References

BLESBOIS, E., LIPTOI, K., WOELDERS, H., BERNAL-JUAREZ, B., NANDI, S., THELIE, A., PIZZI, F., VARKONYI, E., GUIGNOT, F., MCGREW, M., SANTIAGO-MORENO, J. (2019). Breakthrough to improve the reproductive capacity of gene bank material. Meeting of the European Federation of Animal Science. 26-30 August 2019, Ghent Belgium. Session 33. Burning issues in biodiversity 1: what are the benefits from animal gene banks? Presentation 33.05.

BORTOLUZZI, C., CROOIJMANS, R.P.M.A., BOSSE, M., HIEMSTRA, S.J., GROENEN, M.A.M., MEGENS, H.J. (2018). The effects of recent changes in breeding preferences on maintaining traditional Dutch chicken genomic diversity. Heredity **121**:564-578. https://doi.org/10.1038/s41437-018-0072-3

DE OLIVEIRA SILVA, R., VOSOUGH AHMADI, B., HIEMSTRA, S.J., MORAN, D. (2018). Optimizing ex situ genetic resource collections for European livestock conservation. Volume 136, Issue 1, January 2019, Pages 63-73. https://doi.org/10.1111/jbg.12368.

FAO (2018). Status and trends of animal genetic resources. Intergovernmental Technical Working Group on Animal Genetic Resources for Food and Agriculture. Rome, 27 – 29 June 2018 http://www.fao.org/3/CA0121EN/ca0121en.pdf

LEROY, G., BOETTCHER, P., BESBES, B., DANCHIN-BURGE, C., BAUMUNG, R., HIEMSTRA, S.J. (2019). Cryoconservation of animal genetic resources: a gap analysis. Submitted.

KUGLER, W., CHARVOLIN, E., KHAYATZADEH, N., MESZAROS, G., BERGER, B., MARTYNIUK, E., SÖLKNER, J., WURZINGER, M., A. DORE (2019). Stakeholder involvement and ethics in conservation of animal genetic resources in gene banks. Meeting of the European Federation of Animal Science. 26-30 August 2019, Ghent Belgium. Session 33. Burning issues in biodiversity 1: what are the benefits from animal gene banks? Presentation 33.01.

PASSEMARD, A.S., JOLY, L., DUCLOS, D., DANCHIN-BURGE, C. (2017). Inventory and mapping of European animal genetic collections. Report. Institut de L'Elevage. http://www.imageh2020.eu/deliverable/WP2.pdf

RESTOUX, G., ROGNON, X., VIEAUD, A., GUEMENE, D., CHIRON, G., PETITJEAN, F., SEIGNEURIN, F., VASILESCU, A., TIXIER-BOICHARD, M. (2018). Genetic characterization of French local chicken breeds. 11th World Congress of Genetics Applied to Livestock, Auckland, New-Zealand.

SCHURINK, A., HULSEGGE, I., WINDIG, J.J., HOVING, A.H., HIEMSTRA, S.J. (2019) Dutch gene bank for farm animals: beyond conserving genetic material. 70th Annual Meeting of the European Federation of Animal Science. 26-30 August 2019, Ghent Belgium. Session 33. Burning issues in biodiversity 1: what are the benefits from animal gene banks? Poster 33.12.

ZOMERDIJK, F, HIEMSTRA, S.J., D'ARBAUMONT, M., TIXIER-BOICHARD, M., BOETTCHER, P. (2019). Global survey on the quality management practices of gene banks for livestock. Submitted.

From wild populations to global diversity – Insight into the Synbreed Chicken Diversity Panel

Dorcus Kholofelo Malomane[1,2], Steffen Weigend[2,3], Armin Otto Schmitt[2,4], Annett Weigend[3], Christian Reimer[1,2], Henner Simianer[1,2]

[1]Animal Breeding and Genetics Group, Department of Animal Sciences, University of Göttingen, Göttingen, Germany; [2]Center for Integrated Breeding Research, Department of Animal Sciences, University of Göttingen, Göttingen, Germany; [3]Institute of Farm Animal Genetics, Friedrich-Loeffler-Institut, Neustadt, Göttingen; [4]Breeding Informatics Group, Department of Animal Sciences, Georg August-Universität Göttingen, Göttingen, Germany

Corresponding author: dmaloma@gwdg.de

Keywords: Synbreed Chicken Diversity Panel, global chicken diversity, genetic isolation by distance

