FINAL PUBLISHABLE SUMMARY REPORT

Introduction

The main objectives of the project NEUTRADAPT were to develop statistical procedures for the estimation of adaptive/selective and neutral genetic variation, and to apply molecular information to the conservation of genetic resources and inbreeding control/reduction in farm animal breeding programmes. Genotyping data, whole-genome sequences and phenotypes from cattle, sheep and goat populations were available.

Estimation of neutral and adaptive genetic variation

Genetic variation can be differentiated into variation related to environmental adaptation or to response to selection (i.e. variation underlying traits of agricultural interest like milk yield), and variation which is neutral with respect to the objective of selection or adaptation. Several statistical methods to detect genetic signals exist. Within the framework of this project, the basic idea was to look for signals of adaptation and domestication by comparing signatures of selection between populations from different environments and between ancestral and domestic species. Sheep and goat populations from different geographic areas in Morocco (e.g. coast, mountains, desert …) were used for the detection of signals of adaptation. Domestic sheep and goats from the putative domestication centre in Northern Iran and their wild counterparts mouflons (Ovis orientalis) and bezoars (Capra aegagrus) were used to detect signals of domestication. The remaining polymorphisms in the genome could be considered as neutral genetic variation with respect to adaptation and domestication. The EHH/iHS method for the detection of signatures of selection (Voight et al., 2006) was adopted. Details of the method can be found in the scientific report of the ESF exchange at the University of Cardiff where a collaboration on the topic was strengthened (see ESF_ScientificReport.pdf). A few intermediate results can be found in the attached pdf file IntermediateResults. Final results will soon be compared with results obtained with different statistical approaches by partners of the Nextgen project (http://nextgen.epfl.ch/), within which the data were generated, and presented in peer-reviewed publications on the adaptation to different environments and domestication of sheep and goats. During the course of the project, signals of selection were also detected in cattle populations using different statistical methods: composite log-likelihood (Stella et al., 2010), for which a web platform has been presented (DeSSign: see attached EAAP Abstract 2012), and relative EHH (a paper is soon to be submitted: Bomba et al., see attached ManuscriptsInPreparation.pdf). In addition, signals of genotype-phenotype association were detected using either standard single-GWAS associations studies (published paper: Biffani et al., 2014, Livestock Science) or exploring the possibility of using runs of homozygosity (ROH) as alternative or complement to GWAS (published paper: Biscarini et al., 2014, Elares; conference papers: AIDA 2013, Livestock Genomic Resources 2014, WCGALP 2014: see attachments).

Management of genetic diversity

The conservation of genetic resources and inbreeding control are of fundamental importance for effective and sustainable breeding programmes in farm animal populations. The use of molecular information to optimise breeding programmes was explored and compared with traditional methods based on pedigree information (published paper: Gandini et al., 2014, Small Ruminant Research). Molecular information was used either through the estimation of genomic coancestry (see Deliverable 2.1; publication in progress) or through the analysis of ROH for the estimation of inbreeding or the detection of signatures of selection (see publications and conferences).

Computation strategies

Not only SNP genotyping data, but also whole-genome sequences were analysed, consisting of millions of polymorphisms (about 30 million polymorphisms in the case of goats and sheep). Effort was devoted to the storage, processing and analysing of such big data. The potential of using graph databases to store and effectively retrieve big data was explored (EAAP Abstract 2013). Parallel computing strategies and the use of the graphical processing unit (GPU) were adopted to speed up computations involving large coancestry matrices (conference paper at UseR 2013). A combination of Python and C++ (Cython, http://cython.org/) was applied to the computations of genetic signals over millions of polymorphisms (a paper on the subject is in preparation).

PhD student supervision

The tasks of the Marie-Curie ERG comprised teaching activities and the supervision of a PhD students. The PhD project started in autumn 2012 and is about the genetics and breeding of sugar beet (*Beta vulgaris*) populations. This so far resulted in a number of conference presentations and two peer-reviewed published papers (Stevanato et al., 2013, Plant Molecular Biology Reporter; Biscarini et al., 2014, BMC Genetics).

References

B. F. Voight, S. Kudaravalli, X. Wen, and J. K. Pritchard. A map of recent positive selection in the human genome. PLoS biology, 4(3):e72, 2006.

Stella, A., Ajmone-Marsan, P., Lazzari, B., Boettcher, P.: Identification of selection signatures in cattle breeds selected for dairy production. Genetics 185(4), 1451–1461 (2010)