# A genomic breed assignment test for traceability of meat of Dual-Purpose Blue

Q1



H. Wilmot[a,b,*], helene.wilmot@uliege.be, G. Glorieux[c], X. Hubin[c], N. Gengler[b]

[a]National Fund for Scientific Research (F.R.S.-FNRS), Rue d'Egmont, 5, B-1000 Brussels, Belgium

[b]TERRA Teaching and Research Centre, Gembloux Agro-Bio Tech, University of Liège, Passage des Déportés, 2, B-5030 Gembloux, Belgium

[c]Walloon Breeders Association (Elevéo), Rue des Champs Elysées, 4, B-5590 Ciney, Belgium

*Corresponding author.

## Abstract

Assigning meat to its breed of origin for traceability purposes is not always straightforward if the breed from which products are derived is closely related to another one. The objective of this study was to determine if a genomic breed assignment test could distinguish meat of Dual-Purpose Blue, a local endangered breed, from meat of Beef Belgian Blue, a heavily used breed in the Belgian meat industry which is related to Dual-Purpose Blue. For this purpose, a genomic breed assignment test based on a panel of 2,005 SNPs and the nearest shrunken centroids method was used to classify 32 meat samples from Dual-Purpose Blue ($n = 16$), Beef Belgian Blue ($n = 8$) and Holstein ($n = 8$) into their breed of origin. From this SNP panel, 167 SNPs allowed to detect meat of Dual-Purpose Blue and 173 SNPs allowed to detect meat of Beef Belgian Blue. The genomic breed assignment test correctly allocated all the meat samples to their breed of origin with a probability of one. Therefore, the use of the genomic breed assignment test in routine as one step of the certification process of Dual-Purpose Blue meat seemed possible.

**Keywords**:

Local breed, Breed assignment, Traceability, SNP

**Abbreviations**

No keyword abbreviations are available

## 1 Introduction

For the last years, consumers have paid more and more attention to what they consume, the origin of their food and how it is produced, searching for healthy and environmentally-friendly food conform to animal health and well-being.

The demand for local food with low ecological footprint has increased to meet these new needs. Meanwhile, the importance of the transparency of the food production and of the traceability of food products from farm to fork have been emphasized by policymakers, e.g. in the European Regulation N° 178/2002. Traceability of food does not only target the individual identification or the place of production but can also focus on the breed used for production (Dalvit et al., 2007).

When labelled breed derived-products are set up to maintain an endangered breed, the consumer trust is the key for the viability of the products on the market. To gain this trust, the applied certification process must be reliable and precise. However, assigning the product to its breed of origin is not always straightforward if the local breed from which labelled products are derived is closely related to a mainstream breed.

Using markers to determine the breed of origin of individual animals has been widely studied in the literature for pigs (e.g., Schiavo et al., 2019), sheep (e.g., Baumung et al., 2006), horses (e.g., Putnová and Štohl, 2019) and especially cattle (e.g., Hulsegge et al., 2013). A lot of different methodologies have been proposed for selection of markers as, e.g., using $F_{st}$ values (e.g., Wilkinson et al., 2011), absolute allele frequencies differences, called Delta (e.g., He et al., 2018) or a principal component analysis combined (e.g., Bertolini et al., 2015) or not (e.g.,Wilkinson et al., 2011) with another method of selection of SNPs. Different methods have also been proposed for breed assignment as five nearest neighbours classification (Iquebal et al., 2014) or regression based on partial least squares (Funkhouser et al., 2017). Recently, Wilmot et al. (2022) compared different methodologies for selection of SNPs combined with different breed assignment methods to determine which model would be the most appropriate for differentiating 12 reference breeds. In the specific case of meat traceability, different strategies have also been tested to determine the breed of origin. For example, to differentiate Hereford and Angus meat, Judge et al. (2017) elaborated two SNP panels of 300 SNPs each. These SNPs were selected based on an Index method combining Delta statistic and pairwise $F_{st}$ values. Then, they estimate breed proportion with the help of ADMIXTURE program (Alexander et al., 2009). Another option is to use the markers related to the KIT gene to detect variability related to colour patterns as Funkhouser et al. (2017) did for pig breeds. However, this is not a feasible option when one or more colour patterns are shared with a related breed.

