



Genomic Open-source Breeding informatics initiative

*GOBii: Connecting genotyping data, breeding data
and analysis tools to facilitate breeding decisions*

BILL & MELINDA
GATES foundation



BTI BOYCE
THOMPSON
INSTITUTE



CIMMYT
International Maize and Wheat
Improvement Center

ICRISAT

IRRI

The James
Hutton
Institute

Diversity
Arrays
Technology
IT'S A BULLSEYE WITH DART

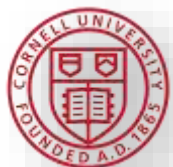
GOBii Mission



To transform breeding by enabling the implementation of **genomic and marker-assisted selection** as part of **routine** breeding programs in developing countries

gobiiproject.org

GOBii A Global Community



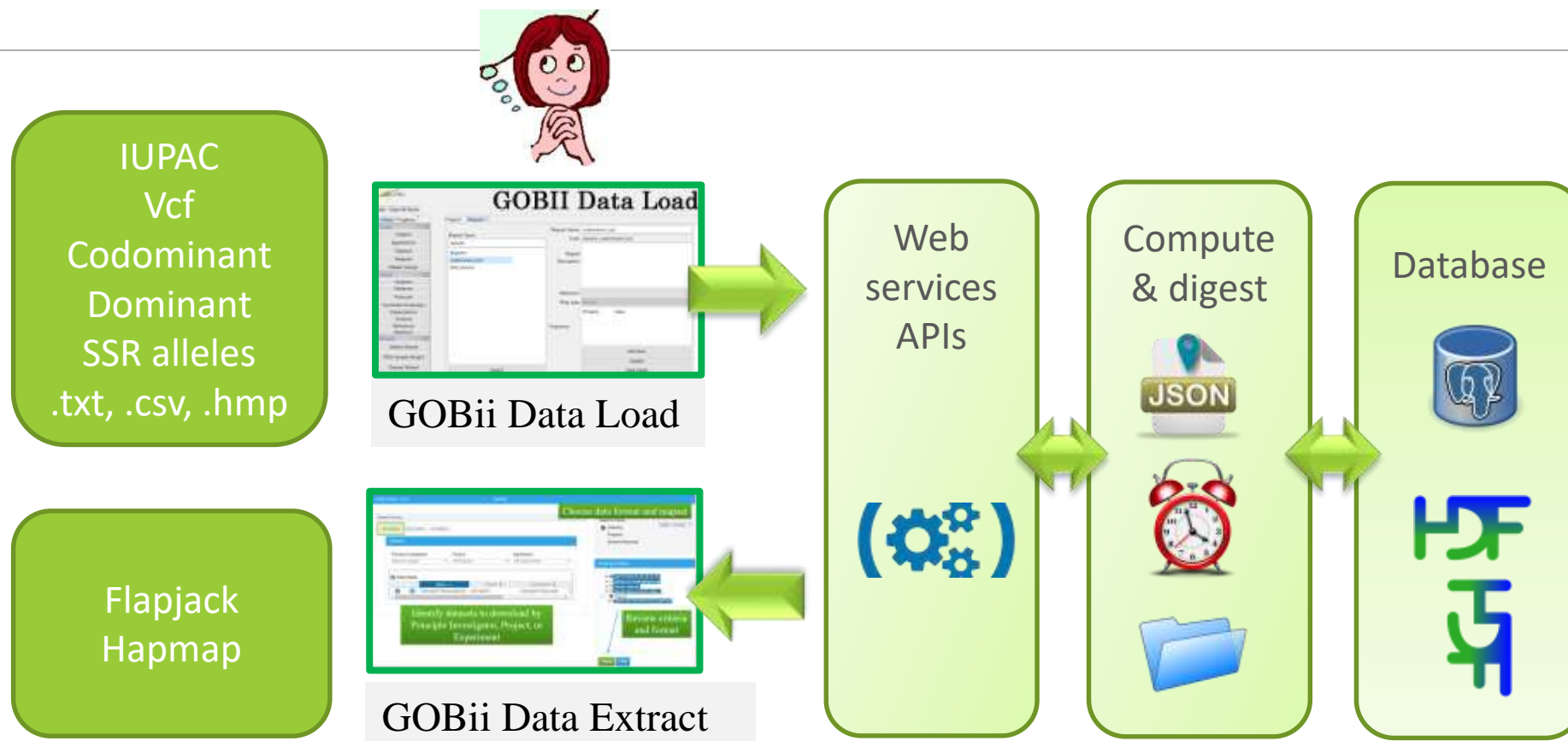
The GOBii Project



- Provide database tools to manage genotyping data from *multiple* genotyping platforms for any crops
- Provide user interfaces to *query across datasets* by samples and markers
- Provide analyses and visualizing tools to *support data curation and breeding decisions*
- Provide solutions to *integrate* genotyping data management to adjacent breeding management, and sample tracking systems, and downstream tools
- Develop a *community* of knowledge through training and consulting

