EAAP, Varsaw, Poland, 31 August – 4 September 2015

Population genomics and signatures of selection in honey bee drones

D. Wragg, M Marti, B Basso, JP Bidanel, E Labarthe, O Bouchez, Y Leconte, A Vignal

The Seqapipop project

Acronym

Seqapipop : Sequencing apis mellifera populations

Goals

- **Estimate genetic diversity between and within the most commonly used honey Bee populations in France**
- **Comparison with major honey bee population in Europe and worldwide**
- **Investigate signatures of selection**
- **Prepare the implementation of experiments to decipher the genetic architecture of economically important traits.**

The Seqapipop project

Material & methods (1/2)

- **Whole genome sequencing of 1000 drones (haploid males)**
- **About 30 drones / population**
  - ***So far : 21 populations sampled,***
    13 sequenced
    4 analysed : Royal Jelly (RJ), Ouessant (OUE), Corsica (AOC), HN (close to carnica)
- **Paired-end Sequencing on an Illumina HiSeq 2000 platform**
- **Reference population of 39 workers (Harpur et al, 2014) downloaded from European nucleotide archive**

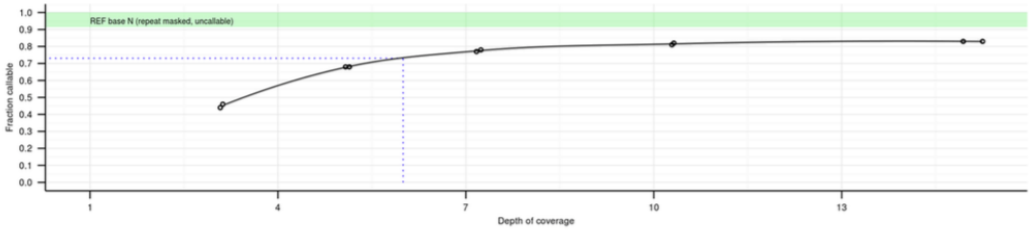
The Seqapipop project

Material & methods (2/2)

- **A bioinformatics pipeline developed for processing the NGS data (Wragg et al, BMC Genomics, submitted)**
- **2 drones sequenced at a moderately (high) coverage (15k) to get target depth of coverage for 70% of the genome to be considered callable**
- **Other methods mentioned later**

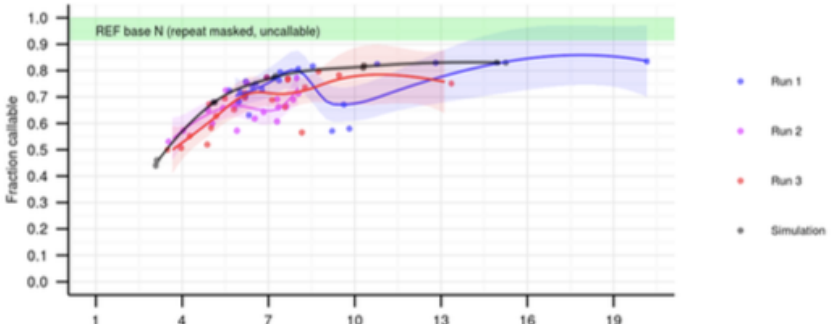
Sequencing depth (DP) vs callability (GX)

DP=5.3 X achieved GX = 0.7

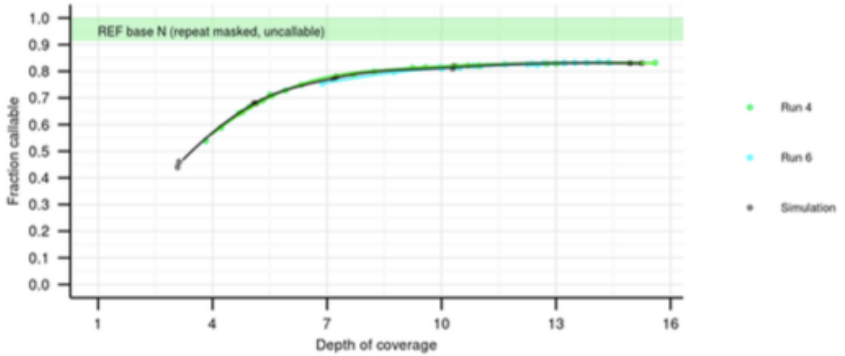


Target DP of 6X

HN – RJ populations
DP ranged from 3.49 to 13.36 (μ_{DP} = 6.78)
and GX between 0.5 and 0.83 (μ_{GX}= 0.68)



