One Hundred Years of Statistical Developments in Animal Breeding

Daniel Gianola[1–5] and Guilherme J.M. Rosa[1,2]

[1]Department of Animal Sciences, [2]Department of Biostatistics and Medical Informatics, and [3]Department of Dairy Science, University of Wisconsin-Madison, Madison, Wisconsin 53706; email: gianola@ansci.wisc.edu, grosa@wisc.edu

[4]Institute of Advanced Studies, Technical University of Munich, 85748 Garching, Germany

[5]Chair of Plant Breeding at TUM-Weihenstephan, 85354 Freising, Germany

Annu. Rev. Anim. Biosci. 2015. 3:19–56

First published online as a Review in Advance on November 3, 2014

The *Annual Review of Animal Biosciences* is online at animal.annualreviews.org

This article's doi:
10.1146/annurev-animal-022114-110733

Keywords

animal breeding, statistical methodology, quantitative genetics, selective breeding, complex traits

Abstract

Statistical methodology has played a key role in scientific animal breeding. Approximately one hundred years of statistical developments in animal breeding are reviewed. Some of the scientific foundations of the field are discussed, and many milestones are examined from historical and critical perspectives. The review concludes with a discussion of some future challenges and opportunities arising from the massive amount of data generated by livestock, plant, and human genome projects.

INTRODUCTION

Statistical methodology has played an important role in transforming animal breeding from an art into a science, and many scholars have contributed to the history of this process. Most animal breeding problems have a quantitative dimension, but not all are inferential, so when or where statistics enters into the picture is sometimes difficult to decide. For example, mate allocation is a decision problem (i.e., no inference) that can be solved deterministically with linear programming (1, 2). Likewise, organizing gene expression data into biologically meaningful subsets is a pattern recognition exercise that can be carried out by ignoring statistics via clustering (3). However, prediction of future phenotypes, clearly a statistical problem, can be carried out from noninferential perspectives (4, 5). Most animal breeders are concerned with predicting breeding value and do this by using theory developed by Henderson (6–11) with a clearly inferential objective.

Statistical methods have been an important feature of the World Congress on Genetics Applied to Livestock Production, the first of which was held in Madrid (1974) and the tenth of which was held in Vancouver in August 2014. Also, statistics has had a visible place in the International Conference on Quantitative Genetics, which was held first in Ames, Iowa (1976), and most recently held in Edinburgh (2012) and will take place next in Madison, Wisconsin (2016). Many innovative statistical ideas can be found in the corresponding proceedings. Reviews of statistical methods in animal breeding can be found, for example, in References 12–14 and in Hill (15), the latter in a broader context. Further, many statistical approaches grew out of collective efforts and informal scientific exchanges made in animal breeding workshops and scientific meetings (e.g., the American Dairy Science Association), regional projects, and even in coffee or beer breaks. Examples of hot spots of free scientific give and take of ideas are the (defunct) US Department of Agriculture (USDA) NCR21 regional project in quantitative genetics, the Edinburgh and Cornell coffee breaks, the Iowa State QTL Lounge, and the Wisconsin breakfast.

Our aim here is to review the history of the process indicated above, often critically, by offering an account of what we perceive are milestones, as it is impossible to discuss every contribution. We shall not attempt to provide a chronology of institutional developments connected with animal breeding. It is assumed that readers have been exposed to basic ideas of quantitative genetics, and recent discussions of fundamental principles and outcomes are available in References 16 and 17. The organization of the paper is as follows: First, some of the scientific foundations of the field are presented from a historical perspective. Subsequently, the stage is developed further through an overview of several animal breeding problems treated statistically. The review continues with a description of milestones followed by sections on best linear unbiased prediction, estimation of genetic parameters, Bayesian ideas as applied to the field, nonlinear models and longitudinal data, statistical treatments of selection bias, and genomic selection. A concluding section suggests some future challenges.

FOUNDATIONS

Application of statistical ideas to animal breeding and genetics traces back to the Victorian times of Galton (1822–1911) and Pearson (1857–1936), both of whom worked before Mendel's laws were rediscovered. Galton (18) found out that, on average, descendants from tall parents were smaller than their parents, whereas progeny from shorter parents were taller. This "regression to the mean" is intimately related to the concepts of heritability and expected response to selection. The difference in trait average between extreme groups of parents is akin to the selection differential; the difference between the corresponding progeny group means is equivalent to selection response.

The statistical regression of offspring on parent is the well-known heritability parameter (19), and the ratio between observed response and effective selection differential is realized heritability, a term coined by Falconer (1913–2004). Statistical problems in connection with inferring realized heritability are discussed in References 20 and 21. Galton's work provided a major impetus for the use of linear models, which are still employed in the twenty-first century. Galton's data suggested that the regression line was linear, but a reanalysis using nonparametric methods (22) found that the father-daughter, father-son, mother-daughter, and mother-son regressions exhibit a bend at approximately 67–68 inches of height. This bend suggests a hidden structure and illustrates how a statistical genetic model can provide a good description of a pattern and produce useful quantities, such as heritability, without necessarily explaining phenomena. Pearson (23, 24) wrote extensive papers on the evolution of traits and derived formulae, which Henderson (10) eventually used in an influential paper on estimation and prediction biases resulting from selection. Pearson's formula on how selection modifies a variance-covariance structure was crucial for Henderson's development and retrieves, as a special case, the standard expression for computing reduction in variance, given certain selection intensity under normality assumptions (25–27). The effect of some sort of stylized selection on genetic variance, termed the Bulmer effect, has been used in many studies. However, the simplest version of Pearson's formula is only an approximation, as it requires that all candidates for selection be independent and identically distributed. In animal breeding, candidates are related (thus, not independent) and have an unequal amount of information; for example, a progeny-tested bull may possess thousands of progeny with records, whereas a young sire may not. The stylized formula gives some rough basis for comparing idealized selection schemes.

After the particulate basis of genetics was established, one issue was how to reconcile continuously varying traits with Mendelian inheritance. Toyama Kametaro (1867–1918) found the first case of Mendelian inheritance in animals while studying the silkworm (28), and Yule (1871–1951) (29) reconciled Mendelism with continuous variation, although Pearson did not accept his work. Two monumental contributors to the foundations of modern quantitative genetics (with a major influence in animal breeding) were Fisher and Wright, in a series of papers. In an extraordinarily difficult paper, Fisher (30) introduced the most widely used model of quantitative genetics (the infinitesimal model) together with an analysis of variance for decomposing genetic variability into pieces. Wright (31), using path analysis and correlation, derived the inbreeding coefficient (F) and, in a series of papers (such as Reference 32), derived properties of Mendelian populations, including the astonishing equilibrium distribution of allelic frequencies in a finite population under random mating and mutation. Wright presented the result without derivation by invoking the Fokker-Planck equation from diffusion in physics, also known as the Kolmogorov forward equation in stochastic processes (33). Arriving at the equilibrium distribution requires many pages of algebra.

Fisher's infinitesimal model has been central in animal breeding. In a caricature of this model, suppose there are K loci at play, such that, at locus k ($k = 1, 2, \ldots, K$), adding an A allele adds the fixed (i.e., not random) quantity a_k to the genotypic value u (the additive value) of an individual, which can then be written as $u = W_1a_1 + W_1a_2 + \ldots + W_Ka_K$. Here, W is a random indicator variable taking the value 0, 1, or 2, depending on whether the individual is aa, Aa, or AA at the appropriate locus; if the population is in Hardy-Weinberg (HW) equilibrium, these genotypes appear with probabilities $(1 - p_k)^2$, $2(1 - p_k)p_k$, and p_k^2, respectively, where p_k is the probability of randomly drawing an A allele at locus k. The marginal distribution of u depends on the joint probability distribution of the genotypes at the K loci. Because u is a linear combination of random variables, if the W's are mutually independent (linkage equilibrium among genotypes), then as K increases the distribution of u converges to a Gaussian one, but linkage disequilibrium (LD) slows

down the rate of convergence. For u to have a finite mean and variance, the individual loci effects and frequencies must become infinitesimally small as $K \to \infty$. At the limit, one ends up with the distribution $u \sim N(m, \sigma_u^2)$, where the mean m is typically set to 0 and σ_u^2 is the additive genetic variance (often called polygenic). Henderson introduced the cow model in the 1960s (L.R. Schaeffer, personal communication), which then became known as the animal model (34). The specification was a vectorial extension of Fisher's model: The additive effect u became a vector of breeding values $\boldsymbol{u}$, and the additive genetic variance σ_u^2 was replaced by $\boldsymbol{A}\sigma_u^2$, where $\boldsymbol{A}$ is a matrix of additive relationships if there is no inbreeding (there are technicalities here) among individuals. The $\boldsymbol{A}$ matrix conveys kinship relationships: An entry of this matrix is twice the probability that an allele at a random locus is identical by descent in a pair of individuals. Such work was built on foundational notions of gametic similarity due to Sewall Wright, using correlation, and Malécot (35), who employed probability instead.

Another idea from Fisher (30, 36) impacting the notion of breeding value was the average effect of a gene substitution at a locus. This idea was taught in animal breeding courses by Lush and was later assisted by an outstanding text by Falconer (37). Assuming that all K loci above are in HW equilibrium, introducing a dominance effect d_k, and letting $1 - p_k = q_k$, the mean value of u is $E(u) = \sum_{k=1}^{K} [a_k(p_k - q_k) + 2d_k p_k q_k]$. The average effect of a gene substitution at locus k is $\alpha_k = a_k + d_k(q_k - p_k)$. The breeding value at that locus is $2q_k\alpha_k$, $(q_k - p_k)\alpha_k$ and $-2p_k\alpha_k$ for AA, Aa, and aa genotypes, respectively, and the breeding value of u for an individual with a certain genotypic configuration is obtained by summing over loci. Breeding values are both frequency and dominance deviation dependent, and the expressions rely on the HW assumption. Under idealized assumptions, only additive effects are transmitted; therefore, a_k is of great interest in breeding, so u is more narrowly defined as the sum of the a_k over loci. Thus, u became a random variable involving additive effects only (the infinitesimal breeding value), and, prior to the advent of genomics, it was the focal point in inference because genes and allelic effects could not be observed.

Until recently, quantitative genetics as applied to animal breeding was essentially a gene-less science. This explains the enormous influence of statistical procedures and, in particular, of Henderson's methods, where the sole genetic input was the matrix $\boldsymbol{A}$, defined earlier. In spite of genomics, little is known about individual genes affecting complex traits because the search for quantitative trait loci (QTL) using markers has not returned dividends consistent with the enormous resources invested in the quest; Reference 17 discusses this issue.

OVERVIEW OF SOME ANIMAL BREEDING PROBLEMS

Genetic selection, a crucial tool of animal breeding, aims to maximize the rate of increase (with respect to time) of some aggregate merit function, e.g., profit, reflecting the combined economic value of traits for which improvement is sought. Traits potentially related to environmental impact, such as methane production or use of energy resources, are increasingly considered as part of this function. Merit can be linear or nonlinear on the unobserved genetic values for the target traits. These traits are called complex because their inheritance remains unresolved, and it is accepted that environmental influences are large and possibly interact with genetic factors. Whenever there is a genetic basis underlying the components of the merit function, i.e., if there is genetic variability among individuals, it may be possible to produce genetic improvement from selection, but this depends on the kind and amount of variance available. Typically, animals deemed to bear the highest additive merit are kept as parents, and those with the lowest merit are culled, although selection can also aim at some intermediate optimum.

In the absence of detailed genetic knowledge, abstraction is required to glean genetic merit from observed data, and quantitative genetics theory provides some of the basis. This theory, albeit useful, is mechanistically simplistic in light of emerging knowledge on metabolic pathways, gene networks, and genome organization. Statistical methods put the theory into action, and complicated multivariate analysis is often required, as genetic or environmental associations among traits must be accounted for so that manifold effects of selection can be assessed reasonably. Ronald Fisher's (1890–1972) fundamental theorem of natural selection and a secondary theorem by Alan Robertson (1920–1989) motivated by animal breeding problems state that the rate of progress under selection is proportional to additive genetic variance and covariance, respectively; Crow & Kimura (33) and Edwards (38) give readable accounts. These two metrics are statistical, and models used to estimate such parameters are based largely on the assumption of additive inheritance. If there is nonadditive genetic variability, much of the theory is based on unrealistic assumptions to accommodate complex interactions about unknown genes. The pervasive presence of LD in livestock, owing to small population size and selection, makes it difficult to partition genetic variance (39). If a network of genes is in LD, inferring how much a specific gene contributes to variance is messy. Variability comes either through direct paths or indirectly through correlations stemming from LD (40). Sewall Wright (1889–1988), one of the founders of the theory of population genetics, introduced path analysis to separate direct from indirect effects, but this requires knowledge of LD relationships among causal genes.

Input information available for assessing the genetic basis of traits and for inferring genetic merit of animals has consisted of on-the-farm records of performance and genealogical data, now supplemented with a massive number of molecular markers known as SNPs (single-nucleotide polymorphisms). These genetic markers are used as a shotgun aimed at picking up associations between genomic regions and phenotype and for what is called genomic selection (41). The records of performance may be growth rate and feed intake in meat animals; fleece weight and quality in sheep and goats; milk yield, composition, reproductive performance, and survival in dairy animals; and egg production or litter size in polytocous species, such as chickens or pigs. Records on reproductive events or diseases, e.g., mastitis in dairy cattle, are more difficult to obtain, and often proxy variables are used, for instance, somatic cell counts (SCC) in cows' milk and number of ticks in the skin of cattle under tropical or subtropical conditions. Other traits, e.g., survival or length of productive life in dairy cows, are plagued by statistical censoring: Sometimes all is known is that an animal is alive at time t, but nothing is known beyond t. Also, many traits are recorded as counts (litter size) or categorical assignments (e.g., calving difficulty or stages of some disease). Thus, statistical modeling often requires going beyond assuming a Gaussian distribution, although the latter often provides a reasonable and useful approximation. Even if a trait is continuous, specifications other than the Gaussian, such as the double exponential or t-distribution, may confer robustness to an analysis.

Contrary to simple Mendelian traits, the genes affecting a complex trait are possibly many, as stated earlier. Lush (1896–1982), arguably the father of scientific animal breeding, often expressed the view (19, 42) that possibly all genes affect all such traits. In spite of spectacular advances in genomics, we still do not know the number of genes, the form of gene action, and the allelic frequencies and effects on most complex traits. The statistical approaches used in animal breeding lump the entire impact of the genome on a phenotype into something called genotypic value. The phenotype is represented using some mathematical model with one of its parts being the additive genetic value, termed breeding value. However, the genetic value or any of the model components are not observable and must be inferred from data on individuals with records or from data on relatives. The inference process most often (but not always) uses a linear model because this is analytically tractable; less taxing to compute than a nonlinear one; and can produce useful,

easy-to-interpret results. Possessing records on relatives is crucial in some cases; for example, a valuable beef bull cannot be destroyed for the purpose of obtaining carcass information, so data from genetically related steers are used. Given the abundance of genome-based markers, it is tempting to conjecture that relatives might not be needed, as it would seem sufficient that individuals be molecularly similar without possessing explicit genetic relatedness. Unfortunately, genomic similarity at markers does not translate well into genetic similarity at causal variants unless markers are in strong LD with QTL. The latter is another abstraction used to denote a genomic region that does some statistically significant thing to the phenotype. Marker-assisted inference in animal breeding probably has its origins in a paper by Neiman-Sorensen & Robertson (43), who intended to relate variants at blood groups to production in cattle.

Although many traits seem to have a polygenic mode of inheritance, standard genome-wide association methods based on regression of phenotypes on a single marker (GWAS) often fail to pick up many statistically significant variants, and the latter explain a tiny fraction of trait variability (44). Failure of rejecting the null hypothesis in a GWAS is often construed as evidence in support of a polygenic model, but this is insufficient from a causal perspective.

