR, Ogawa T, Avarbock MR, Brinster RL, Russell LD. 1998. Germ Cell Genotype Controls Cell Cycle during Spermatogenesis in the Rat. Biol Reprod 59: 1371-1377.

Ogawa T, Arechaga JM, Avarbock MR, Brinster RL, 1997. Transplantation of testis germinal cells into mouse seminiferous tubules. Int J Dev Biol, 41:111-122.

Ogawa T, Dobrinski I, Avorbock M, Brinster RL, 1999. Xenogeneic spermatogenesis following transplantation of hamster germ cells to mouse testes. Biol. Reprod. 60: 515-521.

Russell LD, Ettlin RA, Hikim AP, Clegg ED. 1990. Mammalian spermatogenesis, In Histological and Histopathological Evaluation of the Testis. Cach River Press, Clearwater, pp. 1-40.

Tegelenbosch RAJ and de Rooij DG. 1993. A quantitative study of spermatogonial multiplication and stem cell renewal in the C3H/101 F1 hybrid mouse. Mutat. Res. 290: 193-200.

Spermatogenic Cells as Male Gametes

ATSUO OGURA
Department of Veterinary Science, National Institute of Infectious Diseases
Tokyo 162-8640, Japan

Abstract

Spermatogenic cells are male germ cells that undergo mitosis, meiosis, and spermiogenesis in the seminiferous tubules. Due to recent advances in microfertilization techniques, spermatogenic cells at certain stages have been used to construct diploid zygotes that subsequently develop to term. Using round spermatids, the youngest haploid germ cells, normal births have been reported in the mouse, rabbit, and human. Secondary and primary spermatocytes have also been demonstrated to support full term development after incorporation into mature or immature homologous oocytes. The techniques not only provides oppotunities of treating infertile males, but also provides valuable information on gametogenesis including spermatogenesis, meiosis, and genomic imprinting.

Introduction

In mammalian seminiferous tubules, there are five types of germ cell (spermatogenic cells): spermatogonia, primary and secondary spermatocytes, spermatids and spermatozoa. Although the testicular spermatozoa have already completed the morphological changes that are needed to carry paternal genomes safely, they acquire the ability to fertilize oocytes as they pass through the epididymis and mature. Due to advances in microfertilization techniques, however, these days the term 'fertilizing ability' means the ability of male germ cells to form normal diploid zygotes after transfer into oocytes. In this context, elongated and round spermatids, and secondary and primary spermatocytes have fertilizing ability, as demonstrated by recent microfertilization studies. The clinical and biological significances of microfertilization with spermatogenic cells have been published elsewhere (Ogura and Yanagimachi, 1995; Tesarik *et al.*, 1998; Ogura and Yanagimachi, 1999). In this paper, I review microfertilization with spermatogenic cells in mammals, with special emphasis on its technical aspects that differs remarkably from those of microfertilization with mature spermatozoa.

Technical aspects

1. Collection of spermatogenic cells

Spermatogenic cells should be collected from testicular tissues, although sometimes in humans, round spermatids could be collected from the ejaculates (Tesarik *et al.*, 1995). In general, we can readily obtain suspension of testicular cells containing >80% spermatogenic cells by either enzyme treatment or mechanical separation with pipetting (Ogura and Yanagimachi, 1993). The remaining nucleated cells in the suspension include Sertoli cells, Leydig cells, and leukocytes. Although there are several ways to prepare spermatogenic cell suspensions with more purity, we can correctly identify spermatogenic cells before microfertilization by their morphology or chromosome analyses, as described below.

2. Identification of spermatogenic cells

Most spermatogenic cells can be identified by their size and morphology, but some experience is needed. For more precise identification, presumptive cell types are injected into Met II-arrested oocytes (the mouse oocyte is most suitable) and then the chromosome ploidy is determined (Kimura and Yanagimachi, 1995a).

3. Method of nuclear transfer

Unlike mature spermatozoa, most spermatogenic cells are round cells with a decondensed nucleus. Therefore, they can be introduced into oocytes by either intracytoplasmic injection or electrofusion. Electrofusion is thought to be less traumatic than intracytoplasmic injection, but generally less efficient (compare Ogura *et al.* 1994 with Kimura and Yanagimachi, 1995a). Each nuclear transfer method has inherent advantages and disadvantages (Table 1).