OUE – AOC populations
DP ranged from 3.8 to 15.6 X
Much better fit

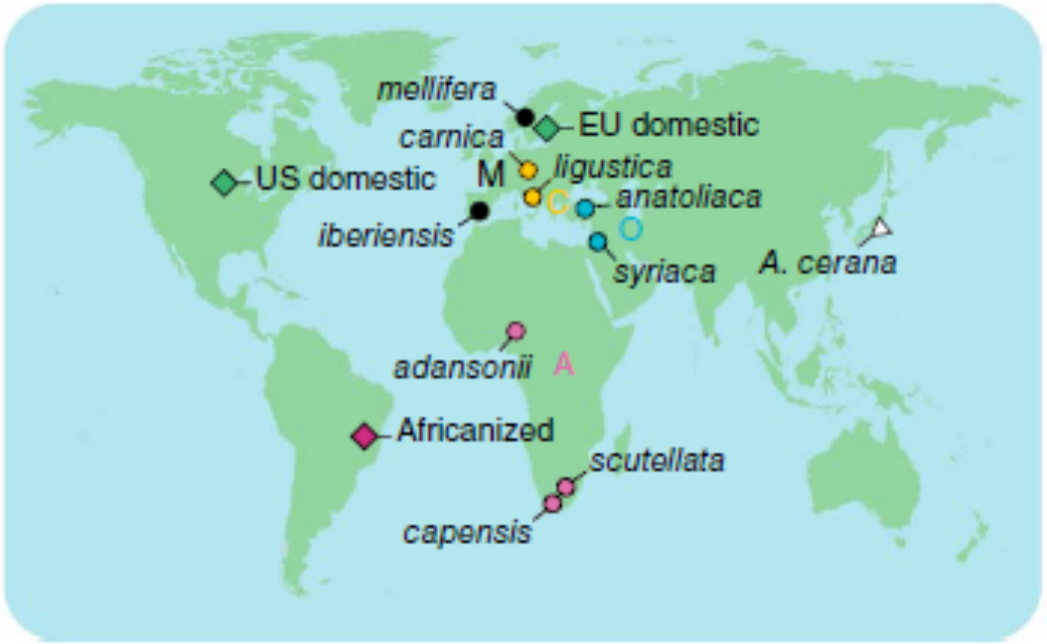


Diversity & population structure

Only 2 populations => comparison with main *A. Mellifera* lineages

Main A. Mellifera lineages

A – Africa
M – Northern Europe
C – Eastern Europe
O – Asia



Diversity & population structure

Within population diversity : number of SNPs

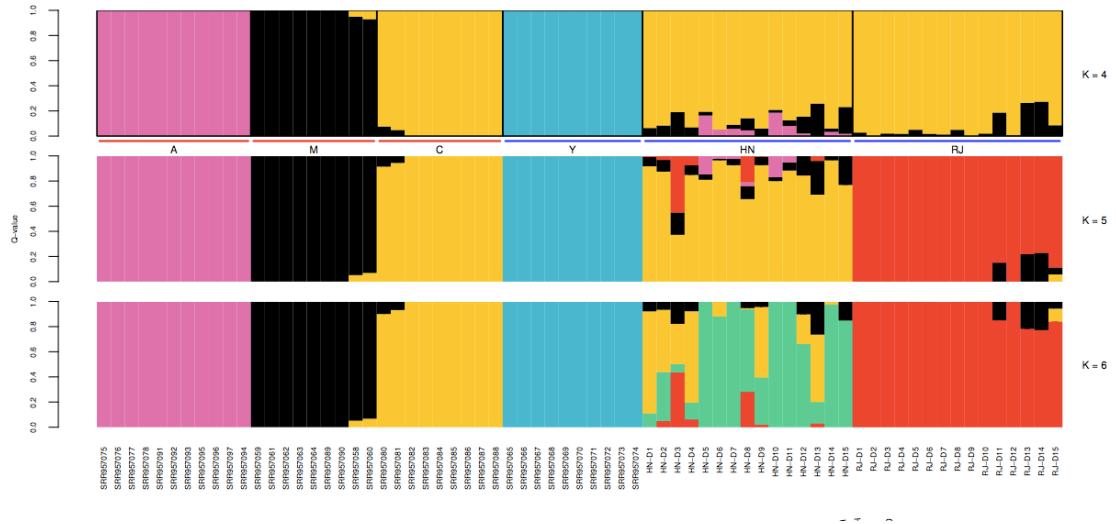
A	4 959 569				
M	121 471	369 556			
C	36 333	49 048	206 753		
O	579 801	14 720	8 334	1 066 987	
Haploid	229 843	187 520	203 013	37 319	798 349

