



## Practical applications of genomic assignment methods in the avian breeding industry

Romuald Rouger, Marc Teissier, Sophie Brard-Fudulea, Noémie Thébault, Juliette Riquet, Sophie Leroux, Alain Vignal, Bertrand Bed'Hom, Xavier X Rognon, Sophie Lumineau, et al.

### ► To cite this version:

Romuald Rouger, Marc Teissier, Sophie Brard-Fudulea, Noémie Thébault, Juliette Riquet, et al.. Practical applications of genomic assignment methods in the avian breeding industry. XIth European symposium on poultry genetics (ESPG), Oct 2019, Prague, Czech Republic. hal-02460150

**HAL Id: hal-02460150**

**<https://hal.science/hal-02460150>**

Submitted on 29 Jan 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Practical applications of genomic assignment methods in the avian breeding industry

Romuald Rouger<sup>1</sup>, Marc Teissier<sup>1</sup>, Sophie Brard-Fudulea<sup>1</sup>, Noémie Thébault<sup>2</sup>, Juliette Riquet<sup>2</sup>, Sophie Leroux<sup>2</sup>, Alain Vignal<sup>2</sup>, Bertrand Bed'hom<sup>3</sup>, Xavier Rognon<sup>3</sup>, Sophie Lumineau<sup>4</sup>, Christian Diot<sup>5</sup>, Bernard Alletru<sup>6</sup>, Maude Cornil<sup>6</sup>, Pascal Bouleau<sup>6</sup>, Magali Blanchet<sup>7</sup>, Guillaume Le Mignon<sup>7</sup>, Olivier Demeure<sup>7</sup>, Daniel Guémené<sup>1</sup>

<sup>1</sup>SYSAAF, INRA, Centre Val-de-Loire, 37380 Nouzilly, France, <sup>2</sup>GenPhySE, Université de Toulouse, INRA, ENVT, 31326 Castanet Tolosan, <sup>3</sup>GABI, INRA, AgroParisTech, Université Paris-Saclay, Domaine de Vilvert, 78352 Jouy en Josas, <sup>4</sup>Univ Rennes, Normandie Univ, CNRS EthoS UMR6552, 35042 Rennes, <sup>5</sup>PEGASE, INRA, Agrocampus Ouest, 35590 Saint-Gilles, France <sup>6</sup>Gourmaud Sélection, Groupe Orvia, 85260 Saint Andre 13 Voies, France <sup>7</sup>Grimaud Frères Sélection, Groupe Grimaud, La Corbière, 49450 Roussay, France

Corresponding author: romuald.rouger@inra.fr

**Keywords:** Assignment, Parentage, Hybrids, Genomics, SNP

Genomic data of any kind of organism has never been easier to obtain. For poultry breeders, introducing genomic information in their selection scheme offers direct benefits. Besides genomic selection, assignment methods are also useful to address current questions in poultry breeding. In this communication, we present two practical uses of genomic assignment methods. In the first example, we detail the development of a parentage assignment panel in ducks. In the second example, we used molecular information in a population assignment test to discriminate Common quail individuals from hybrids with the Japanese quail. Further than these two cases, assignment methods are promising tools in order to trace genealogies in alternative breeding systems or to validate the origin of commercial crossbred products.

## Introduction

Recent advances in sequencing technologies have made easy and cheap the gathering of large genomic datasets for non-model organisms. In poultry, genomic data are becoming available for species that had limited molecular resources so far (e.g. Guinea fowl: Vignal *et al.* 2019, Japanese quail: Recoquillay *et al.* 2015). As a result, the prospects brought by genomic selection are not restricted to the broiler and layer industry only. Poultry species with lower sales volumes may also be entitled to this technology.

Besides genomic selection, other use of genomic information may help solving current challenges faced by the avian breeding industry. Among those, genomic assignment methods have direct applications. By genomic assignment method, we designate any method using probability theory applied on molecular data to assign the genotype of an individual to (1) its parents or (2) a reference population. The goal of this communication is to present examples of the use of genomic assignment methods to address two hot topics in non-model poultry species.

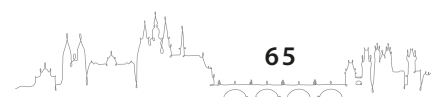
The first example addresses the growing interest of the society concerning ethical views on confined rearing systems. For decades, cage rearing was the preferred system in the breeding industry as it permits to follow accurately the pedigree of any individuals. The impact of such system on animal welfare is now well understood (Duncan 2001) and consumer views regarding this are rapidly changing. We present here the results of an ongoing project aiming to develop a SNP panel for parentage assignment within the Muscovy duck, the Pekin duck and their hybrid the Mule duck. The expected outcome of this project is to propose a tool useable in alternative breeding systems without penalizing the activity of selection.

