transPLANT Newsletter Spring 2014 Issue 1

View this email in your browser

transPLANT

trans-National Infrastructure for Plant Genomic Science

INSIDE THIS ISSUE

TRAINING AND EVENTS

- 13-14 October 2014: 3rd transPLANT User Training Workshop on "Exploiting and Understanding Solonaceous Genomes". To be held at the DLO, Wageningen, Holland. Registration details will follow.

COMMENTS?

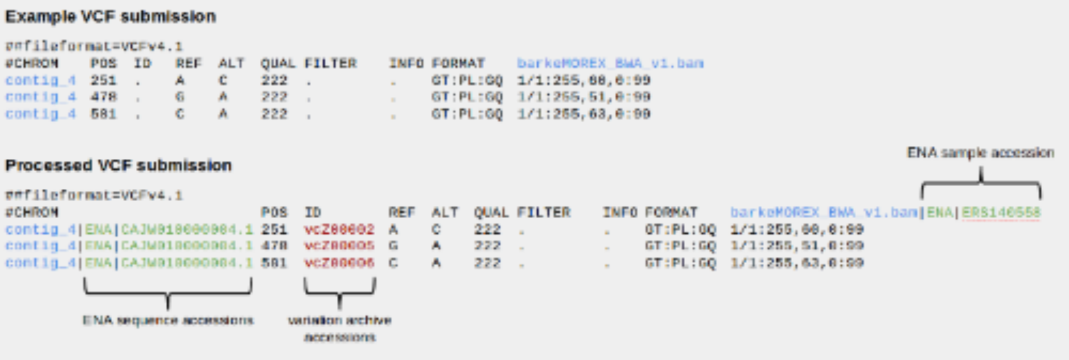
- We want to hear from you! Any comments or suggestions please contact us at transplant_help@ebi.ac.uk

About the transPLANT Project

transPLANT is a consortium of 11 European partners gathered to develop a trans-national infrastructure for plant genomic science. Bringing together groups with strengths in data analysis, plant science, computer science and from the academic and commercial sectors, transPLANT is developing integrated standards and services and undertake new research and development needed to capitalise on the sequencing revolution, across the spectrum of agricultural and model plant species.

transPLANT is committed to establishing the broadest international collaborations for data and standards. Explore the project's aims in more detail at: http://www.transplantdb.eu.

Variation Archive Accepting Submissions



Example VCF submission

```
##fileformat=VCFv4.1
#CHROM    POS  ID   REF  ALT  QUAL FILTER    INFO FORMAT   barkeMOREX_BWA_v1.bam
contig_4  251  .    A    C    222  .         .    GT:PL:GQ 1/1:255,60,0:99
contig_4  478  .    G    A    222  .         .    GT:PL:GQ 1/1:255,51,0:99
contig_4  581  .    C    A    222  .         .    GT:PL:GQ 1/1:255,63,0:99
```

Processed VCF submission

```
##fileformat=VCFv4.1
#CHROM                        POS  ID       REF  ALT  QUAL FILTER    INFO FORMAT   barkeMOREX_BWA_v1.bam|ENA|ERS146558
contig_4|ENA|CAJW010000084.1  251  vcZ00002 A    C    222  .         .    GT:PL:GQ 1/1:255,60,0:99
contig_4|ENA|CAJW010000084.1  478  vcZ00005 G    A    222  .         .    GT:PL:GQ 1/1:255,51,0:99
contig_4|ENA|CAJW010000084.1  581  vcZ00006 C    A    222  .         .    GT:PL:GQ 1/1:255,63,0:99
```

transPLANT variation archive

The transPLANT variation archive stores, accessions and updates plant variation data! We are now accepting submissions in VCF format on public reference sequences. Submit a VCF or browse the archive.

Read more...

Overview

1. Sign up
2. Upload VCF
3. Enter meta-data
4. Submit

Genome-scale resequencing is now widespread in plants, to identify variation in natural populations and crop stocks. This not only provides insight into evolutionary mechanisms and functional processes, but additionally, allows for the rapid statistical correlation of variant loci with quantitative traits, and the identification of probable causative genes.

The transPLANT variation archive has been developed to provide a system for the persistent storage and analysis of variation data from plant species. We are now accepting submissions of variant data in Variant Call Format (VCF) on known reference sequences, i.e. sequences present in the databases of the International Nucleotide Sequence Database Collaboration (INSDC), i.e. the European Nucleotide Archive (ENA), GenBank, and the DNA Database of Japan.

Submitted data will be:

- Persistently archived via the ENA.
- Uniquely accessioned per submission and at *each variant locus*. Data from different submissions referencing the same locus will be assigned a common, non-redundant identifier in the context of the given reference sequence.
- Propagated to new versions of the reference sequence.When a new version of a reference sequence is submitted to the INSDC, all accessioned variants will be mapped onto the new reference sequence (where possible).
- Made available for download. Both the original (and accessioned) submission and any subsequent updates will be made available, e.g. here.
- Made available in the Ensembl Plants interface. Where the submitted data is located on a reference sequence that is used in the Ensembl Plants database, it will be visible there.

This work is being carried out in the context of the development of the European Variation Archive, a new resource to organize and process genomic variation data for all species.

The archive is now open for submission!