Animal breeding data sets can be very large (e.g., millions of lactation records in dairy cattle), multivariate (several traits must be modeled simultaneously), seemingly Gaussian in some instances (e.g., logarithm of concentration of somatic cells in milk, an indication of udder disease), or nonnormal in others (such as with discrete traits). Data structure can be cross-sectional or longitudinal (for example, growth curves in broilers) and extremely unbalanced, often exhibiting a pattern of nonrandom missingness. For instance, not all first-lactation cows produce a second lactation, owing to sequential selection for higher production, reproductive failure, or disease. Some sires are used more intensively than others, because of perceived differences in genetic value, so there is genetic selection as a consequence of variation in contribution to offspring born. Because of this, observed data are seldom a random sample, which thus introduces biases. Another issue creating havoc in genetic evaluation is undeclared preferential treatment of progeny of valuable sires; this produces statistical confounding: The effect of true genetic merit cannot be disentangled from that of the environment. In these respects, there is similarity with problems encountered with observational data in human medicine. Unfortunately, randomized experiments are seldom feasible with humans, and the suitable extent of replication required in animal breeding experiments is often impossible to attain even with laboratory animals. Our opinion, contrary to that of many contributors to the book edited by Robertson (45) or of Hill (15), is that selection or crossbreeding experiments have done little to advance biological understanding beyond what was known prior to the experiment. As Hill (15, p. 8) stated, "There was no real upset to the models and arguments of Wright and Lush, or indeed of Fisher." However, numerous experiments served in many cases to produce genetically distinct lines (e.g., in terms of mean body weight or backfat thickness in pigs) used for detailed biological analysis and for inspiring important theoretical work. Examples of the latter are the probability of fixation of a favorable allele in connection to selection limits (46, 47), or connections between genetic drift, variability of selection response, and genetic relatedness (20, 48, 49). Many of these theoretical ideas have been reenergized owing to the advent of genomic markers because genotypes (although not necessarily the relevant ones) can now be identified.

At least two types of statistical problems are encountered in the process of learning genetic values. First is assessing whether traits have a genetic basis, known as estimation of genetic parameters. Second is developing reasonably accurate methods for inferring merit or genetic evaluation, a term coined by Charles Henderson (1911–1989). A third type of problem, not dealt with here, is that of deciding what to do with animals that have the best evaluations; examples are inbreeding avoidance schemes, mate allocation, and exploitation of heterosis if there is nonadditive genetic

variation. Important contributors to the latter problems were Gordon Dickerson (1912–2000), who worked on uses of crossbreeding, and Ralph Comstock (1912–1999), who introduced a selection scheme based on crossbred progeny known as reciprocal recurrent selection (50). Also, considerable work on beef cattle breed use and crossbreeding, including estimation of heterosis and breed maternal effects, was carried out by many scientists working at Clay Center, a USDA experiment station in Nebraska; Dickerson's influence here was marked. This experimental work did not lead to methodological breakthroughs or inform much about the nature of inbreeding depression or heterosis, phenomena already known by maize breeders. However, it was useful for characterizing breeds quantitatively under controlled conditions.

Additional difficulties are posed by sex-limited traits, such as milk production in females and scrotal circumference in bulls, believed to have a positive genetic correlation with fertility in cows. In dairy cattle, it is relevant to infer genetic merit of males accurately, because of the impact these have on gain from selective improvement. Cows are evaluated as well, but at lower levels of precision than sires. As a result of refinements in artificial insemination techniques and of widespread availability of frozen semen and embryos, some dairy bulls can produce thousands of daughters in several countries, creating an opportunity for international sire evaluation via genetic connectedness, albeit at the expense of complicated statistical modeling and implementation (51). International dairy sire evaluation has been carried out in Sweden by a multicountry organization (Interbull) since the early 1990s, and open meetings of this organization have fostered advances in methods for genetic evaluation, especially ingenious computing procedures for big data.

Arguably, most methodological developments in animal breeding have been inspired by post hoc analyses of data, and Henderson, working at Cornell, was a giant here. This approach is patent in today's era of genomics, where data mining (viewed as a fishing expedition) often supersedes the standard protocol of hypothesis formulation, statistical power calculations given some experimental design, and then analysis. Henderson (personal communication) seldom took statistical significance too seriously, either because conceptual repeated sampling of the same animal breeding scheme (needed to calibrate long-run frequencies) is hard to envisage or because obtaining standard errors in a realistic model with unknown variance components cannot be done without recourse to approximations. The emergence of massive amounts of data as a result of genome projects catalyzed the development of a new field: bioinformatics. Here, data-mining algorithms interrogate and analyze large, complex data sets often generated outside of statistically specified protocols. Most of the ideas emerged from computer science, but statisticians complemented their perspective by creating an interphase field called statistical learning (52–54), in which a probabilistic framework enters into the picture, thus helping to assess uncertainty. Hypotheses are formulated after the data have been observed, an anathema in classical statistical thinking. Perhaps the training of animal breeders will place more focus on bioinformatics and artificial intelligence in the future, a direction that few groups in animal breeding have taken. In short, a vast suite of statistical difficulties is encountered in animal breeding data, and analytical paradigms have evolved vastly in the past hundred years. Foundational ideas leading to statistical genetic models are discussed in the following section.

MILESTONES

Use of Mathematical Models

Use of mathematical models in animal breeding traces back to Lush, who was influenced by Wright and perhaps to a lesser extent by Fisher. Lush's (55, 56) use of path diagrams based on correlations concealed an underlying linear model, as path analysis is a regression model with

standardized coefficients. Oscar Kempthorne (1919–2000), a British statistician who worked at Iowa State University, expressed this view often (e.g., Reference 57). It took a few years after Henderson's (6) thesis before associating an analysis to a model became routine, even though assumptions were not always stated precisely. Ideas from Eisenhart (58), such as the mixed model, were incorporated in Henderson's (7) paper on (co)variance components, where he presented estimators for purely random and for mixed effects models.

Models in animal breeding typically include a mathematical function relating observations to fixed location (defining the mean of a distribution) parameters and random effects, such as $\boldsymbol{u}$, under the assumptions of the infinitesimal model or additional genetic components, e.g., dominance and genetic interactions (epistasis). Additional factors declared random may be herds (flocks), permanent environmental effects in repeated measures applications, and environmental effects common to littermates. Random effects contribute to correlations between phenotypes (owing to genetic and environmental similarity) or between longitudinal records of performance. The distributions of random effects are indexed by genetic and environmental dispersion parameters, e.g., components of variance and covariance; the latter appear in multivariate models or when a multivariate structure is embedded into a model for a single response variate, as with genetic maternal effects. This setting led to the development of many procedures for estimating variance and covariance components (7, 59–61). The fixed effects represent unknown constants that take the same value in every hypothetical repetition of an experiment; declaring an effect to be random implies that it is drawn from a statistical distribution, so realized values change over repetition, such as in a coin-flipping experiment. However, the breeding value of a sire is a fixed entity, as explained earlier, but a sample of distinct alleles from this sire is drawn every time a gamete is formed (unless the sire is completely inbred), resulting in genetically different progenies. There are some reasons for a random effect treatment of sires: It makes breeding value estimable when it may not be so under the fixed effect assumption (11, 62), it produces more stable estimates in the mean squared error sense, it tempers overfitting in prediction of future records, and it allows probabilistic statements about breeding values even for individuals without records. In today's era of genomic markers, treating the latter effects as random is a must because the number of quantities to be inferred (p marker effects) vastly exceeds sample size (n). This is not distinct from the situation that arose with the animal model, where the number of breeding values also exceeded sample size. Nevertheless, there are delicate inferential problems when $n < p$ that animal breeders often ignore in a somewhat naïve quest to unravel genetic architecture via statistical modeling (5). A contradiction is that the basic model of quantitative genetics assumes random genotypes and fixed substitution effects, but whole-genome prediction models use realized genotypes (thus fixed) and random marker effects. This contradiction must be kept in mind for properly interpreting concepts such as genomic heritability and genomic correlations between traits (63).

Most often, model functional forms have been linear. Actually, after Henderson (10), the following specification (more or less) became a sort of panacea, irrespective of trait and species: fixed contemporary groups, fixed genetic groups, and random additive effects, plus a residual. Reasons for this straight-line view are that the basic quantitative genetic model is linear on effects of unknown loci, that a linear model is analytically and computationally tractable, and that an additive model leads to an algebra producing a sparse inverse of the $\boldsymbol{A}$ matrix (64). The additional assumption of normality yields closed-form likelihood functions, facilitating variance component inference (59, 65). With some exceptions, proper modeling yielded to the need of computing a massive number of linear equations for breeding value assessment, and the more equations, the better. Feldman & Lewontin (66) stated that a linear model for genetic data should not be taken as more than a local approximation, a view strongly rebutted by Kempthorne (67). Although convenient, linearity is not always a sensible specification, for example, in analysis of growth and

lactation curves (68), an area with heavy transit in animal breeding. However, nonlinear trajectories can be reproduced using linear splines or reproducing kernels (69), albeit at the expense of dispensing with beautiful theories on lactation and growth (70). Feldman & Lewontin (66) offer a reminder that a statistical genetic model and state of nature are often distinct entities.

Assumptions on the form of the joint distribution of the observations and of the random effects can affect an analysis crucially. The most widely used assumption in animal breeding model construction has been normality. This is so because of the view that there are a large number of genes acting additively with infinitesimally small substitution effects. However, this model can be refuted: With an infinite number of loci and alleles, the probability of finding a significant effect should be 0. Marker-assisted selection (MAS) (71, 72), introducing the QTL abstraction into animal breeding, provides a refutation. Subsequent molecular information, however, has indicated that the assumption of many genes acting together on quantitative traits is not unreasonable, at least in many instances. For example, using genetic markers, Zhang et al. (73) reported QTL affecting fat percentage in milk in chromosomes 2, 6, 14, 26, and 28, and there have been similar reports in many species. The QTL industry has been productive: A release (dated April 23, 2014) of the AnimalQTLdb (**http://www.animalgenome.org/cgi-bin/QTLdb/index**) reported the following numbers of curated QTL: 10,497 (pigs), 9,180 (cattle), 4,282 (chicken), and 798 (sheep), with 0 QTL in the horse because no papers have been published on equines. Some statistical models for marker-assisted genetic evaluation require knowledge of recombination rates between the unobserved QTL and the markers (74, 75). It is unclear what these QTL are (a genomic region?), and it may be useful to adapt Koch's postulates before graduating a QTL: (*a*) The bad variant of a QTL must be found in abundance in cases (diseased or extreme phenotypes) but not in controls. (*b*) The QTL must be molecularly identified in cases and cloned. (*c*) The bad alleles should cause disease when introgressed into controls. (*d*) The bad alleles should be recloned from such controls and verified to be identical to those identified in *a*. These standards, developed in the nineteenth century, have not been applied, save for exceptional cases (76). It is not always clear when and how a putative QTL becomes a legitimate QTL, so caution should be exercised when used in models for genetic evaluation. Juxtaposing QTL maps to massive scans of sequences from analysis of selective sweeps may help in this process (77). That so many QTL affecting so many traits have been found in so many species suggests that Fisher got it approximately right. However, Fisher never recommended that one should jump from univariate to multivariate normality (as in a multiple-trait analysis) automatically, unless some conditions are met. At any rate, most estimates of genetic correlation have been obtained from multivariate analysis of traits with obviously distinct marginal distributions (e.g., calving ease scores and gestation length). Perhaps this violation of theory is of minor importance if the objective is merely descriptive (e.g., if x increases, then y decreases) or predictive, but it is not the best possible use of statistical science.

Epistasis

Theory makes reference to two types of nonadditive (i.e., statistical interaction between allelic effects) gene action: dominance and epistasis. Dominance, an interaction between alleles at the same locus, has been discussed mainly in the context of hybridization schemes; Gowen (78) gave a review of the state of the art by the middle of the twentieth century. In this book, Comstock and Robinson presented the North Carolina designs I, II, and III for estimating average dominance of genes, but not much useful theory was developed thereafter. The book also contains a chapter of historical value by Henderson because it seems to be the first published account of his mixed model equations (in scalar form); Rohan Fernando made the discovery when he was a student at the University of Illinois, and even Henderson was surprised. Practical exploitation of dominance is

mainly a mating scheme issue. Further, when dominance effects are treated as random, it is difficult to obtain precise estimates of the corresponding variance within a population because the data must contain relatives receiving alleles from two lines of descent, such as full-sibs or double first cousins. The dominance variance can be estimated using a dominance relationship matrix constructed (in the absence of inbreeding) from entries of $\boldsymbol{A}$. Construction is complicated under inbreeding, as shown by Smith & Mäki-Tanila (79). Animal breeders have often attempted to account for heterosis or inbreeding depression by using an additive model, resulting in a contradiction. Models for breed crosses typically have been of fixed effects. Recently, Sun et al. (80) revisited mate allocation using SNP markers and also estimated dominance genomic variance, as in Reference 81. How this marked dominance variance relates to genetic variance must be clarified, because markers are not QTL.

Much has been written about epistasis, and discussions have often been on the semantic side (82). Fisher (30, p. 408), in a discussion of two-locus epistasis, wrote:

> Such dual epistacy, as we may term it, is the only kind of which we shall treat. More complex connections could doubtless exist, but the number of unknowns introduced by dual epistacy alone, four, is more than can be determined by existing data. In addition it is very improbable that any statistical effect, of a nature other than that which we are considering, is actually produced by more complex connections.

An intuitive motivation of epistasis assuming no dominance, HW, and linkage equilibrium is obtained by posing a linear regression model of the phenotype (y) on the number of alleles at each of two loci $E(y \mid X_1, X_2) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{12} X_1 X_2$, where X_1 and X_2 denote the number of copies of the A allele in a given locus, and $E(.|.)$ is a conditional expectation. If the regression coefficient β_{12} is null, one has an additive model (no epistasis). The impact of an allelic substitution at locus 1 corresponds to $\dfrac{\partial E(y \mid X_1, X_2)}{\partial X_1} = \beta_1 + \beta_{12} X_2$, indicating that it depends on the number of copies at locus 2, i.e., epistasis. The trait mean, over the entire population, is $\left[E(y|X_1, X_2)\right] = \beta_0 + 2p_1\beta_1 + 2p_2\beta_2 + 4p_1p_2\beta_{12}$, so $\dfrac{\partial E(y)}{\partial p_1} = 2\left(\beta_1 + 2p_2\beta_{12}\right)$. Similar to breeding value, epistasis is also frequency dependent. In agreement with Hill et al. (83), the preceding illustrates that when an allele is rare, the change in mean value from modification of frequencies is largely dominated by the additive term, unless β_{12} is huge. Actually, most of the genetic variance is often additive, even when epistasis is involved in the biology of the trait. However, the latter must be the case because complex traits are outcomes of concerted metabolic reactions whose enzymes are coded by distinct genes, and Michaelis-Menten kinetics dictates nonlinearities between substrate concentration and reaction rates, thus affecting gene product output in some nonlinear manner. Recent evidence of abundant epistasis for quantitative traits is given by References 84 and 85, using genomic data. Also, it does not seem to be true that higher-order epistasis is always negligible. For example, Taylor & Ehrenreich (86) reported interaction among a five-gene system in yeast. However, the results of Hill et al. (83) indicate that considerable epistasis translates into little epistatic variance. A corollary might be that, if epistasis is important biologically but does not translate into much more than additive variance, a plateau is placed on the limits of variance components as detectors of genetic architecture. Such an analysis would deny what it is and explain what it is not. Lush (19) provided comfort to breeders by stating that selection for epistatic effects is like building a sand pile on the seashore: The waves (e.g., recombination) eventually flatten the pile. Animal breeders have been able to make progress by concentrating on breeding value and ignoring epistasis altogether.

Although Fisher (30) mentioned epistasis, it was not until Cockerham (87) and Kempthorne (88) that the variance from such interactions was partitioned into what are called epistatic components. Cockerham employed orthogonal polynomials, and Kempthorne used probability of identity by descent, as developed by Malécot (35). Their assumptions included a large panmictic population and absence of linkage. The epistatic variance then breaks into several orthogonal components, depending on the number of loci involved in the expression of the trait. For example, with two loci, the epistatic variance is the sum of additive × additive, additive × dominance, dominance × additive, and dominance × dominance. Henderson (89, 90) used this result to infer dominance and epistatic genetic effects, as well as to predict total genetic value via BLUP, a topic discussed later.

Models for Other Genetic Effects

In the 1960s, there was much interest in maternal genetic effects in several livestock species, such as beef cattle, swine, and even chickens, with many papers published thereafter. The basic idea is that even though a maternal influence is environmental with respect to the progeny possessing the record of performance, differences among dams may be partly genetic. Willham (91) and Falconer (92) presented two important models. Willham's (91) specification embeds a multivariate within a univariate structure. This result is obtained by positing a direct genetic effect on the phenotype that is peculiar to the offspring and a maternal genetic effect that, although specific to the dam, acts as an environmental factor. If direct and maternal genetic effects are correlated, the genetic variance includes a covariance between direct and maternal genetic effects, thus the multivariate structure. Many beef cattle genetic evaluation schemes use variants of Willham's. In Falconer (92), the phenotype of the dam enters as a covariate in the model for the offspring record, which allows linking of the phenotype of the offspring, e.g., to the dam's and grand-dam's phenotypes. Koerkhuis & Thompson (93) suggested an integrated model. Van Vleck (94) wrote a readable account of models with maternal and grand-maternal effects and used the term embedded characters when referring to these or similar sources of variation. Reference 95 provides a useful review.

