J. Dairy Sci. 84:256–265

Evaluation of Sire Predicted Transmitting Abilities for Evidence of X-Chromosomal Inheritance in North American Sire Families

P. J. Boettcher,*[,1] L. K. Jairath,* and P. M. VanRaden†
*Centre for Genetic Improvement of Livestock,
Department of Animal and Poultry Science,
University of Guelph, Guelph, Ontario, Canada N1G 2W1
†Animal Improvement Programs Laboratory,
Agricultural Research Service, USDA, Beltsville, MD 20705

ABSTRACT

This study tested for differences between genetic merits of sons and daughters of sires and for evidence of segregating quantitative trait loci on the X chromosomes of North American Holsteins. Son PTA adjusted for sire PTA was used as the dependent variable to test for biases and for genes that were passed from sire to daughter but not to son. The test of variability across sires of sons merely indicated an unaccounted source of variation, for which genes on X chromosomes might be responsible. Critical values for this test and power were determined by simulation for a variety of populations and traits differing in heritability, size of the X chromosome effect, and allelic frequency. Simulated genes on the X chromosome were detected with high power at intermediate frequencies of the favorable allele. The power of the test increased as the size of the effect increased and as genetic variance attributed to autosomes decreased. The test was then applied to recently evaluated data from US and Canadian Holstein populations. Genetic evaluations for >17,000 bulls from the US and >9000 from Canada were included. Results suggested that little extra variation was present for some traits formally evaluated in North America, but that genes on the X chromosome were unlikely to be the cause.
(**Key words:** x-linked inheritance, predicted transmitting abilities)

Abbreviation key: DGD = daughter-granddaughter design, **MGS** = maternal grandsire, **PA** = parent average EBV or PTA, **PL** = productive life, **XEFF** = effects of segregating genes on the X chromosome.

INTRODUCTION

Studies (Ashwell et al., 1997; Georges et al., 1995; Spelman et al., 1996) involving the search for specific QTL affecting economically important dairy traits have usually targeted genes on the autosomes rather than the sex chromosomes. This concentration on the autosomes is reasonable, given that they outnumber sex chromosomes by 29:1. Furthermore, the widespread use of AI has created large families of paternal half-sibs, within which autosomal variation can be detected by applying the daughter and granddaughter designs of Weller et al. (1990). By contrast, no variation (other than very recent mutation) exists among paternal half-sibs of the same sex for much of the paternally inherited X chromosome. The sex chromosomes merit some attention because they could include segregating QTL that affect production traits (**XEFF**). Several genes have been mapped to the bovine X chromosome (Sonstegard et al., 1997). Certain diseases in humans, including colorblindness and hemophilia, are known to be sex-linked (McKusick, 1966). Hagger and Stranzinger (1992) and Harris et al. (1984) observed sex-linked effects on body weight and some livability traits of poultry, respectively. If identified, information about XEFF could be used to help select bulls within full-sib families for entry into progeny testing (Cowan et al., 1997). Also, the use of bulls with a favorable XEFF as sires of sons could be restricted, because these effects would not be passed to their sons.

The daughter and granddaughter designs used to find evidence of QTL on autosomes could in theory be modified to detect XEFF. To compare two alleles on the X chromosome, large maternal families, created by embryo transfer, would be needed rather than large paternal families. In reality, the use of genetic markers would be less effective than our approach for three reasons: 1) developmental costs would be required in identifying a set of useful markers and obtaining DNA samples, 2) analytical costs for genotyping one to several thousand animals, and 3) families of the size necessary to obtain

Received February 25, 2000.
Accepted September 6, 2000.
[1]Corresponding author: P. J. Boettcher; e-mail: paulboettcher@anafi.it. Current address: ANAFI, Via Bergamo 292 26100 Cremona, Italy.

equivalent statistical power simply do not exist in the current populations.

The traditional granddaughter design requires large paternal half-sib families. Weller et al. (1990) reported estimates of powers for the granddaughter designs with various family sizes. The smallest design for which they reported estimates included 5 families, each with 40 sons with 10 daughters each. Based on the fall 1998 genetic evaluations, no cow in the US or Canada has had 40 sons with at least 10 daughters. Only one cow has had >30 (31 sons). Nine cows have had more than 20 sons, and all of these cows have been dead for several years, precluding acquisition of new samples of DNA. Furthermore, the expected powers of the tests using the (unattainable) level of 5 dams with 40 sons were very low (0.01 to 0.38) unless QTL effects were very large (Weller et al., 1990). Finally, as low as these estimates of power were, they were the maximum powers for designs of the various sizes, attainable only when the frequency of heterzyogotes for the QTL was 0.50, which will rarely be true for a biallelic QTL.

Daughter designs are also not well suited to detecting XEFF. The smallest daughter design evaluated by Weller et al. (1990) required 5 families of 200 daughters each. For Canada, no cow has had even half that many daughters, and only one cow had more than one-third of the required 200 daughters. Also, the daughters in these large families might have received preferential treatment (Kuhn and Freeman, 1995), which could potentially bias the results and decrease the power of any test.