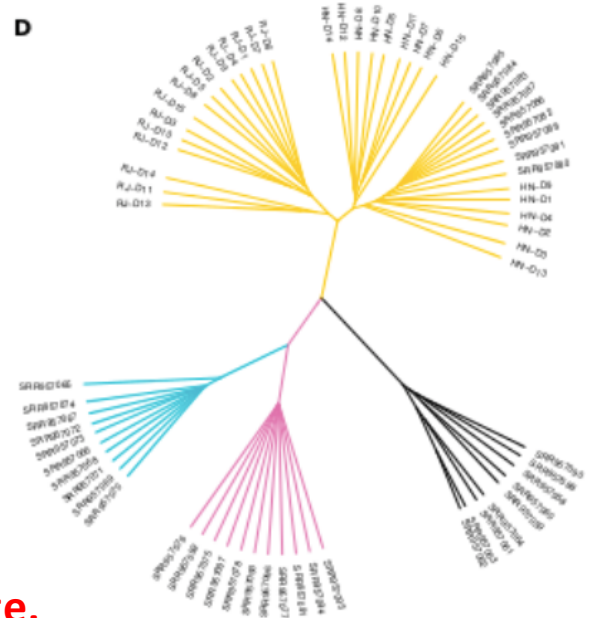
Diversity & population structure

« Diploid » individuals created by randomly grouping pairs of drones

=> 15 « diploid » individuals for HN and RJ populations

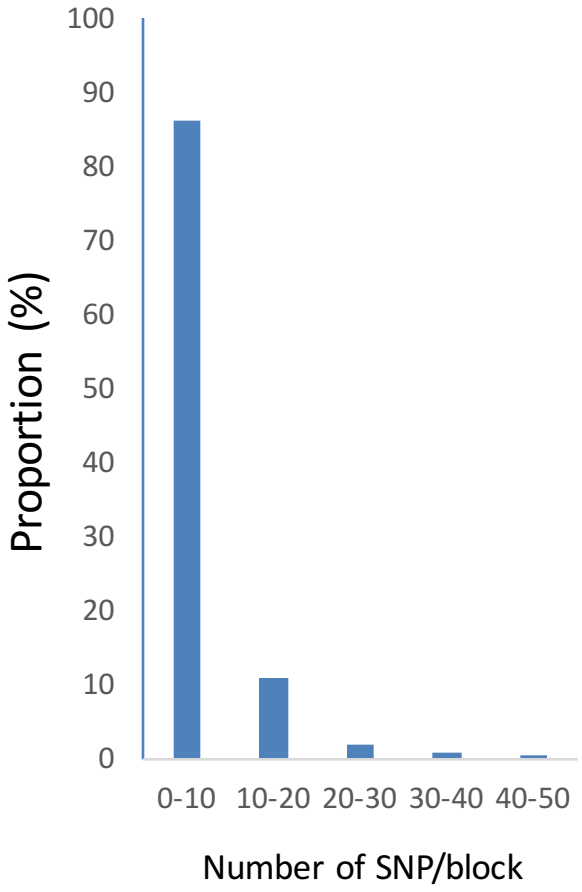
Mixed with the « reference data set of individuals from the 4 lineages





HN and RJ populations mainly originate from the C lineage, with a fraction of admixture from the A lineage

Haplotype block size



Estimated from pair-wise LD values of SNPs less than 2 Mb apart

99% of the blocks < 2.5 kb

99.9% of the blocks < 20.9 kb

Signatures of selection

Use of a 2.53 M SNP haploid dataset

FST and absolute differences in MAF (D_{MAF}) between HN et RJ populations computed in 2.5, 12 and 20 kb windows with a 75% window overlap

Windows in 99th percentile considered as significant

Results presented for the 12kb window

114 and 124 significant intervals with FST and DMAF, respectively, of which 35 intervals were common

Strongest signal on chromosome 6

Signatures of selection

Chromosome 6

Signatures of selection
Chromosome 6

Interval (chr6:6.08-6.11 Mb) includes most significant result

Hosts 2 cytochrome P450 genes, each with non-synonymous variants

Previously suggested that royal jelly might be synthesized by bees' P450-dependant enzyme system

Another genes in the vicinity

- Gene involved in the elongation of long chain fatty acids (FA). Main lipid in the royal jelly is a medium chain FA

- Member of the yellow gene family, encoding major royal jelly protein 7 (MRJP7 - not a major component of royal jelly)

Conclusions

Modelled and validated optimal to achieve 70% genome callable

(Very) preliminary analyses of population diversity

Preliminary investigation on signatures of selection

Future work

- Analysing the (vast majority of) data not yet analysed
- More thoroughly investigated selection signature footprints
- Investigating genotype – phenotype relationships on the subset of individuals with phenotypic data

Acknowledgements

Bioinformatics analyses performed by D Wragg under the direct supervision of A Vignal

This work has been financially supported by :

- INRA animal Genetics Division (first project)
- France Agrimer (Seqapipop project)