Digital phenotyping technologies: an emerging tool in genomic selection and animal welfare science?

Eugenia V. Solodneva[1,2*], *Aleksei K*. Piskunov[1], *Yury* Stolpovsky[1,*], and *Sergey* Kuznetsov[1]

[1]Vavilov Institute of General Genetics, Russian Academy of Sciences, 119333, Moscow, Russia
[2]Department of bioinformatics and systems biology, Moscow Institute of Physics and Technology (MIPT), 141700, Moscow region, Dolgoprudny, Russia

Abstract. Farm animals provide about 30% of humanity's food requirements [1]. The world's population's demand for animal products is expected to keep growing. New technologies in animal husbandry can make it possible to meet this demand and to reduce its impact on the environment. Genomic selection plays a key role in this process as a promising and safe method for improving genetic qualities of farm animals. Before the widespread of genomic selection, high cost of genotyping was the main factor holding back an improvement of animal farming. However, today, the bottleneck in genomic selection is a limited amount of high-quality phenotypic data, which is still either difficult or too costly to obtain [2]. New technologies in animal phenotyping become of special attention. The current paper reviews the next-generation technologies in animal phenotyping and revels the main challenges limiting its widespread in the example of dairy cattle.

1 Advances in phenotyping technologies

Herd books and manual performance records were the principal tool in breeding programs in pre-genomic era. In 1905 the first association to record the volume and the fat content of the milk obtained from cattle was formed in the United States. Currently, manual assessment and entry of phenotypic data tends to be replaced by automated recording with radio sensors and systems for collection and analysis of biological samples making it possible an automized recording of productivity, feed conversion, health, fertility, pregnancy, calving ease, temperature adaption, potential impact on climate and even of indicators of an experienced stress [3]. This information is subsequently used to reveal possible relationships between genomic data and phenotypic traits in order to improve the accuracy of genomic-estimated breeding values.

These methods can also serve for improving the management of high-performance industrial dairy herds [4]. The new sensors are recording health indicators (lameness, disease symptoms), animal welfare and their productivity (fat, protein, lactose, urea nitrogen in milk) and environmental data (feed composition, temperature) in real time [5]. This information

---

[*] Corresponding author: stolpovsky@mail.ru

can be of great importance for consumers providing them an opportunity to track the quality of food at the level of raw materials [6]. The experience of the "Red Tractor" company (UK) can serve an example of successful monitoring of the quality and the traceability (production history) of foods (https://www.redtractor.org.uk/).

Currently, manufacturers in Canada record 8 health indicators: mastitis, metritis, delayed placenta, cystic ovaries, lameness, hypocalcemia, ketosis and abomasum displacement. In order to standardize disease reporting, disease definition guidelines were provided to manufacturers, being adapted from David F. Kelton's Recommendations for Recording and Calculating the Incidence of Selected Clinical Diseases of Dairy Cattle [7].

Mastitis is recognized as the most cost demanding health problem in dairy cattle. The incidence of mastitis is still high worldwide despite a great attention being paid to the problem. In Canada in 2008, the incidence of mastitis was recorded at 23% [8], in the UK - at 14, 21 and 26% in the first, second and third lactations, respectively [9] and in France it is observed as 10, 13 and 14% in the first 3 lactations in the Montbéliard, Norman and Holstein breeds respectively [10]. The relatively slow decline in the incidence of mastitis, as compared to other economically relevant diseases, may be, at least partially, explained by an intense selection for milk production which negatively correlates with udder health.

Since the late 1970s, the Somatic Cell Count (SCC) in milk samples was the main indicator of mastitis. An increase in SCC occurs mainly due to the leukocyte's recruitment in response to udder infection and is closely related to the extent of the inflammatory process. Recently, the electrical conductivity of milk samples, the udder skin temperature obtained from thermal imaging camera, along with the "omics" data were proposed as the new methods in diagnostics of mastitis in cattle [6]. Another effective method for recording a variety of phenotypic data is an analysis of milk spectrum obtained in the mid-infrared range (MIR - mid-infrared). An absorption spectrum obtained by transmitting infrared light through a milk sample serves as predictor of various phenotypes. In addition to the traditionally used indicators of milk quality, the analysis determinates fatty acids composition, lactoferrin, minerals and key technological properties of milk, such as coagulability, protein composition, acidity and citrate content [11]. Besides, MIR spectrum analysis was shown to be an effective tool in assessing fertility, pregnancy, health scores, energy balance, feed conversion and methane emissions [12]. A high fat-to-protein ratio, for example, has been suggested as an indicator of negative energy balance, which leads to the development of ketosis, ovarian cystosis, mastitis, lameness, and abomasal displacement [13].