Fortunately, approaches exist to test for evidence of X chromosomal inheritance that do not require genotyping (Ali et al., 1992; VanRaden, 1988). Because no DNA samples are required, much more information can be included in the test, which increases the power to detect XEFF. These procedures take advantage of the fact that for the sex chromosomes a large portion of the X chromosome does not recombine with the Y during meiosis. Therefore, sex can be used as the marker to identify which animals received the genes on the X chromosome from a given sire. Thus, an approach that combines daughter and granddaughter design (**DGD**) can be used. The performance of a sire's daughters that received the X chromosome of the sire can be contrasted to the performance of the sons (or its paternal granddaughters for sex-limited dairy traits), which did not receive the X chromosome, to determine if variability exists across the population among genes on the X chromosome that affect traits for which genetic evaluations exist.

To perform his test, VanRaden (1988) assumed that a bull PTA [actual predicted differences from the now obsolete Modified Contemporary Comparison (Norman, 1976)] represented the average genetic contribution of the sire to its daughters. So, for each sire of sons, VanRaden (1988) created contrasts to account for the difference in X inheritance by subtracting one-half its own PTA (and one-fourth of the maternal grandsire (**MGS**) PTA to account for merit of mates) from the PTA of each of its sons. He then analyzed these contrasted PTA of the sons for variability across sires of sons. On average, the only difference between the genetic contributions of a sire to his sons versus his daughters is in the sex chromosomes, so any variability in these contrasted son PTA across sires would indicate the presence of XEFF (or some other unknown, systematic source of bias). For example, if the haplotype on the X chromosome for a given sire was relatively superior to the average X haplotype of the population, then its daughters would be expected to perform at a higher level than the daughters of the sire's sons, resulting in a negative contrast. The opposite would be true if the X haplotype of the sire was relatively inferior to that of the population. Other sources of bias, such as preferential treatment, could also inflate the sire variance of the contrasts, even if no XEFF were present, but additional simple tests can be used to distinguish such effects.

One disadvantage of this DGD approach, relative to approaches based on genotyping, is that it provides no information about the location of any potential QTL on the X chromosome. However, because of its relatively low cost and ability to include all sire families, a logical and cost-efficient strategy is to first apply the DGD to determine whether XEFF are likely to exist and which sires of sons (those with particularly large sire contrasts) carry the extreme alleles. Then the more expensive, genotyping-based strategies, such as interval mapping (Haley and Knott, 1994), could be used within these families to help determine the location of the QTL. Without first applying the DGD, a population-based QTL search may be expensive and risky. With evidence of XEFF, breeding companies might then be able to economically justify a designed study involving genotyping, even if such a study required the creation of large maternal families.

By applying this DGD approach to US sires, VanRaden (1988) obtained results suggesting that from 5 to 10% of the genetic variation in milk production could be due to XEFF. By applying a slightly different procedure to data from Canadian sires, Ali et al. (1992) arrived at a conflicting conclusion. They found no evidence of XEFF when they analyzed data from Canada. Thus, the primary objective of this study was to reapply the DGD to data from the US and Canada to reassess whether significant XEFF are likely to exist. If so, future efforts in marker-assisted selection based on the

X chromosome can either concentrate on traits with the strongest evidence of XEFF or continue to emphasize the autosomes if little evidence for major XEFF is observed.

One uncertainty about the DGD is that its power to detect QTL is unknown and will likely vary with the size of the XEFF, frequency of alleles, and the polygenic variance for the trait of interest. Thus, a secondary objective of this study was to initially use simulation to evaluate the power of the DGD and its relationship with changes in various population parameters.

MATERIAL AND METHODS

Simulated Data

General population structure. A population was simulated that was similar in selection structure to, but smaller in size than, US and Canadian Holstein populations. The base population consisted of 125,000 cows and 1000 sires. Cows were equally dispersed across 500 herds and could remain in the herd for up to five lactations. Phenotypic records of cows were generated according to the following model:

$$Y = CG + X + A + PE + e, \quad [1]$$

where Y is the phenotypic record, CG is the effect of contemporary group (distributed normally), X is the XEFF, A is the polygenic effect; PE is the permanent environmental effect; and e is the residual effect.

Two models were considered for X. The first was a biallelic model; the second assumed that XEFF were normally distributed. For the biallelic model, sires in the base generation were randomly assigned one of the two alleles, based on an initial allelic frequency, which was systematically varied across replicates. Females were randomly assigned two alleles. With the normal model, bulls in the base generation were randomly assigned one XEFF, and cows were assigned two XEFF. The size or standard deviation of the XEFF was varied across groups of replicates. In future generations, sires passed their single XEFF to their daughters and not to their sons. Dams randomly passed one of their two XEFF to their sons and daughters.

Inactivation of the X chromosome (Lyon, 1961) was random with respect to parental origin (Takagi, 1978), and we assumed that cows expressed alleles on paternally and maternally derived X chromosomes at a ratio of 50:50 (De La Fuente et al., 1999). With this logic, $X = (x_s + x_d)/2$ where x_s and x_d are the individual XEFF received from the sire and dam, respectively.

For base generation animals, the polygenic effect A was distributed normally with mean = 0, and variance = 5 or 35% of the total variance of A + PE + e ($= \sigma^2_T = 1000^2$). For nonbase animals, A was the average of parental A effects, plus a random Mendelian sampling effect, $\sim N(0, \frac{1}{2}\sigma^2_A)$. Mates were chosen to avoid inbreeding. The PE and e were drawn from normal distributions with mean = 0. The variance of PE was equal to 15% of σ^2_T. The variance of e was equal to 50% of σ^2_T when σ^2_A was 35 and was 80% of σ^2_T when σ^2_A was 5%.

