Modeling identical twins and clones in genetic evaluations

Paul M. VanRaden[1] Gary C. Fok[1], Sajjad Toghiani[1], and Ezequiel L. Nicolazzi[2]

[1] USDA-ARS Animal Genomics and Improvement Laboratory, Beltsville, MD, USA

[2] Council on Dairy Cattle Breeding, Bowie, MD, USA

Abstract

Identical animals cause more complex relationships to model in genetic evaluations. The USA evaluation currently includes 4,762 pairs of natural identical twins, 1,776 split embryos, and 530 nuclear transfer clones from cells of other embryos, calves, or adults, plus seven million other genotyped animals. Genetic effects for the 7,068 animals reported to be a clone or copy of another animal were linked to the source animal, and their own effects were removed from the relationship matrix. The model retained separate permanent environmental effects for each cow. For progeny of clones, the source animals are substituted as their sires and dams. After completing the evaluation, the reverse process restored the actual sires and dams and duplicated the evaluations of source animals to their clones for publication. Pedigree inbreeding coefficients increased slightly for animals with a paternal ancestor and a maternal ancestor that were clones of each other. Genomic predictions improved by estimating just one polygenic effect instead of modelling the copies as full sibs. Milk production of adult clones was not significantly affected, but their fertility and health traits were below expected. Several AI companies now market cloned bulls. The revised model better evaluates identical twins, cloned animals, and their progeny.

Key words: Clones, nuclear transfer, relationship matrix, genomic selection

Introduction

Many elite animals have been cloned in recent years. Examples include a cow named Apple that sold for $1 million in 2008, was a Grand Champion at the World Dairy Expo in 2011 but was Reserve Grand at that show in 2013 when one of her nine clones (Apple-3) beat her to become Grand Champion (Malcolm, 2019). A heifer named Liz was Junior All-American Winter Yearling in 2001 and her clone Liz-2 was Junior Champion at World Dairy Expo in 2004 (Nauman, 2011; Figure 1). More recently, very young calves with the highest genomic predictions are being cloned.

In past generations, artificial insemination (AI) bulls were mature when selected and were mated to thousands of cows per year. With genomic selection, elite bulls are discovered at very young ages, well before they reach puberty, and new animals quickly replace even the most elite animals. In recent generations, a top young bull plus several clones born nine months later may have less direct impact than many famous bulls had in the past. In Canadian evaluations, data for identical bulls was merged since 2011 to solve for just one genetic effect as recommended by Kennedy and Schaeffer (1989) but those procedures were not implemented for identical cows.

In U.S. evaluations, identical bulls were given identical predicted transmitting ability (PTAs) and daughter counts since 2008 (VanRaden and Fok, 2008) following research by Norman et al. (2004) to verify their identical inheritance. Those methods were also not extended to cows or to monthly or weekly genomic evaluations. Clones that are genotyped get combined but not identical PTAs in current genomic evaluations. Their genotypes can be from different chips that are merged before the evaluation to ensure that all use the same genomic information, but the

marker effects in the model only account for 90% of the genetic variance. The other 10% is modelled by polygenic effects using pedigree relationships as if the clones are full sibs for that 10% portion. Those slightly different PTAs are used in parent averages for their progeny before the final published PTAs of bulls are forced to be identical.

Figure 1. Example cow Liz and her clone Liz-2.

Many identical twins have been discovered or confirmed by genomic testing, and elite bulls and cows might have 1 or several clones available for use in breeding. For daughter-proven Holstein bulls actively marketed in April 2023, the top five for lifetime net merit included two clones, and the top 20 included another clone of a different bull. The top 50 marketed young bulls also included a clone. Thus, updated methods were examined to properly model these more complex relationships.

Identical animals have been reported using pedigree format one (CDCB, 2023a) for decades. Bytes 54-70 can report either a second ID for the same animal or the ID of an identical animal. The pedigree record type in byte 88 indicates whether the second ID is a cross-reference (X) or clonal record (C). The multiple birth code (CDCB, 2023b) in byte 91 can report how the identical animal was created (embryo splitting or nuclear transfer) and can report embryo transfer, twin births, or pedigrees for genotyped embryos not born yet (Table 1).