Introduction

When a small group of individuals migrate from a single large founder population to establish their own colony, this results in reduction of genetic diversity by genetic drift (Provine 2004). Furthermore, theory of genetic isolation by distance (Malécot 1969) predicts that the new population will genetically differentiate from the founder population. The theory of genetic isolation by distance refers to cases where the genetic differentiation increases with the geographic distance between populations. This is because the exchange of genetic material between the populations (i.e. mating opportunities) is confined by the distance (Cavalli-Sforza *et al.* 1964, Malécot 1969). The consequences of genetic drift act more rapidly to differentiate the populations than the potential or chances of an individuals' interaction under dispersal (Aguillon *et al.* 2017). Consistent with that, a model of expansion from a single founder predicts that patterns of genetic diversity in populations can be well explained by their geographic expansion from the founders, which is correlated to the genetic differentiation (Prugnolle *et al.* 2005, Ramachandran *et al.* 2005, Deshpande *et al.* 2009). Starting from the chicken wild populations, we aimed at investigating the patterns of genetic diversity in the global set of chicken breeds, represented by the Synbreed Chicken Diversity Panel (SCDP) (Malomane *et al.* 2019). The SCDP is a collection of various chicken breeds with various backgrounds from different parts of the world. In the SCDP the geographical location of the sampling often does not coincide with the geographical location of the breed development, since e.g. many samples of Asian breeds were collected from German fancy breeders. Therefore, we used the genetic distance of the sampled breeds to the wild ancestors as a proxy for geographic distance, and verified, whether the reduction of diversity also can be found with increasing genetic distance to the domestication center.

Material and Methods

In this study, data taken from the SCDP consisted of 3,235 chicken individuals from 172 domesticated chicken populations and two wild populations (*Gallus gallus gallus* and *Gallus gallus spadiceus*). The chicken samples were collected in Asia, Africa, South America and Europe, and were genotyped with the 600K Affymetrix® Axiom™ Genome-Wide Chicken Genotyping Array (Kranis

et al. 2013). The populations were classified into twelve breed categories which were based on their continent of origin and/or type. After quality control and filtering of the data, including LD based SNP pruning to account for ascertainment bias (Malomane *et al.* 2018), a total of 156,753 SNPs were left to conduct our analyses. We estimated the pairwise Reynolds' genetic distances between the two wild type populations (*G. g.* ssp.) and the domesticated populations, and then calculated the mean genetic distance of each domesticated population to the two wild populations. Furthermore, observed heterozygosity was estimated within each population. Then, we estimated the relationship between the overall genetic diversity within the domesticated populations and their mean genetic distances to the two wild type populations.

Results and Discussion

The relationship between the overall genetic diversity of domesticated chickens and genetic distance to wild populations

We observed a strong negetive relationship between the genetic diversity within populations and their genetic distances to the wild populations (Figure 1). The genetic distance to the wild populations explained 87.5% of the total variation in the heterozygosity within the domesticated chickens. This figure is slightly higher than those obtained in several human studies when using geographic distances (Prugnolle *et al.* 2005, Ramachandran *et al.* 2005, Li *et al.* 2008). Furthermore, studies in humans have shown that there is a high correlation between the genetic distances (using different genetic distance measures) and geographic distance. However the correlations were not as high in domesticated cattle as in humans. For example, a correlation of 0.624 was reported by Wang (2015), and, while Scheu *et al* (2015) reported a correlation of 0.750 for ancient cattle samples, the correlation was 0.540 in modern cattle samples. The weakening relationship between geographic and genetic distances in modern domesticated cattle was suggested to be due to the human manipulation of the genetic diversity among other reasons, as it is with many domesticated livestock (Scheu *et al.* 2015). In this study we were not able to validate the relationship between the geographic and genetic distances due to the lack of geographic sampling coordinates. However, our results show that the geographic distance alone may not well predict the observed genetic variations in the chickens because:

i. breeds of the same geographic origin are found scattered across the genetic diversity spectrum, with the main difference being the breeding and management practices and presumably different population sizes. This is the case for Asian (red symbols) and European (green symbols) type breeds. The Asian and European chickens sampled from the German fancy breeders (denoted with prefix DE_ and triangle shapes) have highly reduced genetic diversity as well as higher genetic distance to the wild chickens (*G.gallus*) than their respective local breeds. Therefore, when considering the sampling areas, the genetic diversity may correlate with the geographic distances to the origin of *G.gallus* in the Asian breeds but not in the European breeds. Not only have many of the fancy breeds been started with a small number of breeding birds imported from Asia to Europe, but they have also undergone subsequent selection, with small effective population sizes and intended inbreeding to keep the desired traits. Therefore, such practices could be responsible for most variation in genetic diversity of these Asian and European type breeds.