Cows were culled at random, except that probability of culling on an individual basis increased with age. Selection of bulls and bull dams was based on EBV. The EBV for bulls were calculated based on the EBV of their respective sires and dams and the average production of their daughters. The weights applied to each source of information were derived from selection index methodology (Van Vleck 1993). For cows, EBV were based on the EBV of their respective parents and their own performances. The primary feature of the selection structure was a two-stage approach to sire selection in which young sires were progeny tested with a limited number (~80) of daughters, prior to being activated for widespread use as a sire of cows in the population. In any given year, the bull population comprised 300 progeny test bulls and 100 active sires, of which 15 were eligible to be used as sires of sons. This model resulted in a greater proportion of daughters from progeny test bulls than in the current North American population, but this fact was not expected to strongly affect application of results to real data. The primary result was a decreased number of daughters per active sire compared with real data, but little difference was found in accuracy of EBV. Bull dams were simply the 300 females with the greatest EBV in a given year.

Each replicate of the simulation generated 20 yr of data (cycles of lactations). Across replicates, the autosomal polygenic variance and the magnitude of XEFF were systemically varied to investigate how changes influenced the power of tests for XEFF. For the biallelic model, XEFF were varied according to two factors: 1) the size of the effect of a gene substitution of X alleles, and 2) the relative frequency of the superior X allele. Two magnitudes of the XEFF (the effect of a gene substitution) were considered, large = $0.3\sigma_T$ and small = $0.1\sigma_T$. These QTL effects were the highest and lowest effects considered by Weller et al. (1990). A wide range of initial gene frequencies for the superior X allele were simulated, including 0.0001, 0.05, 0.20, 0.50, 0.80, and 0.99. Changes in allelic frequency across years were monitored. Twenty replicates were generated for each combination of polygenic variance, XEFF, and initial gene frequency.

Testing for XEFF. To perform the DGD on the simulated data, EBV for sires and their parents were required. The EBV were BLUP and were calculated by

using the MAGGIC software of Janss (1998). Each replicate included approximately 2.2 million records from slightly more than 800,000 cows. The model used for analysis of the data included the same factors as equation [1] used to generate the data, except that inheritance was assumed to be strictly additive, so XEFF were not considered.

Contrasts were formed by subtracting one-half the EBV of the sire and one-half the EBV the dam (to account for merit of mates) from the EBV of each bull. Bulls were required to have paternal brothers and an identified MGS. Approximately 3200 bulls met these criteria in each replication. Variability in these contrasts across sires were analyzed with following model:

$$Y = \mu + SIRE + MGS + e, \quad [2]$$

where Y is the contrast (EBV of the bull − mean of parent EBV); μ is the overall mean; SIRE and MGS are the random effects of sire and MGS, respectively, that remained after adjustment for parent average (**PA**); and e is random residual. The effects of SIRE, MGS, and e were assumed to be distributed normally. The model is similar to the one used by VanRaden (1988), except for exclusion of effects for birth year and AI company. The MGS term was included in the model to account for an additional source of variance. The variances for the SIRE (σ^2_S) effects were estimated using the REML VCE 4.0 software of Groeneveld and Garcia-Cortez (1998).

Establishment of critical values for significance tests. Simulation was used to establish empirically the approximate critical values for significance tests. Two hundred populations were simulated with no XEFF. Then the 20th greatest values for σ^2_S were used as critical values for $\alpha = 0.10$ (probability of Type 1 error), and the 10th and 2nd greatest values were used for $\alpha = 0.05$ and 0.01, respectively. Ideally, more replications would have provided more precise estimates for these critical values, but this possibility was limited by computing resources. Approximate values were considered to be suitable for the purposes of this study. Two sets of 200 populations were simulated, one for each level of polygenic variance.

Real Data

The DGD was applied to actual sire PTA and EBV from the US and Canadian national genetic evaluations. Sire PTA (US) and EBV (Canada) for milk, fat, and protein yields were available from the US and Canada. In addition, PTA for productive life (**PL**) (VanRaden and Wiggans, 1995) and somatic cell score (Schutz, 1994) were available for the US. Although sire ETA for these traits are also calculated in Canada, they could not be used in the analysis because ETA for females are not calculated and, therefore, PA were not available to calculate the contrast.

The PTA for production traits of US sires were from the fall 1998 genetic evaluation. The initial data were obtained from the USDA Animal Improvement Programs Laboratory (Beltsville, MD) and contained PTA from 46,532 bulls. Data included only Holstein bulls registered in the US or Canada, born between 1980 and 1992, with <20 daughters, and from sires with at least three sons. Data were for 17,501 bulls from 308 sires.

The data for Canadian production traits were 10,218 sire EBV from the November 1998 evaluation. The EBV were from the Canadian Dairy Network (Guelph, Canada). Edits were consistent with those applied to the US data. The data included 9593 bulls from 290 sires.

Sire ETA for 29 Canadian type traits were also available. The 29 traits were composite traits for overall conformation, frame and capacity, rump, feet and legs, fore udder, rear udder, mammary system, and dairy character and individual measures of stature, front end, size, chest, body depth, loin, rump angle, pin width, foot angle, bone quality, rear legs from the side view, udder depth, udder texture, median suspensory ligament, fore udder attachment, front teat placement, front teat length, height of rear udder attachment, width of rear udder attachment, and desirability of pin setting and rear legs.

