

A novel method allows accurate identification of key ancestors within populations

M. Neuditschko¹, R. Von Niederhäusern¹, H. Signer-Hasler², C. Flury², M. Frischknecht^{1,3}, T. Leeb³, E. Jonas^{4,5}, M.S. Khatkar⁵, H.W. Raadsma⁵ and S. Rieder¹

¹Agroscope Liebefeld-Posieux (ALP-Haras), Swiss National Stud Farm, Les Longs-Prés, 1580 Avenches, Switzerland, ²School of Agricultural Forest and Food Sciences, Länggasse 85, 3052 Zollikofen, Switzerland, ³University of Bern, Institute of Genetics Vetsuisse Faculty, Bremgartenstrasse 109a, 3001 Bern, Switzerland, ⁴SLU, Department of Animal Breeding and Genetics, P.O. Box 7070, 750 07 Uppsala, Sweden, ⁵University of Sydney, Reprogen, Animal Bioscience, Faculty of Veterinary Science, 410 Werombi Road, 2570 Camden, Australia; markus.neuditschko@agroscope.admin.ch

Identification of key ancestors is important for population genetics studies and conservation. Based on Principal Components Analysis (PCA), we present a novel approach using genome wide SNP data which identifies key ancestors within populations without using any prior ancestry information. We demonstrate with a sheep (1,430 individuals and 44,693 SNPs) and a horse (1,077 individuals and 38,124 SNPs) dataset that our approach effectively allocates key ancestors within a population, whilst defining fine-scale population structures. Our novel method successfully identified four influential foundation (F1) sires in an Awassi × Merino resource population as the most informative individuals, despite the close relationship between animals and multiple generations of back- and intercrossing. Our method also performed exceptionally well in the horse dataset, and stallions descending from main lineages have been assigned with highest informative scores. Our approach allows both the selection of informative individuals and the characterization of fine-scale population structures. This method can also be applied to analyse fine-scale population structures in indigenous and wild populations, where ancestry information is not readily available. Furthermore, our approach will be useful for assembling most informative resource populations for full genome sequencing initiatives, to facilitate accurate genotype imputation across populations.

Identification of candidate polymorphism in a QTL region by combining eQTL mapping with NGS data

P.F. Roux^{1,2}, Y. Blum^{1,2,3}, C. Désert^{1,2,3}, A. Djari⁴, D. Esquerre⁵, E. Le Bihan-Duval⁶, B. Bed'hom⁷, P. Le Roy^{1,2}, F. Lecerf^{1,2}, M. Moroldo⁸, S. Marthey⁸, C. Klopp⁴, S. Lagarrigue^{1,2,3} and O. Demeure^{1,2}

¹Agrocampus Ouest, UMR1348 Pegase, 35000 Rennes, France, ²INRA, UMR1348 Pegase, 35590 Saint-Gilles, France, ³Université Européenne de Bretagne, UEB, 35000 Rennes, France, ⁴INRA, SIGENAE, 31320 Castanet Tolosan, France, ⁵INRA, GENOTOUL, 31320 Castanet Tolosan, France, ⁶INRA, UR83 Recherches Avicoles, 37380 Nouzilly, France, ⁷INRA, UMR1313 GABI, 78350 Jouy-en-Josas, France, ⁸INRA, UMR1313 GABI, CRB GADIE, 78350 Jouy-en-Josas, France; pierre-francois.roux@rennes.inra.fr

In this study, we propose a genetical genomic approach aiming at characterizing QTL affecting adiposity in a chicken model, combining hepatic gene expression and whole genome re-sequencing data. After performing a classical linkage analysis on our meat-type chicken design with abdominal fat weight as the target phenotype, we identified 3 QTL regions. Carrying out hepatic eQTL mapping on those QTL regions, we identified no local acting eQTL, suggesting that causal mutations underlying QTL were located on coding region. Using NGS data, we then selected SNPs having an impact on coding region. Secondly, we selected SNPs impacting highly conserved regions among vertebrates, which is conditional to a functional role. These successive filters allowed us to reduce the number of candidate genes in the 3 QTL regions. One of them was coding for acetyl-CoA carboxylase (ACACA), well known for its implication in fatty acids biosynthesis. Moreover, the two SNPs associated to this gene were missense and one of them was located in the protein carboxyl catalytic domain. These two observations make these two SNPs in ACACA strong candidate mutations responsible for adiposity in our scheme. This original approach, combining NGS and QTL and eQTL mapping, could be applied to identify candidate mutations for any complex trait in any species. S. Lagarrigue and O. Demeure contributed equally to this work.