Skjervold & Fimland (96) presented a model similar to that of Willham in which the objective was to take into account the effect of the fetus that a female carries on her milk production, a phenomenon reported in dairy cattle. Van Vleck (97) adapted Willham's model such that the phenotypic variance included an additive direct genetic variance, an additive variance for fetal effects, and a covariance between additive direct and fetal effects. Models have been also proposed for cytoplasmic effects (males do not transmit mitochondrial genes to their offspring) (for relevant literature, see 94, 98, 99). More recently, Muir (100) and Bijma et al. (101) addressed the problem of group selection and introduced a model discussed in detail by Bijma (102) in whose paper, for a group of size n, the phenotype of an individual is affected by its genetic value and by effects from the $n - 1$ comembers of the group, called associative effects. This literature also employs the terms social and interaction effects (in lieu of associative), but this is misleading: There is no interaction in the model. The problem has garnered interest from animal welfare perspectives, but a more realistic representation of competition or interaction among genotypes is needed. Wright's (103) work on reciprocal effects may serve as an inspiration, and there is econometric literature on what was the subject of a Nobel Prize in Economics (104), addressing mutual effects using the concept of simultaneity between variables.

A problem of great interest in the 1980s was that of heterogeneous variances across environments. Hill (105), for example, found that if environmental variance is heterogeneous over herds, say, and genetic evaluation assumes homoscedasticity, then an excess of individuals is selected from the most variable environments. This launched much research into models for accommodating

heterogeneous dispersion, but the interest faded away, although the problem persists and may be important in genomic selection as well (most models assume homogeneous residual variance).

Epigenetics has received little attention in statistical animal breeding, and its relevance in livestock is documented in Reference 106. Epigenetics refers to genetically transmissible changes in gene expression that are not due to DNA structure, such as differential methylation of cytosine. One of the best-known epigenetic mechanisms is genomic imprinting: If a gene is imprinted in the female or male germ line (owing to differential methylation), the embryo carries a single effective copy of the gene. It is still unknown what percentage of the variance of complex traits is associated with imprinting (possibly owing to a dearth of data), but early evidence suggests that it is perhaps not much. Neugebauer et al. (107) developed a pedigree-based model with paternally and maternally imprinted additive effects, allowing for their covariance, and analyzed data for 10 traits on more than 65,000 German Simmental bulls. They found that, at most, imprinting explained 25% of the additive variance, assuming their model does what it intends to do. How knowledge of imprinting will improve prediction of phenotypes or breeding values remains to be seen.

BEST LINEAR UNBIASED PREDICTION (BLUP)

Overview

"Prediction" or "estimation" of breeding values is very important in livestock improvement. This wording created confusion in the field as, statistically, it is nonsense to estimate a random quantity, in our case u, the infinitesimal additive effect. Prediction conveys a futuristic connotation, but in animal breeding one is often interested in ranking candidates (e.g., bulls) that already exist, together with some that may not exist yet, such as the future outcome of a mating between bull A and cow B. Lush (55), using path coefficients, gave formulae for assessing the genetic merit of dairy sires, assuming that means and genetic and environmental components of variance were known. He found that some regression to the mean, i.e., shrinkage, was needed. Robertson (108) showed that Lush's statistic was a weighted average between population information and data, anticipating a Bayesian interpretation. Briefly, let s be the transmitting ability (half of the breeding value, TA) of a sire, and suppose that the distribution of TA in a certain population is $s \sim N(m, v_s)$, where the quantities in parentheses are the mean and variance of a normal distribution. Assume further that the model for a record measured in the progeny is $y = \mu + s + e$, where μ is the known population mean and e is a residual, supposed to have a normal distribution independent of that of the TA, and with mean 0 and variance v_e. If this sire has n progeny with average production (deviated from μ) $\bar{y} - \mu$, the weighted average of this deviation and of m produces as estimated TA $\hat{s} = \left[\frac{1}{\vartheta_s} + \frac{n}{\vartheta_e}\right]^{-1}\left[\frac{1}{\vartheta_s}m + \frac{n}{\vartheta_e}(\bar{y} - \mu)\right] = m + n(n+\alpha)^{-1}(\bar{y} - \mu - m)$. Here, the ratio $\alpha = \frac{4 - h^2}{h^2}$, where h^2 is heritability in the narrow sense; the 4 arises because the intrasire correlation between records is one-quarter of heritability. The preceding equation is the mean of the conditional distribution of the TA of the sire, given the progeny records, and $b = n\left(n + \frac{v_e}{v_S}\right)^{-1}$ is a regression coefficient (not in the least-squares sense) bounded between 0 and 1. The value b takes depends on the amount of information on the sire (n) and a measure of uncertainty is $v_{cond} = \left[\frac{1}{v_S} + \frac{n}{v_e}\right]^{-1} = v_e\left[n + \frac{v_e}{v_S}\right]^{-1}$; this is equal to $Var(s - \hat{s})$, which Henderson (9) termed prediction error variance. The expression also yields the variance of the conditional distribution

mentioned above. Dempfle (109) and Gianola & Fernando (110) later showed that $\hat{s}$ and v_{cond} are the mean and variance, respectively, of the posterior distribution (Bayesian) of the TA of the sire, assuming known variances and normality. Bayesian methods were anathema in animal breeding until the 1990s because such ideas were not taught in biological statistics courses, mainly owing to the influence of Fisher and notably Kempthorne, who viewed these methods adversely. Blasco (111) includes a history of some of the issues from an animal breeding perspective, and Grosu & Schaeffer (112) describe the historical progression of simpler (but obsolete) methods, such as daughter-dam comparisons, contemporary comparisons, and cumulative differences.

The preceding provides a canonical example of a general theory of prediction developed by Henderson (8, 9, 11), although it is mainly suitable for linear models. Henderson (9) introduced the best predictor (BP), the best linear predictor (BLP), and the best linear unbiased predictor (BLUP). BP is the function of the data (linear or not) that minimizes mean squared error of prediction; the answer is the conditional expectation function. Calculation of BP requires knowledge of all parameters of the joint distribution of phenotypes and genetic values. BP was shown (26, 113) to provide an optimum ranking rule: If a set of r out of n individuals is ranked with BP, the mean value of the unobservable quantity one seeks to improve by selection is maximized. A simple example of BP is $\hat{s}$ above, but the procedure also applies to prediction of vectors of genetic values, given vectors of observed phenotypes. In BLP, the search for bestness is restricted to linear predictors, and the same answers can be obtained simply by assuming multivariate normality. A special case of BLP is the selection index of Smith (114) and Hazel (115) in plant and animal breeding. Suppose one wishes to predict the linear aggregate genetic value $\boldsymbol{Mu}$, where $\boldsymbol{u} \sim (\boldsymbol{m}, \boldsymbol{G})$ is a vector of additive genetic values for one or more traits in a set of animals (or plants) and $\boldsymbol{M}$ is a matrix of appropriate order containing economic weights reflecting profit per unit of genetic input; $(\boldsymbol{m}, \boldsymbol{G})$ represents the mean vector and covariance matrix of the distribution of $\boldsymbol{u}$. The information comes from phenotypes distributed as $\boldsymbol{y} \sim (\boldsymbol{\mu}, \boldsymbol{V})$. The result is $BLP(\boldsymbol{Mu}) = \boldsymbol{M}BLP(\boldsymbol{u})$, with $BLP(\boldsymbol{u}) = \boldsymbol{m} + Cov(\boldsymbol{u}, \boldsymbol{y}')\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{\mu})$. Setting $\boldsymbol{B} = Cov(\boldsymbol{u}, \boldsymbol{y}')\boldsymbol{V}^{-1}$ produces the Smith-Hazel equation $\boldsymbol{VB} = Cov(\boldsymbol{u}, \boldsymbol{y}')$, developed by Smith (114) for the special case in which all individuals are genetically unrelated and each is measured for p traits; Hazel (115) used path coefficients instead. Henderson termed their setting equal information case. The classical selection index is a special case of BLP, and further, the economic weights $\boldsymbol{M}$ intervene in the calculation only after $BLP(\boldsymbol{u})$ is calculated. This is a linear invariance property (also holding for BLUP), and its first reported application was to select lambs by using data from Wisconsin farms (116). This was a recording program organized by A.B. Chapman (1908–2004), who was one of the first two students of Lush and who was markedly influenced by Sewall Wright. At any rate, BP requires knowledge of $\boldsymbol{V}$, of $Cov(\boldsymbol{u}, \mathbf{y}')$, of $\boldsymbol{m}$ (usually set to 0), and of the mean value of $\mathbf{y}$.

A crucial breakthrough was the development of BLUP. Henderson assumed that the variance covariance structure of the predictands ($\boldsymbol{u}$) and of the phenotypes was known, but that $E(\boldsymbol{y}) = \boldsymbol{X\beta}$, the mean vector of phenotypes, was a linear combination of some unknown fixed vector $\boldsymbol{\beta}$ and of a known incidence matrix $\boldsymbol{X}$. BLUP makes a search for the linear predictor with a minimum variance of prediction error in the class of linear unbiased predictors, in the sense that the expected value of the predictor is equal to the expected value of $\boldsymbol{u}$, typically a random vector of breeding values. The general linear model results from linking breeding values and phenotypes via $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{e}$, where $\boldsymbol{X}$ and $\boldsymbol{Z}$ are known incidence matrices, $\mathbf{u} \sim (\mathbf{m}, \mathbf{G})$ and $\mathbf{e} \sim (\mathbf{0}, \mathbf{R})$ are uncorrelated random vectors, and $\mathbf{G}$ and $\mathbf{R}$ are variance-covariance matrices that are a function of (known) dispersion parameters. The vector $\boldsymbol{u}$ can also include herd effects and nonadditive genetic effects, permanent environmental deviations common to all records of the same animal, and the incidence matrices allow for any type of covariates, such as the time variable in longitudinal models. The setting holds for any linear model, univariate or multivariate, cross-sectional or

longitudinal. Under this model, the variance-covariance matrix of the phenotypes is $\boldsymbol{V} = \boldsymbol{ZGZ}' + \boldsymbol{R}$, and $\text{BLUP}(\boldsymbol{u}) = \boldsymbol{m} + \boldsymbol{GZ}'\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})$, where $\hat{\boldsymbol{\beta}}$ is the generalized least-squares estimator of the fixed vector. Here, $\boldsymbol{m}$ is assumed known, but it can be given some linear structure (as in genetic group models) and included in the model as an unknown. Note that if $\boldsymbol{Z}$ is an identity matrix and $\boldsymbol{m} = 0$, $\text{BLUP}(\boldsymbol{u}) = \boldsymbol{H}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})$, where $\boldsymbol{H} = \boldsymbol{GV}^{-1}$ is a sort of heritability matrix. This is the celebrated animal (additive) model, and genetically, it is a vectorial representation of Fisher's infinitesimal model that provides for some optimal estimation, in some sense, of fixed and random effects.

The Mixed Model Equations

In a fortuitous mistake, Henderson discovered the celebrated mixed model equations that can often be used to advantage for computing BLUP; an account is in Reference 9. Briefly, by using normality assumptions, the joint distribution that underlies the standard mixed effects model is that of $\mathbf{u}$ and $\mathbf{y}$, given that the dispersion matrices $\mathbf{G}$ and $\mathbf{R}$ are known. If the joint density is maximized simultaneously with respect to the fixed and random effects, one obtains the linear system of equations:

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix},$$

where the $\boldsymbol{\beta}$-solution and $\mathbf{u}$-solution were later shown to be the maximum likelihood (ML) estimator of $\boldsymbol{\beta}$ (under normality) and the BLUP of $\boldsymbol{u}$, respectively. The latter is also an estimator of the BP (under normality) with the fixed vector replaced by a ML estimate, given the variance components. The mixed model equations are particularly advantageous when n is large, so brute force inversion of $\boldsymbol{V}$ is not feasible, or when the inverse of $\mathbf{G}$ is easy to obtain, as the number of $\mathbf{u}$-effects can exceed n, the sample size. Henderson et al. (117) mistakenly thought that a likelihood function was being maximized, so these solutions were termed ML estimators of $\boldsymbol{\beta}$ and of $\mathbf{u}$. The latter vector cannot be estimated, as it is random, and its number of levels can exceed sample size, so the likelihood function is underidentified. Today, it is known that the objective function maximized by Henderson is a joint posterior density, in a certain Bayesian setting, or a penalized or extended likelihood. Henderson (personal communication) presented his derivation at a statistical meeting in North Carolina in the early 1960s. One of the attendants, the famous statistician C.R. Rao, pointed out that the objective function was not likelihood, and Kempthorne was bothered by the shrinkage feature of BLUP, which induced estimation bias in identified models. However, another participant, Irwin Bross (1921–2004), observed that it was a legitimate Bayesian procedure. Henderson commented, "I almost fell from the podium" (C.R. Henderson, personal communication).

Henderson's theory, and especially BLUP, represented the first time that a comprehensive prediction paradigm appeared in animal breeding. Prior to BLUP, dairy sire evaluation was based on the herd-mate method, regressed least-squares, and contemporary comparisons, or variants thereof (112), with data corrected via least-squares and with the resulting deviations regressed through selection index theory. These methods were advocated by Walter Harvey, a biometrician with the USDA, and by Alan Robertson. Harvey wrote a least-squares program (118) that was used extensively in animal breeding; subsequently, a mixed-model package (119) called LSML76 included a BLUP option, incorrectly called maximum likelihood.

Animal breeders often misinterpret the unbiasedness property of BLUP. This method is unbiased under conceptual repeated sampling over the distribution $[\mathbf{y}, \mathbf{u} \,|\, \beta, \mathbf{G}]$ but not over

$[\mathbf{y} \mid \mathbf{u}, \beta, \mathbf{G}]$ (5). The latter is the distribution that practitioners have in mind, that is, one where $\mathbf{u}$ is a vector of realized breeding values. BLUP gives biased predictions of specific breeding values, although the bias can vanish, e.g., in sire models, by letting the number of progeny per sire go to infinity. This can be seen in Robertson's weighted average: As n goes to infinity, the regression coefficient goes to 1, and at the limit, one retrieves the true transmitting ability of the sire. This asymptotic property is hard to justify when the number of breeding values to be inferred exceeds sample size or when some individuals do not even possess a record of performance.

Solving the Mixed Model Equations

An obvious difficulty, at least in animal breeding, was that of inverting $\mathbf{G}$ in the MME (unless this matrix had an exploitable pattern, such as block-diagonality) when the order of $\mathbf{u}$ was huge, as in routine genetic evaluation of dairy cows in the United States. For example, if $\mathbf{u}$ is a vector of additive effects in a multiple-trait model, then $\mathbf{G} = \mathbf{G}_0 \otimes \mathbf{A}$, where $\otimes$ is the Kronecker product, $\mathbf{G}_0$ has order equal to the number of traits (a dozen, say), and the matrix of additive genetic relationships among animals is very large. Here, $\mathbf{G^{-1}} = \mathbf{G}_0^{-1} \otimes \mathbf{A^{-1}}$. In a remarkable breakthrough, Henderson (64) discovered that $\mathbf{A}^{-1}$ could be written directly from a list of parents of the animals. This enabled use of all available relationships in genetic evaluation, which led to more precise inferences about genetic values and to the possibility of correcting some biases owing to selection or to ignoring relationships in variance component analyses. The mixed model equations for calculating best linear unbiased estimator and BLUP have been employed worldwide for genetic evaluation of livestock mainly because much work has been done in the area of computing algorithms. The order of the linear system can be huge, especially for multivariate models, so iterative methods needed to be developed. This dimensionality issue also occurs when a random additive genetic effect is fitted for each animal with a record of production and when animals without records are included, to account properly for genetic covariances between relatives. Early implementations of iterative methods were made at Cornell University using the Northeastern Sire Comparison Method (9). This was essentially a cluster model in which diagonal terms dominated the off-diagonals in the mixed model equations, so the Gauss-Seidel algorithm had guaranteed convergence. Later, more suitable methods were proposed, such as iteration on data (120, 121).

