

# **Draft Landscape Study**

**Ruth Bastow and David Marshall**

# Timeline

- April 2015 – Ruth begins landscape study as per instructions from SC
- May 2015 – Landscape study presented to SC; SC request widening of the study
- June 2015 – Request for additional DivSeek related projects emailed out to DivSeek Partners
- December 2015 – 2<sup>nd</sup> version of study presented to SC
- January 2016 – 1<sup>st</sup> online version of study generated by David

# Reasons for Study

- Try to uncover who is doing what where
  - ❖ Understand the landscape in which DivSeek sits
  - ❖ Uncover potential new DivSeek Partners
  - ❖ Identify potential exemplar projects
  - ❖ Assist with boundary setting for DivSeek
  - ❖ Assist with sharing of information and expertise
  - ❖ Facilitate interactions amongst projects

# Landscape Statistics to date

- 95 Projects identified to date
- 43 - main focus on Informatics
- 3 - only focus on Genome sequencing
- 23 - main focus on Diversity Assessment; molecular/phenotypic or both
- 26 Combination of Informatics, Genome sequencing and Diversity analysis
- [www.divseek.org/](http://www.divseek.org/)

# Landscape Statistics to date

- Projects from across the globe
- Majority in North America and Europe
- Currently working with colleagues in China, Brazil and India to begin to identify projects in these countries
- Welcome input from others to fill gaps and inform us about other projects

# 1<sup>st</sup> Online Version of Study

DivSeek Project Landscape Study  
57 views

All changes saved in Drive

Add layer Share

Projects

- Individual styles
- 1000 Plant Genomes Project
- The UK Barley Genome Sequ...
- 3,000 Rice Genome Project
- African Orphan Crop Consort...
- AMAIZING
- BamYIELD
- BreedWheat
- BRIDGE**
- Cacao Genome Database
- CerealsDB
- Conservation and Divergenc...
- CottonGen
- EURISCO
- EU Tree4Future (2012-2016)
- Genomes To Fields (GZF)
- Geospatial database of colle...

**BRIDGE**

Aims at evolving the Federal ex situ Genebank of IPK Gatersleben from a 'storage facility' of genetic resources to an integrated resource and information centre, representing a one stop shop for facilitated and informed utilization of crop plant biodiversity. <http://bridge.ipk-gatersleben.de/>

51.82666, 11.27726

Google My Maps

# Conclusions/ Questions Raised

- Large number of groups interested in using contemporary molecular technologies to study crop diversity
- Wide variety of approaches
  - Genotyping
    - Array methods
    - Genotyping by sequencing
  - Sequencing
    - Whole genome
    - Exome capture
- High throughput phenotyping is increasingly relevant

# Data comparability/integration

- Multiple projects, even within a species such as maize and wheat, may be using different genotyping or sequencing technologies anchored on different references
- This raises questions of data integration and comparability which may be addressed to some extent by awareness and sharing of common germplasm, ideally at the time of project planning



# Definition of the operational unit

- There is often a tension between the practical germplasm handling unit of an “accession” in genetic resource collections and the association of genotypic and phenotypic data with a genetically defined entity
- This raises a number of issues of strategy such as the purification of lines by inbreeding or the genotyping of what are effectively “populations”
- In turn this results in logistical challenges for genebank managers through increasing the size of their collections in order to store and distribute purified lines
- It can also complicate analytical options if genotype and phenotype data now longer map to the same genetic individual

# Shared experience/resources

- There are considerable options for sharing experience/best practice for DivSeek projects
- This may relate to:
  - overall strategy
  - experience of “what not to do” or particular problems with particular crop types e.g. in cases of high diversity or polyploidy
  - The sharing of software tools and analysis pipelines not only between “DivSeek projects” but also with the more general breeding and genetic communities eg GOBII or BMS
  - the issues in choosing traits that can be meaningfully phenotyped in exotic germplasm

# What are the strategy issues?

- With given resource limitation - as all potential projects will have - how should this best be allocated?
- Should we sequence or genotype?
- How much and what phenotyping should be done?
- What portion of a collection should/could be characterised?
- Do we wait till technology is cheaper or better? Will we have to do it again to replace low density genotyping with high density genotyping or sequencing?
- For a given species can or should we use the same technology or reference sequences across projects?