The second example deals with the genetic impact of releasing captive individuals in wild game bird populations. In Spain or Italy, restocking using captive birds is a common practice to supplement native populations of the Common quail *Coturnix coturnix*. These restocking birds are often hybrids between the wild Common quail and the domestic Japanese quails, *Coturnix japonica*, which is easier to maintain in captivity. High proportions of hybrids in the wild may have detrimental effects on native populations of the Common quail (Puigcerver *et al.* 2014). We report here the development of a SNP panel aiming to indicate the genetic identity of any given individual (*C. coturnix* / *C. japonica* / Hybrid between the two species). This panel may ultimately serve as a test to control the genetic integrity of the individuals used for restocking.

## Parentage assignment in duck species

In France, the production of foie gras is largely made using Mule ducks. This duck is a single-cross hybrid between a Muscovy drake (*Cairina moschata*) and a Pekin duck (*Anas platyrhynchos*). A high-density SNP chip has recently been developed for these two parental species yielding 240 656 and 245 192 SNPs for the Pekin duck and the Muscovy duck respectively. The genotype of 240 *A. platyrhynchos*, 180 *C. moschata*, and 79 Mule ducks were available. The aim of this work was to use this genomic dataset to validate the minimal set of markers powerful enough for parentage assignment within each species. The same task was done using polymorphic markers shared between the two species (687 SNPs) with the aim to design a multi-species (Muscovy, Pekin and Mule duck) parentage assignment panel. A three-step approach was used to achieve these goals.

The first step was to select autosomal SNPs based on their intrinsic qualities. Markers were chosen according to (1) a threshold value concerning the minor allele frequency (MAF) and (2) their distribution along the genome. Sets of respectively 48, 96 or 192 markers were constituted to test the effect of the number of marker on assignment results. As a second step, the statistical power of each set of markers was determined using their probability of parentage exclusion and their probability of identity (Tortereau *et al.* 2017). The probability of parentage exclusion is the probability of excluding one or two parents knowing they are not the



real parents. This probability must be close to one. The probability of identity is the probability that two random individuals share the same genotype. This probability must be close to zero. In the final step, validation of each set of markers is made by testing their efficiency on real data. Parents-offspring trios were tested within both species and for single cross hybridization (68 trios for *A. platyrhynchos*, 36 trios for *C. moschata* and 79 hybrid crosses). The performance of three assignment methods were compared, the Likelihood-Ratio test (LR, Tortereau *et al.* 2017) the Empirical Distribution of Likelihoods (EDL, Boichard *et al.* 2014) and the Auto-adaptive Parentage Inference method (APIS).

Within each species and for any panel sizes, probabilities of exclusion are higher than 0,99 and probabilities of identity are very close to 0. Probabilities of exclusion are very often higher than 0,999 when panels of 96 or 192 markers are used. The MAF threshold and the repartition of markers along the genome doesn't have a strong effect on these probabilities. Validation on real data shows that assignment success is very close to its maximal possible value when panels of 96 markers are used. Very little differences are observed concerning the performance of each tested assignment method. The only noticeable difference resides in the lower probability by the EDL method to wrongly assign an individual when one parental genotype is missing.

The number of polymorphic markers shared between *A. platyrhynchos*, *C. moschata* and Mule duck was small (687 SNPs). Stringency levels over the choice of markers were therefore adjusted in order to keep enough markers in a multispecies panel (lower MAF threshold, no filter upon the repartition of the markers along the genome) without affecting the assignment performance. Probability of identity is above 0,999 and probability of identity is very close to 0 when a panel of 96 markers is used. Validation on real data showed that performance of assignment was almost maximal with 96 markers.

Overall, this work indicates that a single multispecies panel of 96 markers is the most flexible tool for parentage assignment in Muscovy, Pekin and Mule duck. Tests were carried out to include two more breeds of common ducks in a single parentage assignment panel: a mallard breed used as game bird and the Rouen duck. Unfortunately, the number of polymorphic markers conserved across all breeds was too small to get a single parentage assignment panel.