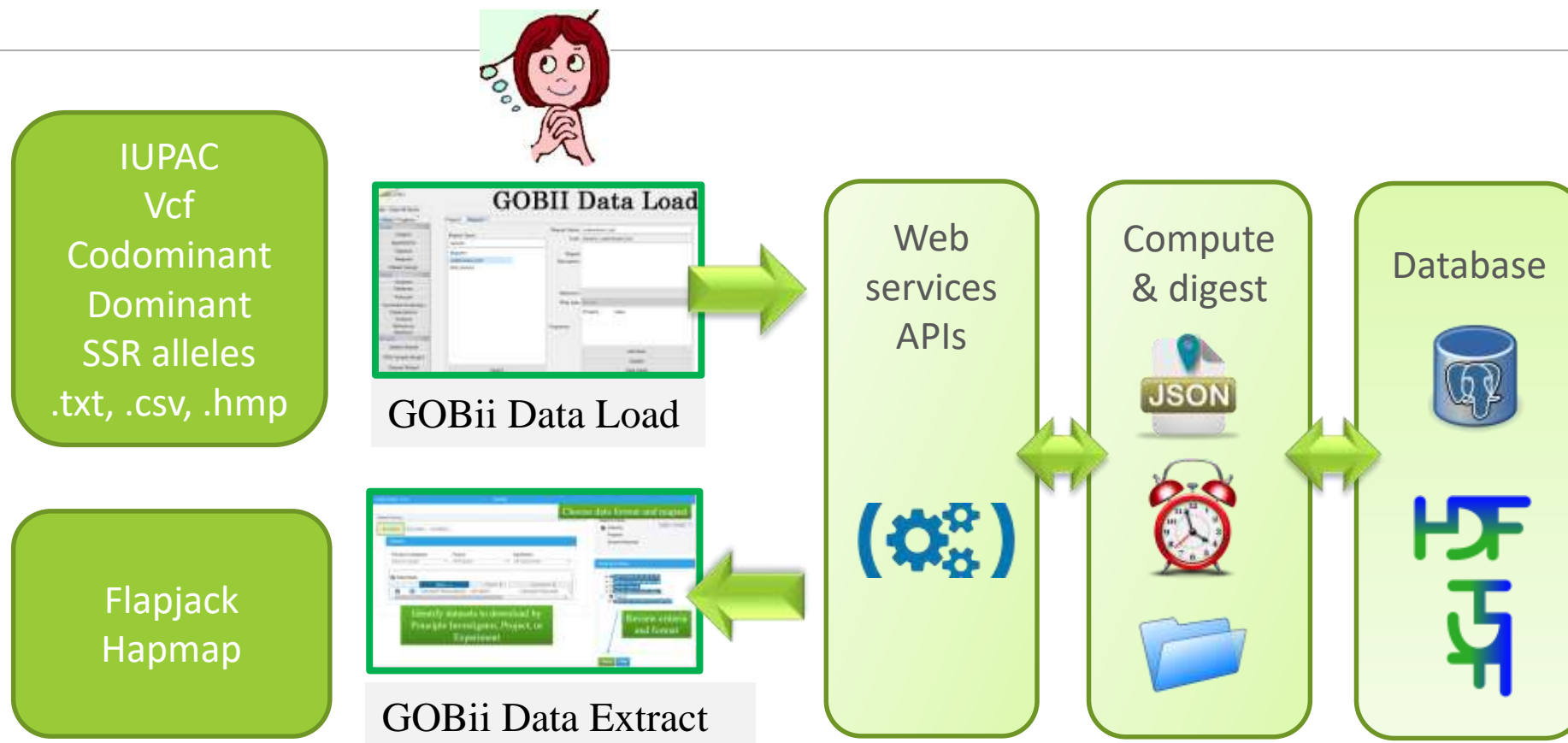
GOBii Genomic Data Management (GDM)

‘Designed with extract performance in mind’



GOBii Genomic Data Management (GDM)

‘Designed with extract performance in mind’



Bioinformatics for Large Scale Genotyping Data Management and Analytics: Mon Jan 14th 4pm