The objective of this study was therefore to test the ability of a genomic breed assignment model to distinguish meat of two closely related breeds sharing the same colour patterns: the endangered Dual-Purpose Blue (DPB) breed and the Beef Belgian Blue (BBB) breed. Meat samples of Holstein (HOL) animals were used as a “benchmark” of the study.

# 2 Material and methods

R v.4.1.2 (R Core Team, 2021) and R Studio (R Studio Team, 2020) programs were used for computations.

## 2.1 Breeds

This study focused on the DPB, a local dual-purpose breed located in southern Belgium and north-east of France, across the border of these two countries. In the DPB breed, a partially recessive mutation of an allele “mh” (Charlier et al., 1995), leading to muscular hypertrophy, is co-existing with the wild type of this allele “+”. This means that, according to the breeding goal of the breeder, a disparity is found in the DPB population regarding this allele. In farms with a more meat orientated-breeding goal (this case is most likely to be encountered in Belgium), most “mh/mh” DPB animals are found. On the other hand, in the case of a more dairy-orientated breeding goal (as most French Breeders), breeders are raising mostly “+/+” DPB animals whereas the heterozygous “mh/+” genotype can be found in both types of farms. Moreover, breeders can also raise animals from the three muscular types in the same farm (Mota et al., 2017).

However, the DPB breed is closely related to BBB. This latter breed is mainly raised in Belgium, even if it is internationally used as a terminal sire because of its muscular properties. The DPB and BBB breeds are derived from a previously called Mid and Upper Belgium breed but diverged during the seventies because of different breeding goals. While the DPB has kept the old type of the breed, the BBB has been selected for superior meat properties, leading to the fixation of the “mh/mh” genotype in this breed (Mota et al., 2017). Because the Belgian market was really demanding for lean meat, the DPB breed was left aside for several decades and has now an endangered breed status ( Colinet, 2010).

At the end of the last century, agro-environmental measures were settled in Wallonia and DPB breeders were eligible to subsidies if their animals were registered to the herd book and their cows milk recorded (Colinet, 2010). Following

these first measures, different projects were launched to preserve the DPB breed. The last one, the BlueSter project [Instruction: I do not succeed to update the reference at the end of the document. The year should be 2022. Thanks.](~~BlueSter, 2021~~BlueSter, 2022), particularly targeted the development of labelled DPB derived-products like meat. In this context, the traceability of DPB meat products is crucial. It is of main importance to ensure the consumers they are effectively buying meat from this breed and not from the relatively close BBB.

### 2.2 Genotyping of meat samples

The EuroG MD v2 (Illumina, San Diego, CA, USA) was used for genotyping 32 meat samples. The samples were coming from three breeds: the BBB (~~n~~ n = 8), the DPB (~~n~~ n = 16) and the HOL (~~n~~ n = 8). Fattening of BBB animals was done at the same place but they were coming from four different farms with the following distribution: one animal from Farm 1, four animals from Farm 2, two animals from Farm 3 and one animal from Farm 4. The transboundary diversity of the DPB breed was considered as an equal number of French and Belgian animals (~~n~~ n = 8 animals from each country) were slaughtered. Moreover, Belgian animals were all carrying the “mh/mh” genotype. On the other hand, the French breeder provided “mh/+” and “+/+” animals. Therefore, the genotypic diversity of DPB regarding the muscular hypertrophy was also accounted for. The HOL samples were used as a “benchmark” for the study and animals from this breed were coming from different farms. For samples with missing values, a PCA-related procedure (Josse and Husson, 2012) was used to impute unknown genotypes, as in Wilmot et al. (2022).