Non-invasive assessment tools can shorten assessment time and minimize animal stress. In particular, an exterior parameter can be obtained using a depth sensor and processing the photographic image of the animal. The main body measurements: height at the withers, height of the croup, depth of the chest, width of the chest, width of the croup, length of the croup, length of the body, girth of the metacarpus, can be estimated with an accuracy of 1 mm and an error of about 2% [14]. When using the Light Detection and Ranging sensor and subsequent obtaining a three-dimensional digital model of the animal, an accuracy of the main parameters is up to 2 mm, and the error is close to 2% [15].

An amount of phenotypic information is constantly growing. Meanwhile, the data collection might be irregular and differs in its format and structure. All digital data needs to be integrated into a unique system for further analysis, interpretation and visualization. Thus, there is a growing demand for the development of statistical analysis methods and visualization tools, as well for the usage of artificial intelligence and machine learning tools for the data transmission, storage and processing [16].

Moving on to other problems in high-performance phenotyping, it should be noted that there is no legal framework that could regulate an effective collection and usage of this information to improve the efficiency of animal husbandry in general. In order to avoid confusion in the definitions used to describe phenotypes, it is recommended to use carefully

selected vocabularies such as Ontology of Animal Product Traits (https://www.animalgenome.org/bioinfo/projects/lpt/) [16].

One of the main issues in collecting phenotypic data is the question of its recoupment, since the supplier of the product must reimburse the costs of introducing the necessary technologies. Many traits can have a direct economic value, for example, dairies often pay extra to producers for low levels of somatic cells in milk, for high fat and protein contents [17]. In an example of A2-kasein containing milk [18] becoming generally recognized as much safer product, the factories are ready to pay extra to producers for such milk, and producers, in turn, are ready to introduce technologies for identifying allelic variants of the beta-casein gene. However, it may not have a direct economic value. The challenge is to encourage dairy producers to provide such data [19]. It is of the fundamental importance that the manufacturers understand that the progress in selection depends on the registration of both traditional and new phenotypes. One option about encouraging the breeders to collect new data on phenotypes is to provide significant discounts on animal genotyping. An example of such a program is the "Milk recording program" used in the USA (national milk-recording programs). The members receive discounts on animal genotyping [19]. The US government costs for production data collection is about $ 50 million per year [20]. Some companies create their own phenotype supplier network through farms that have commercial agreements with them (Immunity Plus from Semex and Wellbeing from Zoetis). The phenotype database plays a special role in determining the economic indicators of the value of an animal.

2 Economic indices and breeding strategies

Estimates of the economic value of animals are widely used throughout the world. Specifically, in the US, the first selection index introduces was by the US Department of Agriculture was Predicted Difference Dollars (PD $). In 1926, it included milk and fat yields only [21]. Today, the Net Merit Dollars (NM $) includes a combination of 36 individual traits [21]. Selection indices are usually used to combine information on many traits associated with profitability in a single value [19]. It is worth noting that selection indices should be periodically updated to include new traits and relevant information about the relationships between them obtained from genomic research, and should also reflect changes in economic conditions and technologies (i.e., industrial or organic).

Selection indices differ both between countries and within one country due to the difference in economic conditions, registered traits and breeds used. For example, the Australian Dairy Herd Improvement Program prefers performance indicators, while Canada favors the exterior ones. In Denmark, Finland and Sweden the most attention is paid to somatic cell count and calving traits [17]. Economic indices for different breeds take into account different indicators; in addition, the same indicators may have different weights. For example, Ideal Commercial Cows for Holsteins estimates milk yield at 5 points, while the national overall NM $ index estimates this at 1 point [21]. Currently, in the United States, Holstein cows are evaluated according to 42 economic indicators, including 5 indicators of productivity, 8 indicators of health, fertility and longevity, 6 direct measures of health, 5 signs of calving and 18 exterior traits [19].

There is a growing interest in individualized indices at the farm level. An example is the North Florida Holsteins dairy farm (Bell, Florida, USA) that uses its own individual index to rank the animals (10000 cows). Differences in breeding purposes should be also taken into account. In order to provide farmers with valuation options tailored to their needs, Council on Dairy Cattle Breeding (CDCB) journal publishes 4 different indices (lifetime net merit, fluid merit, cheese merit, and grazing merit). An introduction of a general economic index, as a weighted sum of sub-indices, is now increasingly discussed [21].