4. Oocyte activation and the timing of nuclear transfer

Round spermatids and younger spermatogenic cells in mouse and hamster have little or no capacity to activate oocytes (Ogura and Yanagimachi, 1993; Kimura and Yanagimachi, 1995a). Therefore, oocytes should be activated artificially at or around the time of transfer of spermatogenic cell nuclei. Mammalian oocytes can be activated by chemical treatment or electric stimulus. The method of oocyte activation, as well as its timing, may affect subsequent embryo development. The combination of spermatogenic cells and oocytes that would result in the formation of diploid zygotes is shown (Table 2). These protocols work best in mouse experiments because premature chromosome condensation of the spermatogenic cell nuclei can readily be induced in mouse oocytes. This character allows the chromosome pairs to segregate correctly and,

Table 1. Comparison of electrofusion and intracytoplastic injection for introduction of the spermatogenic cell neuclei into oocytes

Method	Efficiency	Cell viability	Incorporation	Intermingling of donors and oocytes
Electrofusion	Low to moderate	High	Intact cells	Slow
Intracytoplastic injection	High	Moderate to high	Isolated nuclei or ruptured cells	Quick

Table 2. Combination of spermatogenic cells and oocytes for microfertilization

Spermatogenic cells	Stage of oocytes	Species in which normal birth have been reported
Round spermatid	Metaphase II	Human, Rabbit
	Metaphase II (PCC)	Mouse
	Telophase II	Mouse
Secondary spermatocyte	Metaphase II (PCC)	Mouse
Primary spermatocyte	Metaphase II (PCC)	Mouse
	Metaphase I (PCC)	Mouse

PCC: Premature chromosome condensation

very recently, we demonstrated that the spermatid nuclei could support full term development following premature condensation of their chromosomes in Met II-arrested oocytes (Ogura *et al.*, 1999). In the case of round spermatids from human and cynomolgus monkeys, oocytes are activated at the time of spermatid injection because they probably contain oocyte-activating factor(s) (Sousa *et al.*, 1996; Ogura *et al.*, 1997a).

Historical overview

Since the first attempt to fertilize hamster oocytes with round spermatids (Ogura and Yanagimachi,1993), microfertilization with spermatogenic cells has become an important technique for studying the mechanisms of mammalian fertilization and for treating male factor infertility in humans (Tesarik *et al.* 1998) and animals (Ogura *et al.*, 1996a; Tanemura *et al.*, 1997). Its history is summarized in Table 3. Very recently, normal mice were born after microinsemination with primary spermatocytes, the male germ cells before the first meiotic division (Ogura *et al.*, 1998). This indicates that, at least in some male germ cells, the epigenetic modifications needed for normal embryo development are completed before meiosis I. Cryopreservation of spermatogenic cells has become possible in some mammalian species including humans (Ogura *et al.*, 1996b; Ogura *et al.*, 1997a; Antinori *et al.*, 1997). This avoids the repeated collection of testicular samples and may increase the chance of rescuing endangered species.

Recently we have succeeded in obtaining fertilized oocytes in mastomys, an African rodent, using round spermatids as male gametes (unpublished). Although induction of superovulation and embryo culuture in vitro is relatively easy in this species (Ogura *et al.*, 1997b), it is extremely difficult to fertilize oocytes in vitro (Nohara *et al.*, 1998) or to inject spermatozoa into oocytes due to technical problems (e.g., large sperm head) and the character of their oocytes. After round spermatid injection, however, about 90% of oocytes survived the injection procedure and most (>90%) of them formed the male pronucleus. These oocytes fertilized with round spermatids developed to the blastocyst stage, although so far no pups have been obtained after embryo transfer.

Future prospects

The fact that round spermatids and spermatocytes can be used to fertilize oocytes has several implications for basic research as well as for human clinical

Table 3. Microinsemination with spermatogenic cells by intracytoplasmic injection (ICI) or electrofusion (EF)

Year	Species	Spermatogenic cells used	Method	Development	Reference
1993	golden hamster	round spermatid	ICI	2-cell	Ogura and Yanagimachi
1993	hamster, mouse	round spermatid	EF	2-cell	Ogura *et al.*
1994	mouse	round spermatid	EF	term	Ogura *et al.*
1994	rabbit	round spermatid	ICI	term	Sofikitis *et al.*
1995a	mouse	round spermatid	ICI	term	Kimura and Yanagimachi
1995	human	elongated spermatid	ICI	term	Fisahel *et al.*
1995	human	round spermatid	ICI	term	Tesarik *et al.*
1995b	mouse	secondary spermatocyte	ICI	term	Kimura and Yanagimachi
1996	bovine	culture-derived spermatid	ICI	blastcyst	Goto *et al.*
1996b	mouse	frozen-thawed spermatid	ICI	term	Ogura *et al.*
1997	human	frozen-thawed spermatid	ICI	ongoing preg.	Antinori *et al.*
1997c	mouse	primary spermatocyte	EF	blastocyst	Ogura *et al.*
1998	mouse	primary spermatocyte	ICI	term*	Sasagawa *et al.*
1998	mouse	primary spermatocyte	EF	term	Ogura *et al.*
1998	mouse	Met I spermatocyte	ICI	term	Kimura *et al.*
1999	mouse	round spermatid**	ICI	term	Ogura *et al.*
1999	porcine	round spermatid	ICI	blastocyst	Kim *et al.*
	mastmys	round spermatid	ICI	blastocyst	

* Pups died before wearing. ** After premature chromosome condensation.

practice. Its application for generating transgenic animals is an example. The offspring derived from spermatogenic cells with exogenous DNA must carry the expected genes. The culture of spermatogenic cells in vitro, which is still extremely difficult, would facilitate the use of spermatogenic cells as male gametes in future studies.