Demo GOBii Database: Digital Tools and Resources Tues Jan 15, 12:00pm

GOBii Tools



Curator tools: Data QC

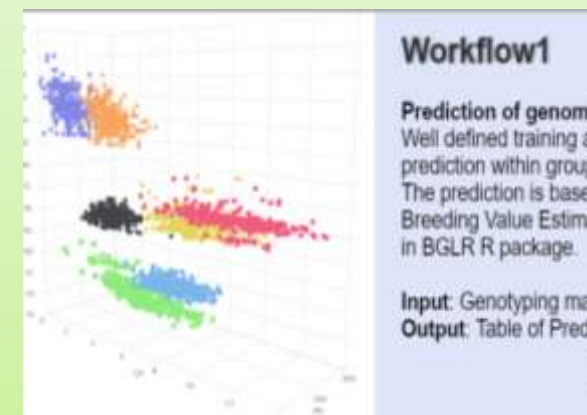


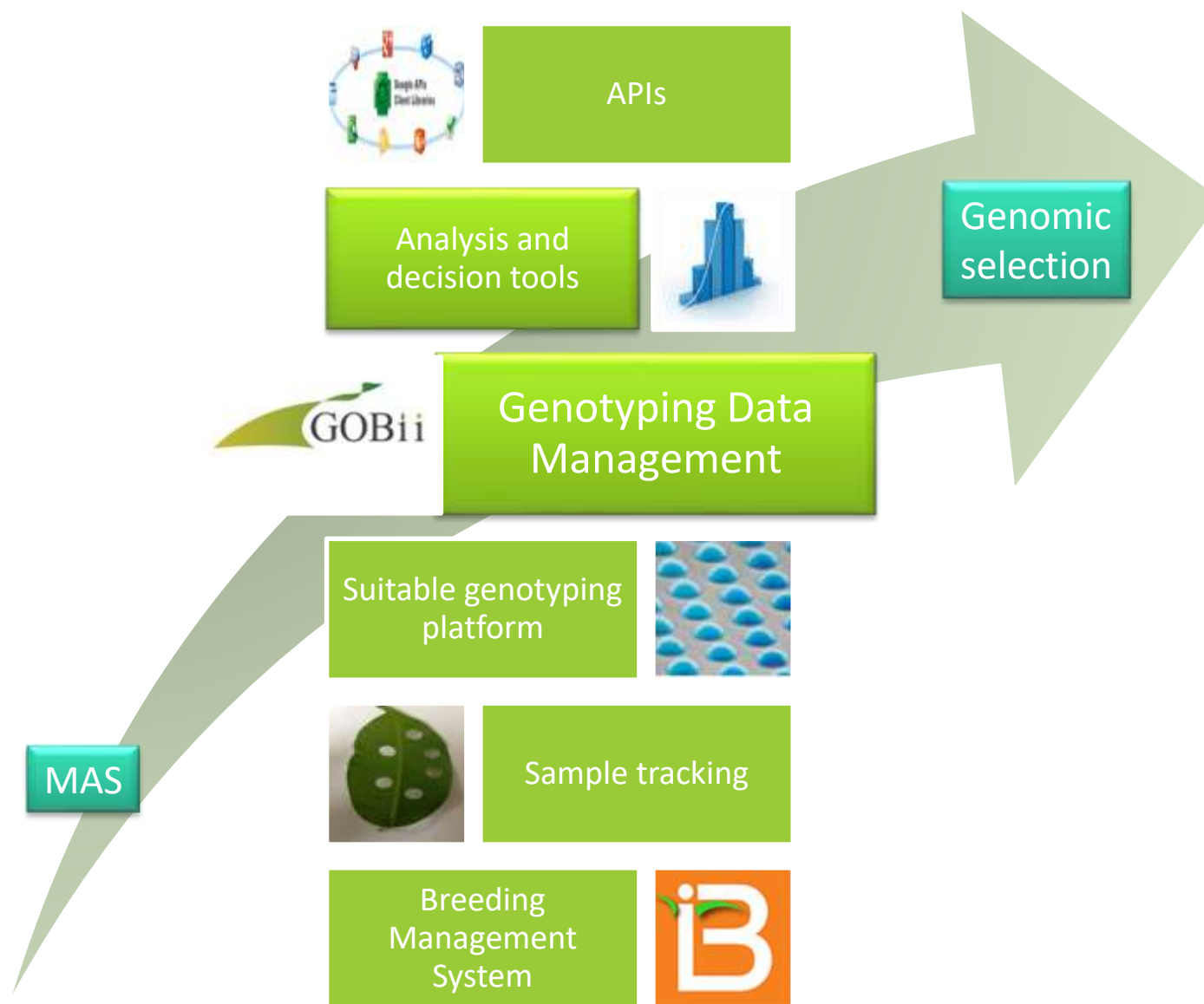
*MABC
Ped ver
Forward Breeding*

Par2	100	58.97	100	C	C	C	C	T	C
Par5	58.97	100	100	C	C	A	C	T	C
Testline1	82.5	76.92	100	C	C	A	C	T	C
Testline2	71.05	61.54	86.84	A		A	C	T	C
Testline3	52.5	45	69.23	A	G	C	G	T	G
Testline4	32.43	48.65	58.33	A	A	CG	C	G	G

