Comparison of Domestic and Foreign Genotypes by Country and Continent

M.E. Tooker[1], T.A. Cooper[1], J. Dürr[2] and P.M. VanRaden[1]
[1]Animal Genomics and Improvement Laboratory, Agricultural Research Service,
USDA, Beltsville, MD 20705-2350, USA
[2] Council on Dairy Cattle Breeding, Bowie, MD 20715, USA

Abstract

Genomic evaluations for foreign animals are easily computed, and reliabilities are highest for animals well connected to the domestic reference population and managed in similar environments. Genomic and pedigree relationships, inbreeding, pedigree completeness, pedigree accuracy and genomic merit were examined for 880 797 genotyped animals from 44 countries in the U.S. national database as of April 2015. For genotyped Holsteins from continents other than North America, >60% of sires were North American, and relationships to the reference population were nearly as high for the 203 637 foreign animals as for the 677 160 domestic animals. Pedigrees from all continents were fairly complete. About 94% of sires were reported by breeders, and half of the 6% of sires that were missing were discovered from genotypes. From 2 to 16% of animals had incorrect sires, and >80% of those errors were corrected. Genomic net merit across all continents was higher than conventional parent averages and the domestic genetic base of cows born in 2010. Countries with smaller dairy populations or without advanced data collection and breeding systems can obtain predictions from larger databases and for additional traits by exchanging genotypes across country borders, allowing animals from many countries to be evaluated together on the same scale.

Key words: genomic evaluation, international exchange, genomic relationship, pedigree completeness

Introduction

International exchange of genotypes allows gains in reliability to be documented within partner countries (Schenkel *et al.,* 2009; Lund *et al.,* 2011), but predictions are harder to test for genotypes contributed from countries with less advanced conventional evaluation systems. If pedigree and genomic relationships of foreign and domestic animals are similar, predictions should be equally accurate on the domestic scale but may be less accurate for foreign performance because of genotype-by-environment interaction.

Demand for genomic testing is growing rapidly among dairy producers in many countries. Those in smaller markets or without advanced data collection and breeding systems can obtain predictions from larger databases by exchanging genotypes across country borders (Cromie *et al.,* 2010). Advantages are larger reference populations, more traits and more efficient service compared with predictions computed from local data. Goals of the current study are to compare genetic connections, data quality and predictions using foreign genotypes and emphasizing continents with less documented dairy cattle populations.

Materials and Methods

Genomic information from April 2015 was examined by country code and continental region. Central and South America were grouped into Latin America, and animals from western and eastern Europe were examined separately to determine if any differences remained from regional breeding programs prior to 1989. The U.S. genomic database currently contains >1 million genotypes from 44 countries and 18 different types of genotyping chips. Country codes for animals and their sires were based on country of registration and were grouped into seven continental regions (Table 1). Most genotyped animals were female: 84% for North America, 93% for Latin America, 51% for western Europe, 67% for eastern Europe, 81% for Asia, 98% for Africa and 64% for Oceania.

Table 1. Countries and numbers of genotyped animals by continental region.

Continental region	Country	Genotyped animals (no.)	Continental region	Country	Genotyped animals (no.)
North America	United States	677 160	Oceania	Australia	5 156
	Canada	99 303		New Zealand	1 126
	Total	776 463		Total	6 282
Western Europe	Italy	21 403	Eastern Europe	Hungary	796
	Germany	14 922		Czech Republic	647
	France	13 336		Poland	406
	Netherlands	12 715		Slovenia	283
	United Kingdom	10 989		Serbia	26
	Switzerland	4 750		Romania	8
	Denmark	2 524		Russia	3
	Spain	1 601		Total	2 169
	Austria	1 227	Asia	China	295
	Finland	947		Iran	174
	Ireland	838		Japan	144
	Sweden	588		Saudi Arabia	120
	Belgium	554		Thailand	25
	Luxembourg	279		Kazakhstan	8
	Norway	5		Taiwan	3
	Total	86 678		India	3
Latin America	Chile	2 599		Total	772
	Mexico	2 496	Africa	South Africa	303
	Argentina	1 773		Total	303
	Brazil	876			
	Ecuador	170			
	Peru	98			
	Uruguay	47			
	Costa Rica	42			
	Colombia	29			
	Total	8 130			

Phenotypic, pedigree, chip type and genomic inbreeding information was combined by breed and edited to include only genotyped animals. If sire was unknown, country was designated as unknown. Because pedigrees can change, initial sire status code is recorded when the genotype is first submitted to allow determining if the correct sire was reported by the breeder or discovered from the genotype. Animals were genotyped using a variety of chip types and densities. All were imputed using findhap to the 60 671 markers currently used in the United States for routine genomic evaluation. Genomic inbreeding and average relationship to genotyped, progeny-tested bulls born in the last 10 years were then computed.