One of the main challenges in modern breeding is to find a balance between increasing productivity while monitoring health and fertility indicators (as in above-mentioned example of balancing milk productivity and mastitis incidence). Economic goals are serving as a guideline in these cases [22].

Main genes and corresponding traits that can be used or are already being used to derive the economic indicators in cattle are listed below.

Health:
- FANCI - the development of a short spine (brachyspina)
- ITGB2 - reduced resistance to bacterial infections (leukocyte adhesion deficiency) [18]
- APOB HCD - cholesterol deficiency
- SLC35A3 - complex vertebral malformation
- UMPS - deficiency of uridine monophosphate synthase
- LRP4 - syndactyly (mulefoot)
- ASS1 - metabolic disorder (citrullinemia)
- HLZ and LSS - mastitis resistance [23], [24]
- HBD3 and SP110 - tuberculosis resistance
- APOL1 - trypanosomiasis resistance [25]
- BoLADRB3 - leukemia resistance.

*It should be noted that climate change entails a change in the habitats of some cattle pests [26], which may cause a demand for the inclusion of new traits in selection programs, for example, resistance to ticks [17].

Productivity:
- MSTN, MRC2, RNF11, WWP1 and ATP2A1 - muscle development
- DGAT1, GHR, CSN2, CSN3, CSN1S1, CSN1S2, LALBA, LGB, FAT1 and ABCG2 - milk yield and milk composition.

*Allelic variants of some genes which cause an increase in some indicators of productivity may negatively correlate with other indicators. For example, the K232A variant of the diacylglycerol O-acyltransferase 1 (DGAT1) gene leads to an increase in fat yield in cows, but at the same time to a decrease in protein yield in milk amount [27]. The myostatin gene (MSTN) increases muscle mass in heterozygous animals, but causes difficulties in giving birth to mothers of homozygous calves.

Fertility
- TTF1, RABGGTB, RNF20 - miscarriage
- ARMC3 - shortening of sperm flagellomania.

Separately, haplotypes associated with cow fertility can be distinguished. For example, the haplotypes in the Holstein breed are:
- APAF1 - HH1, SMC2 - HH3, GART - HH4, TFB1M - HH5, SDE2 - HH6, CENPU - HH7

Adaptation:
- PRLR [28], ATP1A1 [29], HSPB7 [30] - heat resistance.
- MSANTD4, GRIA4, AQP5, RETREG1, RPL7, HDAC3 и SYK - cold resistance [31], [32]

3 Next-generation phenotyping systems in animal welfare science

An attention paid to farm animals' welfare increased significantly in the last few years. Meanwhile one can consider the economic purposes to be the main source of this interest, it should be also noted that a notable proportion of relevant publications was made in the area of fundamental sciences. Herein, its likely to be inspired by recent works that confirmed that

farm animals' cognition is much more complex than it was generally believed before and by the new finding in genomics that showed convergent evolution of human and farm animals neurochemical systems [33]. Just recently a homologous mechanism underlying such a complex traits as human autism-like behavior was found in cattle [34].

Currently a great effort is being made to develop reliable and valid methods to access farm animals' welfare. They include measuring cortisol and its metabolites, acute phase proteins, serotonin and some other molecules (in plasma, hairs or saliva), monitoring of activity via GPS and accelerometer sensors, infrared thermography, electroencephalography, heart rate variation and other measures of sympathetic-vagal activity [35].

Notably, various researchers emphasize that its important to track positive states and emotions along with negative ones moving toward the consideration of "animal happiness" where emphasis has shifted from concern about negative aspects of animal welfare to the positives [35], [36]. The last ones appear to be more difficult to measure, especially due to the absence of reference indicators. However, evidence of pleasure can be based on behaviors such as social interaction, reproductive activity, play, self-grooming, anticipatory hyperactivity, and exploration. Facial emotions recognition is also proposed as a method to track positive states, with a wild animal behavior serving as a reference tool. A drug usage could be somewhat doubtful ethically but physiologically reasoned method to define the signs of positive state and pleasure in animals.

4 Conclusion

The amount of genomic, transcriptomic and metabolomic data has risen exponentially in animal science over the last decade. However, in order to serve for breeding purposes these data has to be combined with the reliable descriptions of phenotypes. Lack of phenotypic information represents now the main bottleneck in genomic selection. Gathering such a data is still considered to be an expensive, time- and labor-consuming process. Recently, a number of next-generation methods were developed to facilitate phenotypes recording. Most of them implement an automatized intellectual processing of information obtained from various biosensors, photo and video cameras, and robotized equipment for biological samples collection and analysis. These systems are extremely useful to monitor animals health and welfare that may be affected by an intense selection aimed at improving economically important traits that still are the main indicators in breeding strategies. Non-invasive assessment tools can shorten assessment time and minimize animal stress. The main challenges to be overcome in the widespread of these technologies include the development a unified methodology for retrieving, transferring and storing information, including both genotypes and phenotypes. Herein, a special attention should be paid for artificial intelligence methods based on machine learning as well as for providing support for those producers who introduce new evaluation methods.