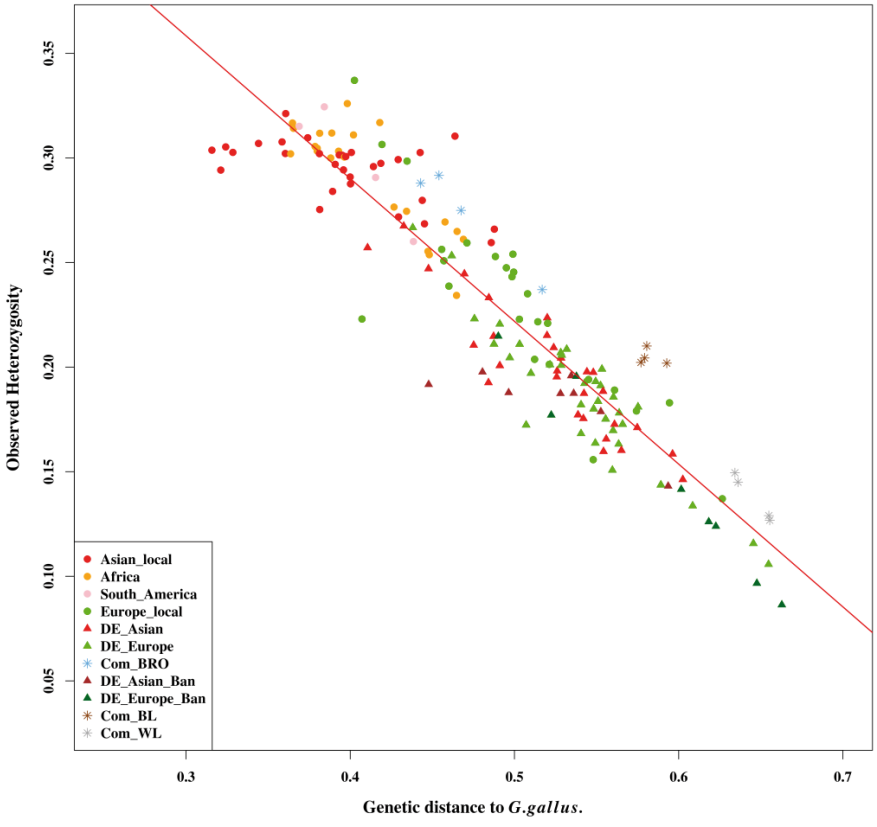

Figure 1. The relationship between the overall genetic diversity within populations and their genetic distance to *Gallus gallus*. The fitted regression line to the data with the equation heterozygosity = 0.563 – 0.683 x (genetic distance to *G. gallus*) is drawn in red. The R^2 for the linear regression is 0.875.

ii. the concept of genetic isolation by distance assumes that individuals from nearby locations are likely to be related due to mating possibilities. This is often the case in traditional breeding systems but it is not the case with the fancy and commercial breeding and management practices. Individuals within a commercial breeding herd are more related to each other than to other lines despide the geographic distances. In fancy breeds, there may be gene flow between small stocks based on personal contacts or personal relationships of breeders but not related to geographic distance forming a substructure within the breed. Actually such gene flow between fancy breeds is also very limited. Furthermore, if geographic distance was solely the main predictor for the loss of genetic diversity and increased differentiation of breeds to the wild populations, then the African and South American breeds might be expected to have highly reduced genetic diversity due to geographic distances. They also would be expected to have high genetic distances to the wild populations as well as to the rest of the Asian populations, in fact, both expectations are not fulfilled, and some of the African populations were found to be clustered with the wild type breeds (Malomane *et al.* 2019).

Therefore, the observed variations in genetic diversity may not well be predicted only by geographic expansion but rather by a combination with other aspects or subsequent events e.g. effective population sizes, types of breeding programmes, and possibly subsequent series of founder events following the geographic expansion, as previously suggested (Ramachandran *et al.* 2005, Hunley *et al.* 2009). Such events which have taken place after geographic expansion have definitely contributed to the variations in allele frequencies and thus the genetic distances of domestic chickens to the wild populations. In addition, the theoretical expansion models are also based on 'natural' expansion through migration, while chickens and other livestock were actively transported by humans (e.g. with ships) to distant places.