### Population assignment: How wild is this quail (Rouger *et al.* 2019)?

For this work, genomic data generated in a previous project for 20 *C. japonica* and 2 *C. coturnix* were used (ANR 009-GENM-008, Recoquillay *et al.* 2015). Identified polymorphic markers were filtered to keep positions along the genome for which an allele appeared as fixed in the 20 *C. japonica* while the other appeared as fixed in the 2 *C. coturnix*. Within this subset, 192 markers were targeted based on (1) genotyping quality controls and (2) their position along the genome. Samples of the two species originating from a wider panel of origins were genotyped at these 192 markers (*C. japonica*: 91 individuals from 8 populations, *C. coturnix*: 86 individuals from 4 populations). These samples served as reference for the two species in subsequent analysis.

The genetic differentiation between the two species estimated with these 192 markers was high ( $F_{ST}=0.80$ ). The 96 markers showing the most differentiation were selected to constitute the assignment panel. Among these 96 markers, 32 of them were diagnostics (i.e. an allele is fixed *C. japonica* while the other allele is fixed in *C. coturnix*). A probabilistic approach was developed to determine the genetic status of any given individual knowing its genotype. Briefly, this method gives the Bayesian probability that an individual belongs to the following different categories: *C. japonica*, *C. coturnix*, F1 hybrids, any generation of F hybrids, 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> generations of backcrosses to *C. coturnix*, 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> generations of backcrosses to *C. japonica*. This method only uses allelic frequencies as observed in the reference samples of the two species. An assignment probability higher than 0,9 for an individual in a given group is considered as true assignment.

The robustness of the method was first evaluated using simulated data. The genotype of individuals belonging to each possible category were simulated (100 individuals per category). The assignment method was then asked to re-assign simulated individuals by calculating their probability of belonging to each category. Assignment success was maximal for simulated individuals known to be *C. japonica*, *C. coturnix*, F1 hybrids or any generation of F hybrids. Concerning backcross individuals, the method failed to make a clear distinction between 2<sup>nd</sup> or 3<sup>rd</sup> generations of backcrosses. However, the method was able to distinguish 3<sup>rd</sup> generation backcrosses to *C. coturnix* from reference *C. coturnix* with an error rate of only 1%.

The assignment panel was subsequently tested on sixty real individuals of known pedigree. These individuals were hybrids obtained at the INRA Val-de-Loire, France (20 F1, 20 F2 and 20 F7). The panel was able to assign all these individuals to the category they belong, apart from one F7 individuals for which no assignment probability was above the 0,9 threshold.

Overall, this work has produced a molecular panel able to tell apart *C. coturnix* individuals from *C. japonica* and hybrids between the two species. This panel has already been used to characterize the genetic status of 22 quails reared in Spain and designated as Common quails. Our analysis showed that only 13 of them could be considered as *C. coturnix* while the others were hybrids. This panel may ultimately be used as a control tool in countries where regulations apply over the rearing of Common quails.

### Conclusion

Of course, the transition towards alternative breeding system will require to solve many questions (relative contribution of males, effects on fertility, effects of competition between males on animal welfare...). However, the use of parentage assignment methods, as presented here in ducks, is a mandatory step of this transition. Similarly, knowing the population of origin of a commercial product is of primary importance for some certification (e.g. Label Rouge). Applications of population assignment tests, like the one presented here, is perfectly suited to validate/control the genetic identity of any given samples. To conclude, the idea of using assignment method in the breeding industry is not new, however, the ease with which large amount of genomic information can be obtained allows the extensive use of these methods, even in species considered as "non-model".

## Acknowledgments

Authors thank Ronan Griot to make the Auto-adaptive Parentage Inference software available for our study. The parentage assignment panel in duck was part of the CanArray project, supported by the competitiveness clusters Valorial and Agri-Sud Ouest, and funded by Région Pays de la Loire-Bretagne, Région Nouvelle Aquitaine, and AgenAvi. Thank you to Eric Poullain for providing samples. AgenAvi funded parts of the quail project (SNP-Cailles).

**BOICHARD, D., BARBOTTE, L. and GENESTOUT, L.** 2014. AccurAssign, software for accurate maximum-likelihood parentage assignment. August 2014, Proceedings of the 10th world congress on genetics applied to livestock production, pp. 17–22. .

**DUNCAN, I. J. H.** 2001. The pros and cons of cages. *World's Poultry Science Journal* **57**: 381–390.

**PUIGCERVER, M., SANCHEZ-DONOSO, I., VILÀ, C., SARDÀ-PALOMERA, F., GARCÍA-GALEA, E. and RODRÍGUEZ-TEJEIRO, J. D.** 2014. Decreased fitness of restocked hybrid quails prevents fast admixture with wild European quails. *Biological Conservation* **171**: 74–81.