References

Antinori, S., Versaci, C., Dani, G., Antinori, M., Selman, H. A. 1997. Successful fertilization and pregnancy after injection of frozen-thawed round spermatids into human oocytes. Human Reproduction 12: 554-556.

Fishel, S., Green, S., Bishop, M., Thornton, S., Hunter, A., Fleming, S., Al-Hassan, S. 1997. Pregnancy after intracytoplasmic injection of spermatid. Lancet 345: 1641-1642.

Goto, K., Kinoshita, A., Nakanishi, Y., Ogawa, K. 1996. Blastocyst formation following intracytoplasmic injection of in-vitro derived spermatids into bovine oocytes. Human Reproduction 11: 824-829.

Kimura, Y., Yanagimachi, R. 1995a. Mouse oocytes injected with testicular spermatozoa or round spermatids can develop into normal offspring. Development 121: 2397-2405.

Kimura, Y., Yanagimachi, R. 1995b. Development of normal mice from oocytes injected with secondary spermatocyte nuclei. Biology of Reproduction 53: 855-862.

Kimura, Y., Tateno, H., Handel, M. A., Yanagimachi, R. 1998. Factors affecting meiotic and developmental competence of primary spermatocyte nuclei injected into mouse oocytes. Biology of Reproduction 59: 871-877.

Kim, N.-H., Shin, J.S., Kim, C., Jun, S.H., Lee, H.T., Chung, K.S. 1999. Fertilization and in vitro development of porcine oocytes following intracytoplasmic injection of round spermatid or round spermatid nuclei. Theriogenology (in press).

Nohara, M., Hirayama, T., Ogura, A., Hiroi, M., Araki, Y. 1998. Partial characterization of the gametes and development of a successful in vitro fertilization procedure in the mastomys (*Praomys coucha*): a new species for reproductive biology research. Biology of Reproduction 58: 226-233.

Ogura, A., Yanagimachi, R. 1993. Round spermatid nuclei injected into hamster oocytes form pronuclei and participate in syngamy. Biology of Reproduction 48: 219-225.

Ogura, A., Yanagimachi, R., Usui, N. 1993. Behavior of hamster and mouse round spermatid nuclei incorporated into mature oocytes by electrofusion. Zygote 1: 1-8.

Ogura, A., Matsuda, J., Yanagimachi, R. 1994. Birth of normal young following fertilization of mouse oocytes with round spermatids by electrofusion. Proceedings of the National Academy of Science, U.S.A. 91: 7460-7462.

Ogura, A. and Yanagimachi, R. 1995. Spermatids as male gametes. Reproduction, Fertility and Development 7: 155-159.

Ogura, A., Yamamoto, Y., Suzuki, O., Takano, K., Wakayama, T., Mochida, K., Kimura, H. 1996a. In

vitro fertilization and microinsemination with round spermatids for propagation of nephrotic genes in mice. Theriogenology 45: 1141-1149.

Ogura, A., Matsuda, J., Asano, T., Suzuki, O., Yanagimachi, R. 1996b. Mouse oocytes injected with cryopreserved round spermatids can develop into normal offspring. Journal of Assisted Reproduction and Genetics 13: 431-434.

Ogura, A., Matsuda, J., Mochida, K., Suzuki, O., Takimoto, K., Nakayama, K., Naiki, M. 1997b. Preimplantation embryo development in mastomys (*Praomys coucha*) in vivo and in vitro. Journal of Reproduction and Development 43: 65-71.

Ogura, A., Matsuda, J., Suzuki, O., Sankai, T., Ogonuki, N., Yoshikawa, Y., Yanagimachi, R. 1997a. Cryopreservation of mammalian spermatids. Journal of Reproduction and Development 43(Suppl): 103-104.

Ogura, A., Wakayama, T., Suzuki, O., Shin, T.-Y., Matsuda, J., Kobayashi, Y. 1997c. Chromosomes of mouse primary spermatocytes undergo meiotic divisions after incorporation into homologous immature oocytes. Zygote 5: 177-182.

Ogura, A., Suzuki, O., Tanemura, K., Mochida, K., Kobayashi, Y., Matsuda, J. 1998. Development of normal mice from metaphase I oocytes fertilized with primary spermatocytes. Proceedings of the National Academy of Science, U.S.A. 95: 5611-5615.

Ogura, A. and Yanagimachi, R. 1999. Microinsemination using spermatogenic cells in mammals. In: Male Sterility for Motility Disorders (ed. Hamamah, S.), Heidelberg, Germany: Springer-Verlag, pp.189-202.

Ogura, A., Inoue, K., Matsuda, J. 1999. Spermatid nuclei can support full term development after premature chromosome condensation within mature oocytes. Human Reproduction 14: 1294-1298.