Acknowledgments

We acknowledge the members of the SCDP consortium as well as the breeders of the "Bund Deutscher Rassegeflügelzüchter e.V." and the "Gesellschaft zur Erhaltung alter und gefährdeter Haustierrassen e.V" in Germany for providing samples or SNP data, or for giving access to their animals for sampling. We also thank the German Federal Ministry of Education and Research (FKZ 0315528E) for funding the "Synbreed – Synergistic Plant and Animal Breeding" project. DKM also acknowledges the financial support from the IMAGE project which enabled her to conduct the study as part of her PhD work.

References

AGUILLON, S. M., FITZPATRICK, J. W., BOWMAN, R., SCHOECH, S. J., CLARK, A. G., COOP, G. and CHEN, N. (2017) Deconstructing isolation-by-distance: The genomic consequences of limited dispersal. *PLoS Genetics* **13**: e1006911.

CAVALLI-SFORZA, L. L., BARRAI, I. and EDWARDS, A. W. F. (1964) Analysis of human evolution under random genetic drift. *Cold Spring Harbor symposia on quantitative biology* **29**: 9–20.

DESHPANDE, O., BATZOGLOU, S., FELDMAN, M. W. and LUCA CAVALLI-SFORZA, L. (2009) A serial founder effect model for human settlement out of Africa. *Proceedings of the Royal Society B: Biological Sciences* **276**: 291–300.

HUNLEY, K. L., HEALY, M. E. and LONG, J. C. 2009. The global pattern of gene identity variation reveals a history of long-range migrations, bottlenecks, and local mate exchange: Implications for biological race. *American Journal of Physical Anthropology* **139**: 35–46.

KRANIS, A., GHEYAS, A. A., BOSCHIERO, C., TURNER, F., YU, L., SMITH, S., TALBOT, R., PIRANI, A., BREW, F., KAISER, P., HOCKING, P. M., FIFE, M., SALMON, N., FULTON, J., STROM, T. M., HABERER, G., WEIGEND, S., PREISINGER, R., GHOLAMI, M., QANBARI, S., SIMIANER, H., WATSON, K. A., WOOLLIAMS, J. A. and BURT, D. W. (2013) Development of a high density 600K SNP genotyping array for chicken. *BMC genomics* **14**: 59.

LI, J. Z., ABSHER, D. M., TANG, H., SOUTHWICK, A. M., CASTO, A. M., RAMACHANDRAN, S., CANN, H. M., BARSH, G. S., FELDMAN, M., CAVALLI-SFORZA, L. L. and MYERS, R. M. (2008) Worldwide human relationships inferred from genome-wide patterns of variation. *Science* **319**: 1100–1104.

MALÉCOT, G. 1969. The mathematics of heredity. (San Francisco, CA USA, Freeman).

MALOMANE, D. K., REIMER, C., WEIGEND, S., WEIGEND, A., SHARIFI, A. R. and SIMIANER, H. (2018) Efficiency of different strategies to mitigate ascertainment bias when using SNP panels in diversity studies. *BMC Genomics* **19**: 22.

MALOMANE, D. K., SIMIANER, H., WEIGEND, A., REIMER, C., SCHMITT, A. O. and WEIGEND, S. (2019) The SYNBREED chicken diversity panel : A global resource to assess chicken diversity at high genomic resolution. *BMC genomics* **20**: 345.

PROVINE, W. B. (2004) Ernst Mayr: Genetics and speciation. *Genetics* **167**: 1041–1046.

PRUGNOLLE, F., MANICA, A. and BALLOUX, F. (2005) Geography predicts neutral genetic diversity of human populations. *Current Biology* **15**: R159–R160.

RAMACHANDRAN, S., DESHPANDE, O., ROSEMAN, C. C., ROSENBERG, N. A., FELDMAN, M. W. and CAVALLI-SFORZA, L. L. (2005) Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proceedings of the National Academy of Sciences* **102**: 15942–15947.

SCHEU, A., POWELL, A., BOLLONGINO, R., VIGNE, J. D., TRESSET, A., ÇAKIRLAR, C., BENECKE, N. and BURGER, J. (2015) The genetic prehistory of domesticated cattle from their origin to the spread across Europe. *BMC Genetics* **16**: 54.

WANG, Y. (2015) Genetic and Geographic Diversity of Gyr (Bos Indicus) Cattle in Brazil (University of Natural Resources and Life Sciences, Vienna).