### 2.3 Breed assignment model

To assign each of the 32 meat samples to their breed of origin, the second best breed assignment model described in Wilmot et al. (2022) was chosen. This model had the advantage of using less SNPs than the best model described in the study of Wilmot et al. (2022) while having a very similar performance of breed assignment. In summary, the chosen model used a panel of 2005 SNPs to allow to differentiate twelve reference breeds from each other. Among the reference breeds, 60 DPB, 60 BBB and 120 HOL were used as reference animals. The SNPs to be contained in the panel were selected using the partial least squares-discriminant analysis (PLS-DA). The PLS-DA built a model for each of the twelve reference breeds and SNPs having coefficients exceeding a certain threshold for at least one of the breeds were selected to be part of the panel. It therefore happened that some selected SNPs were in common between several breeds (i.e., that SNPs dedicated to differentiate one breed can overlap with SNPs dedicated to differentiate another breed). The classification of animals to their breed of origin was then based on the nearest shrunken centroids method (Tibshirani et al., 2002). This method can be seen as a corrected version of the classical nearest centroid method where the probability of an animal to belong to a breed was based on the distance of its genotype to overall and shrunken class centroids. The highest probability was used as the criterion to assign each sample to the breed they were supposed to belong to. This allowed to build a confusion matrix to compare the known breed of origin and the breed predicted by the model.

## 3 Results

From the used SNP panel, 160 SNPs allowed to exclusively detect DPB animals, 166 SNPs allowed to exclusively detect BBB animals and seven additional SNPs, in common between both breeds, allowed to distinguish both breeds from other reference breeds. These seven SNPs at the overlap were also used to differentiate DPB meat from BBB meat because these SNPs segregate between reference animals of both breeds as shown in Table 1. This was a better than expected result because of the genetic proximity between DPB and BBB.

alt-text: Table 1

**Table 1**

The table layout displayed in this section is not how it will appear in the final version. The representation below is solely purposed for providing corrections to the table. To preview the actual presentation of the table, please view the Proof.

Allelic frequencies of B allele in reference Beef Belgian Blue (BBB, $n$ = 60) and Dual-Purpose Blue (DPB, $n$ = 60) for overlapped SNPs that allow to discriminate both BBB and DPB from other reference breeds.

| SNP | Allelic frequencies of |
|---|---|

| | B allele | |
|---|---|---|
| | BBB | DPB |
| [Instruction: For any of the element of the column SNP, there should be no comma in the number because it is a SNP name. Thanks.]BTA-47-105 | 0.417 | 0.892 |
| BTA-110-789 | 0.492 | 0.892 |
| BTB-01-089-169 | 0.517 | 0.058 |
| BTA-85-612 | 0.442 | 0.875 |
| ARS-BFGL-NGS-85-952 | 0.075 | 0.433 |
| ARS-BFGL-NGS-33-483 | 0.375 | 0.058 |
| ARS-BFGL-NGS-22-284 | 0.492 | 0.875 |

Table 2 shows the confusion matrix of the breed assignment test for meat samples. The specificity and the sensitivity of the test were equal to one, meaning that the breed of origin of all meat pieces was correctly detected by the test. Moreover, all samples were assigned with a probability of one, which is the maximum probability.

alt-text: Table 2

**Table 2**

The table layout displayed in this section is not how it will appear in the final version. The representation below is solely purposed for providing corrections to the table. To preview the actual presentation of the table, please view the Proof.

Confusion matrix of the breed assignment test for 32 meat pieces.

| Predicted breed | Breed of origin | | |
|---|---|---|---|
| | Beef Belgian Blue | Dual-Purpose Blue | Holstein |
| Beef Belgian Blue | 8 | 0 | 0 |
| Dual-Purpose Blue | 0 | 16 | 0 |
| Holstein | 0 | 0 | 8 |

# 4 Discussion and conclusion

This study focused on the ability of a breed assignment test to differentiate meat of DPB, a local transboundary breed, from the meat of BBB, a genetically related beef breed. Results showed the high accuracy of this test as the breed of origin of all meat pieces was perfectly determined. Even for DPB meat pieces corresponding to the “mh/mh” genotype that is shared with the BBB breed, the probability of belonging to the predicted breed reached the maximum value of one. This result was better than those obtained by Dalvit et al. (2008) where only 52.5% of the tested individuals were correctly breed-assigned with a probability of at least 90%. However, they were using microsatellites and not SNPs in their study. Another example of a study that targeted the differentiation of meat products of closely related breeds is provided by Kuehn et al. (2011) who used a regression of allelic frequencies of 16 breeds to determine breed composition of animals. In their study, the distinction between Angus and Red-Angus samples was not straightforward and both breeds were pooled together to increase the accuracy of the model. However, Angus and Red-Angus started to be recorded in different herd books in the 1950s (Márquez et al., 2010), that is about two decades before the divergence of DPB and BBB. Moreover, there is a clear distinction concerning the colour pattern of Angus and Red-Angus while black-pied, blue roan and white animals exist in both DPB and BBB. It should be highlighted that the results obtained by Kuehn et al. (2011) may not only be related to the used methodology but also to the fact that breeding goals of Angus and Red Angus have been similar (as they are both beef breeds) whereas DPB and BBB have had divergent breeding goals since their separation.

