

# Statistical challenges on the 1000€ genome sequences in plants

Marco Bink<sup>1</sup>, Thomas Schiex<sup>2</sup>, and David Marshall<sup>3</sup>

[www.statseq.eu](http://www.statseq.eu)



## Situation

**Next-Generation DNA Sequencing (NGS)** technologies has the potential to dramatically accelerate **biological research**. Ongoing **reduction in cost** of sequencing will **democratize** the field to individual investigators

## Twist

The near-term **challenges** will shift from data generation technologies to the **analysis and integration** of NGS data

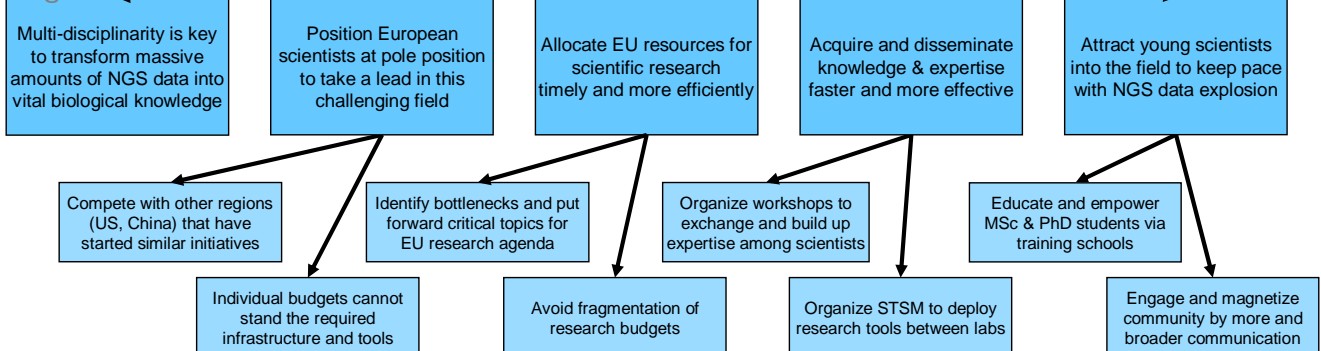
## Question

What are these **challenges** and what do **individual investigators** need to accept them?

## Message

Establish an **inter-disciplinary European-wide community** that collaborates to **accelerate** assessment, deployment & dissemination of **statistical tools and expertise** crucial to unlock NGS data into **biological knowledge**

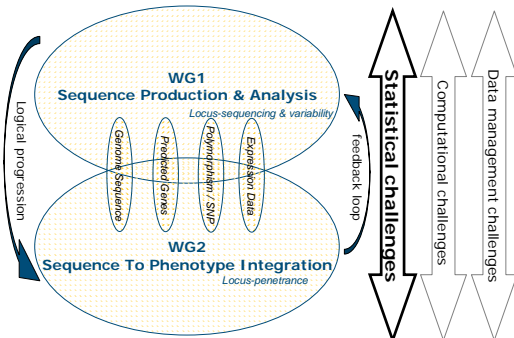
## Arguments



## Objective of EU COST Action TD0801

Establish a **network** of European researchers developing and/or using **efficient statistical and bioinformatics tools and strategies** to produce, assemble, analyze, and integrate high-throughput **Next Generation Sequence** data to increase our **knowledge** of crop and model **plant system**.

Figure 1: Two Working Groups of the COST Action and their putative interactions



### Box 1: Next-Generation Sequencing (NGS) Technology platforms

Embodying a complex engineering of enzymology, chemistry, high-resolution optics, hardware and software. Providing significant time savings and minimal requirement for associated equipment a streamlined sample preparation prior to DNA sequencing. Examples are Roche/454 FLX, Illumina Genome Analyzer; and ABI SOLiD System.

#### This NGS technology will change the way on

- (Re-) Sequencing of (fragments of) genomes
- Profiling transcriptomes
- Elucidating DNA – protein interactions
- Studying Epigenomics, Metagenomics, etc.

This EU COST Action is open to global participation – ready to collaborate with other ongoing initiatives outside Europe. The end date of the Action is anticipated to be early 2013. After installation of a Management Committee, Working Groups will be enforced that will organize the activities within the Action. **Facts and figures:**  
[www.cost.esf.org/domains\\_actions/fa/Actions/statistical\\_challenges](http://www.cost.esf.org/domains_actions/fa/Actions/statistical_challenges)

### Box 2: Uniqueness of plants

- repeat density in plants
- complications of polyploidy
- Immortalized (inbred/mutant) populations
- Extensive phenotyping in wide range of conditions