References

1. Yang, W., Kang, X., Yang, Q. et al. Review on the development of genotyping methods for assessing farm animal diversity. J Animal Sci Biotechnol 4, 2 (2013).
2. Coffey, M. Dairy cows: in the age of the genotype,# phenotypeisking. Animal Frontiers, 10(2), 19-22 (2020).
3. Weigel, K. A., VanRaden, P. M., Norman, H. D., & Grosu, H. A 100-Year Review: Methods and impact of genetic selection in dairy cattle—From daughter–dam

comparisons to deep learning algorithms. Journal of dairy science, 100(12), 10234-10250 (2017).

4. Kaniyamattam, K., & De Vries, A. Agreement between milk fat, protein, and lactose observations collected from the Dairy Herd Improvement Association (DHIA) and a real-time milk analyzer. Journal of dairy science, 97(5), 2896-2908 (2014).

5. Halachmi, I., Guarino, M., Bewley, J., & Pastell, M. Smart animal agriculture: application of real-time sensors to improve animal well-being and production. Annual review of animal biosciences, 7, 403-425 (2019).

6. Martin, P., Barkema, H. W., Brito, L. F., Narayana, S. G., & Miglior, F. Symposium review: novel strategies to genetically improve mastitis resistance in dairy cattle. Journal of dairy science, 101(3), 2724-2736 (2018).

7. Kelton, D. F., Lissemore, K. D., & Martin, R. E. Recommendations for recording and calculating the incidence of selected clinical diseases of dairy cattle. Journal of dairy science, 81(9), 2502-2509 (1998).

8. Riekerink, R. O., Barkema, H. W., Kelton, D. F., & Scholl, D. T. Incidence rate of clinical mastitis on Canadian dairy farms. Journal of dairy science, 91(4), 1366-1377 (2008).

9. Pritchard, T., Coffey, M., Mrode, R., & Wall, E. Genetic parameters for production, health, fertility and longevity traits in dairy cows. Animal, 7(1), 34-46 (2013).

10. Govignon-Gion, A., Dassonneville, R., Baloche, G., & Ducrocq, V. Genetic evaluation of mastitis in dairy cattle in France. Interbull bulletin, 46, 121-126 (2012).

11. Gengler, N., Soyeurt, H., Dehareng, F., Bastin, C., Colinet, F., Hammami, H., ... & Dardenne, P. Capitalizing on fine milk composition for breeding and management of dairy cows. Journal of dairy science, 99(5), 4071-4079 (2016).

12. Friedrichs, P., Bastin, C., Dehareng, F., Wickham, B., & Massart, X. Final OptiMIR Scientific and Expert Meeting: From milk analysis to advisory tools. Palais des Congrès, Namur (Belgium), 16-17 April 2015. BASE. (2015).

13. Friggens, N. C., Ridder, C., & Løvendahl, P. On the use of milk composition measures to predict the energy balance of dairy cows. Journal of Dairy Science, 90(12), 5453-5467. (2007).

14. Batanov, S., Baranova, I., & Starostina, O. Innovative methods in study of animal's conformation. Bulgarian Journal of Agricultural Science, 26(6), 1286-1291. (2020).

15. Huang, L., Li, S., Zhu, A., Fan, X., Zhang, C., & Wang, H. Non-contact body measurement for qinchuan cattle with LiDAR sensor. Sensors, 18(9), 3014 (2018).

16. Koltes, J. E., Cole, J. B., Clemmens, R., Dilger, R. N., Kramer, L. M., Lunney, J. K., ... & Reecy, J. M. A vision for development and utilization of high-throughput phenotyping and big data analytics in livestock. Frontiers in genetics, 10, 1197 (2019).

17. Cole, J. B., & VanRaden, P. M. Symposium review: Possibilities in an age of genomics: The future of selection indices. Journal of dairy science, 101(4), 3686-3701 (2018).