As explained previously, in our study, 167 SNPs out of 2005 SNPs were used by the developed model to distinguish the DPB from other reference breeds, and especially from the closely related BBB breed. The fact that the SNP panel used was specifically fitted to differentiate both breeds partially explained the high accuracy obtained with our model. However, as illustrated by the different allelic frequencies for the seven SNPs that were used to discriminate both DPB and BBB, the divergence of both breeds and of their breeding goals for more than 50 years has also led to an accumulation of phenotypic and genetic differences between DPB and BBB, and this, even for the "mh/mh" subpopulation of DPB. For example, figures of milk production of DPB animals support the fact that the "mh/mh" genotype and dairy performances are not mutually exclusive (Colinet, 2010). Actually, there is only a difference of 1000 litres of milk yield between cows with "+/+" genotypes, producing on average 5000 litres/year, and cows with "mh/mh" genotypes, producing on average 4000 litres/year (~~BlueSter, 2021~~BlueSter, 2022). Another example is the frequency of caesarean sections which was demonstrated to occur heavily for calving of BBB, while it was less common for calving of DPB (Mota et al., 2017). The rationale is that, even for "mh/mh" carrying animals, DPB breeders have been selecting for lighter calves and easier calving.

Even if the high probabilities obtained in this study were an important aspect to focus on, it should not be forgotten that the used model is not able to detect breeds that did not fall under the scope of study (Maudet et al., 2002). Other elements, namely pedigree, herd book records and visual appraisal of an animal's phenotype on farm, will ensure that the animal is belonging to one of the Blue breeds: DPB or BBB. Then, these elements will be combined with the genomic breed assignment test, aiming to distinguish DPB from BBB breeds, to ensure the consumer that the meat is really drawn from DPB. This was possible with a high accuracy, as shown by the results. Therefore, it seems feasible to implement a certification process of DPB meat based on the developed breed assignment test.

Even if the number of SNPs dedicated to differentiate the DPB meat samples is lower than the number of SNPs recommended by Judge et al. (2017) (i.e., 167 *vs.* 300), it is thought that the number of SNPs can decrease even more when the selection of SNPs is based on a high-density chip (Hulsegge et al., 2013). ~~However, t~~The use of high-density chip is currently still limited, especially for local breeds, ~~even if~~but the cost decrease of high-density genotyping would presuppose it would be feasible in the next decade. However, the proposed test model is reliable and adapted to the closely related breeds under study. It is easy to use and to interpret thanks to a confusion matrix. It can be easily adapted to other breeds and it is recommended for further research considering the same topic to use the PLS-DA for selection of SNPs and the nearest shrunken centroids method to assign meat samples to their breed.

## Ethical approval

The SNP data in this study were sampled from meat pieces. Animals were slaughtered following the European regulation 1099/2009 on the protection of animals at the time of killing.

## Consent for publication

Not applicable.

## Data availability statement

The data supporting the findings of this study cannot be made available as a whole. The corresponding author, upon reasonable request, will forward request to relevant data owners.

## Orcid

Hélène Wilmot https://orcid.org/0000-0002-1075-088X

Géry Glorieux https://orcid.org/0000-0003-0380-2837

Nicolas Gengler https://orcid.org/0000-0002-5981-5509

## CRediT authorship contribution statement

**H. Wilmot:** Conceptualization, Methodology, Software, Investigation, Data curation, Writing – original draft. **G. Glorieux:** Conceptualization, Investigation, Writing – review & editing. **X. Hubin:** Conceptualization, Writing –

review & editing. **N. Gengler:** Conceptualization, Writing – review & editing, Supervision.

## Declaration of Competing Interest

The authors declare that they have no competing interests.