Type data were from the January 1999 monthly genetic evaluation. Data initially included EBV from 8577 bulls. After edits based on requirements similar to those for production traits, records of 7351 bulls from 213 sires remained.

These three data sets were of different sizes and all included many more bulls than did the sets of simulated data. Therefore, the empirical critical values that were established by using the simulated data were not expected to be appropriate. The size of these data and computing restrictions precluded the use of simulation to determine significance levels and powers for the real data. Therefore, appropriate tests were created by dividing randomly (based on the final digit of bull registration number) these large sets of data into smaller sets, each with approximately the same number of bulls and sires as in the simulated data sets. Then results from each subset within country and trait were combined to yield a single overall test. The US production, Canadian production, and Canadian type data were divided into five, three, and two smaller sets, respectively. These subsets each contained EBV for approximately 3100 to 3600 different bulls from 250 sires.

The σ^2_S was estimated with REML and the following model:

$$Y = YEAR + STUD + SIRE + MGS + e, \quad [3]$$

which was the same model [2] as used for simulated data, except that fixed effects of birth year of the bull (YEAR) and the breeding company that owned the bull (STUD) were also included in the model. VanRaden (1988) reported that these factors accounted for some of the variability in Y.

The estimates of σ^2_S from each subset of data were compared with the critical values established by the simulation. Then, for each source of data, results from all subsets were combined into an overall test by assuming that the number of significant tests was distributed binomially (n, α), where n is the number of subsets and α is probability of type I error. Tests were based on the number of individual tests with $P < 0.10$. Although this level of significance may seem very liberal for each subset, the test became more stringent when results across subsets were combined. For example, with three subsets of data, the probability of two or more significant ($P < 0.10$) estimates of σ^2_S was 0.028. Also, liberal levels of significance were appropriate for the primary purpose of this experiment, which was to identify traits for which XEFF may be important and eliminate the inefficiency of potential future application of more expensive studies to traits for which XEFF are not likely to exist.

As mentioned earlier, significant variance among sires in the adjusted PTA of their sons may not necessarily indicate the presence of XEFF. Other sources of bias in PTA could cause an increase in σ^2_S. To help determine whether significant estimates of σ^2_S were due to XEFF or some other source, we used an additional test. Correlation between sire and MGS effects should be zero or slightly negative in the presence of XEFF. VanRaden (1988) stated that negative correlations between solutions for sire and MGS that he observed favored the possibility that significant estimates of σ^2_S were an indication of XEFF. A negative correlation would result from sires that carry a favorable XEFF, transmitting the effect to their maternal grandsons but not to their sons. Other sources of variance, such as preferential treatment of the daughters of a sire of sons would result in positive correlations. Thus, correlation coefficients between sire and MGS solutions were calculated for every trait in the three different genetic evaluations and were tested to see if they were significantly greater than zero. For these calculations, all data were used rather than the subsets.

RESULTS AND DISCUSSION

Simulation

Table 1 has the estimated critical levels for *P*-values of 0.01, 0.05, and 0.10 for σ^2_S (expressed as a percentage of total variance) when true proportions of polygenic variance were 0.05 and 0.35 and no XEFF were simulated. Estimates of σ^2_S were more variable when the proportion of polygenic variance was $0.05\sigma^2_T$ versus $0.35\sigma^2_T$. As a result, the critical values for significance tests were relatively greater at lower heritability.

Table 1. Estimates[1] of critical values, for a range of *P*-values, of sire variance[2] in the contrasts of son EBV and parent average EBV when no X-chromosomal effects were simulated and true proportions of polygenic variance were 0.05 or 0.55.

True polygenic variance	*P*-value		
	0.10	0.05	0.01
	(%)		
0.05	1.7	2.2	2.7
0.35	1.3	1.7	2.3

[1]Based on 200 replicates of simulation.

[2]Expressed as percentages of total variance.

Means across 20 replicates of σ^2_S for a range of initial frequencies of the favorable X haplotypes, small and large XEFF, and low and high levels of true polygenic variance are in Table 2. Corresponding powers to detect XEFF are also given.

Several distinct trends are apparent in the results. First, estimates of σ^2_S increased when the XEFF was increased and its ratio to polygenic variance increased. Second, σ^2_S (and power) were greatest at intermediate (between 0.20 and 0.50) initial frequencies for the favorable X haplotype. These results indicated that, when holding all other factors constant, σ^2_S increased as the variance of XEFF in the population increased. Specifically, σ^2_S increased as the expected X genotype of the daughters of sons became more different from the expected X genotype of daughters of the sire. Daughters of a sire always received his chromosome and another X from their dam. Sons of a sire (and the daughters of these sons) received an X chromosome from their dams only. When the frequency of one haplotype was high, the sons (and their daughters) often had XEFF that were identical to those of their female half-sibs simply by chance, rather than by descent from the sire.

