Genotypes are useful for more than genomic evaluation

P.M. VanRaden[1], C. Sun[2], T.A. Cooper[1], D.J. Null[1], & J.B. Cole[1]

[1]Animal Improvement Programs Laboratory, Agricultural Research Service, United States Department of Agriculture, Beltsville, MD 20705-2350, USA,
[2]National Association of Animal Breeders, Columbia, MO 65205, USA

Abstract

New services that provide pedigree discovery, breed composition, mating programs, genomic inbreeding, fertility defects, and inheritance tracking all are possible from low-cost genotyping in addition to genomic evaluation. Genetic markers let breeders select among sibs before their phenotypes became available, and 50,000 markers combined with larger reference populations greatly increased the reliability of genomic evaluations since 2008. With genotypes available for most sires, grandsires, and great-grandsires, fairly complete pedigrees can be discovered in dairy herds that had poor or no pedigree recording. For 289,390 females genotyped in 2013, 11% had no sire reported, 15% had an incorrect sire reported, and 6% had a non-genotyped sire reported. Of the 75,905 females with incorrect or missing sires, a true sire was suggested for 50,538 (67%). Breed composition for crossbred animals can be accurately estimated using large differences among breeds in allele or haplotype frequencies. An animal's own genomic inbreeding is easy to estimate, and genomic instead of pedigree inbreeding in the next generation can be controlled by computing genomic relationships for each potential mate pair. Use of genomic instead of pedigree inbreeding gives an estimated benefit of $30 per female calf produced. A total of 18 recessive defects and simply inherited conditions are tracked or imputed for each genotyped animal. Web displays allow visual inspection of chromosomal breeding values and marker effects for each trait. Cost of genotyping is greatly reduced by reading a subset of the markers and imputing the rest, with a small reduction in accuracy. All of these services are now being or can be provided routinely for >500,000 genotyped animals in the North American database and are also available and routinely used by dairy producers from 35 other countries that submit genotypes to the database.

Keywords: pedigree discovery, genomic inbreeding, mating programs, recessive defects

Introduction

Producers and breeders now genotype thousands of animals to improve selection decisions, and the resulting genotypes are used for several other purposes. Similarly, producers have recorded phenotypes for many traits of millions of animals to improve management decisions, and selection decisions often are a secondary use of the resulting phenotypes when combined with pedigrees and genotypes. Because large-scale genotyping is relatively new, producers and breeders have not yet thought of many potential secondary uses of the genotypes being collected. North American breeders have genotyped more than half a million dairy animals, mostly with low-density genotyping chips (Boichard *et al.*, 2012). Although those chips have slightly reduced accuracy, they have decreased the cost of genotyping substantially, from >$100 per animal with >50,000 markers to about $40 per animal with >6,000 markers. Other

uses that may be worth just a few dollars per animal can help stimulate further adoption of genomic testing.

Potential uses for genotypes beyond predicting additive genetic merit (breeding values) are 1) discovering unknown ancestors or correcting inaccurate pedigrees, 2) estimating breed composition for crossbred animals or confirming breed purity, 3) avoiding inbreeding at a genomic level and estimating nonadditive genetic effects, and 4) discovering and tracking new defects or individual genetic effects on traits of interest. Many of these new uses require only additional computing and very little investment other than labor, organization, and a growing database of genotypes.

Pedigree discovery and breed composition

Fairly complete pedigrees can be discovered and mistakes corrected for herds that had poor or no pedigree recording if the ancestor bulls were genotyped. For young U.S.-sampled Holstein bulls used in artificial insemination, genotypes are available for 100% of the bulls, 99.4% of their sires, and 98.9% of their maternal grandsires (MGS). Percentages are lower for young females because many owners now genotype heifers in commercial herds to determine their ancestry. When any new animal's genotype is first loaded into the U.S. database, codes are set to indicate if the genomic sire and dam matched the pedigree sire and dam already on file. Then, pedigrees usually are changed by the breeder to match the genomic report, but the initial codes are kept to quantify how accurate the pedigrees were before the genomic information arrived. For 289,390 females genotyped as of December 2013, 11% had no sire reported, 15% had an incorrect sire reported, and 6% had a non-genotyped sire reported. Of the 75,905 females with incorrect or missing sires, a true sire was suggested for 50,538 (67%). For 43,931 males genotyped, 3% had no sire reported, 6% had an incorrect sire reported, and 6% had a non-genotyped sire reported. Of the 4,011 males with incorrect or missing sires, a true sire was suggested for 2,968 (74%).

The suggested MGS provided by genomics are about 95% accurate. Suggested maternal great grandsires (MGGS) also could have about 90% accuracy (VanRaden *et al.*, 2013) but are not yet provided to owners. Timing of service is also important to customers, and methods that do not require imputation can return results immediately upon receiving the genotype instead of waiting for monthly processing steps (van Kaam *et al.*, 2013). Currently in the U.S. database, 87% of males and 49% of females have a "likely" MGS confirmed by genotype. Another 9% of males and 41% of females have a reported MGS that is not genotyped and cannot be confirmed, leaving 4% of males and 10% of females with a MGS that is missing or confirmed incorrect. These statistics are after pedigree correction; initial status codes are kept for parents but not for MGS.

Breed composition for crossbred animals can be accurately estimated using large differences among breeds in allele or haplotype frequencies (Kuehn *et al.*, 2011; VanRaden *et al.*, 2011b). Multiple regression on 50,000 markers with breed treated as the trait can predict the percentage contribution of each breed with a standard error of < 3%. Smaller sets of markers produce larger standard errors, but accuracy may remain high by first imputing from lower density and then using the full set of regressions as is done in other genomic evaluations. Predictions of breed composition are useful because owners might not know the pedigrees and because actual fractions of inheritance differ from those expected, just as genomic and pedigree relationship matrices differ. All genotypes in the U.S. database are tested for breed purity, but breed composition percentage is not yet provided for crossbred animals because genomic evaluations are computed only within breeds.

