transPLANT milestone report

MS10 (work package 4): 2nd transPLANT training workshop

The 2nd transplant user training workshop was held in Poznań, Poland, from 27-28 June 2013 at the Adam Mickiewicz University campus, Dept. of Biology. Local organization was carried out by the Institute of Plant Genetics, Polish Academy of Sciences (Paweł Krajewski and colleagues).

This second workshop was planned as a follow-up of the first transPLANT workshop held in Versaille, France, and focused on current developments in plant data resources at transPLANT partner sites, with a special reference to Triticeae, as a lot of exciting but complex data was and is generated here (barley, bread wheat, rye, Aegilops; Brenchley et al., 2013, Nature 491: 705-710; IBSC, 2013, Nature 491: 711-716). Topics related to standardisation and annotation of plant phenotypic data were also introduced and discussed during the workshop. The workshop was targeted at (experimental) biologists and breeders who have needs to use these resources in everyday work to interpret own observations.

To ensure appropriate workshop conditions and resources and to prevent misunderstandings about the workshop's objectives and contents, user registration required an application explaining the researcher's background and motivation for the workshop.

A total of 21 participants were selected and taught about transPLANT services and resources during the course, coming mainly from eastern-Europe countries such as Poland. Teaching sessions were contributed by the transPLANT partners: INRA Versailles, EMBL-EBI, Helmholtz Center Munich, IPG PAS and the Adam Mickiewicz University, Poznań (as an invited guest talk). The workshop was announced over transPLANT partner websites, several mailing lists (including transPLANT and TriticeaeGenome mailing lists), and cooperating partners and within connected communities (specifically Triticeae communities in Europe and the US).

The workshop program is attached below. All sessions consisted of short database/resource introductions followed by extensive "hands-on" training involving a number of exercises. All workshop presentations as well as training materials were uploaded to a dropbox server where participants were able to access it in a convenient way.

All participants were asked to evaluate the workshop, and to provide feedback on how to improve organization, content and maximize impact on their daily work. We received back a total of 18 evaluation forms. A detailed assessment is provided in WP4 yearly report.

Program

Day 1

9:00 Opening

9:05 - 13:00 Plant data resources at HMGU: PlantsDB (HMGU)

MIPS PlantsDB (http://mips.helmholtz-muenchen.de/plant/genomes.jsp) is a database framework for integrative and comparative plant genome research and provides data and information resources for individual plant species (including Medicago, Arabidopsis, Brachypodium, Sorghum, maize, rice, barley and wheat). Building up on that, state-of-the-art comparative genomics tools such as CrowsNest are integrated to visualize and investigate syntenic relationships between monocot genomes. Results from novel genome analysis strategies targeting the complex and repetitive genomes of Triticeae species (wheat and barley) are provided and cross-linked with model species. The MIPS Repeat Element Database (mips-REdat) and Catalog (mips-REcat) as well as tight connections to other databases, e.g. via web services, are further important components of PlantsDB.

Specific topics:

1. Concept of and interactive data access to the barley and wheat genome zippers

2. The barley genome: integration of physical and genetic map, data access and use cases

3. UK wheat 5x WGS+analysis results: concepts to use this new data resource

4. Comparative genomics – from models to crops: exploring synteny, visualization tools

(CrowsNest)

14:00 – 14:45 Application of new web technologies in biological research and databases (Adam Mickiewicz University, Poznań)

As biologists venture into bioinformatics, they often have to trade their favorite graphical computer desktop environments for a cumbersome command line versions of very useful scientific programs. Most recently the Web has rapidly evolved into a platform suitable for user-friendly applications that exhibit a level of richness and interaction that could barely be envisioned several years ago. The new web technologies are closing the gap to native applications and fill an important need for the development of data analysis software that provides bioinformatics functionalists to biologists without requiring prior knowledge of programming and scripting languages. We will present desing and key technologies underlying some of the recent biological databases and web applications. We show their roles in integration, management and visualization of biological data in mirEX, our latest database of expression levels of Arabidopsis pre-miRNAs. We also show the use of different common programming languages to build an Ajax-driven biologist-friendly web applications. Finally, we demonstrate application of our recently-developed web service for annotation of highly-divergent tryptophan-containing Argonaute-binding proteins in grass genomes.

15:00 - 18:00 Genomic, genetic and phenomic plant data at the INRA URGI: GnpIS.

GnpIS is an integrative information system for plant and pest genomics hosted and developed at the URGI, INRA. It stores and allows data mining of genomic, genetic and phenotype data for several species such as grape, wheat, maize, Arabidopsis, rice, poplar (apple, in progress) but also fungi. During this training, we will focus on some specific use cases of the URGI information system. We will explore wheat portal and genomic resources, then genetic and phenotype data for Vitis and Hordeum and finally explore integrated data set through BioMarts.

Day 2

9:00 - 12:00 Plant data resources at EMBL-EBI: Ensembl Plants (EMBL-EBI)

Ensembl Genomes (http://www.ensemblgenomes.org) is an integrative resource for genome-scale data from non-vertebrate species, including plants and plant pathogens. The project exploits and extends the vertebrate-focused Ensembl technology, originally developed for the human genome. Ensembl Plants (http://plants.ensembl.org) provides a complementary set of resources for plant species, providing a consistent set of interactive (web) and programmatic (API and MySQL) interfaces. We currently host 22 plant genomes and provide access to data including: reference sequence, gene models, transcriptional data, polymorphisms and comparative analysis. Since its launch in 2009, Ensembl Genomes has undergone rapid expansion, with the goal of providing coverage of all major experimental organisms, and additionally including taxonomic reference points to provide the evolutionary context in which genes can be understood. Against the backdrop of a continuing increase in genome sequencing activities in all parts of the tree of life, we seek to work, wherever possible, with the communities actively generating and using data, and are participants in a growing range of collaborations involved in the annotation and analysis of genomes.

12:00 – 14:00 Methods of standardisation and annotation for plant phenotypic data (IPG PAS)

One of the transPLANT aims is to develop community-accepted standards for data description and submission, covering format, and content and policy. In this part of the workshop we will concentrate on the topics of biological data formatting and annotation possibilities. Currently available tools and Internet resources helpful for these tasks will be presented.

14:00 End of workshop