



Biscarini F.<sup>1</sup>, Caprera A.<sup>1</sup>, Nazzicari N.<sup>1</sup>, Cozzi P.<sup>1</sup>, Strozzi F.<sup>1</sup>, Rossini L.<sup>1,2</sup>, Stella A.<sup>1,3</sup>, Nicolazzi E.L.<sup>1</sup>

<sup>1</sup> Fondazione Parco Tecnologico Padano, V. Einstein, Loc. Cascina Codazza, 26900 Lodi (Italy)

<sup>2</sup> Università degli Studi di Milano – DiSAA, Via Celoria 2, 20133 Milano (Italy)

<sup>3</sup> Istituto di Biologia e Biotecnologia Agraria, Consiglio Nazionale delle Ricerche, V. Einstein, Loc. Cascina Codazza, 26900 Lodi (Italy).

### INTRODUCTION

In recent years, the use of genomic information in plant and animal species for genetic improvement and related fields has become routine. In order to accommodate market requirements (i.e. genotyping cost), manufacturers of single nucleotide polymorphism (SNP) arrays, private companies and international *consortia* have developed a large number of arrays with different content and different SNP density. The number of currently available SNP arrays differs among species.

There is a strong need to standardize and integrate array-specific (e.g. SNP IDs, allele coding) and species-specific (i.e. assemblies) SNP information. After the very positive response to **SNPChImp** ([bioinformatics.tecnoparco.org/SNPchimp](http://bioinformatics.tecnoparco.org/SNPchimp)), where we store and provide tools for the 6 major livestock species and more than 20 SNP arrays, we are now extending our family of tools to plant species. **SNPGreen** ([bioinformatics.tecnoparco.org/SNPgreen](http://bioinformatics.tecnoparco.org/SNPgreen)) currently includes 3 SNP arrays for Rice and Maize.

### SNPGreen, UP CLOSE

This tool allows researchers to easily "jump" between different SNP arrays and types of information in, currently, two plant species: Maize (2 chips currently available) and Rice (1 chip currently available). Following the experience of SNPChImp, we plan to include more arrays, species (in collaboration with species *consortia*) and data (genome assemblies, etc..).

**Feel free to contribute!**



The two main menus are:

- A. **Download:** for full array information. If you need to update your chip data, this is where you want to interrogate our database.
- B. **Browse:** for a subset of SNPs. When the species is available in Ensembl, **SNPgreen** provides a direct link to annotate a group of SNPs (e.g. results of a GWAS) using Ensembl Biomart and Variant Effect Predictor tools.

### CHARACTERISTICS AND FUNCTIONALITY

**SNPGreen** is an open access web-tool that stores only original (e.g. coming directly from producers and *consortia*) and IP-free (e.g. publicly available) information.

With **SNPGreen**, you're only a few clicks away from:

- Updating your SNP map information to any available assembly;
- Getting a list of SNPs in common (or not) across any SNP arrays combination you want (useful for imputation) [Not currently available for any species in **SNPGreen**];
- Have complete access to the "sender/ss" linked information contained in dbSNP (useful when working with sequence data);
- Obtain direct access to Illumina's Forward/Reverse or Top/Bot allele coding and to Affymetrix (AB to Forward) allele coding;
- Use the open-source tools for basic SNP management developed specifically to interact with **SNPGreen** outputs;
- Browse your genomic analyses results: explore your results directly, and annotate your desired genomic regions using Ensembl BioMart and Variant Effect Predictor [Not currently available for any species in **SNPGreen**].

### CONCLUSIONS

This platform allows easy integration and standardization, and it is aimed at both industry and research. It also enables users to easily link the information available from the array producer with data in public databases, without the need of additional bioinformatics tools or pipelines.

Further developments integrating *EnsemblGenomes* information (e.g. updated assemblies, BioMart and VeP functionalities) are under study.

Following the great response from the animal genetics community to **SNPChImp**, we trust to enhance the awareness of plant scientists towards standardization of data, with the final objective of boosting the exchange of genomic data.

Collaboration with species-specific *consortia* will be essential to obtain the best results possible.

### ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement n° 289592 - *Gene2Farm*. Ali Pirani, Fiona Brew for providing information and updates on Affymetrix products. Cindy Lawley and Chandrasen Soans for providing information and updates on Illumina products.