Genomic evaluations in one country may accurately predict merit in other environments, especially if many foreign bulls have their multitrait across-country evaluations included in the reference population. The genetic merit is expressed on the U.S. scale, but many foreign phenotypes contribute to estimating the marker effects for each breed (Table 2). For Brown Swiss and Red Dairy Cattle (Ayrshires), numbers of foreign reference bulls are much larger than domestic reference bulls. For Holsteins and Jerseys, domestic cows now contribute much more to the reference population.

Results and Discussion

Genomic and pedigree relationships of foreign animals to the domestic reference population were similar across continents (Table 3). Expected future inbreeding is half the average pedigree relationship of an animal to recent

Table 2. Numbers of domestic and foreign reference animals and total numbers of genotyped animals by breed.

	Reference population			
Breed	U.S. bulls	Foreign bulls	U.S. cows	Total genotypes
Holstein	15 288	11 941	124 666	764 029
Jersey	3 005	1 529	31 233	99 212
Brown Swiss	970	5 155	1 367	18 805
Red Dairy Cattle	188	523	95	4 228

Table 3. Average pedigree completeness, inbreeding, expected future inbreeding, and genomic future inbreeding by continental region.

Continent	Pedigree completeness (%)	Inbreeding (%)	Expected future inbreeding (%)	Genomic future inbreeding (%)
Western Europe	97.6	6.1	6.1	6.7
Oceania	93.6	6.0	5.8	6.2
Eastern Europe	88.0	5.5	5.8	6.2
Africa	87.4	6.3	6.6	7.3
North America	86.1	6.5	6.4	6.9
Asia	71.9	5.8	6.0	6.3
Latin America	67.7	5.9	6.0	6.0

U.S. cows, and genomic future inbreeding is half the average genomic relationship to recent reference bulls. A previous comparison of progeny-tested bulls from each country also showed highly related pedigrees among nearly all countries (VanRaden, 2005). Pedigree completeness was highest for western Europe (98%) and lowest for Latin America (68%). Differences may reflect testing only elite animals versus testing commercial females with missing pedigrees for which gains in reliability are larger (Weigel *et al.*, 2012). Inbreeding levels were similar across continents.

When a genotype is first submitted, status of the sire is documented (confirmed, conflict, missing, or not genotyped) to keep track of how many conflicts are being corrected because of genomics. Nominators may submit genotypes from commercial herds without knowing the identity of the sire. Codes in the database for initial sire status report if the pedigree sire was confirmed by genotype, if the animal's genotype conflicted with the sire's genotype, if pedigree was not available or if the genotype was received before these codes were implemented. For those animals with a sire that was unknown or that conflicted with the genotype, a second code reports if the sire was discovered or a correct sire was found in the database. North American and foreign genotyped animals were similar, and most sires are accurately submitted by nominators. Sires were reported for 94% of animals, and about half of the 6% of sires that were missing were discovered from genotypes. From 2 to 16% of genotyped animals had incorrect sires, and >80% of those errors were corrected. Corresponding statistics for each continent are in Table 4.

Countries of origin for sires show the use of genetics worldwide (Table 5). North American genetics account for 61 to 94% of sires by continent and average 90% worldwide. Breeders on other continents are not primarily using local domestic sires but are testing progeny of well-known international sires. Statistics for western Europe include the exchange of Holstein bull genotypes between North America, Italy and the United Kingdom.

The sires used most heavily based on number of genotyped progeny are fairly consistent across continents (Table 6). Well-known older sires such as Shottle and Goldwyn were used most in some regions, and younger sires such as Mogul and Supersire were used in other regions. Sires labeled as

Table 4. Percentages of sires in identity and discovery categories by continental region.

Continent	Confirmed (true sire known)	Conflict		Sire not reported		Sire not genotyped	
		True sire found	True sire not found	True sire found	True sire not found	True sire found	True sire not found
North America	84.5	7.0	0.5	4.1	2.5	0.2	1.3
Western Europe	89.1	1.6	0.4	1.0	1.0	0.2	6.7
Eastern Europe	88.9	1.6	0.1	0.0	2.0	0.2	7.2
Oceania	71.2	1.6	0.9	0.0	2.9	0.4	23.1
Latin America	61.5	8.0	2.9	2.9	9.9	5.4	9.4
Asia	62.0	9.0	2.7	0.1	14.8	0.3	11.2
Africa	87.7	6.6	0.3	0.0	0.0	1.3	4.0

Table 5. Percentages of sires by country of origin and continental region.