Sasagawa, I., Kuretake, S., Eppig, J., Yanagimachi, R. 1998. Mouse primary spermatocytes can complete two meiotic divisions within the oocyte cytoplasm. Biology of Reproduction 58: 248-254.

Sofikitis, N. V., Miyagawa, I., Agapitos, E., Pasyianos, P., Toda, T., Hellstrom, W. J. G., Kawamura, H. 1994. Reproductive capacity of the nucleus of the male gamete after completion of meiosis. Journal of Assisted Reproduction and Genetics 11: 335-341.

Sousa, M., Mendoza, C., Barros, A., Tesarik, J. 1996. Calcium responses of human oocytes after intracytoplasmic injection of leukocytes, spermatocytes and round spermatids. Molecular Human Reproduction 2: 853-857.

Tanemura, K., Wakayama, T., Kuramoto, K., Hayashi, Y., Sato, E., Ogura, A. 1997. Birth of normal young by microinsemination with frozen-thawed round spermatids collected from aged azoospermic mice. Laboratory Animal Science 47: 203-204.

Tesarik, J., Mendoza, C., Testart, J. 1995. Viable embryos from injection of round spermatids into oocytes. New England Journal of Medicine 333: 525.

Tesarik, J., Sousa, M., Greco, E., Mendoza, C. 1998. Spermatids as gametes: indications and limitations. Human Reproduction 13: 89-107.

Use of Freeze-Thaw and Freeze-Dry Spermatozoa for Intra-Cytoplasmic Sperm Injection

T.WAKAYAMA and R.YANAGIMACHI
Department of Anatomy and Reproductive Biology,
University of Hawaii Medical School,
1951 East-West Road. Honolulu, Hawaii 96822, USA

Introduction

Genetically engineered animals could be the subjects for study of human diseases and therapy (Sharp and Mobraaten, 1997). Thus, the number of genetically engineered animals, especially mice, has increased to the point where it became a serious financial burden to maintain them. Maintaining all genetically engineered animals in breeding colonies has potential disadvantages: the loss of animals due to (a) impaired reproductive efficiency, (b) disease or catastrophic accidents and (c) genetic drift. The most commonly used alternative to the maintenance of numerous mouse strains and mutants is to cryopreserve preimplantation embryos and later used them when they became needed. Many potentially important strains of mouse, as well as genetically engineered mice, have been "banked" using embryo cryopreservation technique. While embryo cryopreservation is an alternative to maintain breeding colonies, it still requires significant physical facilities and financial backup. Many females are required to obtain sufficient numbers of embryos to be frozen. Thus, many laboratories are looking for more cost effective means of preserving important mouse germplasm.

To maintain a particular strain of mouse, approximately 100 females are required to provide sufficient numbers of embryos. In contrast, only one or two males would be enough to preserve that strain. Although sperm cryopreservation is well advanced in animal husbandry and human reproductive medicine, cryopreservation of mouse spermatozoa has not been very successful. This contrasts to the highly successful cryopreservation of mouse oocytes and preimplantation embryos. Poor sperm survival after freeze-thawing is due, at least in part, to the fact that mouse spermatozoa are extremely sensitive to osmotic (Willoughby *et al.*, 1996) and mechanical shock damages (Wakayama *et al.*, unpublish data). At present, the rate of sperm survival after freeze-thawing is rather low, yet it is possible to use these spermatozoa for *in vitro* fertilization to obtain

normal offspring (Nakagata *et al.*, 1996; Songsasen *et al.*, 1997).

We report that mouse spermatozoa that had been frozen or freeze-dried without cryoprotectants are all dead (or almost all are dead) in the conventional sense, yet they are capable of supporting normal embryo development when they are injected microsurgically into oocytes (Wakayama *et al.*, 1998; Wakayama and Yanagimachi, 1998).

Materials and methods

Sperm collection

For each experiment, two caudae epididymides of a mature B6D2F1 male were used. While applying finger pressure to each epididymis, its distal portion was punctured with sharp forceps. A dense sperm mass oozing out of the epididymis was transferred into a 1.5 ml polypropylene tube containing 1 ml of a test medium. After incubation for 30 min at 37.5 °C, the upper 0.3-0.5 ml of the sperm suspension was removed from the tube and used for each experiment. Over 90% of spermatozoa in this suspension (approximately 3-10×10^6/ml) were actively motile.

Freeze-thaw experiment

Sperm suspensions were placed in CZB, PBS or isotonic NaCl with or without raffinose. Aliquot (50 μl) of each suspension were dispensed into 1 ml cyrotubes and place directly into a –15 °C or –50 °C freezer or liquid nitrogen (-196 °C). All samples were stored for 1 day to 18 months. Just before experiment, tube were removed from freezer or liquid nitrogen and placed in water or air at 24-26 °C for about 10 min. Then 5 μl of sperm suspension was thoroughly mixed with 50 μl Hepes-CZB containing 12% (w/v) polyvinylpyrrolidone (PVP, average Mol. wt. 360,000; Sigma).