Table 1. Multiple birth codes used for reporting twin or clone status at birth or as embryo.

Code	Description
1	Single
2	Multiple birth (not from embryo transfer)
3	Birth from embryo transfer
4	Split embryo (artificially)
5	Clone from nuclear transfer
6	Embryo pedigree (implantation date stored as birth date)

Cloning had a limited impact on livestock breeding until recently, because some reproductive technologies can result in large offspring syndrome (Center for Veterinary Medicine, 2008), and cloning remains expensive. Vegetative cloning is much simpler and is used in some plant breeding programs, where genomic prediction methods were tested for simulated cloned trees but without inbreeding in the model (Stejskal et al., 2022). Goals of the present study were to examine clone reporting methods, develop more precise modelling for clones, and apply revised programs to the national genetic evaluation of dairy cattle.

Materials and Methods

The National Cooperator Database used for April 2023 official evaluations of CDCB included 7,068 animals reported to be a clone or copy of another animal. Those include 4,762 natural identical twins, 1,776 split embryos, and 530 nuclear transfers of cells from other embryos, calves, or adults. Number of source animals was 6,625 including 5,871 females and 754 males. Copies per source animal ranged from one to 11 but averaged only 1.07. For identical twins, usually the first one in the database is considered the source animal and the other is counted as a copy. Animal names were reported for 4,416 copies and for 4,442 source animals. For nuclear transfer clones, the clone names often indicate their status by repeating the source animal's name plus a clone count suffix.

The pedigree file included 94,499,373 animals of many dairy breeds and crossbreds.

Among all animals, 88,793 were sired by copies and those sire IDs were replaced by the source animal's ID. Similarly, 7,956 reported dams were copies and were replaced by the source ID. The reduced pedigree file had 94,492,305 (94,499,373 minus 7,068) animals after also removing the IDs of copies. Producers can report identical twins without genotyping, but nearly all are discovered by genotyping and then confirmed by the producer. Nearly all nuclear transfer clones were genotyped.

Edits for cloning attempted to separate real identical animals from other cases of duplicate IDs that should instead be cross-references for the same animal. The latter cases were often caused by multiple forms of ID, typos, or reidentification of calves after export to another country. Some mistakes were easy to identify, such as those from large batches of nearly consecutive IDs with multiple birth code reported as 2 (twins) instead of 3, 4, or 5 (embryo transfer), but were not marked as twins in the name. About 80-90% of the 7,068 animals initially reported as identical twins or clones appeared to be valid.

Examples of two genotypes for the same animal but with different IDs included: many animals reported with both a USA and 840, 982, or metal ear tag number that should probably be cross-references instead of clones; calves with nearly sequential ID numbers that had two different country codes (such as CHN and USA) sent by two different companies; animals whose ID numbers were the same but with inconsistent ID format; and a few obvious typos. Some identified mistakes were changed from clones to cross-references, but not all as the owner must first agree to such changes.

Modeling for clones was improved by removing clone copies from pedigree files and by using different IDs for genetic vs. permanent environmental effects. The clone copies were removed from both the full pedigree file used in phenotypic modelling, and reduced pedigree files used in weekly or monthly predictions from subsets of genotyped animals. For progeny of clones, the source animals are substituted as their sires and dams. After completing the evaluation, the reverse process is then used to restore the actual sires and dams and duplicate the evaluations of source animals to their clones.

For females with records in each phenotypic trait group, the new code now links genetic effects of each clone to the source animal and links each permanent environment effect to the cow's own ID, recognizing that clones are different animals with different environmental effects. A previous program had merged genotypes from the source animal and its clones because they might be genotyped with different chips or have different missing loci within each genotype. The revised program now outputs only one row for the source animal instead of duplicating the merged genotype to its clones so that the model can solve for just 1 genetic effect.