## Q2 Acknowledgements

H. Wilmot, as a current Research Fellow, and N. Gengler, as a former Senior Research Associate, acknowledge the support of the Fonds de la Recherche Scientifique – FNRS (Brussels, Belgium). The authors thank the INTERREG V France-Wallonie-Vlaanderen program and the Walloon Government (Service Public de Wallonie – Direction Générale Opérationnelle Agriculture, Ressources Naturelles et Environnement, SPW-DGARNE; Namur, Belgium) for their financial support through the BlueSter project and previous projects. The authors also acknowledge the technical support by the Walloon Breeders Association (elevéo). The University of Liège – Gembloux Agro-Bio Tech (Liège, Belgium) supported computations through the technical platform Calcul et Modélisation Informatique (CAMI) of the TERRA Teaching and Research Centre, partly supported by the Fonds de la Recherche Scientifique – FNRS under grant no. T.0095.19 (PDR "DEEPSELECT"). Genotyping was facilitated through the support of the Fonds de la Recherche Scientifique – FNRS under grant no. J.0174.18 (CDR "PREDICT-2″).

## References

ⓘ The corrections made in this section will be reviewed and approved by a journal production editor. The newly added/removed references and its citations will be reordered and rearranged by the production team.

Alexander, D.H., Novembre, J., Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 19 (9), 1655–1664. doi:10.1101/gr.094052.109.

Baumung, R., Cubric-Curik, V., Schwend, K., Achmann, R., Sölkner, J., 2006. Genetic characterisation and breed assignment in Austrian sheep breeds using microsatellite marker information. J. Anim. Breed. Genet. 123 (4), 265–271. doi:10.1111/j.1439-0388.2006.00583.x.

Bertolini, F., Galimberti, G., Calò, D.G., Schiavo, G., Matassino, D., Fontanesi, L., 2015. Combined use of principal component analysis and random forests identify population-informative single nucleotide polymorphisms: application in cattle breeds. J. Anim. Breed. Genet. 132 (5), 346–356. doi:10.1111/jbg.12155.

BlueSter. (2021). BlueSter. https://www.projet-bluester.eu/

Charlier, C., Coppieters, W., Farnir, F., Grobet, L., Leroy, P.L., Michaux, C., Mni, M., Schwers, A., Vanmanshoven, P., Hanset, R., Georges, M., 1995. The mh gene causing double-muscling in cattle maps to bovine Chromosome 2. Mamm. Genome 6 (11), 788–792. doi:10.1007/BF00539005.

Colinet, F., 2010. The dual purpose Belgian Blue. Brochure. University of Liege.

Dalvit, C., De Marchi, M., Cassandro, M., 2007. Genetic traceability of livestock products: a review. Meat Sci. 77 (4), 437–449. doi:10.1016/j.meatsci.2007.05.027.

Dalvit, C., Marchi, M.De, Targhetta, C., Gervaso, M., Cassandro, M, 2008. Genetic traceability of meat using microsatellite markers. Food Res. Int. 41, 301–307. doi:10.1016/j.foodres.2007.12.010.

Funkhouser, S.A., Bates, R.O., Ernst, C.W., Newcom, D., Steibel, J.P., 2017. Estimation of genome-wide and locus-specific breed composition in pigs. Transl. Anim. Sci. 1 (1), 36–44. doi:10.2527/tas2016.0003.

He, J., Guo, Y., Xu, J., Li, H., Fuller, A., Tait, R.G., Wu, X.-.L., Bauck, S., 2018. Comparing SNP panels and statistical methods for estimating genomic breed composition of individual animals in ten cattle breeds. BMC Genet. 19. doi:10.1186/s12863-018-0654-3.

Hulsegge, B., Calus, M.P.L., Windig, J.J., Hoving-Bolink, A.H., Maurice-van Eijndhoven, M.H.T., Hiemstra, S.J., 2013. Selection of SNP from 50K and 777K arrays to predict breed of origin in cattle. J. Anim. Sci. 91 (11), 5128–5134. doi:10.2527/jas.2013-6678.