18. Peñagaricano, F. Genetics and genomics of dairy cattle. In Animal Agriculture (pp. 101-119). Academic Press (2020).

19. Cole, J. B., Eaglen, S. A., Maltecca, C., Mulder, H. A., & Pryce, J. E. The future of phenomics in dairy cattle breeding. Animal Frontiers, 10(2), 37-44 (2020).

20. VanRaden, P. M.. An example from the dairy industry: the net merit index. In Proceedings of the Beef Improvement Federation's 37th Annual Research Symposium and Annual Meeting (pp. 96-100). (2005, July).

21. Cole, J. B., Dürr, J. W., & Nicolazzi, E. L. Invited review: The future of selection decisions and breeding programs: What are we breeding for, and who decides? Journal of Dairy Science (2021).

22. Georges, M., Charlier, C., & Hayes, B. Harnessing genomic information for livestock improvement. Nature Reviews Genetics, 20(3), 135-156 (2019).

23. Liu, X., Wang, Y., Tian, Y., Yu, Y., Gao, M., Hu, G., ... & Zhang, Y. Generation of mastitis resistance in cows by targeting human lysozyme gene to β-casein locus using zinc-finger nucleases. Proceedings of the Royal Society B: Biological Sciences, 281(1780), 20133368 (2014).

24. Liu, X., Wang, Y., Guo, W., Chang, B., Liu, J., Guo, Z., ... & Zhang, Y. Zinc-finger nickase-mediated insertion of the lysostaphin gene into the beta-casein locus in cloned cows. Nature communications, 4(1), 1-11 (2013).

25. Genovese, G., Friedman, D. J., Ross, M. D., Lecordier, L., Uzureau, P., Freedman, B. I., ... & Pollak, M. R. Association of trypanolytic ApoL1 variants with kidney disease in African Americans. Science, 329(5993), 841-845 (2010).

26. Bett, B., Kiunga, P., Gachohi, J., Sindato, C., Mbotha, D., Robinson, T., ... & Grace, D. Effects of climate change on the occurrence and distribution of livestock diseases. Preventive veterinary medicine, 137, 119-129 (2017).

27. Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C., Berzi, P., ... & Snell, R. Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. Genome research, 12(2), 222-231 (2002).

28. Littlejohn, M. D., Henty, K. M., Tiplady, K., Johnson, T., Harland, C., Lopdell, T., ... & Davis, S. R. Functionally reciprocal mutations of the prolactin signalling pathway define hairy and slick cattle. Nature communications, 5(1), 1-8 (2014).

29. Liu, Y., Li, D., Li, H., Zhou, X., & Wang, G. A novel SNP of the ATP1A1 gene is associated with heat tolerance traits in dairy cows. Molecular biology reports, 38(1), 83-88 (2011).

30. Zeng, L., Cao, Y., Wu, Z., Huang, M., Zhang, G., Lei, C., & Zhao, Y. A missense mutation of the HSPB7 gene associated with heat tolerance in Chinese Indicine cattle. Animals, 9(8), 554 (2019).

31. Igoshin, A. V., Yurchenko, A. A., Belonogova, N. M., Petrovsky, D. V., Aitnazarov, R. B., Soloshenko, V. A., ... & Larkin, D. M. Genome-wide association study and scan for signatures of selection point to candidate genes for body temperature maintenance under the cold stress in Siberian cattle populations. BMC genetics, 20(1), 5-14 (2019).

32. Yurchenko, A. A., Daetwyler, H. D., Yudin, N., Schnabel, R. D., Vander Jagt, C. J., Soloshenko, V., ... & Larkin, D. M. Scans for signatures of selection in Russian cattle breed genomes reveal new candidate genes for environmental adaptation and acclimation. Scientific reports, 8(1), 1-16 (2018).

33. Bertolini, F., Servin, B., Talenti, A. et al. Signatures of selection and environmental adaptation across the goat genome post-domestication. Genet Sel Evol 50, 57 (2018).

34. Roy Costilla, Kathryn E. Kemper, Enda M. Byrne, Laercio R. Porto-Neto, Roberto Carvalheiro, Deirdre C. Purfield, Jennifer L. Doyle, Donagh P. Berry, Stephen S. Moore, Naomi R. Wray, Ben J. Hayes. Genetic control of temperament traits across species: association of autism spectrum disorder risk genes with cattle temperament. Genetics Selection Evolution, 52 (1) (2020).

35. Graham K. Barrell. An Appraisal of Methods for Measuring Welfare of Grazing Ruminants Front. Vet. Sci., 28 August (2019).

36. Webb LE, Veenhoven R, Harfeld JL, Jensen MB. What is animal happiness? Ann NY Acad Sci. 1438:62–78 (2019).