All trait groups and breeds were tested to ensure a working system and measure impact. New code was developed to copy female PTAs and to do the final reporting of clone PTAs in weeklies or monthlies. The new system could be implemented for the December 2023 evaluation. To estimate if nuclear transfer clones perform as expected from their identical genotype, a regression was added to the clone model with coefficient of one for each embryo nuclear transfer (ETN) cloned cow and zero for all other cows.

Results & Discussion

Pedigree inbreeding coefficients increased slightly for some animals after removing the clone copies and listing the source animal instead as the sire or dam. The cases examined had a paternal ancestor and a maternal ancestor that were clones of each other, which increased descendant inbreeding, versus if the clones were treated as full sibs. Examples were 1) an increase from 7.0% to 7.7% for an animal whose paternal granddam and maternal 2nd-great granddam were identical, and 2) an increase from 9.8% to 10.6% for an animal

whose maternal great grandsire (MAN-O-MAN2) was a clone of his paternal 2nd-great grandsire (MAN-O-MAN).

Genetic evaluations from the pedigree model differed most for cloned animals and for bulls with daughter records if their clones also had daughter records. Reliabilities increased for those animals, as expected. Evaluations for all other animals had almost no change, and estimated genetic trends were nearly identical. Of the 6,625 source animals in the model, 3,241 had no change to their evaluation for milk and 4,725 had no change to their reliability, presumably because the copies had no phenotypes or descendants.

For source animals that did change, average difference in milk estimated breeding value (EBV) was -3.9 pounds, average absolute change in milk EBV (test - official) was 79.4 pounds, and average gain in percent reliability was +2.6. The maximum difference in EBV milk was -3052 pounds for a USA Jersey cow born in 1991 that had 11 clones. Maximum reliability difference was +50, increasing from 47% to 97% for a Holstein bull who had only 1 daughter but whose split embryo twin had 520 total progeny. For bulls, the public will not see those EBV and reliability changes because such evaluations were already superseded by data from the clone member with highest reliability.

Evaluations from the genomic model had much smaller differences because only the polygenic effects had used the full sib instead of clonal relationships, and because bull PTAs had been forced to be identical.

Ancestor discovery (Nani et al., 2020) previously did not detect and add a cloned bull or the original bull because the 1st choice was no better than the 2nd choice. The new model with revised pedigree discovered about 20,000 ancestors that were members of a clone group. The ID of the source maternal grand sire (MGS) or maternal grand grand sire (MGGS) can be automatically added to the pedigree if missing, but to make pedigrees more precise, owners can replace the discovered source ancestor with the clone ancestor if it was used in that mating.

Genomic relationships of 1.0 and singular genomic relationship matrices can cause problems in genomic BLUP algorithms. Those issues can be avoided by solving for marker effects directly (SNP-BLUP), but in both strategies the polygenic effects would remain incorrect for identical animals. Updated models and pedigree inputs to multi-step software will provide further benefits for use in single-step models.

Direct effects of nuclear transfer cloning on phenotypic performance were very small for yield traits but effects were larger and unfavorable for several other traits (Table 2). The estimated phenotypic losses were mostly in the range of one to two genetic SD which are well within normal biologic ranges but too large to justify creating whole herds of cloned cows. Compared to trait means, the unfavorable effects ranged from 27% increase in somatic cell count to 2% increase in age at first calving. However, the 0.34 effect on SCS was only 1.2 genetic standard deviation (SD) whereas the 17 days later calving date was 8.1 genetic SD. Most cloned heifers are used as embryo donors and their phenotypes should probably be edited from the age at first calving dataset.

Many countries are adopting the Cartagena Protocol on Biosafety as recommended by the United Nations. Clones and gene-edited animals are not considered “genetically modified”. Guidelines limiting cloning were proposed to the EU parliament but were not adopted. Private companies sometimes enforce cloning rules that do not exist. For example, importers may demand “clone-free” pedigrees before export. Breed associations such as Holstein USA then must provide such reports and inspect all previous generations to discover any clone. Today about 0.3% of US Holsteins have a clone in their pedigree, but >3% may in 5 generations and >50% in 10 generations (about 20 years). International exchange of breeding stock should not become limited by artificial barriers.