Freeze-drying experiment

Two major test solutions were CZB medium (Chatot *et al.*, 1990) without EDTA and Dulbecco's modified Eagle's medium (DMEM; Sigma Chemical Co, St. Louis, MO) supplemented with 10% (v/v) fetal bovine serum (Hyclone; Logan, UT). An aliquot (100 μl) of the sperm suspension was put in a 2 ml ampul (Wheaton Scientific, Millville, NJ; cat. no. 651506) which was plunged directly into liquid nitrogen. Ten minutes later, ampules were placed in a pre-cooled (-50 °C) freeze-flask attached to a freeze-dry system (Model 10-020, VirTis

Co., Gardner, NY). Inlet pressure was approximately 1 milli torr. About 12 h later, the flask was removed from the system after it had been filled with argon supplied by way of a gas-drying jar (Fisher Scientific, Pittsburgh, PA; cat. no. 09-204). Each ampul was connected to a vacuum pump and frame-sealed after >99% of the gas was pumped out it. Ampules were individually wrapped with aluminum foil and stored at room temperature (about 25 °C) or at 4 °C. All samples were stored for 1 day to 3 months. Just before experiment, the ampul was broken and 100 μl of distilled water was added to the dried spermatozoa. Then 5 μl of sperm suspension was thoroughly mixed with 50 μl Hepes-CZB containing 12% (w/v) PVP.

Sperm viability test

Sperm viability was assessed using a commercially available cell viability test kit (Live/dead FertiLight; Molecular Probes, Eugene, Oregon) which differentiate between plasma membrane-intact (live) and -damaged (dead) cells according to fluorescence staining pattern under a UV microscope. The nuclei of "live" spermatozoa with intact plasma membranes fluoresced green whereas those of "dead" spermatozoa fluoresced bright orange-red.

Oocyte preparation and sperm injection

Mature B6D2F1 females were induced to superovulate by consecutive injections of pregnant mare serum gonadotropin and human chorionic gonadotropin (hCG) 48 h apart. Fourteen hours after hCG injection, cumulus-oocyte complexes were collected from oviducts and treated with bovine testicular hyaluronidase (300 USP units/ml) in Hepes-buffered CZB medium (Hepes-CZB) for 3 min to disperse cumulus cells. Before sperm injection the oocytes were rinsed and stored in CZB medium for up to 4 h at 37.5 °C in an atmosphere of 5% CO_2 in air. A single spermatozoon was aspirated into an injection pipette attached to a Piezo electric pipette-driving unit (Kimura and Yanagimachi, 1995). After the sperm head and tails were separated by applying a single or a few Piezo pulses to the neck region, each head alone was injected into each oocyte. All injections were performed in Hepes-CZB at room temperature within 1 h of sperm rehydration. Sperm-injected oocytes were incubated in CZB at 37.5 °C under 5% CO_2 in air and examined 5-6 h later. Those with two distinct pronuclei and a second polar body were considered normally fertilized.

Embryo transfer

Normal fertilized eggs were cultured for 4 days and those reaching the morulae or blastocyst stages were transferred into the uterine horns of recipient CD-1 females (albino) which had been mated with vasectomized CD-1 males three days previously. A mean number of eight morulae/blastocysts was transferred into each horn. Females were allowed to deliver and raise their surrogate offspring (with black, brown or gray coats). Some mature male and female offspring were randomly selected and mated to examine their fertility.

Results and Discussion

After thawing or rehydrating, almost all (frozen) or all (freeze-dried) spermatozoa were motionless and judged to be dead by live/dead cell staining. Electron microscopy revealed that sperm plasma membranes and the acrosome were extensively disrupted.

The majority of the oocytes fertilized by the spermatozoa frozen in CZB medium developed to blastocysts (about 90%, Fig.1). Embryonic development was poor when the oocytes were fertilized by spermatozoa frozen in PBS and isotonic saline, especially when they were frozen in the absence of raffinose. When spermatozoa were suspend and frozen in CZB medium and stored for 1 month at -15 °C, -50 °C or -196 °C, then injected into oocytes, almost all of the oocytes developed to the blastocyst stage (Fig.2). The proportion of two-cell embryos developed into live offspring after transfer to foster mothers are summarized (Figs.3 and 4). Normal offspring were obtained after all treatments. The best results were obtained when spermatozoa were suspended and frozen at -50 °C or -196 °C in the absence of raffinose (Fig.3). Spermatozoa could participate in the development of normal offspring after 18 months storage at -15 °C (Fig.4).