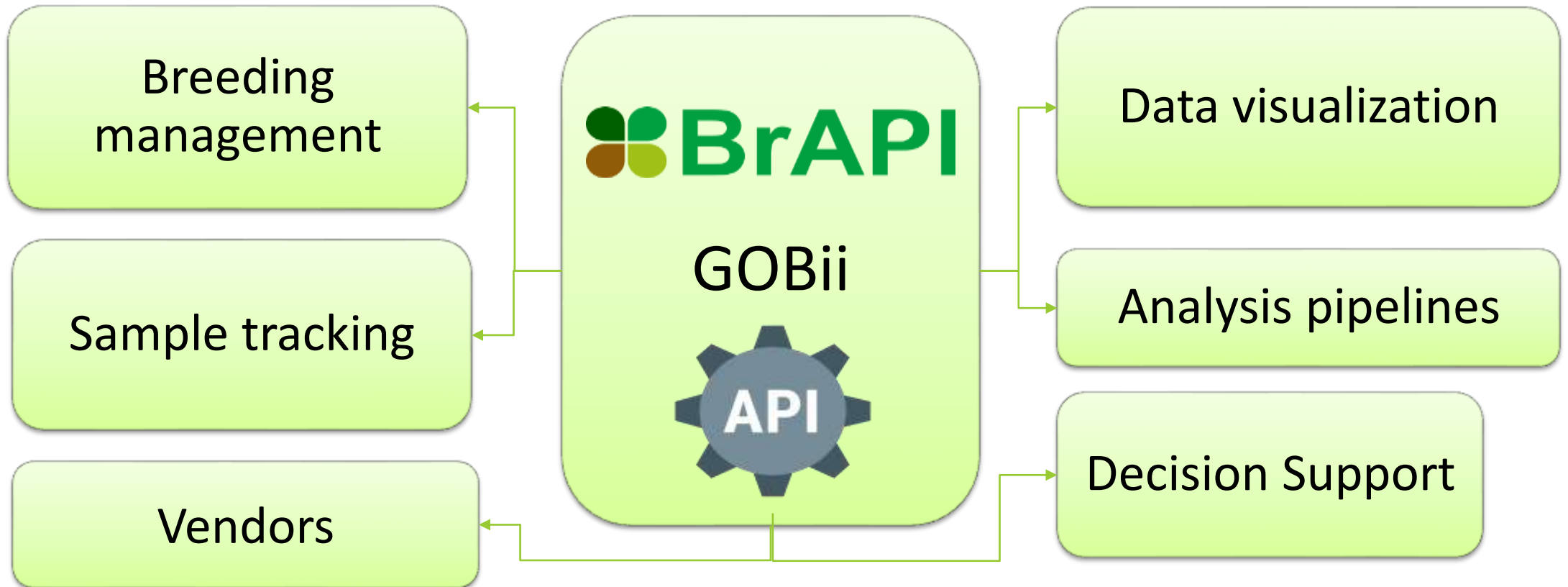


Galaxy
PROJECT

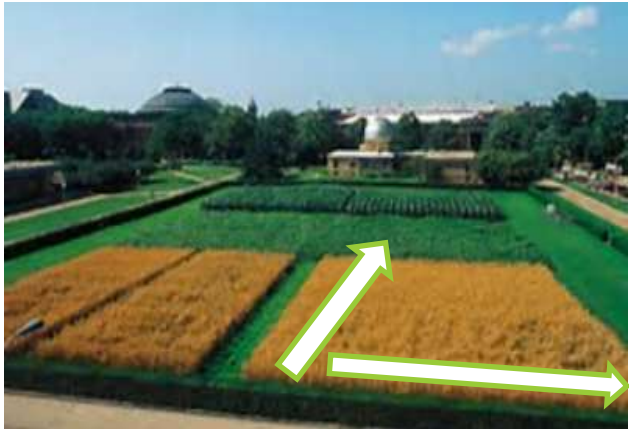




GOBii Integration



The sample tracking use case: from field to lab and back



Sample ID

P1	40.00	true	A	C	A	G	C	A
P2	40.00	true	C	C	C	C	C	G
Line 1: P1xP2	100.00	true	A	C	A	G	C	A
P1	100.00	true	A	C	A	G	C	A
F1-2	100.00	true	A	C	A	G	C	A
F1-3	40.00	false	A	C	A	G	C	A
F1-4	100.00	true	A	C	A	G	C	A
F1-5	94.74	false	A	C	A	G	C	A
F1-6	65.00	false	A	C	A	G	C	A
F1-7	100.00	true	A	C	A	G	C	A
F1-8	100.00	true	A	C	A	G	C	A
F1-9	100.00	true	A	C	A	G	C	A
F1-10	40.00	false	A	C	A	G	C	A
F1-11	100.00	true	A	C	A	G	C	A
F1-12	100.00	true	A	C	A	G	C	A

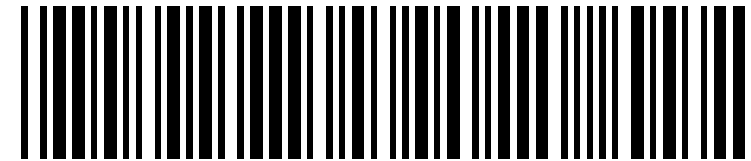
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
C	T	T	G	A	G	A	A	A	A	T	T	C	T	T	A	G	A	T	A
T	C	T	A	A	A	G	A	A	A	T	T	A	T	A	T	A	A	T	A
T	T	T	A	A	G	A	A	A	A	T	T	C	T	T	A	G	A	T	A
T	T	T	A	A	G	A	A	A	A	C	T	T	C	A	C	A	A	T	A
C	T	T	A	A	G	A	A	A	A	T	T	C	T	T	A	G	A	T	A
C	T	C	A	T	G	A	A	A	A	T	T	C	T	T	A	G	A	T	A
C	T	T	A	T	G	A	A	A	A	T	T	C	T	T	C	G	A	T	A
T	T	T	A	A	G	A	A	A	A	T	T	C	T	T	C	A	A	T	A



