Simon Boitard, Origine, Structure et Evolution de la Biodiversité, Museum National d'Histoire Naturelle.

Inferring the past dynamics of effective population size using genome wide molecular data

Inferring the effective size of a given population, and its eventual expansions or reductions in the past, from genetic data, is a long standing question in population genetics. Due to the complexity and the high dimension of the mathematical models that are used in this context, exact inference is impossible and the most popular inference mehtods are based on numerical approaches as Markov Chain Monte Carlo, Importance Sampling or Approximate Bayesian Computaion (ABC).

Until recently, these methods were designed for data sets including a small number of independent markers or non recombining DNA sequences. However, the spectacular progress of genotyping and sequencing technologies during the last decade has enabled the production of high density genome wide data in many species, so new statistical methods are needed to take benefit of this new type of data.

In this study we present an ABC approach for inferring the past effective size of a single population. This approach is based on coalescent simulations and on the use of a large number of summary statistics related to allele frequencies and linkage disequilibrium. We illustrate the performance of this approach using cross validation. We compare different ABC strategies and discuss the influence of the different summary statistics.

We finally apply this method to a set of 25 bovine sequences from the Holstein breed and compare our results with those obtained by the Pairwise Sequentially Markovian Coalescent approach of Li and Durbin (2011).

Christèle Robert-Granié, Station d'Amélioration Génétique des Animaux (SAGA) - UR631, INRA-Toulouse

Integration of genomic information into genetic evaluation model : Is it a good statistical model?

christele.robert-granie@toulouse.inra.fr

The rapid evolution in sequencing and genotyping raises new challenges in the development of methods of selection for livestock, also called genomic selection. With this genomic information, it is now possible to estimate breeding values of selection candidates at birth without waiting for phenotypic data collection. Genomic selection requires the creation of a reference population made of genotyped animals with precise phenotypes. Genomic evaluations consist in predicting phenotypes in this reference population as the sum of molecular markers. The main methodological challenge is the large number of effects to estimate, usually much larger than the number of available phenotypes. We briefly describe and compare the various families of proposed methods: genomic BLUP (Best Linear Unbiaised Prediction) based on relationship computed from marker information, Bayesian methods, variable selection methods and a single step method which combined pedigree and genomic data and raw phenotypes. The precision of genomic selection is done via cross validation among the youngest animals of the reference population. Several parameters (size of the reference population, relationship between selection candidates and the reference population, genetic parameters, etc) have a significant impact on the efficiency of genomic selection methods. Some results on real data will be presented.

Anne-Louise Leutenegger, University Paris Diderot, France and Inserm U946.

Mapping genes in consanguineous and isolated populations in the era of high throughput sequencing

Great progress has been made in the identification of genetic variants for complex human traits thanks to genome-wide association studies (GWAS). However, part of the heritability remains unknown for most complex diseases, suggesting that some genetic factors remain to be discovered. The study of rare variants could help to characterize more exhaustively the genetic background of complex traits and is now facilitated by recent advances in sequencing technologies. However, those technologies remain too expensive for many academic research groups to sequence a large number of subjects from general populations, which is necessary to attain sufficient power to detect effect of such variants.

Population isolates have well-documented characteristics that can aid to identify rare variants associated with complex traits, namely reduced phenotypic, environmental and genetic heterogeneity. Besides, alleles that are rare in general populations may have become more frequent in those isolated populations, which could facilitate their identification.

We will present the strategy that has recently been proposed to study complex traits in isolated populations: selection of subset of individuals for sequencing, imputation of the sequence data in the remaining individuals to obtain sequence data on the entire population, and finally, association analysis with traits of interest.