Conclusions

Clone modelling was improved in the national evaluation. About 67% of the 7,068 copies in the clone file were natural identical twins, 25% were split embryos, and 7% were nuclear transfer clones. The model changes were not complex but required slight revisions to many programs, which led to small positive effects for many downstream analyses. Benefits of the new model were more exact pedigree inbreeding coefficients for descendants of clones, more precise genetic evaluations for clones, identical genomic evaluations for female clones, identical evaluations for bulls in additional trait groups such as type and calving, combined progeny counts for cloned bulls instead of reporting only the daughter count of the clone with the most, and improved ancestor discovery. The new or revised programs better account for cloned animals and identical twins. Milk production of cows obtained by nuclear transfer cloning was as expected, but the clones had poorer performance than the source animals for some other traits. Many AI companies now market cloned bulls, and many dairy cattle may soon have clones in their pedigrees.

Acknowledgments

The authors thank CDCB staff and industry cooperators for data, Tad Sonstegard (Recombinetics) for manuscript suggestions, Pete Sullivan (Lactanet) for information on the Canadian evaluation, and Frank Robinson (Lodi, CA, USA) for permission to use the photograph.

References

Center for Veterinary Medicine. 2008. Animal cloning: a risk assessment. U. S. Food and Drug Administration. https://public4.pagefreezer.com/browse/FDA/06-04-2023T20:24/https://www.fda.gov/media/75280/download

Council on Dairy Cattle Breeding. 2023a. Format 1. https://redmine.uscdcb.com/projects/cdcb-customer-service/wiki/Format_1

Council on Dairy Cattle Breeding. 2023b. Reference Note 24. https://redmine.uscdcb.com/projects/cdcb-customer-service/wiki/REFERENCES#Ref24

Kennedy, B.W. and Schaeffer, L.R. 1989. Genetic evaluation under an animal model when identical genotypes are represented in the population. *J. Anim. Sci.* 67:1946-1955.

Malcolm, D. 2019. KHW Regiment Apple-Red-ET – Everything and more. The Bullvine.

Nani, J.P., Bacheller, L.R., Cole, J.B., and VanRaden, P.M. 2020. Discovering ancestors and connecting relatives in large genomic databases. *J. Dairy Sci.* 103(2):1729–1734.

Nauman, D. 2011. Commitment got Nelsons Pronto Liz across. http://www.dairyagendatoday.com/News.aspx?nid=5557

Norman, H.D., Lawlor, T.J., Wright, J.R., and Powell, R.L. 2004. Performance of Holstein clones in the United States. *J. Dairy Sci.* 87(3):729–738.

Stejskal, J., J. Klápště, J. Čepl, Y. A. El-Kassaby and Lstibůrek, M. 2022. Effect of clonal testing on the efficiency of genomic evaluation in forest tree breeding. Scientific Reports 12:3033.

VanRaden, P. and Fok, G. 2008. Genetic evaluations for clones. August 2008 evaluation changes. USDA.

Table 2 – Performance of nuclear transfer clones for 10 traits.

Trait	Units	Mean	Genetic SD	Clones	Effect	Effect/SD	Effect/Mean
Milk	Pounds	28,071	1134	472	+18	0.0	+0%
Fat	Pounds	1,077	50	467	-8	-0.2	-1%
Protein	Pounds	871	30	467	+7	0.2	+1%
SCS (or SCC)[1]		200k	0.28	460	+0.34	1.2	+27%
Productive life	months	25	3.4	119	-3.3	-1.0	-13%
Dtr. pregnancy rate	%	27	2.8	354	-5.0	-1.8	-19%
Heifer conception rate	%	45	2.6	37	-5.5	-2.1	-12%
Cow conception rate	%	41	3.2	123	-8.3	-2.6	-20%
Age at first calving	months	831	2.1	115	+17.0	8.1	+2%
Cow livability	%	97	3.2	423	-7.3	-2.3	-8%

[1] SCS (somatic cell score) evaluations are on log base 2 scale but were converted and compared to the SCC mean in cells / ml.