For more information and to make a submission, please visit:
http://www.transplantdb.eu/variation/intro

The transPLANT Resource Registry

The number of plant genome sequencing (and re-sequencing projects is increasing and data repositories are becoming more numerous as a consequence. Many of the projects are producing genome-, transcriptome- and marker- based data from a range of technologies. Drawing on the broad expertise of the consortium, we developed and maintain a registry of important sequence-based resources for species of agricultural and economic importance as well as for plant model systems.

A total of **214 distinct plant genome resources** is currently registered at the **transPLANT genome resources registry** with regular updates scheduled every month. Users can query the registry with keywords such as species name or type of data and retrieve a list of resources together with URLs, provider details, associated tools and release version.

We implemented an intuitive user interface to submit new resources or updates to existing resources and like to invite all data providers as well as all interested users to contribute to this community effort with their expert knowledge.

The transPLANT genome resources registry is available for download in multiple file formats and as an online resource for search and browse.

http://www.transplantdb.eu/resources

http://mips.helmholtzmuenchen.de/plant/transplant/genomeResources.jsp

---

transPLANT Training Resources Now Online

The transPLANT training pages now include documents, presentations, tutorials and worked examples that were developed as part of the various transPLANT user training workshops and other transPLANT-related events.

Take a look at the training pages at http://www.transplantdb.eu/training-resources

Future plans include a blog page for transPLANT users, further training tips, "how-to" videos and a search feature. If you have any comments or suggestions about things you would like to see on the transPLANT site please do get in touch with us at transplant_help@ebi.ac.uk

---

Surveying community needs in plant informatics; the results of the transPLANT 2013 user survey

In 2013, transPLANT launched a user survey to collect bioinformatics stakeholder's needs in the field of agronomical research. The aims were:

- to identify potential needs that are not already covered by the TransPLANT project
- to help drawing the landscape of possible overlaps with other projects
- in order to better coordinate developments and avoid redundancies.

This survey was made accessible on the transPLANT web site: http://www.transplantdb.eu/survey

Take the transPLANT survey!

published by dbolser on Thu, 06/06/2013 - 13:08

We are collecting information about the needs of bioinformaticians in the field of agronomic research. Our goal is to chart the current resource landscape to both coordinate its development and to identify gaps that can be filled by transPLANT.

The survey contains 41 questions. Thanks in advance for your involvement.

Following EU review recommendations, this survey aims at collecting bioinformatics stakeholders needs and to describe some existing initiatives launched in the field of agronomical research to be compared to TransPlant scope. The goal is to identify potential needs not already covered by TransPlant project. The second objective is to draw the landscape of possibly overlapping initiative in the field or related fields to better coordinate developments and possibly avoid redundancy.

Based to this survey and on other actions of WP2 and WP4, EU will be given in summer 2013 a review on existing initiatives in the field of transnational research for bioinformatics in crop science and an overview of the needs of the stakeholders in terms of bio-informatic applications.
This survey contains 40 questions. Thanks you in advance for your involvement.

URL:

https://fr.surveymonkey.com/s/83KGSTS

It was addressed to both scientists from academic and private sectors, working on wheat, barley, maize, pea, sunflower, rapeseed…, genomics and genetics. More than 200 people were contacted by email, and the survey contained 41 questions covering topics about biological data types, databases, analysis tools and infrastructure needs.

A first survey analysis was done in autumn 2013 based on 74 responses. The key findings were as follows:

- 80% have a problem of **data integration**. They have difficulties because of the scale of the data (56,3%): algorithms and tools are not well adapted. They have limited information about strengths/weaknesses of available softwares (46,9%).
- However they agree to share with others their elaborated (compute and expertized) data. It concerns genomics sequences (55%), resequencing data (52,5%), genotyping data (50%), RNASeq expression data (52,5%).
- **Problems in analyzing their data** because of species complexity (43,8 %), of rapid technological change (40,6%) and because of problem to have / to hire skilled staff.
- To solve these problems, they would like to have **more user training sessions** (85,7%), to have a **unique web portal** (71,4%) easy to query with guidelines to access tools and data resources. They would like to have also a **forum** (45,7%) to exchange with people and **to receive newsletters** (28,6%) to be able to get synthesized information. Their preferred way to be informed is **on their own** (55%), **by mailing** (51%) and by **newsletter** (41%).

Developing Standards for Plant Phenotyping Data

transPLANT is working on standardisation of plant phenotypic data within its WP3: 'Community standards for interoperability of data resources'. In the field of phenotypic datasets' standardisation our activities concern three layers:

- Standardisation of metadata: Minimum Information about Plant Phenotyping Experiment standard **(MIAPPE)**
- Standardisation of vocabulary: recommendations regarding the use of ontologies
- Format for phenotypic data exchange by flat files and web services **(ISA-Tab for Phenotyping)**

All information about the progress in our work are being published on the portal http://cropnet.pl/phenotypes. You can find there our propositions of documents and formats, together with descriptions of updates and feedback from the community, as well as information about compatible tools for handling phenotypic data.

Partners, community members and groups interested in phenotyping are welcome to follow the portal and get involved in the standardisation process by providing their feedback on the proposed standards.

To achieve a broad agreement on phenotyping standards we are in collaboration with European Plant Phenotyping Network (EPPN), who target similar problems as one of their activities (EPPN standards). Development of MIAPPE and ISA-Tab for Phenotyping takes place in contact with BioSharing and ISA-Tab team.

Back to top

The transPLANT project is funded by the **European Commission** within its 7th Framework Programme, under the thematic area "iInfrastructure", **contract number 283496.**