# Genome sequencing galore: what, how and why

*Lino Ometto[1], Omar Rota Stabelli[1], Alessandro Cestaro[2], Maria Cristina Crava[1], Gianfranco Anfora[1]*

[1] Department of Sustainable Agroecosystems and Bioresources and [2] Department of Computational Biology, Research and Innovation Center – Fondazione Edmund Mach, 38010 S. Michele all'Adige (TN), Italy

The advent of affordable high-throughput Next-Generation Sequencing (NGS) technologies has revolutionised all fields of biology. Numerous technologies, such as Illumina, Roche 454, Ion torrent, and SOLiD are available to generate massive NGS data and are accompanied by a growing set of tools for downstream analyses. Because NGS data are relatively cheap, it is also possible to sequence genomic DNA and transcribed genes (RNA) for non-model species of evolutionary, ecological, agricultural or epidemiological importance, for which very little or no prior genetic information is available. More generally, NGS is becoming a cost effective practice for a variety of non-strictly genomic laboratories. For such data to be useful, however, attention should be paid to the full experimental design approach, from sample choice to material preparation, from sequencing technology to assembly strategy. Here we provide examples of approaches and of bioinformatics tools that we used to obtain (after a small economical investment) good-quality genome data useful to investigate the ecology and evolution of various insect species, including *Drosophila subpulchrella*, *Cydia pomonella* and *Cacopsylla melanoneura*. For instance, we show how it is possible to use assembly tools that account for heterozygosity in the sample, thus obtaining draft genomes even when starting from outbred lines. Furthermore, it is now possible to reduce the impact of heterozygous sequences by extracting genomic DNA from single insects and preparing ad-hoc libraries. These examples illustrate that, although generating NGS data is now straightforward, much effort should be given to experimental design in order to produce good-quality and useful data.