Freeze-dried spermatozoa were rehydrated and their heads were injected individually into oocytes. Most oocytes were fertilized normally regardless of the type of sperm-suspension medium (CZB, Fig.5; DMEM, Fig.6), and the storage temperature (room temperature 24-26 °C; 4 °C), and periods of freeze-dried spermatozoa (1 day to 3 weeks). The majority of fertilized eggs could develop into morulae/blastocysts *in vitro*. When transferred to foster mothers, many embryos developed into normal offspring which grew normally. The effects of storage temperature and period of freeze-dried spermatozoa on the embryonic development to term are shown (Figs.7 and 8). Embryonic development (from morulae

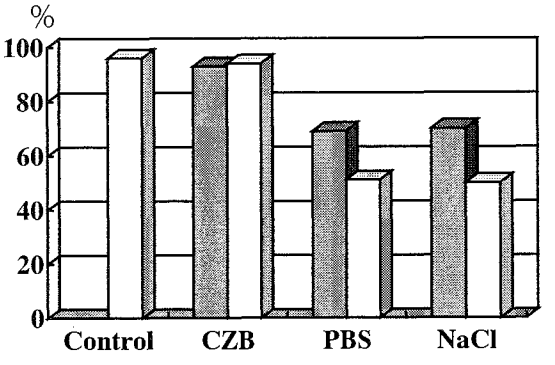


Fig. 1. Effect of sperm freezing media on embryo development. Spermatozoa were stored in LN_2 for 1 month. Sperm were frozen without raffinose (□) or with raffinose (■).

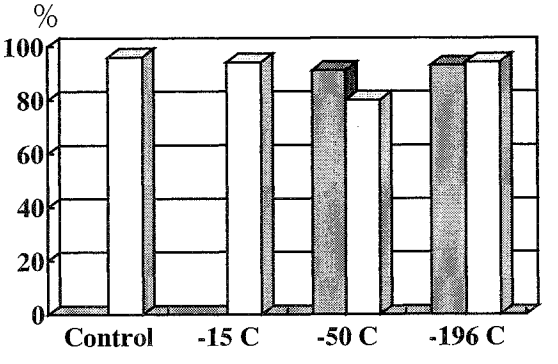


Fig. 2. Effect of storage temperature on embryo development. Spermatozoa were frozeN in CZB medium and stored for 1 month. Sperm were frozen without raffinose (□) or with raffinose (■), and stored at -15 °C, -50 °C or LN_2.

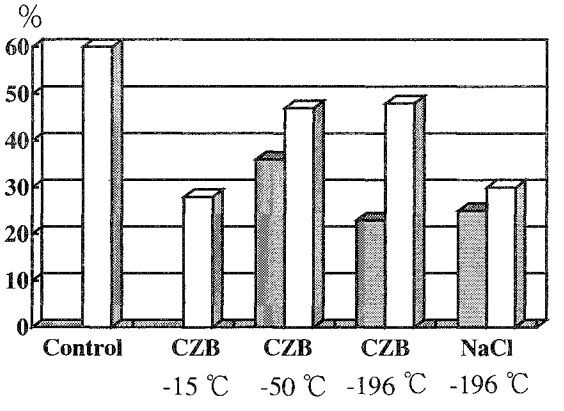


Fig. 3. Effect of sperm freezing media and storage temperature on the development of live young. Spermatozoa were frozen without raffinose (□) or with raffinose (■) and stored for 1 month.

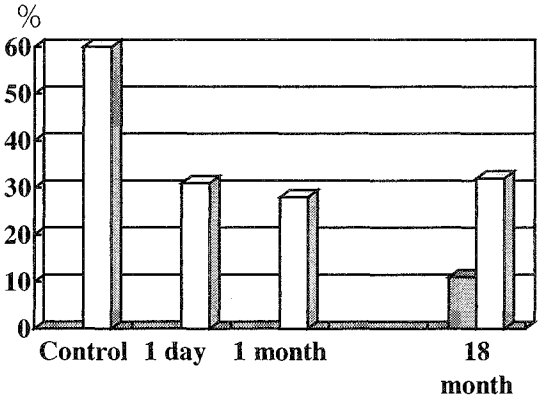


Fig. 4. Effect of storage period at -15 °C on the development of live young. Spermatozoa were frozen in CZB medium and stored at -15 °C up to 18 months. Sperm were frozen without raffinose (□) or with raffinose (■).

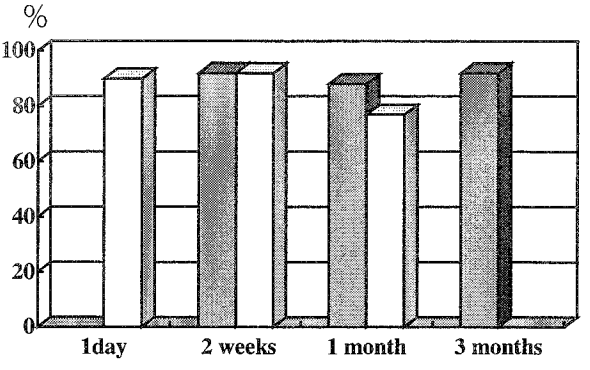


Fig. 5. Effect of the storage period of freeze-dried sperm on embryo development. Spermatozoa were suspended and freeze-dried in CZB medium. Freeze-dried sperm were stored at room temperature (□) or at 4 °C (■).