Power was usually greatest when initial frequency was <0.50, because maximum variability in XEFF occurred when the average (rather than initial) allelic frequency was 0.50, and frequency of the favorable allele increased over time with selection. The maximum σ^2_S was not always observed at the same initial frequency for all scenarios, which was presumably due to differences in the dynamics of selection across scenarios. Finally, when holding all other variables constant, σ^2_S were greater when the variability of polygenic effects were smaller. Seemingly, at high levels of polygenic variance, contributions of XEFF to bull EBV were

Table 2. Mean[1] estimates of the variances[2] of sire effects (σ^2_S) for son EBV minus the mean of parent EBV and corresponding powers to detect effects when a small[3] or large biallelic effect X-chromosomal effect was simulated for a range of initial gene frequencies of the favorable allele and low[4] and high levels of true polygenic variance.

		Size of X-chromosomal effect					
		Small			Large		
True polygenic Variance	Initial allelic frequency		Power			Power	
		σ^2_S	$\alpha = 0.10$	$\alpha = 0.05$	σ^2_S	$\alpha = 0.10$	$\alpha = 0.05$
		(%)					
Low	0.0001	1.06	20	10	6.04	70	55
	0.05	6.27	100	100	9.52	100	95
	0.20	4.82	95	90	11.13	100	100
	0.50	2.14	75	40	17.74	100	100
	0.80	1.38	35	20	5.70	100	95
	0.99	0.82	10	10	1.30	25	10
High	0.0001	0.40	0	0	1.62	25	25
	0.05	2.50	95	85	6.22	100	100
	0.20	4.64	100	90	4.84	100	100
	0.50	1.89	80	55	5.76	100	95
	0.80	0.82	25	15	1.57	50	45
	0.99	0.28	0	0	0.43	5	0

[1]Based on 20 replicates.

[2]Expressed as percentage of total variance in bull EBV adjusted for parent average.

[3]Small and large were 0.1σ and 0.3σ, respectively, where σ is the standard deviation of polygenic, permanent environmental, and residual effects.

[4]Low and high were $0.05\sigma^2$ and $0.35\sigma^2$, respectively, where σ^2 is the total variance of polygenic, permanent environmental, and residual effects.

comparatively masked by differences among bulls in their polygenic effects.

Table 2 also has corresponding estimates of the powers of tests (assuming *P*-values of 0.10 and 0.05) for XEFF resulting in significantly increased σ^2_S. As expected, the powers follow the same trends shown by the means of σ^2_S. For most combinations of the size of XEFF and autosomal polygenic variance, a test based on σ^2_S was quite powerful for detecting when the initial gene frequency of the favorable X haplotype was between 0.05 and 0.50. This result was particularly true for the large XEFF, for which the power within this range of initial frequency was always ≥0.95. Even when the size of XEFF was small, the power was high for frequencies of 0.05 and 0.20 but dropped to around 0.50 for initial gene frequencies of 0.50. For reasons discussed previously, XEFF were essentially impossible to detect based on estimates at extreme initial frequencies.

Table 3 shows means of σ^2_S and associated powers of the test when XEFF were distributed normally with variances equal to 5, 10, and 20% of polygenic variance. As expected, mean σ^2_S and power increased as the variability of XEFF increased. For a trait with high polygenic variance, XEFF were detected with reasonable power (≥0.70), even when variabilities of XEFF were only one-twentieth as large as polygenic variance. Means and powers were much less when true polygenic variance was low, but this result was due to the fact that variances for XEFF, being expressed as a ratio to polygenic effects, were relatively much lower in magnitude than when polygenic variance was high.

In general, relationships between power for detection of XEFF and the size of gene substitution effect and amount of polygenic variance were similar to the rela-

Table 3. Mean[1] estimates of the variances of sire EBV adjusted for parent average EBV and associated power for testing for evidence of bias when effects of genes on the X chromosome (XEFF) were distributed normally with variances equal to 5, 10, and 20% of polygenic variance and when polygenic variance was low or high.

			Power	
Polygenic variance	XEFF[2]	Mean[3]	$\alpha = 0.10$	$\alpha = 0.05$
	(%)			
Low	5	0.5	10	0
	10	0.7	20	10
	20	2.8	80	70
High	5	2.2	90	70
	10	4.0	95	95
	20	6.7	100	100

[1]Based on 20 replicates.

[2]Expressed as percentage of variance due to polygenic effects.

[3]Expressed as percentage of total variance in bull EBV adjusted for parent average.

Table 4. Estimates from five subsets of data from the US and three from Canada of the variances[1] of sire effects on bull PTA for production traits adjusted for parent average and *P*-values for the combined estimates based on a binomial[2] distribution.

	Trait				
Subset	Milk	Fat	Protein	Productive life	SCS
US					
1	1.2	0.9	1.1	3.0**	0.8
2	1.3†	0.5	1.0	2.5*	0.1
3	1.0	1.6†	0.7	2.6*	0.9
4	1.7*	2.3**	1.8*	1.7†	0.2
5	1.4†	0.9	1.2	1.8†	0.0
P-value	0.01	0.08	0.41	<0.0001	1.00
Canada					
1	1.4†	2.3*	2.0*	. . .	. . .
2	0.4	0.8	0.4	. . .	. . .
3	1.2	1.0	0.4	. . .	. . .
P-value	0.27	0.27	0.27	. . .	. . .

[1]Expressed in percentage of total variance in bull EBV adjusted for parent average.

[2]Binomial (5, 0.10).

[3]Significance tests for milk, fat, protein assumed high polygenic variances; and SCS and productive life assumed low polygenic variances.