Tracking samples back to plants



GOOD



ABC1234

BETTER

012459708-4382920202-22938384747-22

BEST

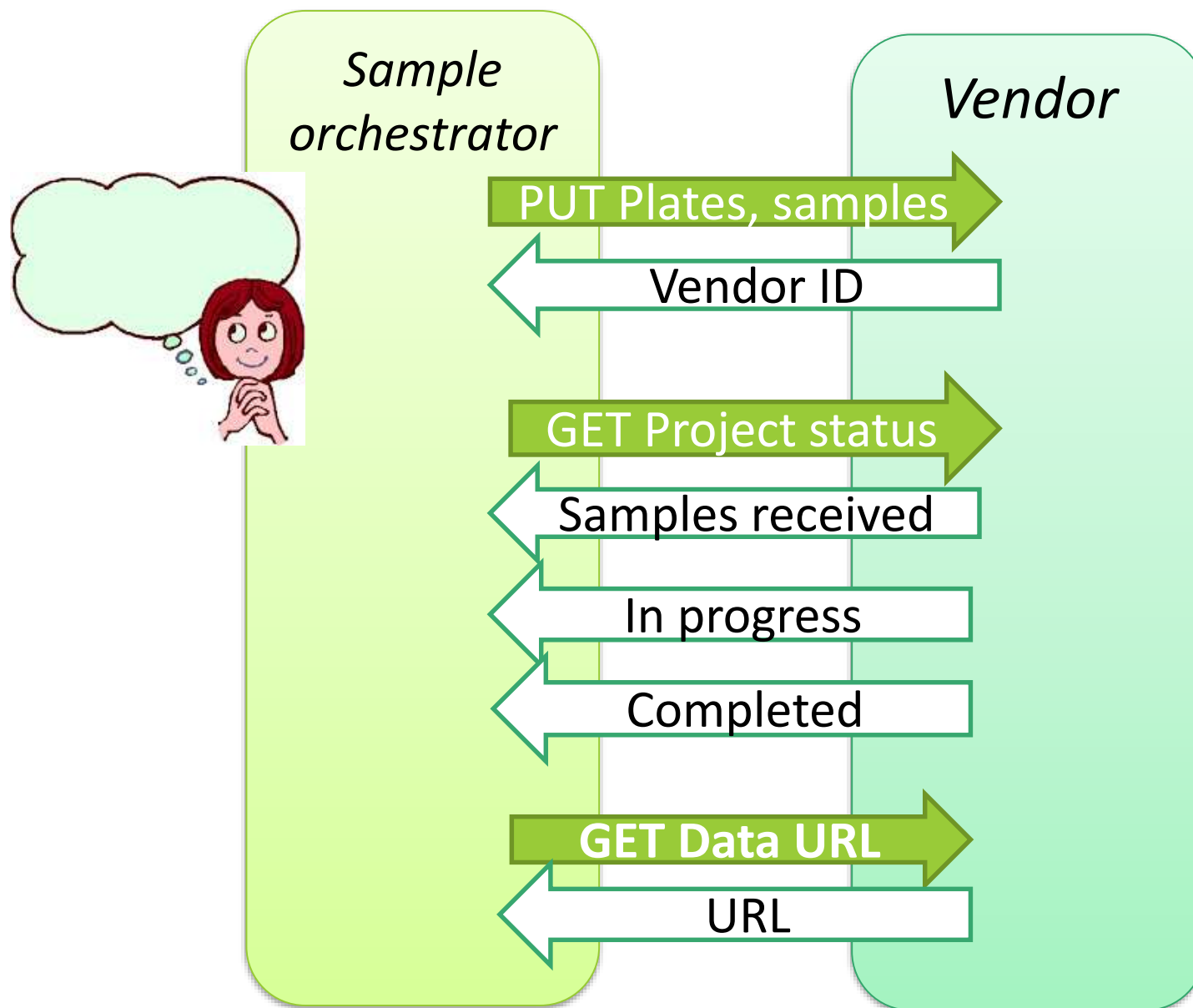
UUID - Universal Unique Identifier

Guaranteed to be different from any other **UUID** generated until 3400 A.D

Sample UUIDs



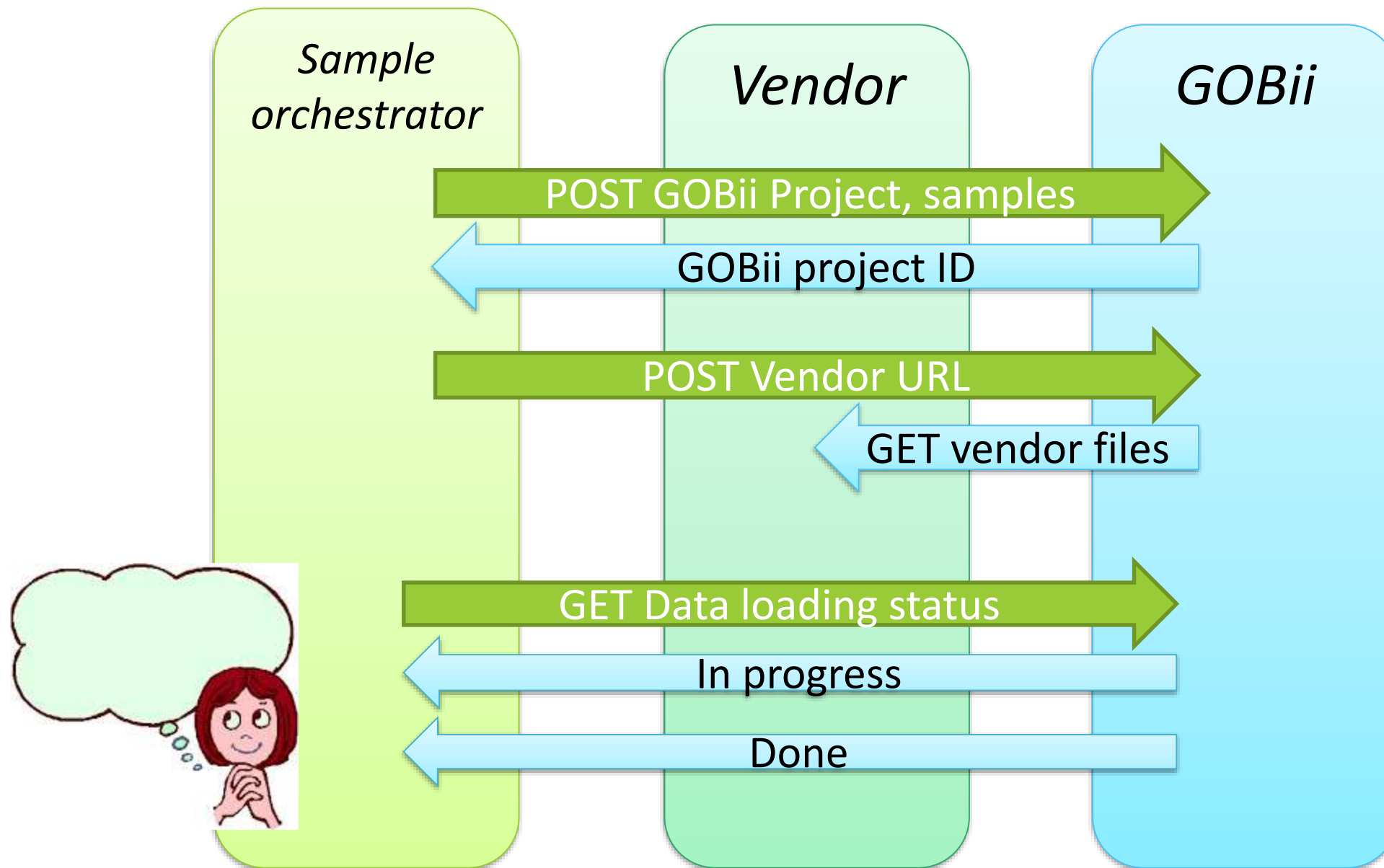
- Uniquely identify samples generated across databases and systems