BLUP was adopted earlier in Europe than in the United States for genetic evaluation of dairy animals. The delay was mainly because the USDA used a method called the modified contemporary comparison, which, although less theoretically appealing than BLUP, could be computed in a feasible manner. In 1988–1989, George Wiggans, a dairy cattle breeder with the Animal Improvement Program (USDA) in Washington, took a sabbatical at the University of Illinois, which had become a National Center for Supercomputing, and joined forces with Ignacy Misztal. These two scientists were the main forces driving use of the animal model in the United States. Other major contributors to computing strategies have been Karin Meyer, Steve Smith, Bruce Tier, Hans Graser, and Arthur Gilmour (Australia); Brian Kennedy (1943–1994) and Larry Schaeffer (Canada); Just Jensen and Per Madsen (Denmark); Esa Mäntysaari, Martin Lidauer, and Ismo Strandén (Finland); Vincent Ducrocq and Andres Legarra (France); Eildert Groeneveld (Germany); Robin Thompson (United Kingdom); Ignacio Aguilar (Uruguay); and Dick Quaas, Paul Van Raden, Curt Van Tassel, Dale Van Vleck, and Keith Boldman (United States), among many others. Curiously, the mixed model equations appeared in the statistical theory literature late and sparingly (65, 122–124). This is surprising because these equations can be used to advantage in computing algorithms for variance component estimation in generalized mixed effects linear models (125–127).

ESTIMATION OF GENETIC PARAMETERS

Important formulae in animal breeding, such as expected direct and correlated response to selection, depend on knowledge of genetic and environmental variance and covariance components, which translate into heritability and genetic correlations. The same holds for prediction of breeding value, as BLUP assumes that genetic parameters are known without error. Many methods for estimating such parameters have been developed over the past six decades, but only a few have stood the test of time because of lack of either generality or statistically optimal properties. Methods based on least-squares regression à la Galton or product-moment correlation between different types of relatives are now viewed as archaic. Hofer (128) reviewed procedures that had been applied in animal breeding close to the end of the twentieth century.

Today, most animal breeding data come from field records as opposed to randomized studies. In animal breeding, data sets are large, unstructured, and unbalanced and contain errors and sources of bias, such as unrecorded preferential treatment. Data are eventually processed at regional or national centers, where pedigree repositories and now massive DNA marker information are stored. Statistical methodology was needed to contemplate this messiness, and one aim was to separate signal (genetic value) from noise (everything else), which was complicated by the presence of a large number of nuisance parameters, such as contemporary groups (e.g., herd-year-season classes). The simpler analysis of variance type methods used by Lush, Hazel, and other pioneers lacked generality. Henderson (7), in a classical paper on variance and covariance component estimation, described three methods for unbalanced data: One was for purely random effects models, and two were for mixed effects models. The most general, Method 3, computes a series (often not unique) of least-squares-based quadratics on the data and equates these to their expected values under the model. If the specification holds, the approach produces unbiased estimators; however, not much is known about their statistical properties, e.g., sampling distribution, making it difficult to construct confidence intervals. Unfortunately, these procedures can produce negative estimates of heritability and, if extended to a multiple-trait setting, estimates of covariance matrices with negative eigenvalues. Estimates outside of the parameter space are ridiculous, whether or not a procedure is unbiased. Method 3 was implemented in software that was used amply in animal breeding (118, 119). Searle, a New Zealand statistician (1928–2013) working at Cornell, clarified Henderson's methods and presented these in matrix form (61, 129). Routine use of matrix algebra in animal breeding started in the 1970s; as Henderson (9, p. 10) predicted, "Much of the presentation that follows is in matrix notation, and for this I offer no apology as this has rapidly become an essential tool of any serious student of animal breeding." This statement was viewed with skepticism by some, but today it is unusual to see a paper in animal breeding that does not use matrix algebra.

Another epoch in genetic parameter estimation started when Rao (60) and LaMotte (130) introduced minimum norm quadratic unbiased estimation and its minimum variance version (under normality), respectively. By using (a lot of) matrix algebra, much work was done at Cornell to express these estimators as solutions to the mixed model equations, because the inverse of the phenotypic covariance matrix was needed in the original representations. Knowledge of the true parameters is needed for these methods to be optimal, in a sort of self-defeating exercise with a happy ending if the procedures are iterated. Although they are more advanced than Method 3, negative values of variance components or embarrassing estimates of covariance matrices could still be obtained with these procedures. Then ML assuming normality, where these problems do not occur, became a focal point. ML dates back to Fisher (131), who introduced likelihood as a rational measure of belief in lieu of probability; the latter was a dangerous concept if cast in a Bayesian fashion (111). Papers by Hartley & Rao (59) and Harville (65) were an overture to ML

estimation of variance components. Subsequently, many algorithms for ML estimation were derived by using the mixed model equations (11, 65, 132), including the celebrated EM algorithm (133). It is unclear whether the move toward likelihood-based methods in animal breeding was a consequence of the availability of something new that could be computed with the mixed model equations (used iteratively, as ML estimators cannot be written explicitly for most models) or because of the appeal of the large sample properties of the method. Getting estimates inside of the parameter space was important conceptually, as researchers were interpreting negative estimates of heritability as evidence supporting the hypothesis that there was no recoverable genetic variability. This is wrong, as one can simulate data with non-null genetic variance and obtain negative estimates, simply because unbiasedness is a weak and perhaps overrated property. Henderson often spoke with reservation about contrived standard errors, such as those presented in Harvey's packages, and not without reason, as the formulae used seldom applied to a mixed model. ML also offered a solution because model-specific asymptotic confidence intervals can be obtained as a by-product.

Although biased, the most appealing property of the ML estimator is its consistency: It converges to the true value when information content is infinite, assuming the model holds plus some extra conditions. However, the well-known downward bias of the ML likelihood estimator of the residual variance in a regression model disturbed some researchers and led to widespread interest in a method called residual or restricted maximum likelihood, or REML for short. The basic ideas date to the early 1950s, but Patterson & Thompson (122) gave a more general description, suitable for a mixed effects model. REML was an advanced attempt at accounting for loss of degrees of freedom incurred in estimating fixed effects, which would reduce bias of the estimates; i.e., a search for lack of bias seemed to be the driver. Patterson & Thompson (122) noted that a modification of the likelihood led to estimating equations that were similar to those in ANOVA, at least in balanced layouts. The logic was not entirely Cartesian: Bias removal always occurs at the expense of reducing precision of estimates. A well-known phenomenon in statistics is the bias-variance tradeoff: Bias reduces variance and often mean squared error of estimation. Meyer & Kirkpatrick (134) recognized this in an animal breeding paper employing penalized ML (REML). The penalty intends to reduce mean squared error, but in practice one seldom knows what the optimum penalty is for a given data set.

It was hard to choose between REML and ML because the two methods have exactly the same asymptotic properties. Harville (135) later made a more compelling case for REML, showing it was the mode of a (Bayesian) posterior distribution of the variance parameters after integrating the fixed effects (over an improper, uniform prior) out of the joint posterior distribution, this being proportional to the likelihood function. In hierarchical or variance component models, both ML and REML are biased, so it would have been unfair to focus the discussion on bias only. In general, whatever favors REML in this sense may be compensated by a loss of precision of the estimator. Its Bayesian interpretation, indicating how uncertainty about fixed effects (acting as nuisance parameters) is accounted for via integration, is compelling (to some) and clearer than many convoluted arguments based on likelihood profiling advanced later. REML gradually became established as a method of choice for estimating genetic parameters, with multiple-trait generalizations coming shortly thereafter. Robin Thompson, a British statistician with an interest in animal breeding, was a major player in the process and contributed much to REML software development. In short, after Henderson's BLUP, the advent of ML and REML was the next major breakthrough in animal breeding statistical methodology. There was also an impact on statistical training because animal breeders realized that understanding of joint and marginal distributions was needed to comprehend a likelihood function. Well-trained animal breeders now take at least two semesters of graduate-level mathematical statistics.

An important question that remained to be answered in the post-REML world was what method of estimation of genetic parameters was best when the end point was prediction. A method that is good for estimating variance components (parameters are never observable, so one is at the mercy of theoretical arguments) may not be so for predicting observables, e.g., progeny averages or phenotypes. In this latter case, predictive quality can be calibrated via a suitably designed cross-validation, such as the one currently employed in genomic selection. Because cross-validation came into the picture later, the arguments remained theoretical. For example, Gianola et al. (110, 136) employed Bayesian ideas to answer this question and argued that REML provided a sensible approximation to the question. In the late 1990s, the state of the art in animal breeding data analysis was the REML + BLUP tandem. REML had a likelihood justification, whereas BLUP had a frequentist one. The question of whether this recombinant produced the best predictive performance remained open. The answer seemed to be negative, as illustrated at least by Harville & Carriquiry (137).

BAYESIAN IDEAS IN ANIMAL BREEDING

Frequentist and likelihood-based approaches dominated statistical views in animal breeding during most of the twentieth century, as the teaching of statistics was focused on such methods. Bayesian ideas reentered into statistics with force by the middle of the century because statistical scientists such as Savage (1917–1971), Lindley (1923–2013), and Box (1919–2013) began to question foundations of the field. The work of James & Stein (138) provided impetus: These authors showed that the ML estimator of a vector with at least two parameters in an orthonormal linear model was always inferior (mean squared error sense) to an estimator that shrunk estimates, later shown to have a Bayesian interpretation. However, this work was too stylized to capture attention from animal breeders, as the area had become dominated by a *pensée unique* represented by BLUP and REML. Nevertheless, Lindley & Smith (139) provided a link between Henderson's mixed models and hierarchical Bayesian approaches, and Box & Tiao (140) gave technical details useful for putting Bayesianism into practice.

The Bayesian method assigns a prior distribution to all unknowns, including the model, the link function in generalized linear models, and the covariance matrices. The prior is combined with data to obtain a revised state of knowledge conveyed by a posterior, typically multidimensional distribution. Marginal and predictive distributions and counterparts of the classical hypothesis tests are based on posterior probabilities. Estimation, forecasting, and model assessment are all based on a single formula, and results are always conveyed in terms of probability, facilitating interpretation. Multidimensional integration is needed to obtain exact results, which constrained use of fully Bayesian methods until sampling algorithms came into the picture and samples could be obtained without knowledge of the posterior distributions. The contentious point is that prior distributions are elicited arbitrarily, especially in problems with many parameters. The flexibility, elegance, and power of the Bayesian construct has its Achilles' heel at the prior, and it suffices to examine literature on genomic selection to see how a battery of models can be produced via choice of more or less arbitrary priors and hyper-priors (coming out of the blue) constructed from somewhat naïve expectations about genetic architecture (5).

Interest in a Bayesian approach to animal breeding perhaps started as a consequence of a seminar given at Cornell by the statistician Daniel Solomon on how the selection index could be viewed as a Bayesian procedure. Henderson (personal communication) stated that this presentation spurred Rönningen (141) to investigate connections between BLUP and Bayesian ideas. Dempfle (109) pursued the issue further, showing that BLUP was a matrix-weighted average between the least-squares estimator and the mean vector of a prior distribution, representing

knowledge about the distribution of genetic effects in a population, along the lines of Robertson (108). Later, Gianola & Fernando (110) suggested the Bayesian approach as a general inferential method for solving many animal breeding problems, linear or nonlinear, even when there was uncertainty about genetic parameters.

Early applications in animal breeding used Gaussian approximations to joint or partially marginalized posteriors, because of technical difficulties in carrying out the needed integrations. With the advent of Markov chain Monte Carlo (MCMC) sampling methods, the power and flexibility of the Bayesian approach could be exploited in full. The most popular MCMC method has been the Gibbs sampler, although it can be used only under certain conditions (142). Gibbs sampling was used first in quantitative genetics by Guo & Thompson (143) and then by Wang et al. (144) in animal breeding. Many papers using MCMC have been published thereafter, especially in genomic selection, a topic discussed later. Use of Bayesian measures for assessing genetic trends in designed experiments was a useful development (145, 146). Trend assessment is difficult in animal breeding, and the likelihood-frequentist tandem approach gives only an approximate answer, even under normality assumptions. Results depend critically on input genetic parameters, as Thompson (147) demonstrated. Further, it is not trivial to obtain standard errors of the estimated trend because selection alters the distribution of phenotypes and genetic values. The Bayesian method estimates posterior distributions of measures of genetic change, given conditions that permit ignoring selection. Sorensen et al. (148) proposed a method for monitoring the evolution of additive genetic variance in the course of selection. Bayesian methods have been used subsequently in many areas of genetics, such as gene mapping, QTL detection, population differentiation, phylogeny analysis, sequence alignment, and genomic selection in animal and plant breeding. Animal breeders were pioneers, but the road was not free of stones.

NONLINEAR MODELS, SURVIVAL ANALYSIS, AND LONGITUDINAL DATA

Principles of mixed linear model methodology were established gradually, but a linear model, albeit useful, is not always a sensible statistical specification, especially for traits that are truncated or subject to censoring. Examples are discrete variables (all or none), commonly used to score fertility and disease traits, and productive life span or survival time. Hence, challenges remained, and these were tackled by building upon the foundations set by the mixed model methodology. In the 1980s and 1990s, animal breeders became more aware of statistical research literature, and better training in mathematical statistics enabled them to go beyond least-squares and BLUP. The next challenge was coping with nonlinearity.

Categorical and Counted Response Variables

Close to the end of the twentieth century, animal breeders were using linear models for discrete variables, even if this produced concern among statisticians and some geneticists. In fact, these models are still in use today. Wright (149, 150) had already developed the idea of a latent scale on which allelic substitutions take place and introduced the probit, an inverse probability transformation for dichotomous data, and the threshold model for ordered categorical responses, such as the number of digits in crosses of guinea pigs. Dempster & Lerner (151) had shown that a linear model analysis of 0–1 data was plagued by problems, such as frequency-dependent estimates, leading to incorrect interpretation of genetic parameters. Falconer (152) developed a simple and elegant method for inferring heritability on such a latent scale. Thompson (153, p. 350) stated, "I have some unease at using linear models for these dichotomous traits" and suggested an intuitively appealing approach to mixed model prediction with binary data, but without formal justification.

Although ML estimation for parameter estimation already had been embedded in the framework of generalized linear models (154), it was not clear how mixed linear models could be paralleled when encountering 0–1 or categorically scored variables.

In the 1980s, three papers presented a solution for the threshold model but without offering a theoretically clean way of estimating genetic parameters concomitantly, as in a BLUP-REML analysis. Harville & Mee (125) and Gianola & Foulley (155) addressed inferences about fixed and random effects in mixed models for binary and ordered categorical responses. The two methods give the same answer when predicting breeding values and estimating fixed effects and yield BLUP when the data are Gaussian, rather than discrete. Their technique was similar to that used by Henderson et al. (117) in a derivation of BLUP, although cast in a Bayesian or semi-Bayesian framework. For ordered categorical data, it was assumed that there was a latent Gaussian variate called liability that followed a mixed effects linear model. If liability falls between two consecutive thresholds, this is equivalent to an observation in the corresponding category of response. Gilmour et al. (126) proposed a different method for binomial data based on ideas from generalized linear models. Here, fixed effects were estimated after genetic effects were integrated out, but assuming a known genetic variance in the liability scale; the genetic variance and genetic effects were inferred by using an analogy with the mixed linear model. This method also yields BLUP when the model is linear and the variance components are known and requires iteration with a reweighted set of MME, as the estimating equations are not explicit. Procedures (125, 126) produce REML when the data are Gaussian. Sorensen et al. (156) provided a more general solution in a Bayesian treatment of ordered polychotomies based on Gibbs sampling. Many countries now use the threshold model for routine genetic evaluation, following this foundational research plus considerable software development. Misztal et al. (157) reported one of the early programs available.

Extensions were for models with Gaussian and categorical responses (158), for multivariate binary responses (159), and for models where a categorical response variable (e.g., survival) depended on a count (e.g., litter size) having a Poisson distribution (160). Tempelman & Gianola (161, 162) presented a nonlinear mixed model implementation for count data that made use of Laplacian integration to approximate the marginal posterior distribution of the genetic variance. Historically, the preceding work indicated that animal breeding had transcended the stricture of linearity imposed by BLUP, although a linear approach is often satisfactory enough for ranking candidates for selection.