Inbreeding and dominance

An animal's own genomic inbreeding is easy to estimate by computing genomic (**G**) and pedigree (**A**) relationship matrices and then standardizing **G** so that average diagonals equal those of **A**. Genomic instead of pedigree inbreeding in the next generation should be controlled by computing genomic relationships for each potential mate pair. Use of genomic instead of pedigree inbreeding in a mating program gives an estimated benefit of $30 per female calf produced (Pryce *et al.*, 2012; Sun *et al.*, 2013). Efficient web-based transfer of elements of **G** may be needed because storage and retrieval are difficult with many animals genotyped.

Non-additive genetic effects such as dominance, additive by additive interactions, imprinting, or epigenetic effects also contribute to phenotypic differences along with the additive effects currently used in genomic prediction and can be used in mating programs even if they do not accumulate across generations. Dominance variance accounted for 5% and 7% of total variance for production traits of Holsteins and Jerseys, respectively (Sun *et al.*, 2013), but was very small for productive life, daughter pregnancy rate, and somatic cell score in both breeds. For production traits, inclusion of dominance in the model for genotyped cows (8,323 Jerseys and 30,583 Holsteins) improved prediction accuracy expressed as average correlations between estimated total genetic effects and phenotypes based on ten-fold cross validation. However, inclusion of cows with genotype probabilities derived from genotyped sires and dams or genotyped sires and maternal grandsires did not improve prediction accuracy.

Defects and inheritance tracking

A total of 18 recessive defects and simply inherited conditions are tracked or imputed for each genotyped animal. These include seven previously known defects, nine fertility haplotypes discovered since 2011, polledness, and red hair color (Cole *et al.*, 2013). Some gene tests had previously been patented, but the U.S. Supreme Court invalidated those patents in 2013. Causative mutations already have been found and added to chips for several of the fertility haplotypes, and addition of the mutation tests slightly improves imputation accuracy for animals previously genotyped before the tests were added. Marker effects for each trait and individual chromosomal breeding values are displayed on the web (Cole and Null, 2013) and have become popular educational tools that allow breeders to visually inspect Mendelian sampling and transmission of effects across generations for their own animals.

Conclusions

Genotypes increase in value when used for more than just genomic evaluation. Pedigree discovery has become a big business because sire-identified females may be worth much more than unidentified females. Genomic mating programs, non-additive effect predictions, genetic defect monitoring, and chromosomal inheritance tracking also become valuable services to offer as genomic databases grow. Genotypes also increase in value when matched to more traits, and investing in additional phenotypes helps maximize benefits from the investment in genotypes. Timely delivery of new services is also important because owners often desire to act immediately on genomic information, and providing such information on demand to thousands of owners requires efficient algorithms and exchange methods. Genomic services have expanded very rapidly.

List of References

Boichard, D., H. Chung, R. Dassonneville, X. David, A. Eggen, S. Fritz, K.J. Gietzen, B.J. Hayes, C.T. Lawley, T.S. Sonstegard, C.P. Van Tassell, P.M. VanRaden, K.A. Viaud-Martinez & G.R. Wiggans, 2012. Design of a bovine low-density SNP array optimized for imputation. PLoS ONE 7: e34130.

Cole, J.B. & D.J. Null, 2013. Visualization of the transmission of direct genomic values for paternal and maternal chromosomes for 15 traits in US Brown Swiss, Holstein, and Jersey cattle. J. Dairy Sci. 96: 2713-2726.

Cole, J.B., P.M. VanRaden, D.J. Null, J.L. Hutchison & T.A. Cooper, 2013. Haplotype tests for recessive disorders that affect fertility and other traits. AIPL Res. Report Genomic3, http://aipl.arsusda.gov/reference/recessive_haplotypes_ARR-G3.html.

Kuehn, L.A., J.W. Keele, G.L. Bennett, T.G. McDaneld, T.P.L. Smith, W.M. Snelling, T.S. Sonstegard & R.M. Thallman, 2011. Predicting breed composition using breed frequencies of 50,000 markers from the US Meat Animal Research Center 2,000 Bull Project. J. Anim. Sci. 89: 1742-1750.

Pryce, J.E., B.J. Hayes & M.E. Goddard, 2012. Novel strategies to minimize progeny inbreeding while maximizing genetic gain using genomic information. J. Dairy Sci. 95: 377-388.

Sun, C., P.M. VanRaden, J.R. O'Connell, K.A. Weigel & D. Gianola, 2013. Mating programs including genomic relationships and dominance effects. J. Dairy Sci. 96: 8014-8023.

van Kaam, J.B. & B.J. Hayes, 2013. Maternal grandsire verification and detection without imputation. Interbull Bull. 47: 120-124.

VanRaden, P.M., T.A. Cooper, G.R. Wiggans, J.R. O'Connell & L.R. Bacheller, 2013. Confirmation and discovery of maternal grandsires and great-grandsires in dairy cattle. J. Dairy Sci. 96: 1874-1879.

VanRaden, P.M., K.M. Olson, D.J. Null & J.L. Hutchison, 2011a. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. J. Dairy Sci. 94: 6153-6161.

VanRaden, P.M., K.M. Olson, G.R. Wiggans, J.B. Cole & M.E. Tooker, 2011b. Genomic inbreeding and relationships among Holsteins, Jerseys, and Brown Swiss. J. Dairy Sci. 94: 5673-5680.