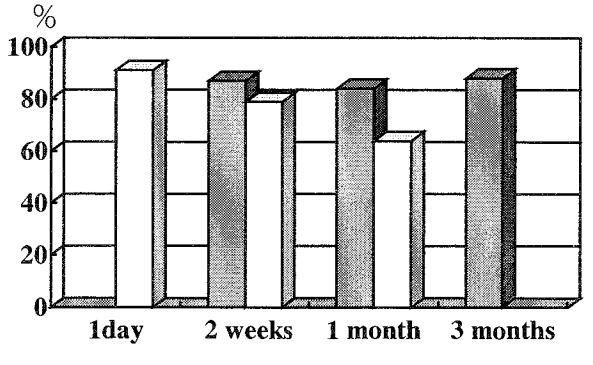


Fig. 6. Effect of storage period of freeze-dried sperm on embryo development. Spermatozoa were suspended and freeze-dried in DMEM medium. Sperm were stored at room temperature (□) or at 4 °C (■).

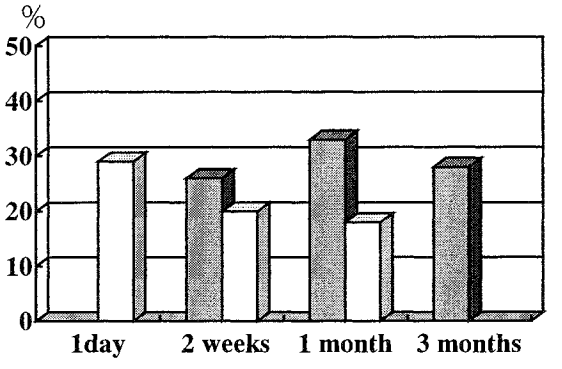


Fig. 7. Effect of storage period of freeze-dried sperm on embryonic development to term. Sperm were suspended and freeze-dried in CZB medium, and stored up to 3 month at either room temperature (□) or 4 °C (■).

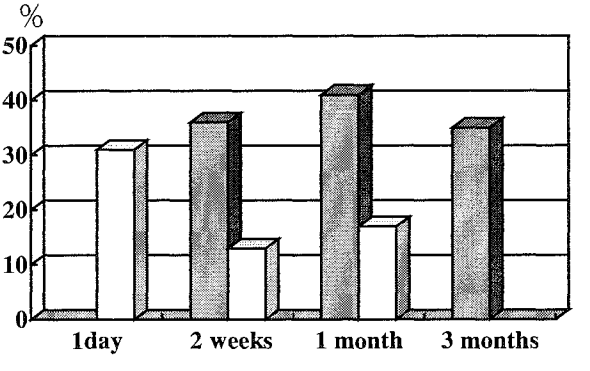


Fig. 8. Effect of storage period of freeze-dried sperm on embryonic development to term. Sperm were suspended and freeze-dried in DMEM medium, and stored up to 3 month at either room temperature (□) or 4 °C (■).

/blastocysts to live young) was better when freeze-dried spermatozoa were stored at 4 °C than at room temperature, regardless of the type of medium used for the suspension of spermatozoa (CZB, Fig.7; DMEM, Fig.8).

Although the techniques require further improvement, the results of above experiments suggest that spermatozoa do not need to be alive, in the conventional sense, to support normal embryonic development when injected into oocytes microsurgically. Conventional sperm preservation is expensive in the long run because of the need for the constant supply of liquid nitrogen. If we can prefect the techniques of sperm freeze-drying and are able to store spermatozoa at ambient temperature or in ordinary refrigerators, the reduction in the cost of the maintenance and shipment costs of spermatozoa would be immense.

References

Chatot, C.L., Lewis, J.L., Torres, I., and Xiomek, C.A. 1990. Development of 1-cell embryos from different strains of mice in CZB medium. Biol. Reprod. 42: 432-440.

Kimura,Y. and R.Yanagimachi, 1995. Intracytoplasmic sperm injection in the mouse. Biol. Reprod. 52: 709-720

Kuretake,S.,Y.Kimura,K.Hoshi, and R.Yanagimachi, 1996. Fertilization and development of mouse oocytes injected with isolated sperm heads. Biol. Reprod. 55:789-793

Nakagata, N. 1996. Use of cryopreservation techniques of embryos and spermatozoa for production of transgenic (Tg) mice and for maintenance of Tg mouse lines. Lab. Anim. Sci. 46: 236-238.

Sharp, J.J. and Monbraaten, L.E. 1997. In: Transgenic Animals: Generation and Use. Harwood Pub. Amsterdam. Pp 525-532.

