JOINT ESTIMATION OF EFFECTIVE POPULATION SIZE AND SELECTION COEFFICIENT WITHOUT NEUTRAL MARKERS: METHOD VALIDATION AND APPLICATION TO EXPERIMENTAL EVOLUTION OF VIRUSES

Elsa Rousseau (1,2,3), Benoît Moury (3), Ludovic Mailleret (1,2), Rachid Senoussi (4), Alain Palloix (5), Vincent Simon (3,6), Frédéric Grognard (1), Sophie Valière (8) and Frédéric Fabre (9)

(1) Biocore Team, INRIA, Sophia Antipolis, France
(2) UMR ISA, CNRS, INRA, Université Nice Sophia Antipolis, Sophia Antipolis, France
(3) UR PV, INRA, Montfavet, France
(4) UR BioSp, INRA, Avignon, France
(5) UR GAFL, INRA, Montfavet, France
(6) UMR BFP, INRA, Villenave d'Ornon, France
(7) GeT-PlaGe, INRA, Genotoul, Castanet-tolosan, France
(8) UAR DEPT GA, INRA, Castanet-Tolosan, France
(9) UMR SAVE, INRA, Villenave d'Ornon, France (frederic.fabre@bordeaux.inra.fr)

BACKGROUND and OBJECTIVES
Experimental evolution studies deserve considerable attention to the estimation of basic evolutionary forces such as selection and genetic drift. With the advent of high-throughput sequencing techniques, these studies gained a renewed attention. However the joint estimation of selection and genetic drift still remain challenging when no neutral markers are available, a common situation with many microbes, such as viruses, due to their small tightly packed genomes.

MATERIAL and METHODS
We have developed a mechanistic-statistical model for estimating effective population size Ne and selection coefficient s from genetic time series data generated in evolve-and-resequence experiments. Numerical simulations of the dynamics of competing genotypes in asexual haploid Wright-Fisher populations subjected to contrasting genetic drift (Ne $\in$ [20, 5000]) and selection (|s| $\in$ [0, 0.1]) regimes were used to test the model. This model was then applied to data for competition between five variants of Potato virus Y (PVY, genus Potyvirus) in 15 doubled haploid lines of pepper (Capsicum annuum, family Solanaceae) that constituted different host environments. The frequencies of the variants were determined at six time-points, in eight plants per time point for each host environment.

RESULTS
The numerical experiments validated the method and made it possible to obtain bias-adjusted estimators of s for each competing genotype and of the dynamics of Ne over the time course of the experiment (R² $\in$ [0.93, 0.95], slopes $\in$ [0.98, 1.07] and intercepts $\in$ [-0.14, 0.06] for the best linear fit between true and estimated parameters). The real-life experiment showed that the virus populations experienced either stochastic or deterministic evolution, depending on host genotype and that Ne and s were genetically controlled by the host genotype.

CONCLUSIONS
This method constitutes an advance in the joint estimation of key population genetic parameters from timesampled data in conditions in which no neutral markers are available and both selection and genetic drift are strong. This is a situation frequently encountered in studies of microorganisms. The observation of highly contrasted genetic drift between plant genotypes may have practical implications for slowing viral emergence through the careful choice of plant cultivars