Iquebal, M.A., Ansari, M.S., Sarika, S., Dixit, S.P., Verma, N.K., Aggarwal, R.A.K., Jayakumar, S., Rai, A., Kumar, D., 2014. Locus minimization in breed prediction using artificial neural network approach. Anim. Genet. 45 (6), 898–902. doi:10.1111/age.12208.

Josse, J., Husson, F., 2012. Handling missing values in exploratory multivariate data analysis methods. J. Soc. fr. Stat. 153 (2), 77–99.

Judge, M.M., Kelleher, M.M., Kearney, J.F., Sleator, R.D., Berry, D.P., 2017. Ultra-low-density genotype panels for breed assignment of Angus and Hereford cattle. Anim 11 (6), 938–947. doi:10.1017/S1751731116002457.

Kuehn, L.A., Keele, J.W., Bennett, G.L., McDaneld, T.G., Smith, T.P.L., Snelling, W.M., Sonstegard, T.S., Thallman, R.M., 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 (6), 1742–1750. doi:10.2527/jas.2010-3530.

Márquez, G.C., Speidel, S.E., Enns, R.M., Garrick, D.J., 2010. Genetic diversity and population structure of American Red Angus cattle. J. Anim. Sci. 88 (1), 59–68. doi:10.2527/jas.2008-1292.

Maudet, C., Luikart, G., Taberlet, P., 2002. Genetic diversity and assignment tests among seven French cattle breeds based on microsatellite DNA analysis. J. Anim. Sci. 80 (4), 942–950. doi:10.2527/2002.804942x.

Mota, R.R., Mayeres, P., Bastin, C., Glorieux, G., Bertozzi, C., Vanderick, S., Hammami, H., Colinet, F.G., Gengler, N., 2017. Genetic evaluation for birth and conformation traits in dual-purpose Belgian Blue cattle using a mixed inheritance model. J. Anim. Sci. 95 (10), 4288–4299. doi:10.2527/jas2017.1748.

Putnová, L., Štohl, R., 2019. Comparing assignment-based approaches to breed identification within a large set of horses. J. Appl. Genet. 60 (2), 187–198. doi:10.1007/s13353-019-00495-x.

R Core Team. (2021). *R: a* language and environment for statistical computing. (3.6.3). R Foundation for Statistical Computing. http://www.r-project.org/

R Studio Team. (2020). RStudio: integrated development for R. RStudio. PBC. http://www.rstudio.com/

Schiavo, G., Bertolini, F., Galimberti, G., Bovo, S., Dall'olio, S., Nanni Costa, L., Gallo, M., Fontanesi, L, 2019. A machine learning approach for the identification of population-informative markers from high-throughput genotyping data: application to several pig breeds. Anim. 223–232. doi:10.1017/S1751731119002167.

Tibshirani, R., Hastie, T., Narasimhan, B., Chu, G., 2002. Diagnosis of multiple cancer types by shrunken centroids of gene expression. PNAS 99 (10), 6567–6572. doi:10.1073/pnas.082099299.

Wilkinson, S., Wiener, P., Archibald, A.L., Law, A., Schnabel, R.D., McKay, S.D., Taylor, J.F., Ogden, R., 2011. Evaluation of approaches for identifying population informative markers from high density SNP Chips. BMC Genet. 12. doi:10.1186/1471-2156-12-45.

Wilmot, H., Bormann, J., Soyeurt, H., Hubin, X., Glorieux, G., Mayeres, P., Bertozzi, C., Gengler, N., 2022. Development of a genomic tool for breed assignment by comparison of different classification models - Application to three local cattle breeds. J. Anim. Breed. Genet. 139 (1), 40–61. doi:10.1111/jbg.12643.

## Highlights

- A genomic breed assignment model was tested on meat of the dual-purpose blue breed.
- Meat of dual-purpose blue was perfectly assigned to its breed of origin.
- It means that this breed could be distinguished from the related beef belgian blue.
- Traceability of meat of dual-purpose blue by using this tool seems feasible.

## Queries and Answers

Q1

**Query:** Please confirm that given names and surnames have been identified correctly.

**Answer:** Yes

Q2

**Query:** [FND001] Correctly acknowledging the primary **funders and grant IDs** of your research is important to ensure compliance with funder policies. Please make sure that funders are mentioned accordingly.

**Answer:** Yes, funders were correctly acknowledged.