**Research project for Master student (Cirad / University of Yaoundé 1) - 2016**

General information:
- Title: **Breeding oil palm to improve the regularity of production along the year**
- Keywords: R programming, linear mixed model, quantitative genetics and genomics
- 6 months, at the **University of Yaoundé 1**
- contact: **david.cros@cirad.fr**

Details:
Oil palm produces bunches all year long but its production follows annual cycles with a peak period. The magnitude of the peak is of economic importance as it determines the size of the oil mill and the scale of the investment. The development of oil palm varieties with an even production along the year is therefore an interesting goal for the industry.

The regularity of bunch production along the year can be summarized by the coefficient of Gini, which will be used as selection criterion. The goal of this study is to acquire the necessary knowledge to decide whether breeding for reduced Gini coefficient is possible, using genomic selection approaches:
- what is the minimum number of individuals per elementary plot to obtain an estimate of the coefficient of Gini representative of the cross value?
- what are the values of the genetic parameters of interest for breeding (in particular genetic variances of the trait and genetic correlations with the other traits)?
- what would be the selection accuracy for progeny-tested individuals?
- what would be the prediction accuracy for non progeny-tested individuals?

This work will use the data provided by PalmElit, a private oil palm breeding company. They were collected in two experimental designs located in Indonesia, involving 700+ oil palm hybrid crosses. In details, the data are:
- the genotype of the parents of the hybrids for 3000+ SNP markers
- the monthly production of bunches for each hybrid individual at young age (designs 1 and 2) and adult age (design 1)

Planning:

| | |
|---|---|
| Literature review<br>(mixed model for genetic evaluations / gini coefficient / oil palm breeding) | 3 weeks |
| . Compute Gini coefficient for all elementary plots in the two experimental designs, varying the number of individual per elementary plot<br>. Measure phenotypic correlation between Gini coefficient and bunch production traits | 3 weeks |
| Analyze the 2 datasets with genomic multivariate mixed models<br>To be done with R-Asreml. This will give genetic and residual correlations, variances, selection accuracy and GCA of progeny-tested parents | 1.5 months |
| Predict the GCA of the individuals progeny tested in the 2$^{nd}$ dataset using genomic multivariate mixed models calibrated with the 1st dataset<br>To be done with R-Asreml for Gini coefficient at young age and production traits, to estimate prediction accuracies | 1.5 months |
| Report writing | 1.5 months |

The student will have to adapt pre-existing R scripts and to write new specific scripts. **Good knowledge of R is required**. **Knowledge of linear mixed model, statistics or quantitative genetics is required.**

The results are expected to give a scientific article.