Survival Analysis

Survival analysis has obvious importance in the medical sciences but did not receive much attention in animal breeding until the 1980s, when focus begun to be placed on functional traits. One such trait is the length of productive life span of an animal (163, 164). Famula (165) presented an exponential survival model with covariates, and the next step was to adapt proportional hazard models employed in biostatistics and engineering to the needs of quantitative genetics. The hazard function is the probability of instantaneous death at time t, given that the individual has survived to time t, and the key here is modeling hazard in a way that allows for correlations between relatives. A common feature of life-span data is the presence of censored observations, and this creates difficulties. For example, it may be known that a certain cow was present in a herd at some time but that she was sold to another herd thereafter for production purposes, without information about the date of termination of her career. It would be incorrect to use the date of sale as termination time. Another difficulty is that animals encounter different culling risks at various stages of their career, so time-changing covariates are needed. Empirical Bayes approaches were used in this

context for inferring breeding values (164), but fully Bayesian analyses are now available (166), although they are computationally more taxing. A contentious area has been that of defining heritability in a survival model. Contrary to the threshold model, in which there is an explicit Gaussian scale on which the mixed effects model holds, the survival model requires a logarithmic scale. Genetic parameters can be derived for such scale, but it is not obvious how these translate into heritability of survival time. However, the concept of heritability, holding for a linear scale where a linear decomposition of variance is available, is not always necessary. If significance tests suggest the existence of additive variability, it is to be expected that selection could modify the population (Fisher's fundamental theorem), irrespective of whether heritability is low or high. The difficulty is developing a selection criterion with reasonable accuracy and precision, but the survival models do provide the counterpart of BLUP. One widely used software program for this purpose has been the Survival Kit (167), which has been updated since its inception.

Linear and Nonlinear Models for Longitudinal Data

Treatment of repeated measurements has always been of interest in animal breeding, and interest on repeatability (the correlation between successive records of an animal) dates back to the 1940s and 1950s (168). However, repeatability does not adequately address the situation in which successive records follow a time series with some trajectory, as in a lactation or growth curve (68, 169). As a consequence of more intensive recording systems (for instance, it is now possible to monitor instantaneous milk flow in dairy cattle), there was a need for more refined statistical methods for longitudinal mixed effects models. Linear (in the parameters, but not necessarily in time) random regression models and similar approaches began to be developed in animal breeding, and a large body of literature on analysis of test-day yields in dairy cattle emerged. Similar applications have been made in meat-producing species. Main instigators were papers by Schaeffer & Dekkers (170) and Kirkpatrick & Lofsvold (171), the latter of which was more mathematically sophisticated but shared a similar spirit.

Briefly, this type of treatment of longitudinal data is as follows. In a randomly drawn sample, each individual is measured longitudinally. For example, male and females from several breeds are weighed at several phases of their development, from near birth to the adult stage. An objective may be to study growth patterns in each of the breeds, while taking into account interindividual variability. Typically, the number of measurements per individual varies, leading to longitudinal unbalancedness. A hierarchical or multistage model assigns a series of nested functional specifications, together with distributional assumptions. At the first stage of the model, a mathematical function (linear or nonlinear) describes the expected trajectory of individuals, and a residual having some distribution reflects departure of the observations from such a trajectory. At the second stage, a submodel describes interindividual variation of parameters of the first-stage specification. A second-stage residual reflects the inability of the submodel to explain completely the variation of the parameters. Additional stages can be imposed in a Bayesian context to describe uncertainty about all parameters. A curve is obtained for each combination of fixed effects of interest and for each individual, and the random treatment of coefficients generates a covariance function driven entirely by time. Time-dependent heritability and genetic correlations can be obtained. However, these have doubtful biological meaning because genetic variation and covariation pertain to the regression parameters, and it is at that level that gene substitutions are postulated to take place. In this type of model, gene effects remain the same over time, and the resulting correlograms and heritability-grams are driven solely by modification of the time variable. Often, the random regression models produce strange results at the end of the time domain.

Meyer (172) described a REML implementation of covariance functions, but only for models that are linear on parameters. Bayesian analyses and some work on semiparametric methods using splines were also done, and many papers were published. Much research was done on random regression models with hundreds of variations of the theme, each offering little additional advantage. Curiously, as it happened during the progression of BLUP, little or no use was made of cross-validation. Researchers continued believing that bigger was better, with a main objective being making massive computations feasible. Many countries eventually adopted the test-day model (jargon for a linear longitudinal mixed model with random coefficients) for dairy cattle, save for the United States owing to constraints posed by a patent held by Cornell.

Use of Robust Distributions

There has been work in fitting thick-tailed instead of normal distributions (173–177) by using pedigreed or genomic data. These researchers addressed how univariate and multitrait mixed effects linear models could be extended to accommodate t-distributions, to attain a more robust analysis in the sense of reducing influences from outlying data. Using Bayesian measures for strength of evidence, these studies found that thick-tailed residual distributions produced more plausible models than when normality was assumed. Gaussian assumptions might be dangerous for entire probabilistic inference, as opposed to just seeking a point prediction of breeding value. For example, calculating the probability of correctly ordering breeding values that are neither independent nor identically distributed, given some data, is an old problem of interest in animal breeding (9). Reber et al. (178) applied this idea to sire rankings, though they employed Gaussian assumptions and Bayesian MCMC. A thick-tailed distribution may deliver a ranking of breeding values (e.g., based on the posterior mean) similar to BLUP but with different probabilities that a particular animal will truly be best, given the ranking of the posterior means.

Mixture Models

Application of finite mixture models to genetics dates back to Pearson (23). These models can uncover heterogeneity owing to hidden structure created by, for instance, unknown loci with major effects. Often, this heterogeneity can be resolved by fitting a mixture, producing, as a by-product, conditional probabilities that a datum is drawn from one of the several putative, but unknown, genotypes. Concealed heterogeneity produces curious phenomena: The offspring-parent regression depends on the mixing proportions, and the genetic correlation between a mixture trait and a Gaussian character is a function of the mixing proportions and of the ratio of genetic variances between mixture components (179). Ignoring the heterogeneity can produce misleading interpretations and unrealistic expectations about response to selection when applied to a heterogeneous trait. Hidden population structure must be accounted for when using whole-genome prediction, and this is a typical mixture model problem.

Many QTL detection procedures have been based on ideas from mixture models, and inference about breeding values using finite mixture models may be warranted in some cases. For example, mastitis is an inflammation of the mammary gland of cows associated with bacterial infection. Recording of mastitis events is not routine in many nations, so SCC has been used as a proxy in genetic evaluation of sires because an elevated SCC is often an indication of disease. SCC on healthy and diseased animals displays different but overlapping distributions, which are hidden in the absence of disease recording. Finite mixture models were applied in this context (180–183). When dealing with counts (e.g., the number of episodes of a disease), the number of observed zeros is often larger than what would be expected under some distribution, such as a Poisson distribution.

Zero-inflated Poisson mixture models can be useful, and Rodrigues-Motta et al. (184) implemented a fully Bayesian zero-inflated Poisson analysis (via MCMC) for the number of mastitis episodes in dairy cattle. Meuwissen et al. (41) also suggested mixture models for whole-genome prediction.

Although finite mixtures can approximate any distribution (e.g., the most widely used method of density estimation is based on mixing N Gaussian distributions, where N is sample size), application is not without its pitfalls. Algorithms often do not converge unless all parameters are identified in the likelihood, producing spurious results because of a phenomenon called label switching, in which the algorithm does not recognize the label of the underlying true component of the mixture. Celeux et al. (185) have warned about this problem, but their advice is often ignored.

Computing Software

Because of the sheer size of animal breeding data sets, much effort has been devoted to making BLUP, REML, and Bayesian methods computationally feasible, even in multivariate models. Examples of widely used packages for mixed effects linear models are available (186–191). Some software for nonlinear models, survival analysis, and limited dependent variables is available, but it is not general. One example is the already mentioned Survival Kit for survival models (167). Mistzal and collaborators (**http://nce.ads.uga.edu/wiki/doku.php**) and Fernando & Garrick (**http://www.biomedcentral.com/content/supplementary/1471-2105-12-186-S1.PDF**) developed programs for large-scale computations using genomic data. Animal breeders are increasingly using R, a free software environment for statistical computing and graphics (**http://www.r-project.org**). For example, Bates & Vazquez's package Pedigreemm (**http://cran.r-project.org/web/packages/pedigreemm/pedigreemm.pdf**) uses R for mixed model analysis, and de los Campos and Perez's package BGLR (**http://bglr.r-forge.r-project.org/BGLR-tutorial.pdf**) uses it to implement genomic BLUP and other Bayesian regression models discussed in the genomic selection section of this paper.

BIASES FROM SELECTION PROCESSES

Animal breeding data seldom arise from a truly random mechanism. Except in designed experiments, the history of the selection process is known incompletely, because field records are used that contain missing data in some statistical sense, as also happens with retrospective data in medical studies. The question of how selection and assortative mating affect estimation of genetic parameters and prediction of breeding values is an important one. Here, Henderson et al. (10, 117) and Curnow (192) made significant contributions.

Kempthorne & von Krosigk, in a section of Henderson et al. (117), and later in Curnow (192), assumed normality and a certain form of sequential selection. They found that the ML estimator of a parameter had the same form with and without selection, provided that all data used for selection decisions had been used in the analysis. Im et al. (193) found this result to hold for any distribution and more general forms of selection. This does not imply that the asymptotic distribution of the ML estimator is unaffected by selection, as one would need to take expectations under the unknown distribution of the observation's given selection, rather than under random sampling. Hence, selection is not completely ignorable if standard errors are sought.

In perhaps the most influential paper on this matter, Henderson (10) assumed known genetic parameters and multivariate normality and derived BLUP of breeding value under a specific selection model. He used a formula from Pearson (24) that forces incidence and kinship matrices to remain constant in conceptual replication. This selection model was received uncritically by

animal breeders, save by Thompson (153). He was the first to point out that a crucial matrix (called **L** by Henderson) must remain constant from repetition to repetition for Henderson's result to hold. This is unrealistic, as selection decisions (at least when generations overlap) are made between and within generations, between and within years, and across and within families, leading to a sampling scenario where **L** changes at random, a situation that is not well represented by Henderson's machinery. The 1975 paper gave conditions for unbiasedness as well as remedies that have been largely followed by the animal breeding community. One of these, for example, states that if selection is based on linear functions of the (unobservable) breeding values, some random elements in the model (e.g., herds) must be treated as fixed to obtain unbiased predictors of breeding values. However, if constructing the linear functions on which selection is based requires knowledge of breeding values, there would be no point in predicting anything. This specific setting does not describe any type of selection encountered in practice and led to the perhaps unfortunate and widespread practice of treating contemporary groups as fixed. This can be criticized for at least two reasons. One is that a fixed treatment of contemporary groups burns away information, because thousands of degrees of freedom are consumed in estimating levels with little information each (e.g., a herd-year-season class in Finland, where herds are typically small), leading to unnecessarily large variances of estimates. The second one follows from the James & Stein (138) result: Treating a vector with a large number of effects as fixed leads to estimates with unnecessarily large mean squared error. Henderson (10) represents a courageous attempt at tackling unbiased prediction of random effects under selection, but his approach has its shortcomings.

Im et al. (193) addressed the selection problem by using a likelihood-based framework, and Gianola & Fernando (110) employed a Bayesian approach. It was shown that, if all data on which selection is based are used for constructing a likelihood or a posterior distribution, selection can be ignored for point inference (likelihood) and more generally for posterior inference (Bayesian). Im et al. (193) used a vector **r** that contains indicator variables denoting whether a record is present or absent (owing to culling) and that could be included as part of the data if one had observed the entire history of the selection process; however, **r** is seldom known. They showed that if (*a*) the conditional distribution of **r** given the observed and missing records does not depend on the latter, and (*b*) the parameters of the distribution of **r** are separable (distinct) from those of the data distribution, then selection can be ignored for likelihood inference. Sorensen et al. (148) adapted this idea to the problem of inferring genetic variance in the course of selection (in the context of a structured experiment) and presented a detailed proof for the Bayesian case. Altogether, these results partially support the claim that a multiple-trait analysis palliates bias, because it often conveys additional (to that from a univariate analysis) information on history of the selection process.

Unfortunately, selection is not always ignorable. For example, a genome-based analysis of carcass traits in a sample of beef cattle ignoring some preselection for growth rate would lead to incorrect inference. In these instances, it is crucial to attempt to model the missing data or selection process or, alternatively, use more robust methods of inference. Results would depend crucially on the missing data process assumed, however.

THE ERA OF GENOMIC SELECTION

In stylized formulae for expected response to selection, the rate of genetic progress is directly proportional to the square root of the additive genetic variance, precision, and intensity of selection and inversely proportional to generation interval, and these factors are interdependent. Early predictors of breeding value had been sought for years, and BLUP makes use of all available information on relatives, so it is possible to obtain a prediction of breeding value of animals

irrespective of their age. The use of DNA MAS had also been of interest for some time, as stated earlier. Blasco & Toro (17) gave an account of the progression (or lack thereof) of MAS and of the associated quest for QTL, an area where results have been below expectation. During the MAS era, some statistical methods were developed, such as scans for QTL detection (75, 194, 195), regression on a limited number of markers (196), and BLUP for MAS (74). One limitation, however, was the lack of a sufficient number of markers spanning the entire genome. As a by-product of projects attempting to sequence the genome of several species, a massive number of biallelic markers emerged thereafter: the SNPs.

In a paper that revolutionized animal and plant breeding, Meuwissen et al. (41) proposed a relatively simple idea: Given a battery of p SNP and a sample of n individuals genotyped for such markers, fit a multiple linear regression on the number of copies of a reference allele at each of the p loci. Because $p>>n$ (this situation will become $p>>>n$ when individual genome sequence data become available, as p will increase and n will decrease, at least at the beginning), the marker incidence matrix $\mathbf{X}$, of order $n \times p$, will have rank n, at most, leading to at least $p - n$ regressions that are not identified in the likelihood. The solution was to introduce restrictions on the size of the coefficients or to use some random effects or Bayesian model, which produces shrinkage of regressions. There is an issue here of how much effective learning from data takes place about individual regression coefficients, but this is not important from a predictive perspective (5). Meuwissen et al. (41) recognized that a BLUP procedure applied to marker effects provided an answer to this type of problem and also suggested two methods, later known as Bayes A and Bayes B, following a Bayesian route via MCMC. Bayes A assigns a t prior distribution to the unknown marker effects, and Bayes B postulates a mixture of a zero-state and a t-distribution as prior, although the original formulation of Bayes B had a different spirit (197). By dividing the data into training (model fitting) and testing (prediction) sets, one can obtain estimates of marker effects or of genetic value in the training set to predict phenotypes in the testing set. To the extent that these predictions could be obtained earlier in life, and perhaps with more accuracy and precision than with pedigree-based BLUP, the rate of genetic progress attained could be accelerated (198). Whether or not more cost-effective genetic progress is attained varies with species (17), but animal breeding industries embraced the concept with enthusiasm, and early results in dairy cattle breeding have been encouraging (199, 200). However, beliefs that this sort of approach would be a panacea for lowly heritable traits, where precision of selection is low, have not yet been corroborated. Perhaps this was a naïve expectation: Low heritability implies a low signal/noise radio. A metaphor might be that even the most sophisticated computers cannot cope with a weak Internet signal.

At the onset, it was expected that a battery of markers (now at 800,000 in the most advanced livestock chips) would capture relevant LD relationships between alleles at markers and at the elusive QTL. However, Habier et al. (201) found that the better predictive ability of genome-enabled regressions was perhaps because markers provided a better representation of genetic relatedness than a pedigree can. Observation of molecular similarity allowed for differentiation of genetic relatedness in, say, a group of full-sibs. Using pedigree, all full-sibs have the same expected relatedness, but the realized relatedness varies. This led Van Raden (202) to suggest that $\boldsymbol{A}$ in BLUP could be replaced by $\boldsymbol{G}$, a marker-based matrix, thus adding extra resolution to the machine. The suggestion was influential, and BLUP evolved into genomic BLUP (G-BLUP), which rapidly became a standard method for genetic evaluation of livestock using SNPs. There is an incorrect perception that, given $\boldsymbol{G}$, the matrix $\boldsymbol{A}$ is redundant. This does not seem to be so for at least two reasons. First, pedigree and genomes can be viewed as different inputs into a predictive machine. Second, $\boldsymbol{G}$ has $\boldsymbol{A}$ as expected value only if, for example, there is no selection or HW equilibrium holds. It is unclear what the best possible manner of estimating molecular similarity via

a relationship matrix is. A crucial point is that the genetic relatedness that matters is that induced by similarity at the level of the QTL, but markers are not QTL so there is a disconnectedness that remains to be closed. Perhaps sequence information will help, but it will introduce additional problems and challenges. It was also recognized (203) that G-BLUP and a BLUP on markers with an assignment of a common variance to a normal distribution of random effects were equivalent. Hence, and for the purpose of obtaining a prediction of the molecularly marked additive genetic value of an animal, G-BLUP sufficed.