- Can be generated in breeding management systems or in mini-apps 'Sample Tracker', 'Simple Tracker'
- Downside - hard to read
- Not always usable in vendor submission systems
- Can use an associated human readable name tracked by a sample tracking system
- 'Sample orchestrator' will connect sample UUID between systems



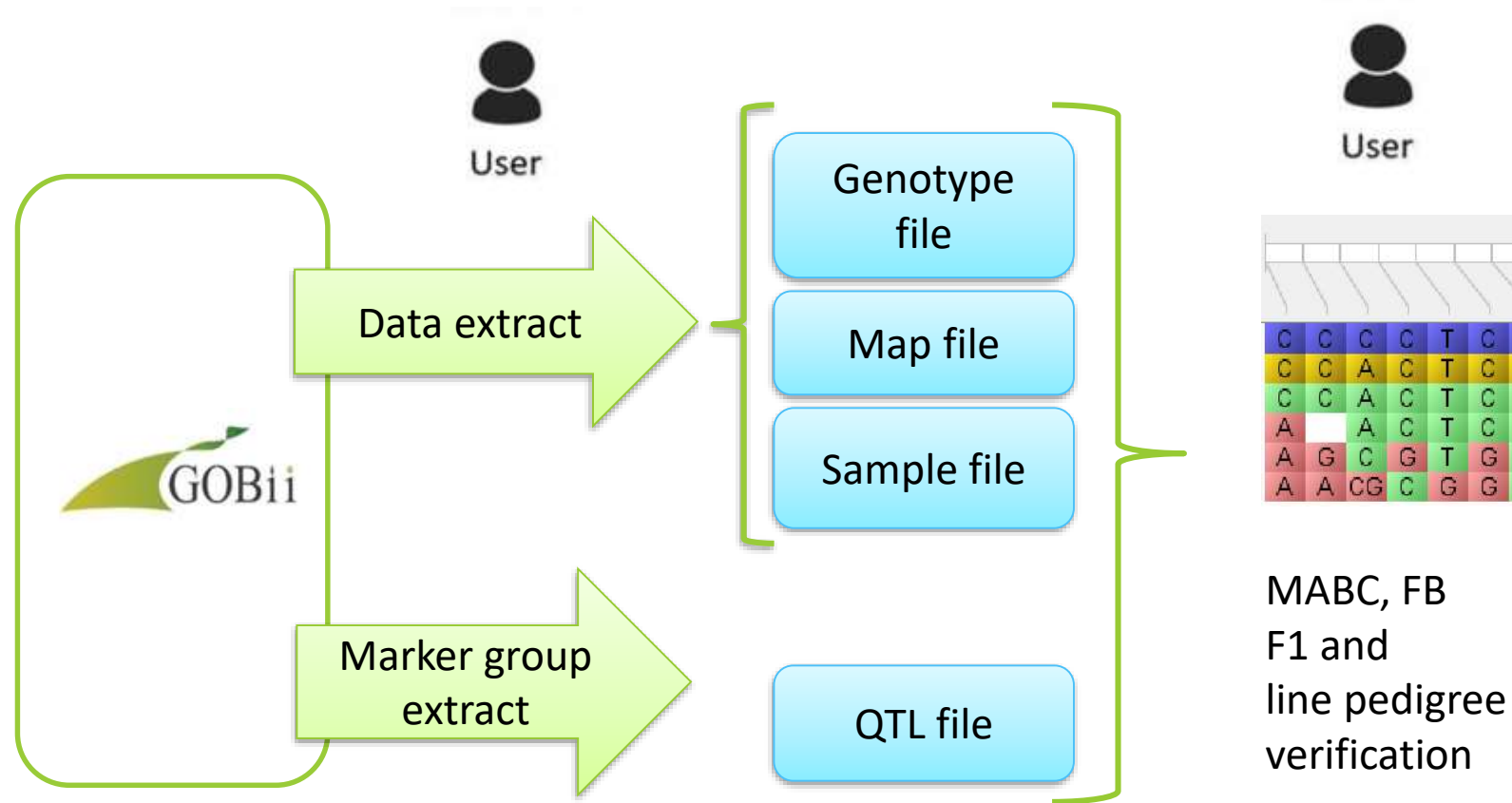
Connecting
sample UUIDs
across vendors
and databases