**RECOQUILLAY, J., PITEL, F., ARNOULD, C., LEROUX, S., DEHAIS, P., MORÉNO, C., CALANDREAU, L., BERTIN, A., GOURICHON, D., BOUCHEZ, O., VIGNAL, A., FARIELLO, M., MINVIELLE, F., BEAUMONT, C., LETERRIER, C. and LE BIHAN-DUVAL, E.** 2015. A medium density genetic map and QTL for behavioral and production traits in Japanese quail. *BMC Genomics* **16**: 10.

**ROUGER, R., BRARD-FUDULEA, S., THÉBAULT, N., PITEL, F., VIGNAL, A., BED'HOM, B., ROGNON, X., LUMINEAU, S. and GUÉMENE, D.** 2019. Développement d'un panel de marqueurs SNP pour détecter l'hybridation entre la caille des blés et la caille japonaise. *13èmes Journées de la Recherche Avicole et Palmipèdes à Foie Gras*, Tours, pp80-84.

**TORTEREAU, F., MORENO, C. R., TOSSER-KLOPP, G., SERVIN, B. and RAOUL, J.** 2017. Development of a SNP panel dedicated to parentage assignment in French sheep populations. *BMC Genetics* **18**: 50.

**VIGNAL, A., BOITARD, S., THÉBAULT, N., DAYO, G.-K., YAPI-GNAORE, V., KARIM, I. Y. A., BERTHOULY-SALAZAR, C., PÁLINKÁS-BODZSÁR, N., GUÉMÉNÉ, D., THIBAUD-NISSEN, F., WARREN, W. C., TIXIER-BOICHARD, M. and ROGNON, X.** 2019. A guinea fowl genome assembly provides new evidence on evolution following domestication and selection in galliformes. *Molecular Ecology Resources* **19**: 997–1014.

# Global biodiversity genomics and birds: a general overview, and case studies of functional diversity in immune genes in ducks

Robert H. S. Kraus<sup>1,2</sup>

<sup>1</sup>Department of Migration, Max Planck Institute of Animal Behavior, 78315 Radolfzell, Germany; <sup>2</sup>Department of Biology, University of Konstanz, 78464 Konstanz, Germany

Corresponding author: rkraus@ab.mpg.de

**Keywords:** evolution, animal breeding, domestication, ornithology, functional variation

## Introduction

The human population has grown dramatically over the past centuries and decades. Globally, food and space are becoming increasingly scarce resources. Without going into the depth of this multi-faceted problem I here want to highlight an aspect of why fundamental biodiversity research has potential to contribute to easing this problem. Due to an increasing world population more food needs to be produced on less space. A simple consequence is a necessary increase in efficiency (I will not touch upon ethical questions concerning animal husbandry and related topics).

For yield to increase, agriculture has tried standardisation via monoculture but suffers the risks of a low diversity, high density strategy. For example, clonal production of banana was highly efficient for a variety called Gros Michel until the 1960s. Then, a fungal disease (Panama Disease) wiped it out nearly completely. Today, the Cavendish banana is the dominant clone. But history repeats itself and today's banana production is immensely threatened by a newly emerging fungal disease called Black Sigatoka (Marín et al., 2003).

The generally accepted opinion among plant and animal breeders is that breeding for diversity, incorporating knowledge and resources from seed banks, fancy breeds and wild ancestors, is a going to be a major factor (e.g., Fu, 2015). A high diversity in global breeding stocks is thought to increase not only the availability of locally adapted breeds with increased yield under local conditions but also higher resilience against change and thereby increased sustainability.

## Biodiversity

Accompanying the human genome project Kruglyak and Nickerson (2001) wrote a seminal roundup on genetic variation: "Variation is the spice of life". In it, they highlight the importance of understanding every single SNP in the human genome and what could be achieved by describing genetic variation down to the level of the nucleotide. This variation stems from the process of mutation and is maintained or acted against by selection. The result is organismal evolution leading to the diversity of form and function. The term biodiversity is a modern fusion of biology and diversity and comes from the nature conservation community. An early promoter of the term and its ideas and context to a broader public was E. O. Wilson with his book "The diversity of life" (Wilson, 1992). Defining biodiversity and what the concept entails is a broad field and there are whole books attempting to do justice to it (Maclaurin and Sterelny, 2008) – but I here refer to it in a rather technical way and in the context of genetic variation, particularly on the genomic level. In the next sections I will explain how genomics of avian model species and non-model species (Vignal and Eöry, 2019) come together via technical advances and what can be gained in the future on both sides.