Bayes A and B of Meuwissen et al. (41) became the Adam and Eve of Bayesian genome-enabled prediction, and a large number of Bayesian linear regression methods emerged thereafter. Examples are the Bayesian Lasso (204), Bayes C (205), and Bayes R (206), in which R stands for genomic regions, to name just a few members of an extensive list nicknamed the Bayesian alphabet (197). These methods were reviewed in Reference 207 and share the same regression model but differ in the assumptions on the prior distribution of marker effects. The end result is that the latter are shrunk with different degrees of severity. For instance, Bayes B and the Bayesian Lasso shrink more than BLUP of marker effects. Most methods present trivial differences in terms of predictive ability unless there is a major gene segregating, in which case procedures based on mixtures, such as Bayes B or perhaps Bayes R, may perform better (208). Other methods, as opposed to assigning independent prior distribution to marker effects, have incorporated correlations between such effects, presumably reflective of some LD structures having an impact on predictions (209–210). In absence of major genes, most methods deliver more or less the same predictive capacity, as shown vividly by Wimmer et al. (211). Using simulation and plant data and a measure of estimation loss function, these authors found that G-BLUP was quite robust, except when there were some genomic regions with strong effects. Our opinion is that a QTL (in a loose sense) detected by using statistical scans seldom graduates into a causal region. A more fruitful approach seems to be that of candidate genes, i.e., a search guided by knowledge of molecular genetics and pathways, an approach of great promise in oncogenomics, although not devoid of assumptions. Biochemical considerations may be in error, but the same applies to any prior distribution. Conversely, given knowledge of a QTL, a good statistical method will likely pick it up. The $p>>n$ condition gives the prior a strong impact on the posterior, and the data modifies the prior state of knowledge little. This was shown theoretically by Gianola (5) and illustrated by Lehermeier at al. (212) with plant data.

Another consequence of Meuwissen et al. (41) was introducing cross-validation as a routine form of calibrating predictive performance. The belief that bigger was better (more traits and parameters in a model) led animal breeders to think somewhat uncritically that this extra baggage gives, necessarily, better predictions. However, it is well established in prediction that this may not be so. A highly parameterized model, even when based on mechanistic considerations (far from being the case for the multiple linear regression methods used by breeders) fitted to a sample of a finite size may produce overfit, hampering predictive performance. Also, a cross-validation distribution reflects all sources of uncertainty, such as errors of specification, errors in parameters, structural differences between training and testing sets, and environmental variance fluctuations. Hence, cross-validation confidence bands are typically much wider than differences in predictive ability between methods (213, 214). Many animal breeders do the cross-validation only once, but this is equivalent to estimating a parameter without producing a standard error. Discussion on differences between methods is often about noise, because cross-validation variability is much larger than such differences.

Goddard (203) and Daetwyler et al. (215) developed formulae for assessing the precision of genomic selection (the term accuracy is misleading) for a G-BLUP model. Here, an important factor is the nominal size of the training set, although a large training sample of size N may span

little genetic variation, so some genotype configurations may not be found in the testing set. However, another set of size N with larger genetic variability may perform better, so the degree of molecular redundancy is an important factor. Goddard (203) attempted to account for this, but his formulae require assumptions about effective population size, a parameter notoriously difficult to estimate well.

A question of practical importance concerning genomic selection for additive effects remained to be answered: Because only a few (and typically elite) animals are genotyped, what do we do with information (pedigree or phenotypic) on nongenotyped individuals? An international group including D. Johnson (New Zealand), A. Legarra (France), I. Aguilar (Uruguay), I. Misztal and T. Tsuruta (United States), and others has proposed a solution, illustrating how animal breeding research has become global. The method is known as single-step BLUP; a representative paper can be found in Reference 216. The perhaps ominously named SS-BLUP has heuristic components, but it represents a valuable attempt at integrating available genotypic, pedigree, and phenotypic information and can be implemented nicely using available BLUP machinery.

The genome-enabled prediction models described above are, in some sense, a finite number of loci counterparts of the infinitesimal specification in Reference 30 but do not accommodate nonadditive genetic variance. Dekkers & Hospital (217) pointed out limitations of genome-wide association studies based on additive assumptions, and some of their observations carry to prediction as well. A challenge is that of positing a functional form relating phenotypes to SNP genotypes (hundreds of thousands or millions of possible configurations) while allowing for interaction. Explicit modeling of interactions produces a construct that requires intensive computing (technological constraint) and has excessive complexity, as the $n < p$ problem is exacerbated further and regression coefficients on epistatic effects turn out to be nearly zero, owing to severe shrinkage. Gianola's (5) warning concerning lack of identifiability applies even more strongly here. However, the genome is more interactive than what standard quantitative genetic analysis indicates, often ending up with the genetic variance being mostly additive (83). For instance, References 84–86 give examples of extensive epistatic interactions. In theory, genetic variance can be partitioned into orthogonal additive, dominance, additive $\times$ additive, additive $\times$ dominance, and dominance $\times$ dominance components, only under highly idealized conditions, as mentioned earlier. These include no linkage, but MAS is supposed to exploit LD, and even chance creates disequilibrium. Hence, theory breaks down.

Evidence from molecular biology on the importance of gene networks affecting pathways, plus a lack of good theory, suggested that a nonparametric treatment of the data could be valuable, as these methods are suitable for complex problems (54). Reproducing kernel Hilbert spaces regression (RKHS) and neural networks (218–221) have been suggested as contenders to the Bayesian alphabet that are capable of exploiting nonadditivity. RKHS uses notions of genetic distance and of similarity between individuals and searches for a function with optimal predictive ability in a rich space of unknown functions. Neural networks are based on their mathematical property of being universal approximators of functions. Although there have been fairly extensive comparisons among members of the Bayesian alphabet (e.g., Reference 212), similar studies involving RKHS are lacking, especially with animals. González-Recio et al. (213, 222) found a slightly better predictive ability of RKHS over parametric methods for early mortality and feed efficiency in broilers, but differences were within the range of the noise stemming from the cross-validation distribution. Heslot et al. (208) compared many prediction methods, including ridge-regression BLUP, Bayes C-pi (another member of the alphabet that uses a mixture model with unknown mixing probabilities as prior of marker effects), and RKHS (neural networks and support vector machines were included as well), using 18 plant breeding data sets. On average, most methods produced the same predictive correlations; however, by using figures in Reference

208, if a scatter plot is made for the 18 pairs of predictive correlations, RKHS can be found to be better than G-BLUP or Bayes C-pi in 16 comparisons. This form of analysis suggests that some methods are consistently better for a specific prediction problem. In the absence of detailed knowledge of the underlying basis of a trait, explanations of why a prediction machine is better than other ones under a set of circumstances are largely based on conjectures.

Jarquín et al. (223) used a reaction norm model with a matrix of similarities among environments entering into the covariance structure. This method was applied to 139 wheat lines genotyped with 2,395 markers with 68 environmental conditions modeled. Genotype × environment interaction was fitted by constructing a Hadamard product matrix, essentially a RKHS representation. Predictive ability was much increased by accommodating the environmental and interaction inputs. Taken collectively, the studies above suggest that RKHS delivers at least as good a predictive performance as parametric methods, but neural networks can be very unstable, unless implemented via MCMC (221, 224). Hill (15) criticized these methods because an estimate of breeding value, which standard theory considers the focus of genetic improvement, is not produced. However, this is not so, as either a neural network or a RKHS can be configured such that a predicted breeding value emerges. In fact, BLUP and G-BLUP are particular cases of RKHS, and the latter can be tailored to capture breeding values per se, plus additional forms of genetic signal. Ornella et al. (225) found that these methods can better classify superior individuals at tails of the distribution. A study addressing the forward predictive abilities of varying methods is lacking. Is a linear regression on additive codes of markers better than RKHS one, two, or three generations ahead?

CONCLUSION

Our historical account indicates that animal breeders have taken up new statistical ideas rapidly and have also contributed to the field of biological statistics, significantly so in the cases of BLUP, REML, Bayesian methods, and whole-genome prediction. The main underlying theoretical foundation in the field is quantitative genetics, which is primarily a descriptive and predictive science, although perhaps not effective enough for discovery of genes, especially when compared with the astonishing record of molecular genetics. However, the explosive availability of genomic and postgenomic data has provided means and opportunities for refining and enhancing prediction of complex traits, an exciting area per se, but not too amenable to reductionist reasoning or experimentation. A tentative forecast of some future developments and issues is outlined below.

Soon, genome sequence information on individuals will be increasingly available (e.g., the 1,000 Bull Genomes Project), and the expectations are huge. Many authors argue that all causal mutations will be present in the sequence, and that this advantage will be exploited fruitfully. There are some caveats here. The first one is that such a view is based on a somewhat linear map of the genome, i.e., that a string of bases can produce an accurate genotype-phenotype mapping. The second one is that more information is better; for instance, instead of 800,000 markers, there will be 10 million. With respect to the first thought, the DNA-protein process is not linear because of, e.g., protein folding and pervasive interaction and feedbacks in metabolism and nonlinear enzyme kinetics. DNA and methylation information may be crucial for breeding value assessment, but appropriate environmental modeling (environmentomics) with supplementary omics-type information should also be considered for building more effective prediction machines. A study from human genetics (226) has indicated that integrating mRNA and microRNA expression data substantially increases predictive performance in the context of personalized medicine; this approach is a special case of RKHS. Concerning the second argument, one difficulty with an explosive increase in potential covariate number from sequence information is that the p/n ratio will

increase markedly. For instance, if 1,000 bulls are sequenced, the *p/n* ratio will easily surpass 1,000–2,000. The implication is that all linear regression coefficients will become minute because of strong regularization. One could still use the Bayesian alphabet (given a huge amount of computing resources) for prediction, but inference about genomic regions must be done with caution because priors will matter even more than with SNP data. Perhaps a deluge of genomic and postgenomic data will further complicate separation of signal from noise because of the temptation to overmodel. Anyhow, large *p/n* ratios confer a strong advantage to $n \times n$ methods such as G-BLUP or RKHS.

It does not seem sensible to expect that all quantitative traits in animal breeding will be described suitably by a linear model with Gaussian residuals. Given the continued growth in computer power and algorithms, there is flexibility for fitting more realistic error distributions, such as a t-distribution. Analysis of cross-validation residuals is also an important diagnosis tool, and use of bootstrap methods (227) will enable us to obtain realistic measures of candidate-specific cross-validation reliability. Animal breeders religiously use theory-derived measures of reliability that convey information content in training data but without regard to how accurate a predictive machine actually is. A very reliable predictor can have a bad cross-validation performance if it is inaccurate with respect to realized target phenotypes in a testing set.

Animal breeders should also be cautious about making overly strong assumptions concerning the dimension of a model. A multivariate analysis is not necessarily better unless some traits enter as part of the prediction machines as valuable covariates, and not through correlations. Selection and ascertainment bias should be an issue of concern in genome-enabled prediction, and sources of bias will need to be assessed more carefully.

The advent of genomic data will also enable us to study relationships between genomic regions using systems and causal perspectives, e.g., gene networks affecting pathways via graphical and structural equation models (228–231). Systems analysis, however, is not new in animal breeding (232), and it is not unfair to state that dividends from this approach have been scarce. The neo-systems view exploits much more refined data, but understanding the dynamics of a system requires knowledge of rate coefficients. To estimate the latter, experiments of a reasonable scale are needed, but these are virtually nonexistent with farm animals, and going from a yeast or a fly to a cow is a huge extrapolation.

There have been many contributors of statistical methods to animal breeding during the past century, so it has not been one of solitude. It would have been neither possible nor entertaining to honor every contribution or to provide a comprehensive bibliography. Any narrative of history depends on the narrators' experiences and perspectives, thus introducing unavoidable subjectivity. We attempted to produce an involved narrative, offering opinions that may not be shared by many animal breeders, as well as introducing personal biases. An alternative recommended narrative, focusing on genetics issues, is that of Hill (15). We plagiarize the last statement of his paper by apologizing for factual errors, misrepresentations, and omissions, but opinions we can debate.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Research was partially funded by a US Department of Agriculture Hatch Grant (142-PRJ63CV) to D.G. and by the Wisconsin Agriculture Experiment Station. The paper was completed at the

Institute of Advanced Studies of the Technical University of Munich (TUM), where D.G. was supported by a Hans Fischer Senior Fellowship. Professor Chris-Carolin Schön, Chair of Plant Breeding at TUM-Weihenstephan, and her group are gratefully acknowledged for discussions and for providing a pleasant working environment. Professor W.G. Hill (University of Edinburgh) contributed many useful and insightful comments, and Professor Larry Schaeffer (University of Guelph) offered constructive suggestions. The funders had no role in study design, data collection and analysis, the decision to publish, or preparation of the manuscript.

LITERATURE CITED

1. Fernández J, Toro MA. 1999. The use of mathematical programming to control inbreeding in selection schemes. *J. Anim. Breed. Genet.* 116:447–66
2. Weigel KA, Lin SW. 2000. Use of computerized mate selection programs to control inbreeding of Holstein and Jersey cattle in the next generation. *J. Dairy Sci.* 83:822–28
3. Brown PO, Botstein D. 1999. Exploring the new world of the genome with DNA microarrays. *Nat. Genet.* 21:33–37
4. Geisser S. 1993. *Predictive Inference: An Introduction*. New York: Chapman & Hall
5. Gianola D. 2013. Priors in whole-genome regression: the Bayesian alphabet returns. *Genetics* 194:573–96
6. Henderson CR. 1948. *Estimation of general, specific and maternal combining ability in crosses among inbred lines of swine.* PhD Thesis, Iowa State Univ.
7. Henderson CR. 1953. Estimation of variance and covariance components. *Biometrics* 9:226–52
8. Henderson CR. 1963. Selection index and expected genetic advance. In *Statistical Genetics and Plant Breeding*, ed. WD Hanson, HF Robinson, pp. 141–63. Publ. 992. Washington, DC: Natl. Acad. Sci. Natl. Res. Counc.
9. Henderson CR. 1973. Sire evaluation and genetic trends. In *Proceedings of the Animal Breeding and Genetics Symposium in Honor of Dr. Jay L. Lush*, pp. 10–41, Champaign: Am. Soc. Anim. Sci., Am. Dairy Sci. Assoc.
10. Henderson CR. 1975. Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31:423–49
11. Henderson CR. 1984. *Application of Linear Models in Animal Breeding*. Guelph, Can.: Univ. Guelph
12. Freeman AE. 1973. Genetic statistics in animal breeding. *J. Anim. Sci.* 1973:1–9
13. Gianola D, Hammond K, eds. 1990. *Advances in Statistical Methods for Genetic Improvement of Livestock*. Berlin: Springer-Verlag
14. Gianola D. 2006. Statistics in animal breeding: angels and demons. *Proc. 8th World Congr. Genet. Appl. Livest. Prod.*, CD Paper 00-03. Belo Horizonte, Brazil: Inst. Prociencia
15. Hill WG. 2014. Applications of population genetics to animal breeding, from Wright, Fisher and Lush to genomic prediction. *Genetics* 196:1–16
16. Rosa GJM. 2013. Foundations of animal breeding. In *Sustainable Food Production*, ed. P Christou, R Savin, B Costa-Pierce, I Misztal, B Whitelaw, pp. 58–78. New York: Springer
17. Blasco A, Toro MA. 2014. A short critical history of the application of genomics to animal breeding. *Livest. Sci.* 166:4–9
18. Galton F. 1886. Regression towards mediocrity in hereditary stature. *J. Anthropol. Inst. G.B. Irel.* 15:246–63
19. Lush JL. 1945. *Animal Breeding Plans*. Ames, IA: Coll. Press. 3rd ed.
20. Hill WG. 1971. Design and efficiency of selection experiments for estimating genetic parameters. *Biometrics* 27:293–311
21. Hill WG. 1972. Estimation of realised heritabilities from selection experiments. I. Divergent selection. *Biometrics* 28:747–65
22. Wachsmuth AW, Wilkinson L, Dallal GE. 2003. Galton's bend: a previously undiscovered nonlinearity in Galton's family stature regression data. *Am. Stat.* 57:190–92