Continent	Sire country of origin		
	United States	Canada	Other countries
North America	84	10	2 (*unknown*)
Asia	74	7	8 (Japan)
Africa	67	15	7 (United Kingdom)
Latin America	61	14	9 (Mexico)
Eastern Europe	61	7	7 (Italy)
Western Europe	49	13	9 (Italy)
Oceania	44	17	18 (New Zealand)
All	80	10	2 (*unknown*)

Table 6. Most used sires by continental region.

Continent	Sire	Progeny (no.)	Continent	Sire	Progeny (no.)
Africa	Shottle	20	North America	*Unknown*	15 790
	Goldwyn	16		Mogul	6 927
	Gold Chip	16		Planet	6 900
	Tbone	14		Shamrock	6 166
	Atwood	13		Supersire	5 899
Asia	Otto	46	Oceania	Man-O-Man	183
	Sudan	43		Supersire	162
	Unknown	40		Goldwyn	152
	Altanato	24		Snowman	141
	Freddie	23		Dauntless	112
Eastern Europe	Shottle	60	Western Europe	Shottle	1 473
	Massey	50		Goldwyn	1 340
	Altaross	39		Man-O-Man	1 312
	Altaiota	38		NumeroUno	1 162
	Superstition	37		Mogul	1 017
Latin America	*Unknown*	685			
	Bogart	172			
	Elias	159			
	Dover	155			
	Jayton	154			

unknown had more progeny than any individual sire in North America and Latin America. Nominators often submit genotypes from commercial herds, where pedigrees are unknown. Such cows may be less related to the current purebred population and, therefore, a sire cannot be determined for many of those genotypes. Most genotypes in the U.S.

Table 7. Average genomic net merit, parent average (PA) for net merit and difference by continental region.

	Net merit		
Continent	Genomic	PA	Genomic minus PA
Africa	48	11	37
Asia	64	51	13
Eastern Europe	111	72	39
Latin America	9	–22	31
North America	191	171	20
Oceania	44	20	24
Western Europe	146	130	16
All	184	165	19

database were from Holsteins, whereas other breeds may not be as closely related across continents (Thomasen *et al.*, 2013).

The genotyped animals on each continent were superior to the average cows born in 2010 that form the U.S. genetic base, and genomic net merit was slightly higher than traditional parent average for net merit (Table 7). This indicates that most breeders are using genomically superior young sires as parents of their genotyped animals.

Use of foreign genotypes also contributes to U.S. genomic evaluations by improving pedigree discovery and identifying harmful recessive haplotypes. Direct evaluation of foreign genotypes allows males and females from many countries to be evaluated together on the same scale, which enables simpler genomic selection.

Conclusions

Breeders in many countries are genotyping animals that are highly related to the North American reference population. Inclusion of foreign animals in the reference population should make predictions more accurate across environments. About 6% of sires were incorrect, and the correct sire was determined for about 80% of those animals. Most of the genotyped animals had U.S. sires.

Acknowledgements

The authors wish to thank dairy farmers throughout the world for submitting DNA samples and the Council on Dairy Cattle Breeding (Bowie, MD) for providing genomic and traditional evaluations and pedigrees.

References

Cromie, A.R., Berry, D.P., Wickham, B., Kearney, J.F., Pena, J., van Kaam, J.B.C.H., Gengler, N., Szyda, J., Schnyder, U., Coffey, M., Moster, B., Hagiya, K., Weller, J.I., Abernethy, D. & Spelman, R. 2010. International genomic co-operation: Who, what, when, where, why and how? *Interbull Bulletin 42,* 72–78.

Lund, M.S., de Roos, A.P.W., de Vries, A.G., Druet, T., Ducrocq, V., Fritz, S., Guillaume, F., Guldbrandtsen, B., Liu, Z., Reents, R., Schrooten, C., Seefried, F. & Su, G. 2011. A common reference population from four European Holstein populations increases reliability of genomic predictions. *Genetics Selection Evolution 43,* 43.

Schenkel, F.S., Sargolzaei, M., Kistemaker, G., Jansen, G.B., Sullivan, P., Van Doormaal, B.J., VanRaden, P.M. & Wiggans, G.R. 2009. Reliability of genomic evaluation of Holstein cattle in Canada. *Interbull Bulletin 39,* 51–58.

Thomasen, J.R., Sørensen, A.C., Su, G., Madsen, P., Lund, M.S., & Guldbrandtsen, B. 2013. The admixed population structure in Danish Jersey dairy cattle challenges accurate genomic predictions. *Journal of Animal Science 91,* 3105–3112.

VanRaden, P.M. 2005. Inbreeding adjustments and effect on genetic trend estimates. *Interbull Bulletin 33,* 81–84.

Weigel, K.A., Hoffman, P.C., Herring, W. & Lawlor, T.J. Jr. 2012. Potential gains in lifetime net merit from genomic testing of cows, heifers, and calves on commercial dairy farms. *Journal of Dairy Science 95,* 2215–2225.