Hidden genes in chickens

Daniel Elleder

Institute of Molecular Genetics of the Czech Academy of Sciences, 142 20 Prague, Czech Republic

Corresponding author: elleder@img.cas.cz

Keywords: chicken genome, GC content, next generation sequencing

Introduction

The completeness of genomic assemblies is an important general issue, and was shown to be highly relevant in avian research (Peona et al. 2018). The chicken genome was published in 2004, and already then the existence of a GC-rich gene subset underrepresented in the assembly was noticed (International Chicken Genome Sequencing Consortium 2004). Similar observations were made later, when the genome sequences of hundreds of additional bird species were released (Zhang et al. 2014, Jarvis 2016). However, at the same time, a consensus has developed that birds lack up to 30% of protein-coding genes compared to mammals, possibly due to their smaller genome size. As one example, at 2014 study reported 274 genes located in conserved syntenic clusters as missing in birds, but present in the genomes of most other vertebrate species (Lovell et al. 2014). New methods of sequencing and assembly, and analysis of extremely large amount of RNA-seq data available, are now demonstrating that most of the genes considered missing are in fact present in avian genomes (Bornelöv et al. 2017, Laine et al. 2019, Yin et al. 2019).

There is a major reason why the genes were considered missing: a subset of bird genes is rich in G and C nucleotides. This makes these genes very difficult to amplify by PCR and to analyze by next-generation sequencing methods, causing extreme biases in sequence coverage (Hron et al. 2015). This applies mainly to the most commonly used Illumina sequencing technology. The Nanopore and PacBio long-read sequencing platforms show smaller coverage biases, but their higher error rate still precludes complete assembly of the difficult GC-rich regions of avian genomes. From an evolutionary perspective, it is still not clear why some bird genes are so exceptional in their GC content compared to other vertebrates. One explanation suggested recently includes the specific GC-biased gene conversion process in birds (Rousselle et al. 2019).

We developed bioinformatic and PCR approaches to identify avian GC-rich genes. The results and specific examples of such chicken genes will be discussed.

Methods

Datasets of next generation sequencing (NGS) reads were downloaded from NCBI sequence read archive (SRA) and probed with sequences of known orthologs for the respective genes, using blastn and tblastn algorithms (Altschul et al. 1990). The sequences obtained from blast searches were assembled with CLC genomics workbench (Qiagen) and Lasergene SeqMan (DNASTAR), usually with multiple iterations. PCR amplifications of GC-rich templates were performed with a mixture of Deep Vent and Taq polymerases (both from NEB) and long (5-10 min) extension times, as described previously (Rohde et al. 2018). Experimental methods relevant for characterization of specific genes are described in the references below.

Examples of GC-rich genes in chicken

Individual cases of 'hidden' avian genes will be presented. The first gene that we identified as belonging to the GC-rich gene set was chicken erythropoietin (chEPO). The functional characterization of chEPO is still ongoing. Further, others have shown that the protein product of chEPO has effect on flight performance in experimental conditions *in vivo* (Yap et al. 2018). Tumor necrosis factor α (TNF-α) was identified independently by our group and by others (Bornelöv et al. 2017, Rohde et al. 2018). We have generated recombinant chTNF-α, which is able to induce NF-κB cellular reporter. Endogenous chTNF-α is also strongly stimulated by lipopolysaccharide (LPS) in macrophages, similarly to the mammalian protein. These results point to the existence of a functional TNF-α/TNF-α-receptor system in birds (Rohde et al. 2018). chTNF-α was also recently mapped to chicken chromosome 16 by a combination of radiation hybrid mapping and fluorescence in situ hybridization (FISH) (Seroussi et al. 2019). Another example of GC-rich chicken gene previously deemed missing is leptin (chLEP), again identified independently by two laboratories (Farkašová et al. 2016, Seroussi et al. 2016), and mapped by Seroussi et al. to chicken chromosome 1 (Seroussi et al. 2017).

Other examples of GC-rich chicken genes currently under study will be discussed. Importantly, we predict that the sequences of a number of additional genes of interest to avian research will be uncovered. The best argument in favor of this was obtained from the analysis of Botero-Castro and others (Botero-Castro et al. 2017). They have shown that for hundreds of genes found in both humans and turtles and still considered missing in birds, the GC content of the turtle genes is skewed to higher values. One possibility is that such GC-rich genes are more prone to be lost in the evolution of avian lineage. However, at least for a subset of these genes, we expect that highly GC-rich orthologs will be found in birds.

Acknowledgments

This work was funded by grant 17-23675S from the Czech Science Foundation and by the Ministry of Education, Youth and Sports of the Czech Republic (project NPU I No. LO1419).