†$P \leq 0.10$.

*$P \leq 0.05$.

**$P \leq 0.01$.

tionships reported by Weller et al. (1990) when trying to detect autosomal QTL using daughter and granddaughter designs.

Real Data

Table 4 has estimates of σ^2_S for the production traits from US and Canada. Estimates of σ^2_S are given for each subset of data along with an indication of their respective test-wise levels of significance. Because multiple tests for each trait were performed, for the US data a combined *P*-value based on a cumulative binomial (5,0.10) distribution is given for each trait. Productive life was the only trait for which significantly ($P < 0.10$) high estimates of σ^2_S (possibly X-chromosomal inheritance) were consistently observed. For all five subsets of data, the estimate of σ^2_S was significant at the $P \leq 0.10$ level, accounting for as much as 3.0% of the variability in the contrasts of bull PTA adjusted for PA. Assuming no sources of bias in the evaluation, the probability of all five tests being significant was <0.0001. Some evidence of bias in PTA for milk was also observed, as estimates of σ^2_S were significantly ($P \leq 0.10$) large for three of the five subsets. This result corresponded to an overall *P*-value of slightly >0.01. For the other three traits, no more than two estimates of σ^2_S were significant ($P \leq 0.10$). For all traits except PL, estimates of σ^2_S were <2.0%, much lower than most of the values in Tables 2 and 3. In particular, estimates of σ^2_S were very low for SCS, <1.0% for all five subsets.

Table 4 also shows the corresponding results for production traits in Canada. In contrast with the results from the US, very little evidence of bias in bull EBV for milk was observed, with only a single significant ($P \leq 0.10$) estimate of σ^2_S observed among the three subsets of data. Estimates of σ^2_S from the same subset were also significant, and at a higher level ($P \leq 0.05$), for fat and protein, but estimates for these traits from the other subsets were very low at ≤1.0%.

Table 5 shows estimates of σ^2_S for the six Canadian type traits for which estimates from both subsets were

Table 5. Estimates from two subsets of data from Canada of the variances[1] of sire effects on bull PTA adjusted for parent average for the type traits for which estimates from both subsets were statistically significant ($P \leq 0.10$).[2]

	Subset of data	
Trait	1	2
Overall conformation	2.0*	3.3**
Frame and capacity	1.3†	2.8**
Rear udder	2.2*	1.6†
Mammary system	1.9*	2.0*
Size	1.4†	1.6†
Height of rear udder attachment	1.6†	1.3†

[1]Expressed as percentage of total variance in bull EBV adjusted for parent average.

[2]Based on critical values for trait with relatively high polygenic variance.

†$P \leq 0.10$.

*$P \leq 0.05$.

**$P \leq 0.01$.

significantly ($P \leq 0.10$) large. Such a result had a binomial probability of 0.01 (0.10^2). Indications of bias were particularly evident for overall conformation, as the *P*-values for both subsets were <0.05. Assuming a binomial (2,0.05) distribution, the probability of observing this result if no bias was present was <0.003. The other five traits have particularly strong relationships with overall conformation. According to an unpublished technical report to the Canadian Genetic Evaluation Board (Z. Liu, L. Jairath, and J. Dekkers, 1994), genetic correlations of overall conformation with frame and capacity and size were approximately 0.60 and were approximately 0.80 with rear udder, mammary system, and height of rear udder attachment.

Because of the power of these tests, nonsignificant tests for σ^2_S are compelling evidence against major XEFF for many of the dairy traits, at least for XEFF segregating at intermediate allelic frequencies, and are large enough to justify the cost of additional studies needed to characterize such effects and use them in breeding programs. For example, the binomial distribution can be used to estimate the probability that these results could have been observed had major XEFF existed. Protein yield is currently the most important dairy trait based on selection indexes used in Canada (Boettcher and VanDoormaal, 1999). Polygenic effects account for approximately 35% of the variability in protein yield. Considering this fact and the results from Table 3 and assuming that an XEFF existed with a variance equal to 5% of the polygenic variance, the probability of obtaining no more than one significant ($P < 0.10$) estimate of σ^2_S when analyzing the three subsets of data from Canada was only 0.028. Under the same assumptions, the probability of observing no more than one significant estimate from the five subsets of US data was only 0.00028.

Although nonsignificant estimates of σ^2_S are strong evidence against major XEFF, as previously mentioned, statistically significant estimates of σ^2_S were not necessarily indicative of XEFF. Rather, they simply indicated that the PTA or EBV of some sires are not consistent with the average EBV of their sons. In other words, some unaccounted source of bias existed in the genetic evaluations.

Table 6 has the correlation coefficients between sire and MGS solutions for all of the production traits from the US and Canada and the Canadian type traits in Table 5. Among the traits listed, only for SCS and height of rear udder attachment were the correlations between sire and MGS solutions not significantly >0 ($P \leq 0.03$). For SCS, estimates of σ^2_S were very low. Thus, among all the traits studied, only height of rear udder attachment showed results consistent with the presence of XEFF. This trait is relatively unimportant economically.

For confirmation, the correlation between sire and MGS solutions when XEFF were present were estimated from 40 simulated populations with different combinations of high and low sizes of XEFF and high and low polygenic variances. The average correlation between sire and MGS solutions from these populations was –0.01. For only one replicate was a correlation coefficient observed that was significantly >0 ($P < 0.05$). Thus, the significant estimates of σ^2_S that were observed in the real data were most likely due to effects other than XEFF.