 **BrAPI**





From GOBii to Flapjack analysis



BrAPI Dataset View in Flapjack



Welcome to Flapjack

Import Data

Via BrAPI Maps and Genotypes Phenotyp

Use this tab to import map and genotype data into a new dataset.

Data files to import:

Map file (optional): k demo file maize MAP

Genotype file: no file

Advanced options:

Edit the advanced options...

Advanced options...

Import from BrAPI (V1.1)

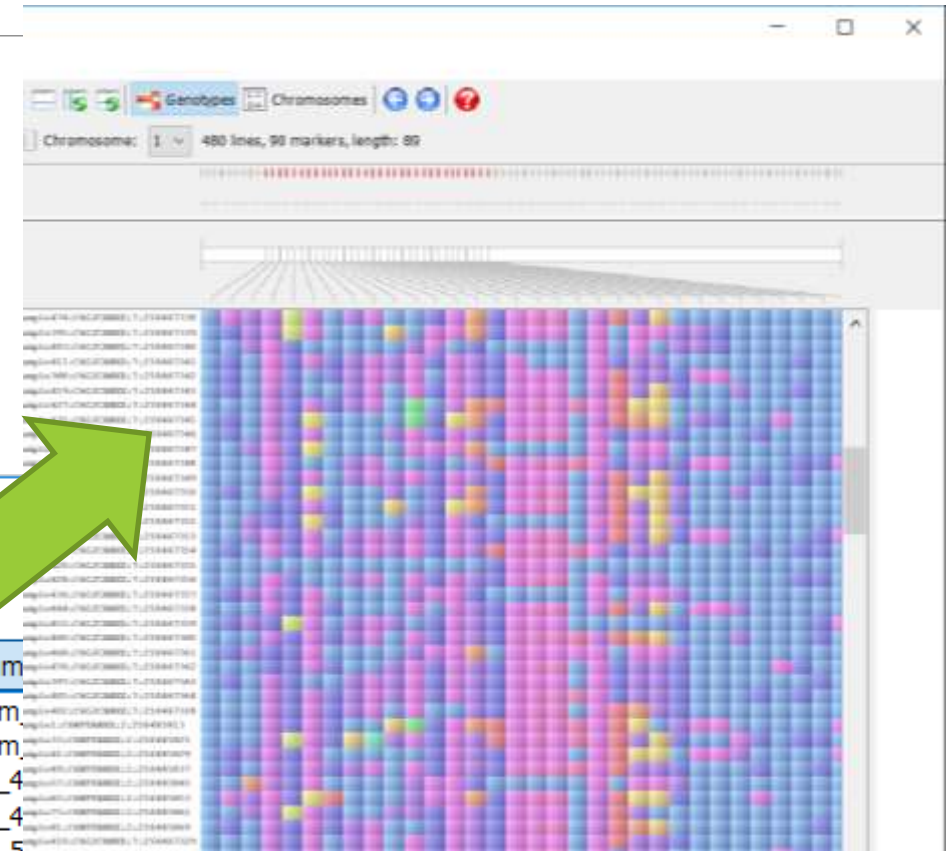
Matrix selection:

Available matrices:

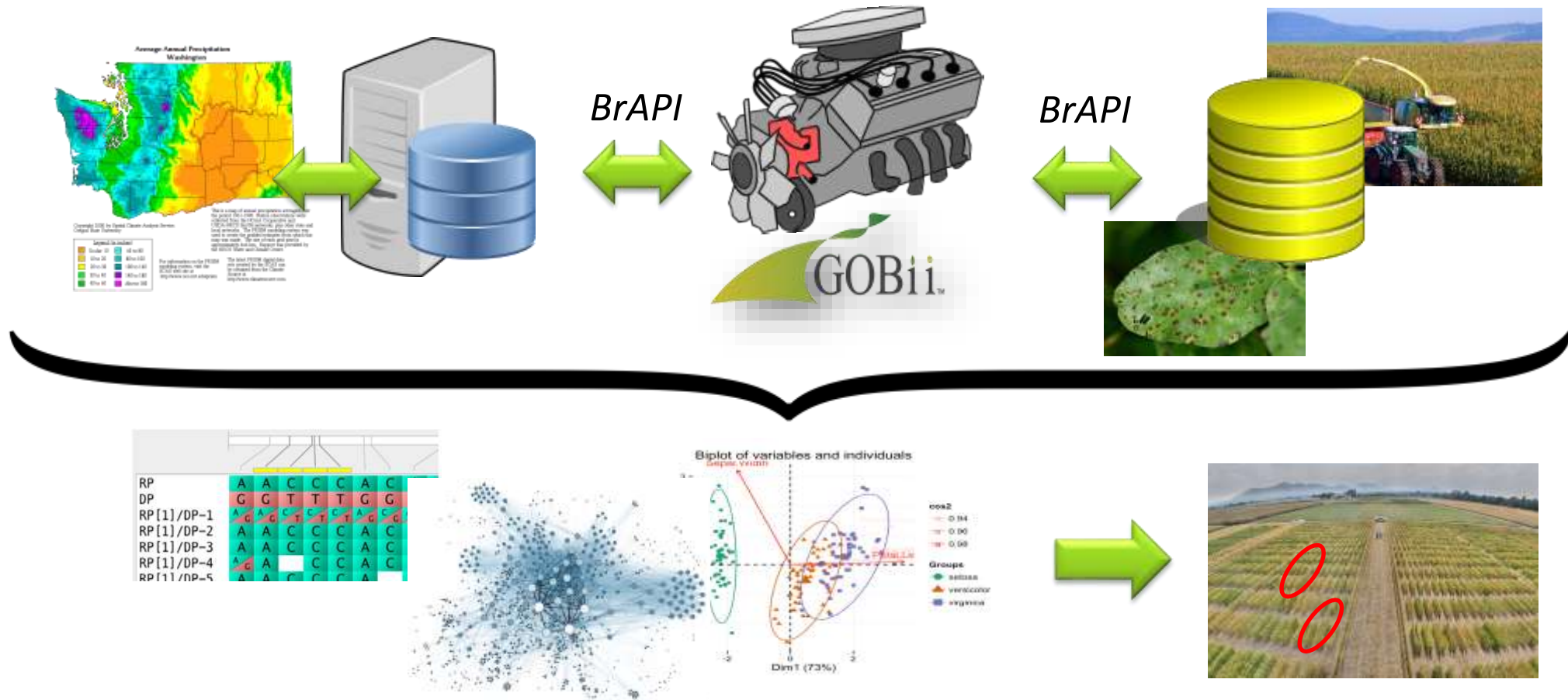
Details:

Name: Deb_GSD450
ID: 21
Sample count: 500
Marker count: 990

Deb_GSD450_Proj_Exp_Illum-Deb_GSD450_Proj_Exp_Illum
Deb_GSD450_Proj_Exp_Illum-Deb_GSD450_Proj_Exp_Illum
Deb_GSD450_Proj_Exp_Illum-Deb_GSD450_Proj_Exp_Illum
Deb_GSD450_Proj_Exp_Illum-Deb_QC_NPE_Illum_IUPAC_4
Deb_GSD450_Proj_Exp_Illum-Deb_QC_NPE_Illum_IUPAC_4
Deb_GSD450_Proj_Exp_Illum-Deb_QC_NPE_Illum_IUPAC_500_DS1 (500 * 500)
Deb_GSD450_Proj_Exp_Illum-Deb_QC_NPE_Illum_IUPAC_501_DS1 (500 * 501)
Deb_GSD450_Proj_Exp_vcf-Deb_GSD450_Proj_Exp_vcf_SSR_2ltr_DS (90 * 480)



GOBii Integration



Acknowledgement



Software Development Team



Angel Manica Raquel



Angel Villahoz-Baketa



Charlanya Samra



Dave Matthews



Faw Ni-Addae



Joshua Lamza-Sumwong



Kevin Palla



Phil Gieser



Selvarajogam Sivasekaramani



Syed Raza



Venice Margarette Juanillas



Viana Carla C. Calaminos



Victor Jun Ulat



Monica Franciscus



Application Team



Rosemary Dineen



Claudio Cesar AYALA HERNANDEZ



Kate Dreher



Renshi Zhang



Unwin Rozario



Roma Das



Anil Kumar



Adishesh Rathore



Manish Rooknath



Hema Doshi



Jeffrey Dattaz



John Carlos Ignacio



Liz Jones



Deb Weigand



Star Yamen Gas

Core System - GDM

Load, store, extract

Tools

Genomic selection
Imputation
Haplotype analysis
MABC, Pedigree Verification
Data QC

Integration

Tools
Vendors
Breeding management
systems

Community

seminars, meetings, hackathons