23. Pearson K. 1894. Contributions to the mathematical theory of evolution. *Philos. Trans. R. Soc. A* 185:71–110
24. Pearson K. 1903. Mathematical contributions to the theory of evolution. XI. On the influence of natural selection on the variability and correlation of organs. *Philos. Trans. R. Soc. A* 200:1–66
25. Bulmer MG. 1971. The effect of selection on genetic variability. *Am. Nat.* 105:201–11
26. Bulmer MG. 1980. *The Mathematical Theory of Quantitative Genetics*. Oxford: Oxford Univ. Press
27. Falconer DS, Mackay TFC. 1996. *Introduction to Quantitative Genetics*. Essex, UK: Longman
28. Onaga L. 2010. Toyama Kametaro and Vernon Kellogg: silkworm inheritance experiments in Japan, Siam, and the United States, 1900–1912. *J. Hist. Biol.* 43:215–64
29. Yule GO. 1902. Mendel's laws and their probable relations to intra-racial heredity. *New Phytol.* 9:193–238
30. Fisher RA. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinb.* 52:399–433
31. Wright S. 1922. Coefficients of inbreeding and relationships. *Am. Nat.* 56:330–38
32. Wright S. 1931. Evolution in Mendelian populations. *Genetics* 16:97–159
33. Crow JF, Kimura M. 1970. *An Introduction to Population Genetics Theory*. New York: Harper & Row
34. Quaas RL, Pollak EJ. 1980. Mixed model methodology for farm and ranch beef cattle testing programs. *J. Anim. Sci.* 51:1277–87
35. Malécot G. 1948. *Les Mathématiques de l'hérédité*. Paris: Masson et Cie
36. Fisher RA. 1930. *The Genetical Theory of Natural Selection*. Oxford: Clarendon
37. Falconer DS. 1960. *Introduction to Quantitative Genetics*. New York: Ronald
38. Edwards AWF. 1977. *Foundations of Mathematical Genetics*. Cambridge: Cambridge Univ. Press
39. Gallais A. 1974. Covariance between arbitrary relatives with linkage and epistasis in the case of linkage disequilibrium. *Biometrics* 30:429–46
40. Gianola D, Hospital F, Verrier E. 2013. On the contribution of an additive locus to genetic variance when inheritance is multifactorial with implications on the interpretation of GWAS. *Theor. Appl. Genet.* 6:1457–72
41. Meuwissen THE, Hayes BJ, Goddard ME. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819–29
42. Lush JL. 1948. *The Genetics of Populations*. Mimeo. Ames: Iowa State Univ. 381 pp.
43. Neiman-Sorensen A, Robertson A. 1961. The association between blood groups and several production characters in three Danish cattle breeds. *Acta Agric. Scand.* 11:163–96
44. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, et al. 2009. Finding the missing heritability of complex diseases. *Nature* 461:747–53
45. Robertson A, ed. 1980. *Selection Experiments in Laboratory and Domestic Animals*. Oxfordshire, UK: Commonw. Agric. Bur., Farnham R.
46. Robertson A. 1960. A theory of limits in artificial selection. *Proc. R. Soc. Lond. B Biol. Sci.* 153:234–79
47. Hill WG, Robertson A. 1966. The effect of linkage on limits to artificial selection. *Genet. Res.* 8:269–94
48. Hill WG. 1974. Heritabilities: estimation problems and the present state of information. *Proc. First World Congr. Genet. Appl. Livest. Prod.*, I:343–51. Madrid: Graficas Orbe
49. Sorensen DA, Kennedy BW. 1986. Analysis of selection experiments using mixed model methodology. *J. Anim. Sci.* 63:245–58
50. Comstock RE, Robinson HF, Harvey PH. 1949. A breeding procedure designed to make maximum use of both general and specific combining ability. *Agron. J.* 41:360–67
51. Schaeffer LR. 1985. Model for international evaluation of dairy sires. *Livest. Prod. Sci.* 12:105–15
52. Wahba G. 1990. *Spline Models for Observational Data*. Philadelphia: Soc. Ind. Appl. Math.
53. Tibshirani R. 1994. Neural networks: a review from statistical perspective. *Stat. Sci.* 9:48–49
54. Wahba G. 2007. *Statistical learning in medical data analysis*. Tech. Rep. 1136. Dep. Stat., Univ. Wisconsin-Madison
55. Lush JL. 1931. The number of daughters necessary to prove a sire. *J. Dairy Sci.* 14:209–20
56. Lush JL. 1935. Progeny test and individual performance as indicators of an animal's breeding value. *J. Dairy Sci.* 18:1–19
57. Kempthorne O. 1967. *An Introduction to Genetic Statistics*. New York: Wiley

58. Eisenhart C. 1947. The assumptions underlying the analysis of variance. *Biometrics* 3:1–21
59. Hartley HO, Rao JNK. 1967. Maximum likelihood estimation for the mixed analysis of variance model. *Biometrika* 54:93–108
60. Rao CR. 1971. Estimation of variance and covariance components: MINQUE theory. *J. Multivar. Anal.* 1:257–75
61. Searle SR. 1971. Topics in variance component estimation. *Biometrics* 27:1–76
62. Searle SR. 1974. Prediction, mixed models and variance components. In *Reliability and Biometry*, ed. F Proschan, RJ Serfling, pp. 229–66. Philadelphia: Soc. Ind. Appl. Math.
63. de los Campos G, Sorensen DA, Gianola D. 2014. Genomic heritability: What is it? *Proc. 10th World Cong. Genet. Appl. Livest. Prod., Vancouver, Can.* **https://asas.org/wcgalp-proceedings**
64. Henderson CR. 1976. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics* 32:69–83
65. Harville DA. 1977. Maximum likelihood approaches to variance component estimation and to related problems. *J. Am. Stat. Assoc.* 72:320–40
66. Feldman MW, Lewontin RC. 1975. The heritability hang-up. *Science* 190:1163–68
67. Kempthorne O. 1978. Logical, epistemological and statistical aspects of nature-nurture data interpretation. *Biometrics* 34:1–23
68. Wood PDP. 1967. Algebraic model of the lactation curve in cattle. *Nature* 216:164–65
69. White IMS, Thompson R, Brotherstone R. 1999. Genetic and environmental smoothing of lactation curves with cubic splines. *J. Dairy Sci.* 82:632–38
70. Taylor St CS. 1980. Genetic size-scaling rules in animal growth. *Anim. Prod.* 30:161–65
71. Soller M, Beckmann JS. 1982. Restriction fragment length polymorphisms and genetic improvement. *Proc. 2nd World Cong. Genet. Appl. Livest. Prod.*, 6:396–404. Madrid: Editor. Ceuta
72. Soller M, Beckmann JS. 1983. Genetic polymorphism in varietal identification and genetic improvement. *Theor. Appl. Genet.* 67:25–33
73. Zhang Q, Boichard D, Hoeschele I, Ernst C, Eggen A, et al. 1998. Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree. *Genetics* 149:1959–73
74. Fernando FL, Grossman M. 1989. Marker assisted selection using best linear unbiased prediction. *Genet. Sel. Evol.* 21:467–77
75. Haley CS, Knott SA. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69:315–24
76. Andersson L, Georges M. 2004. Domestic-animal genomics: deciphering the genetics of complex traits. *Nat. Rev. Genet.* 5:202–12
77. Qanbari S, Pausch H, Jansen S, Somel M, Strom TM, et al. 2014. Classic selective sweeps revealed by massive sequencing in cattle. *PLOS Genet.* 10(2):e1004148
78. Gowen JW. 1952. *Heterosis*. Ames: Iowa State Coll. Press
79. Smith SP, Mäki-Tanila A. 1990. Genotypic covariance matrices and their inverses for models allowing dominance and inbreeding. *Genet. Sel. Evol.* 22:65–91
80. Sun C, VanRaden PM, O'Connell JR, Weigel KA, Gianola D. 2013. Mating programs including genomic relationships and dominance effects. *J. Dairy Sci.* 96:8014–23
81. Nishio M, Satoh M. 2014. Including dominance effects in the genomic BLUP method for genomic evaluation. *PLOS ONE* 9:e85792
82. Carlborg O, Haley CS. 2004. Epistasis: Too often neglected in complex trait studies? *Nat. Rev. Genet.* 5:618–25
83. Hill WG, Goddard ME, Visscher PM. 2008. Data and theory point to mainly additive genetic variance for complex traits. *PLOS Genet.* 4(2):e1000008
84. Huang W, Richards S, Carbone MA, Zhu D, Anholt RR, et al. 2012. Epistasis dominates the genetic architecture of *Drosophila* quantitative traits. *PNAS* 109:15553–59
85. Mackay TFC. 2014. Epistasis and quantitative traits: using model organisms to study gene–gene interactions. *Nat. Rev. Genet.* 15:22–33
86. Taylor MB, Ehrenreich IM. 2014. Genetic interactions involving five or more genes contribute to a complex trait in yeast. *PLOS Genet.* 10(5):e1004324

87. Cockerham CC. 1954. An extension of the concept of partitioning hereditary variance for the analysis of covariances among relatives when epistasis is present. *Genetics* 39:859–82
88. Kempthorne O. 1954. The correlation between relatives in a random mating population. *Proc. R. Soc. B* 143:103–13
89. Henderson CR. 1985. Best linear unbiased prediction of nonadditive genetic merits in noninbred populations. *J. Anim. Sci.* 60:111–17
90. Henderson CR. 1988. Progress in statistical methods applied to quantitative genetics since 1976. In *Proceedings of the Second International Conference on Quantitative Genetics*, ed. BS Weir, EJ Eisen, MM Goodman, G Namkoong, pp. 85–90. Sunderland, MA: Sinauer
91. Willham RL. 1963. The covariance between relatives for characters composed of components contributed by related individuals. *Biometrics* 19:18–27
92. Falconer DS. 1965. Maternal effects and selection response. In *Genetics Today*, ed. SJ Geerts, pp. 763–74. Oxford: Pergamon
93. Koerkhuis ANM, Thompson R. 1997. Models to estimate maternal effects for juvenile body weight in broiler chickens. *Genet. Sel. Evol.* 29:225–49
94. Van Vleck LD. 1993. *Selection Index and Introduction to Mixed Model Methods*. Boca Raton, FL: CRC Press
95. Bijma P. 2006. Estimating maternal genetic effects in livestock. *J. Anim. Sci.* 84:800–6
96. Skjervold H, Fimland E. 1975. Evidence for a possible influence of the fetus on the milk yield of the dam. *Z. Tierz. Zücht.* 92:245–51
97. Van Vleck LD. 1978. A genetic model involving fetal effects on traits of the dam. *Biometrics* 34:123–27
98. Kennedy BW, Schaeffer LR. 1990. Reproductive technology and genetic evaluation. In *Advances in Statistical Methods for Genetic Improvement of Livestock*, ed. D Gianola, K Hammond, pp. 507–32. Heidelberg: Springer-Verlag
99. Van Vleck LD, Snowder GD, Hanford KJ. 2003. Models with cytoplasmic effects for birth, weaning and fleece weights, and litter size at birth for a population of Targhee sheep. *J. Anim. Sci.* 81:61–67
100. Muir WM. 2005. Incorporation of competitive effects in forest tree or animal breeding programs. *Genetics* 170:1247–59
101. Bijma P, Muir WM, van Arendonk JA. 2007. Multilevel selection 1: quantitative genetics of inheritance and response to selection. *Genetics* 175:277–88
102. Bijma P. 2010. Estimating indirect genetic effects: precision of estimates and optimum designs. *Genetics* 186:1013–28
103. Wright S. 1960. The treatment of reciprocal interaction, with or without lag, in path analysis. *Biometrics* 16:423–45
104. Haavelmo T. 1943. The statistical implications of a system of simultaneous equations. *Econometrica* 11:1–12
105. Hill WG. 1984. On selection among groups with heterogeneous variance. *Anim. Prod.* 39:473–77
106. Khatib H. 2012. *Livestock Epigenetics*. New York: Wiley-Blackwell
107. Neugebauer N, Räder I, Schild HJ, Zimmer D, Reinsch N. 2010. Evidence for parent-of-origin effects on genetic variability of beef traits. *J. Anim. Sci.* 88:523–32
108. Robertson A. 1955. Prediction equations in quantitative genetics. *Biometrics* 11:95–98
109. Dempfle L. 1977. Relation entre BLUP (Best Linear Unbiased Prediction) et estimateurs bayésiens. *Ann. Genet. Sel. Anim.* 9:27–32
110. Gianola D, Fernando RL. 1986. Bayesian methods in animal breeding theory. *J. Anim. Sci.* 63:217–44
111. Blasco A. 2001. The Bayesian controversy in animal breeding. *J. Anim. Sci.* 79:2023–46
112. Grosu H, Schaeffer LR. 2014. *History of Genetic Evaluation Methods in Dairy Cattle*. Bucharest: Publ. House Rom. Acad.
113. Fernando RL, Gianola D. 1986. Optimal properties of the conditional mean as a selection criterion. *Theor. Appl. Genet.* 72:822–25
114. Smith FH. 1936. A discriminant function for plant selection. *Ann. Eugen.* 7:240–50
115. Hazel LN. 1943. The genetic basis for constructing selection indexes. *Genetics* 28:476–90

116. Karam HA, Chapman AB, Pope AL. 1953. Selecting lambs under farm flock conditions. *J. Anim. Sci.* 12:148–64
117. Henderson CR, Searle SR, Kempthorne O, von Krosigk M. 1959. Estimation of environmental and genetic trends from records subject to culling. *Biometrics* 15:192–218
118. Harvey WR. 1960. *Least-squares analysis of data with unequal subclass numbers*. Bull. 20-8. US Dep. Agric., Agric. Res. Serv., Washington, DC
119. Harvey WR. 1970. Estimation of variance and covariance components in the mixed model. *Biometrics* 26:485–504
120. Schaeffer LR, Kennedy BW. 1986. Computing solutions to mixed model equations. *Proc. 3rd World Congr. Genet. Appl. Livest. Prod.*, XII:382–93, Lincoln: Agric. Commun., Univ. Neb.
121. Misztal I, Gianola D. 1987. Indirect solution of mixed model equations. *J. Dairy Sci.* 70:716–23
122. Patterson HD, Thompson R. 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika* 58:545–54
123. Wolfinger R. 1993. Laplace's approximation for nonlinear mixed models. *Biometrika* 80:791–95
124. Lee Y, Nelder JA. 1996. Hierarchical generalized linear models. *J. R. Stat. Soc. B* 58:619–78
125. Harville DA, Mee RW. 1984. A mixed model procedure for analyzing ordered categorical data. *Biometrics* 40:393–408
126. Gilmour AR, Anderson RD, Rae AL. 1985. The analysis of binomial data by a generalized linear mixed model. *Biometrika* 72:593–99
127. Foulley JL, Im S, Gianola D, Höschele I. 1987. Empirical Bayes estimation of parameters for n polygenic binary traits. *Genet. Sel. Evol.* 19:197–224
128. Hofer A. 1998. Variance component estimation in animal breeding: a review. *J. Anim. Breed. Genet.* 115:247–65
129. Searle SR. 1968. Another look at Henderson's methods of estimating variance components. *Biometrics* 24:749–78
130. LaMotte LR. 1973. Quadratic estimation of variance components. *Biometrics* 32:793–804
131. Fisher RA. 1922. On the mathematical foundations of mathematical statistics. *Philos. Trans. R. Soc. A* 222:594–604
132. Harville DA, Callanan TP. 1990. Computational aspects of likelihood-based inference for variance components. In *Advances in Statistical Methods for Genetic Improvement of Livestock*, ed. D Gianola, K Hammond, pp. 36–176. Heidelberg: Springer-Verlag
133. Dempster AP, Laird NM, Rubin DB. 1977. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. B* 39:1–38
134. Meyer K, Kirkpatrick M. 2010. Better estimates of genetic covariance matrices by "bending" using penalized maximum likelihood. *Genetics* 185:1097–110
135. Harville DA. 1974. Bayesian inference for variance components using only error contrasts. *Biometrika* 61:383–85
136. Gianola D, Foulley JL, Fernando RL. 1986. Prediction of breeding values when variances are not known. *Proc. 3rd World Congr. Genet. Appl. Livest. Prod.*, XII:356–70. Lincoln: Agric. Commun. Univ. Neb.
137. Harville DA, Carriquiry AL. 1992. Classical and Bayesian prediction as applied to an unbalanced mixed linear model. *Biometrics* 48:987–1003
138. James W, Stein C. 1961. Estimation with quadratic loss. *Proc. 4th Berkeley Symp. Math. Statist. Prob.*, 1:361–379. Berkeley: Univ. Calif. Press
139. Lindley DV, Smith AFM. 1972. Bayes estimates for the linear model (with discussion). *J. R. Stat. Soc. B* 34:1–41
140. Box GEP, Tiao GC. 1973. *Bayesian Inference in Statistical Analysis*. Reading, MA: Addison-Wesley
141. Rönningen K. 1971. Some properties of the selection index derived by Henderson's mixed model method. *Z. Tierz. Zuecht.* 8:186–93
142. Sorensen D, Gianola D. 2002. *Likelihood, Bayesian and MCMC Methods in Quantitative Genetics*. New York: Springer
143. Guo SW, Thompson EA. 1991. Monte Carlo estimation of variance component models for large complex pedigrees. *Biometrics* 48:361–72