For most traits that showed an increased σ^2_S, preferential treatment of the daughters of sires of sons is one plausible cause for this bias. Kuhn and Freeman (1995) demonstrated that preferential treatment of some or all daughters could result in biased sire PTA. Daughters of bulls with high-priced semen may receive better care than their herd mates as the breeder tries to protect his or her investment in semen. The preferential treatment and semen price issue might be particularly important for PL, because producers may be willing to give offspring of expensive sires the best health care and may relax culling criteria. Such effects would both increase σ^2_S and the correlation between sire and MGS solutions.

Preferential treatment of bull dams might have also played a role in the increased σ^2_S. Van Vleck (1986) reported that genetic evaluations for bull dams did not predict the ETA of their sons as accurately as theory dictated. For most traits in our study, the contrast of [bull PTA – PA] was negative; indicating that PA tended

Table 6. Correlation coefficients and significance levels of solutions for sire and maternal grandsire effects for all US and Canadian production traits and for six Canadian type traits for which the greatest amount of bias was observed.

Trait	Number of bulls	Correlation	*P*-value
US production	180		
Milk		0.38	<0.0001
Fat		0.43	<0.0001
Protein		0.42	<0.0001
SCS		–0.03	0.66
PL[1]		0.43	<0.0001
Canada production	179		
Milk		0.38	<0.0001
Fat		0.49	<0.0001
Protein		0.36	<0.0001
Canada type			
Overall conformation	99	0.30	<0.0001
Frame and capacity		0.27	0.0002
Rear udder		0.18	0.02
Mammary system		0.17	0.03
Size		0.25	0.002
Height of rear udder attachment		–0.01	0.90

[1]Productive life.

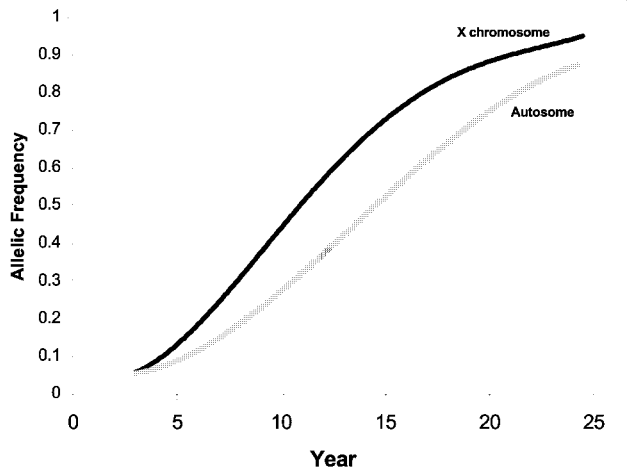

Figure 1. Trends in allelic frequency for favorable X-chromosomal and autosomal effects of size $0.3\sigma_P$ under conventional two-stage selection for a trait when polygenic effects account for 35% of the total variance.

to over predict bull ETA on average. However, effects on σ^2_S of this type of preferential treatment are likely to be small, because on average sires had many sons ($\mu = 56$). Thus, most sires of sons would have been mated to some dams with highly inflated PTA and others with only slightly inflated or unbiased PTA. On average, these effects would have balanced out for most sires and contributed little to σ^2_S, unless some sires of sons were systematically bred to dams with particularly inflated (or unbiased) PTA, which seems unlikely.

Regardless of the reason for the observed biases in bull PTA, the level of bias for most traits was <2% of the variance in adjusted PTA, which was probably too small to be of major concern for sire selection.

VanRaden (1988) reported much higher estimates of σ^2_S in his original study. His study was performed more than 10 yr ago, and his positive results might have been due to the presence of a favorable allele that has since then approached fixation. Figure 1 shows how the frequency of a favorable allele (size = $0.3\sigma_P$) on the X chromosome or on an autosome increases with conventional two-stage selection, according to our simulation for a trait with a polygenic variance of 35%. The favorable allele increases in frequency quickly and remains for only a short span of years at a frequency that would allow detection with a reasonable power. Alternatively, VanRaden's (1988) high estimates might also have occurred because his study used pedigree indexes from the Modified Contemporary Comparison to adjust the ETA of bulls rather than PA from the animal model and, therefore, did not fully account for the contributions of the bull dams. When our analyses of real data were repeated using pedigree index (rather than PA) to adjust bull PTA, estimates of σ^2_S were much higher, between 5.5 and 7.1%, for the production traits.

CONCLUSIONS

Few QTL on the X chromosome with large effects are segregating across many families and at intermediate allelic frequencies for most of the dairy production traits that are routinely evaluated in the US and Canadian populations. These results do not suggest that no genes on the X chromosomes encode for useful proteins, only that variation in such genes is not great enough to be easily detected and exploited in current sire selection programs. Also, XEFF may affect traits that are not currently evaluated and, therefore, were not considered in this study. Investment in marker-assisted selection for genetic variation on the X chromosome is likely not profitable as a breeding tool. Research about genes on the X chromosome should focus on their biological structure and function, rather than their use for marker-assisted selection. Agreement between son and daughter evaluations indicates that an autosomal, additive genetic model describes biology well.