Songsasen, N., Betteridge, K.J. and Leibo, S.P. 1997. Birth of live mice resulting from oocytes fertilized in vitro with cryopreserved spermatozoa. Biol. Reprod. 56: 143-152

Wakayama,T. and R.Yanagimachi, 1998. Development of normal mice from oocytes injected with freeze-dried spermatozoa. Nat. biotechnol. 16:639-641

Wakayama,T., D.G.Whittingham, and R.Yanagimachi, 1998. Production of normal offspring from mouse oocytes injected with spermatozoa cryopreserved with or without cryoprotection. J. Reprod. Fert. 112: 11-17

Willoughby, C. Mazur, P. Peter, A. and Critser, J.K. 1996. Osmotic tolerance limits and properties of murine spermatozoa. Biol. Reprod. 55: 715-727.

Closing remarks

SHOJI MIYAZAKI
Genetic Resources Coordinator, National Institute of Agrobiological Resources
Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan

I wish we could have more time to continue this interesting discussion. But, since we have already used up time, we would like to close this workshop now. On behalf of the organizing committee of the workshop, I should deliver a few words. I heard that worldwide conservation of native breeds and land-races of animals related to agriculture is at a relatively early stage of development compared with plants and microorganisms. But during this workshop, a large amount of data and impressive results from many years of research have been presented by invited speakers.

I am very conscious that in Asia meat consumption is increasing rapidly as people across the region have greater spending power. Rapid economic development is resulting in a rapid increase in demand for livestock production. The research on the conservation and improvement of domestic animals are really important subjects. It will require international efforts not only to conserve animal genetic resources but also to conserve the rural culture, since local breed has strong relationship to the conservation of rural culture. Sometimes modern farming resulted in the environmental problem. We often hear the words 'Sustainable agriculture'. Local breeds can help maintain sustainable agriculture systems.

The MAFF International Workshop on Genetic Resources aims promoting research exchange and collaboration on the development of technologies and global strategies for conservation and use of genetic resources in national programs and international research institutions. This workshop has been of a small scale, but I anticipate that by continuing these efforts in the future, we will contribute to the promotion of research and collaboration in conserving and evaluating animal genetic resources. As with previous workshop in this series, the proceedings of this workshop, including all papers, will be compiled into a book. This will enable a wider audience to benefit from what has been presented here over the past two days.

I have learned much over the last two days of intensive discussions. I thank everyone for your wholehearted participatión. I thank my colleagues here in Japan for their hard work in preparing papers for this meeting. I would like to

thank all the cooperating institutes in Japan and the Agriculture Forestry and Fisheries Research Council for their support.

To friends who have come from abroad it has been marvelous to have you with us. We look forward to collaboration and interaction in the years to come. We wish you a safe journey home when you leave at the end of the week.

It is my honor to close the 6th MAFF workshop.

Thank you very much.

The 6th MAFF International Workshop on Genetic Resources, November 4－5, 1998, Tsukuba, Japan

Photograph of Workshop Participants

Front row left to right:

Sekikawa K (NIAH), Miyazaki S (NIAR), Kanai Y (Tsukuba Univ.), Matsukawa T (Shirakawa Inst. Animal Genetics), Thuy LT (NIAH, Vietnam), Komiyama T (Nippon Agricultural Research Inst.), Katsura N (NIAR), Tanabe Y (Obirin Univ.), Hicks C (NIAI), Harrison R (Babraham Inst., UK), Tash J (Univ. Kansas Med.Center), Xu S (California Univ.), Mrode R (Animal Data Centre, UK).

Second row:

Kaneko H (NIAR), Kikuchi K (NIAR), Shimizu M (NIAH), Ito S (Gifu Univ.), Muramatsu S (Utsunomiya Univ.), Obata T (NIAR), Nomura T (Kyoto Sangyo Univ.), Takahashi H (Gunma Pig Farms, Inc.), Nakanishi Y (Kanazawa Univ.), Ogawa T (Yokohama City Univ.), Ogura A (NIID), Oono K (NIAR).

Third row:

Nakatsukasa E (Azabu Univ.), Yokoyama M (Mitsubishi Kasei Inst. of Life Sciences), Matsubara Y (NIAI), Takeda K (NIAI), Naito M (NIAI), Nagamine Y (NIAI), Hanada H (NIAI), Oishi T (NIAI), Naito K (Tokyo Univ.), Saito H (AFFRC), Ohkawa M (AFFRC).

Forth row:

Kunihiro Y (NIAR), Shirata K (NIAR), Shimada A (NIAR), Kosugiyama M (Ibaraki Univ.), Dr., Nagai T (TNAES), Ueda H (NIAH), Hammond K (FAO), Roberts M (Missouri Univ.), Noguchi J (NIAR).

Fifth row:

Takahashi R (YS New Technology Inst.), Kashiwazaki N (Azabu Univ.), Izaike Y (NIAI), Gotoh H (NIAH), Nambo Y (JRA Equine Research Inst.), Hasegawa T (JRA Equine Research Inst.), Takeda H (NIAR), Yamauchi N (TNAES), Kato K (NIAR), Nirasawa K (NLBC), Kojima T (NLBC), Onishi A (NIAI).

February, 2000

Published by

Research Council Secretariat of MAFF

National Institute of Agrobiological Resources

Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan

ISBN4-931511-02-3