144. Wang CS, Rutledge JJ, Gianola D. 1993. Marginal inference about variance components in a mixed linear model using Gibbs sampling. *Genet. Sel. Evol.* 25:41–62
145. Sorensen DA, Wang CS, Jensen J, Gianola D. 1994. Bayesian analysis of genetic change due to selection using Gibbs sampling. *Genet. Sel. Evol.* 26:333–60
146. Wang CS, Gianola D, Sorensen DA, Jensen J, Christensen A, Rutledge JJ. 1994. Response to selection for litter size in Danish Landrace pigs: a Bayesian analysis. *Theor. Appl. Genet.* 88:220–30
147. Thompson R. 1986. Estimation of realized heritability in a selected population using mixed model methods. *Genet. Sel. Evol.* 18:475–84
148. Sorensen DA, Fernando RL, Gianola D. 2001. Inferring the trajectory of genetic variance in the course of artificial selection. *Genet. Res.* 77:83–94
149. Wright S. 1926. A frequency curve adapted to variation in percentage occurrence. *J. Am. Stat. Assoc.* 21:162–78
150. Wright S. 1934. An analysis of variability in number of digits and in an inbred strain of guinea pigs. *Genetics* 19:506–36
151. Dempster ER, Lerner IM. 1950. Heritability of threshold characters. *Genetics* 35:212–36
152. Falconer DS. 1965. The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Ann. Hum. Genet.* 29:51–76
153. Thompson R. 1979. Sire evaluation. *Biometrics* 35:339–53
154. Nelder J, Wedderburn R. 1972. Generalized linear models. *J. R. Stat. Soc. A* 135:370–84
155. Gianola D, Foulley JL. 1983. Sire evaluation for ordered categorical data with a threshold model. *Genet. Sel. Evol.* 15:201–24
156. Sorensen DA, Andersen S, Gianola D, Korsgaard I. 1995. Bayesian inference in threshold models using Gibbs sampling. *Genet. Sel. Evol.* 27:229–49
157. Misztal I, Gianola D, Foulley JL. 1989. Computing aspects of a nonlinear method of sire evaluation for categorical data. *J. Dairy Sci.* 72:1557–68
158. Foulley JL, Gianola D, Thompson R. 1983. Prediction of genetic merit from data on categorical and quantitative variates with an application to calving difficulty, birth weight and pelvic opening. *Genet. Sel. Evol.* 15:407–24
159. Höschele I, Foulley JL, Colleau JJ, Gianola D. 1984. Genetic evaluation for multiple binary responses. *Genet. Sel. Evol.* 18:299–320
160. Foulley JL, Gianola D, Im S. 1987. Genetic evaluation for traits distributed as Poisson-binomial with reference to reproductive traits. *Theor. Appl. Genet.* 73:870–77
161. Tempelman RJ, Gianola D. 1996. A mixed effects model for overdispersed count data in animal breeding. *Biometrics* 52:265–79
162. Tempelman RJ, Gianola D. 1999. Genetic analysis of fertility in dairy cattle using negative binomial mixed models. *J. Dairy Sci.* 82:1834–47
163. Smith SP, Allaire FR. 1986. Analysis of failure times measured on dairy sows: theoretical considerations in animal breeding. *J. Dairy Sci.* 69:217–27
164. Ducrocq V, Casella G. 1996. Bayesian analysis of mixed survival models. *Genet. Sel. Evol.* 28:505–29
165. Famula TR. 1981. Exponential stayability model with censoring and covariates. *J. Dairy Sci.* 64:538–45
166. Damgaard LH, Korsgaard IR. 2006. A bivariate quantitative genetic model for a linear Gaussian trait and a survival trait. *Genet. Sel. Evol.* 38:45–64
167. Ducrocq V, Sölkner J. 1998. The Survival Kit: a Fortran package for the analysis of survival data. *Proc. 6th World Congr. Genet. Appl. Livest. Prod.*, 22:51–52. Armidale, UK: Anim. Genet. Breed. Unit
168. Lush JL, Shrode RR. 1950. Changes in milk production with age and milking frequency. *J. Dairy Sci.* 33:338–57
169. Brody S. 1945. *Bioenergetics and Growth*. New York: Reinhold Publ. Corp.
170. Schaeffer LR, Dekkers JCM. 1994. Random regressions in animal models for test-day production in dairy cattle. *Proc. World Cong. Genet. Appl. Livest. Prod.*, 18:443–46. Guelph, Can.: Univ. Guelph
171. Kirkpatrick M, Lofsvold D. 1989. The evolution of growth trajectories and other complex quantitative characters. *Genome* 31:778–83

172. Meyer K. 1998. Modeling repeated records: covariance functions and random regression models to analyse animal breeding data. *Proc. 6th World Congr. Genet. Appl. Livest. Prod.*, 25:517–20. Armidale, UK: Anim. Genet. Breed. Unit
173. Strandén I, Gianola D. 1999. Mixed effects linear models with t-distributions for quantitative genetic analysis: a Bayesian approach. *Genet. Sel. Evol.* 31:25–42
174. Rosa GJM, Padovani CR, Gianola D. 2003. Robust linear mixed models with normal/independent distributions and Bayesian MCMC implementation. *Biom. J.* 45:573–90
175. Rosa GJM, Gianola D, Padovani CR. 2004. Bayesian longitudinal data analysis with mixed models and thick-tailed distributions using MCMC. *J. Appl. Stat.* 31:855–73
176. Kizilkaya K, Tempelman RJ. 2005. A general approach to mixed effects modeling of residual variances in generalized linear mixed models. *Genet. Sel. Evol.* 37:31–56
177. Kizilkaya K, Fernando RL, Garrick DJ. 2014. Reduction in accuracy of genomic prediction for ordered categorical data compared to continuous observations. *Genet. Sel. Evol.* 46:37
178. Reber DL, Stern HS, Berger PJ. 2000. Bayesian inference for the mixed linear model with application to selection in animal breeding. *J. Agric. Biol. Environ. Stat.* 5:240–56
179. Gianola D, Heringstad B, Ødegård J. 2006. On the quantitative genetics of mixture characters. *Genetics* 173:2247–55
180. Detilleux J, Leroy PL. 2000. Application of a mixed normal mixture model to the estimation of mastitis-related parameters. *J. Dairy Sci.* 83:2341–49
181. Ødegård J, Jensen J, Madsen P, Gianola D, Klemetsdal G, et al. 2003. Mixture models for detection of mastitis in dairy cattle using test-day somatic cell scores: a Bayesian approach via Gibbs sampling. *J. Dairy Sci.* 86:3694–703
182. Ødegård J, Jensen J, Madsen P, Gianola D, Klemetsdal G. 2005. A Bayesian liability-normal mixture model for analysis of a continuous mastitis-related trait. *J. Dairy Sci.* 88:2652–59
183. Boettcher PJ, Moroni P, Pisoni G, Gianola D. 2005. Application of a finite mixture model to somatic cell scores of Italian goats. *J. Dairy Sci.* 88:2209–16
184. Rodrigues-Motta M, Gianola D, Chang YM, Heringstad B. 2006. A zero-inflated Poisson model for genetic analysis of number of mastitis cases in Norwegian red cows. *Proc. 8th World Congr. Genet. Appl. Livest. Prod.*, CD Pap. 26-05. Belo Horizonte, Brazil: Inst. Prociencia
185. Celeux G, Hurn M, Robert C. 2000. Computational and inferential difficulties with mixture posterior distributions. *J. Am. Stat. Assoc.* 95:957–79
186. Meyer K. 1991. Estimating variances and covariances for multivariate animal models by restricted maximum likelihood. *Genet. Sel. Evol.* 23:67–83
187. Kriese LA, Boldman KG, Van Vleck LD, Kachman SD. 1994. A flexible set of programs to estimate (co)variances for messy multiple trait animal models using derivative free REML and sparse matrix techniques. *Proc. 5th World Congr. Genet. Appl. Livest. Prod.*, 22:43–44. Guelph, Can.: Univ. Guelph
188. Groeneveld E. 1994. VCE: a multivariate multimodel REML (co)variance component estimation package. *Proc. 5th World Congr. Genet. Appl. Livest. Prod.*, 22:47–48. Guelph, Can.: Univ. Guelph
189. Gilmour AR, Thompson R, Cullis BR. 1995. AI, an efficient algorithm for REML estimation in linear mixed models. *Biometrics* 51:1440–50
190. Van Tassel CP, Van Vleck LD. 1996. Multiple-trait Gibbs sampler for animal models: flexible programs for Bayesian and likelihood based (co)variance component inference. *J. Anim. Sci.* 74:2586–97
191. Groeneveld E, Garcia Cortes LA. 1998. VCE 4.0: a (co)variance component package for frequentists and Bayesians. *Proc. 6th World Congr. Genet. Appl. Livest. Prod.*, 27:455–56. Armidale, UK: Anim. Genet. Breed. Unit
192. Curnow RN. 1961. The estimation of repeatability and heritability from records subject to culling. *Biometrics* 17:553–66
193. Im S, Fernando RL, Gianola D. 1989. Likelihood inferences in animal breeding under selection: a missing data theory viewpoint. *Genet. Sel. Evol.* 21:399–414
194. Weller JI. 1986. Maximum likelihood techniques for the mapping and analysis of quantitative trait loci with the aid of genetic markers. *Biometrics* 42:627–41

195. Lander ES, Botstein D. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185–99
196. Lande R, Thompson R. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* 124:743–56
197. Gianola D, de los Campos G, Hill WG, Manfredi E, Fernando RL. 2009. Additive genetic variability and the Bayesian alphabet. *Genetics* 187:347–63
198. Schaeffer LR. 2006. Strategy for applying genome-wide selection in dairy cattle. *J. Anim. Breed. Genet.* 123:218–23
199. Van Raden PM, Van Tassell CP, Wiggans GR, Sonstegard TS, Schnabel RD, et al. 2009. Reliability of genomic predictions for North American Holstein bulls. *J. Dairy Sci.* 92:16–24
200. Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME. 2009. Genomic selection in dairy cattle: progress and challenges. *J. Dairy Sci.* 92:433–43
201. Habier D, Fernando RL, Dekkers JCM. 2007. The impact of genetic relationship information on genome-assisted breeding values. *Genetics* 177:2389–97
202. Van Raden PM. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–23
203. Goddard ME. 2009. Genomic selection: prediction of accuracy and maximisation of long term response. *Genetica* 136:245–57
204. de los Campos G, Naya H, Gianola D, Crossa J, Legarra A, et al. 2009. Predicting quantitative traits with regression models for dense molecular markers and pedigree. *Genetics* 182:375–85
205. Habier D, Fernando RL, Kizilkaya K, Garrick DJ. 2011. Extension of the Bayesian alphabet for genomic selection. *BMC Bioinform.* 12:186
206. Erbe M, Hayes BJ, Matukumali LK, Goswami S, Bowman PJ, et al. 2012. Improving accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. *J. Dairy Sci.* 95:4114–29
207. de los Campos G, Hickey JM, Pong-Wong R, Daetwyler HD, Calus MPL. 2013. Whole-genome regression and prediction methods applied to plant and animal breeding. *Genetics* 193:327–45
208. Heslot N, Yang HP, Sorrells ME, Jannink JL. 2012. Genomic selection in plant breeding: a comparison of models. *Crop Sci.* 52:146–60
209. Gianola D, Perez-Enciso M, Toro M. 2003. On marker-assisted prediction of genetic value: beyond the ridge. *Genetics* 163:347–65
210. Yang W, Tempelman RJ. 2012. A Bayesian ante-dependence model for whole genome prediction. *Genetics* 4:1491–501
211. Wimmer V, Lehermeier C, Albrecht T, Auinger HJ, Wang Y, Schön CC. 2013. Genome-wide prediction of traits with different genetic architecture through efficient variable selection. *Genetics* 195:573–87
212. Lehermeier C, Wimmer V, Albrecht T, Auinger HJ, Gianola D, et al. 2013. Sensitivity to prior specification in Bayesian genome-based prediction models. *Stat. Appl. Genet. Mol. Biol.* 12:1–17
213. González-Recio O, Gianola D, Rosa GJM, Weige KA, Kranis A. 2009. Genome-assisted prediction of a quantitative trait measured in parents and progeny: application to food conversion rate in chickens. *Genet. Sel. Evol.* 41:3–13
214. Erbe M, Pimentel ECG, Sharifi AR, Simianer H. 2010. Assessment of cross-validation strategies for genomic prediction in cattle. *Proc. 9th World Congr. Genet. Appl. Livest. Prod.*, Abstr. 129.
215. Daetwyler HD, Villanueva B, Woolliams JA. 2008. Accuracy of predicting the genetic risk of disease using a genome-wide approach. *PLOS ONE* 3(10):e3395
216. Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S, Lawlor TJ. 2010. A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J. Dairy Sci.* 93:743–52
217. Dekkers JCM, Hospital F. 2002. The use of molecular genetics in the improvement of agricultural populations. *Nat. Rev. Genet.* 3:22–32
218. Gianola D, Fernando RL, Stella A. 2006. Genomic assisted prediction of genetic value with semiparametric procedures. *Genetics* 173:1761–76
219. Gianola D, Van Kaam JT. 2008. Reproducing kernel Hilbert spaces regression methods for genomic assisted prediction of quantitative traits. *Genetics* 178:2289–303

220. Gianola D, Okut H, Weigel KA, Rosa GJM. 2011. Predicting complex quantitative traits with Bayesian neural networks: a case study with Jersey cows and wheat. *BMC Genet.* 12:87
221. Pérez-Rodríguez P, Gianola D, Weigel KA, Rosa GJ, Crossa J. 2013. An R package for fitting Bayesian regularized neural networks with applications in animal breeding. *J. Anim. Sci.* 91:3522–31
222. González-Recio O, Gianola D, Long L, Weigel KA, Rosa GJM, Avendaño S. 2008. Non-parametric methods for incorporating genomic information into genetic evaluations: an application to mortality in broilers. *Genetics* 178:2305–13
223. Jarquín D, Crossa J, Lacaze X, Du Cheyron P, Daucourt J, et al. 2014. A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theor. Appl. Genet.* 127:595–607
224. Tusell L, Pérez-Rodríguez P, Forni S, Wu XL, Gianola D. 2013. Genome-enabled methods for predicting litter size in pigs: a comparison. *Animal* 7:1739–49
225. Ornella L, Pérez P, Tapia E, González-Camacho JM, Burgueño J, et al. 2014. Genomic-enabled prediction with classification algorithms. *Heredity* 112:616–26
226. Wheeler HE, Aquino-Michaels K, Gamazon ER, Trubetskoy VV, Dolan ME, et al. 2014. Poly-omic prediction of complex traits: OmicKriging. *Genet. Epidemiol.* 38:412–15
227. Breiman L. 1996. Bagging predictors. *Mach. Learn.* 24:123–40
228. Gianola D, Sorensen D. 2004. Quantitative genetic models for describing simultaneous and recursive relationships between phenotypes. *Genetics* 176:1407–24
229. Rosa GJM, Valente BD, de los Campos G, Wu XL, Gianola D, Silva MA. 2011. Inferring causal phenotype networks using structural equation models. *Genet. Sel. Evol.* 43:6
230. Valente BD, Rosa GJM, Gianola D, Wu XL, Weigel KA. 2013. Is structural equation modeling advantageous for the genetic improvement of multiple traits? *Genetics* 194:561–72
231. Civelek M, Lusis AJ. 2014. Systems genetics approaches to understand complex traits. *Nat. Rev. Genet.* 15:34–48
232. Cartwright TC, Fitzhugh HA, Long CR. 1975. Systems analysis of sources of genetic and environmental variation in efficiency of beef production: mating plans. *J. Anim. Sci.* 40:433–43

Annual Review of
Animal Biosciences

Volume 3, 2015

Contents