ACKNOWLEDGMENTS

The authors thank the Animal Improvement Programs Laboratory of the USDA and the Canadian Dairy Network for making available the results from genetic evaluations. Financial support for this research was provided by the Cattle Breeding Research Council of Canada (Guelph, Canada).

REFERENCES

Ali, T. E., L. R. Schaeffer, J. P. Gibson, and E. B. Burnside. 1992. Usefulness of pedigree indexes in predicting future transmitting abilities of sons. Can. J. Anim. Sci. 72:409–412.

Ashwell, M. S., Y. Da, P. M. VanRaden, C. E. Rexroad, Jr., and R. H. Miller. 1997. Detection of loci affecting milk production and health traits in an elite US Holstein population using microsatellite markers. Anim. Genet. 28:216–222.

Boettcher, P. J., and B. J. VanDoormaal. 1999. Tools for selection for functional traits in Canada. Proc. Int. Workshop EU Concerted Action Genet. Improvement of Functional Traits in Cattle (GIFT); Longevity. Interbull Bull. No. 23.

Cowan, C. M., O. M. Meland, D. C. Funk,and D. F. Erf. 1997. Realized genetic gain following marker assisted seleciton of progeny test dairy bulls. Page 296 *in* Proc. Int. Plant Anim. Genome V Conf., San Diego, CA (Abstr.).

De La Fuente, R., A. Hahnel., P. K. Basrur, and A. W. King. 1999. X inactive-specific transcript (*Xist*) expression and X chromosome inactivation in the preattachment bovine embryo. Biol. Reprod. 60:769–775.

Georges, M. D., D. Nielsen, M. Mackinnon, A. Mishra, R. Okimoto, A. T. Pasquino, L. S. Sargeant, A. Sorensen, M. R. Steele, X. Zhao, J. E. Womack, and I. Hoeschele. 1995. Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. Genetics 139:907–920.

Groeneveld, E., and A. Garcia Cortez. 1998. VCE4.0, a (co)variance component package for frequentists and Bayesians. Proc. 6th World Congr. Genet. Appl. Livest. Prod. Armidale, Australia 27:455–456.

Hagger, C., and G. Stranzinger. 1992. A method to detect large influences of specific sex chromosomes on quantitative traits. Theor. Appl. Genet. 83:925–929.

Haley, C. S., and S. A. Knott. 1994. Interval mapping. Proc. 6th World Congr. Genet. Appl. Livest. Prod. Guelph, Canada 21:25–32.

Harris, D. L., V. A. Garwood, P. C. Lowe, P. Y. Hester, L. B. Crittenden, and A. M. Fadly, 1984. Influence of sex-linked feathering phenotype of parents and progeny upon lymphoid leukosis virus infection status and egg production. Poult. Sci. 63:401–413.

Janss, L.L.G. 1998. MAGGIC: a package of subroutines for genetic analyses with Gibbs sampling. Proc. 6th World Congr. Genet. Appl. Livest. Prod. Armidale, Australia 27:459–460.

Kuhn, M. T., and A. E. Freeman. 1995. Biases in predicted transmitting abilities of sires when daughters receive preferential treatment. J. Dairy Sci. 78:2067–2072.

Lyon, M. F. 1961. Gene action in the X chromosome of the mouse (Mus musculus L). Nature 190:372–373.

McKusick, V. A. 1966. Mendelian Inheritance in Man. Heinemann, London, UK.

Norman, H. D. 1976. Theoretical background for the USDA-DHIA modified contemporary comparison sire summary procedure. USDA Prod. Res. Rep. 165. USDA.

Schutz, M. M. 1994. Genetic evaluation of somatic cell scores for United States dairy cattle. J. Dairy Sci. 77:2113–2129.

Sonstegard, T. S., N. L. Lopez-Corrales, S. M. Kappes, R. T. Stone, S. Ambady, F. A. Ponce de Leon, and C. W. Beattie. 1997. An integrated genetic and physical map of the bovine X chromosome. Mammalian Genome 8:16–20.

Spelman, R. J., W. Coppieters, L. Karim, J.A.M. van Arendonk, and H. Bovenhuis. 1996. Quantitative trait loci analysis for five milk poduction traits on chromosome *six* in the Dutch Holstein-Friesian population. Genetics 144:1799–1808.

Takagi, N. 1978. Preferential inactivation of the paternally derived X chromosome in mice. Pages 341–360 *in* Genetic Mosaic and Chimeras in Mammals. L. P. Russel, ed. Plenum Press, New York, NY.

VanRaden, P. M. 1988. Evaluations of sires based on sons and maternal grandsons. J. Dairy Sci. 71(Suppl. 1):185. (Abstr.).

VanRaden, P. M., and G. R. Wiggans. 1995. Productive life evaluations: calculation, accuracy and economic value. J. Dairy Sci. 78:631–638.

Van Vleck, L. D. 1986. Evaluation of dairy cattle breeding programs: specialized milk production. Proc. 3rd World Congr. Genet. Appl. Livest. Prod. Lincoln, NE IX:141–152.

Van Vleck, L. D. 1993. Selection Index and Introduction to Mixed Model Methods. CRC Press, Boca Raton, FL.

Weller, J. I., Y. Kashi, and M. Soller. 1990. Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. J. Dairy Sci. 73:2525–2537.