

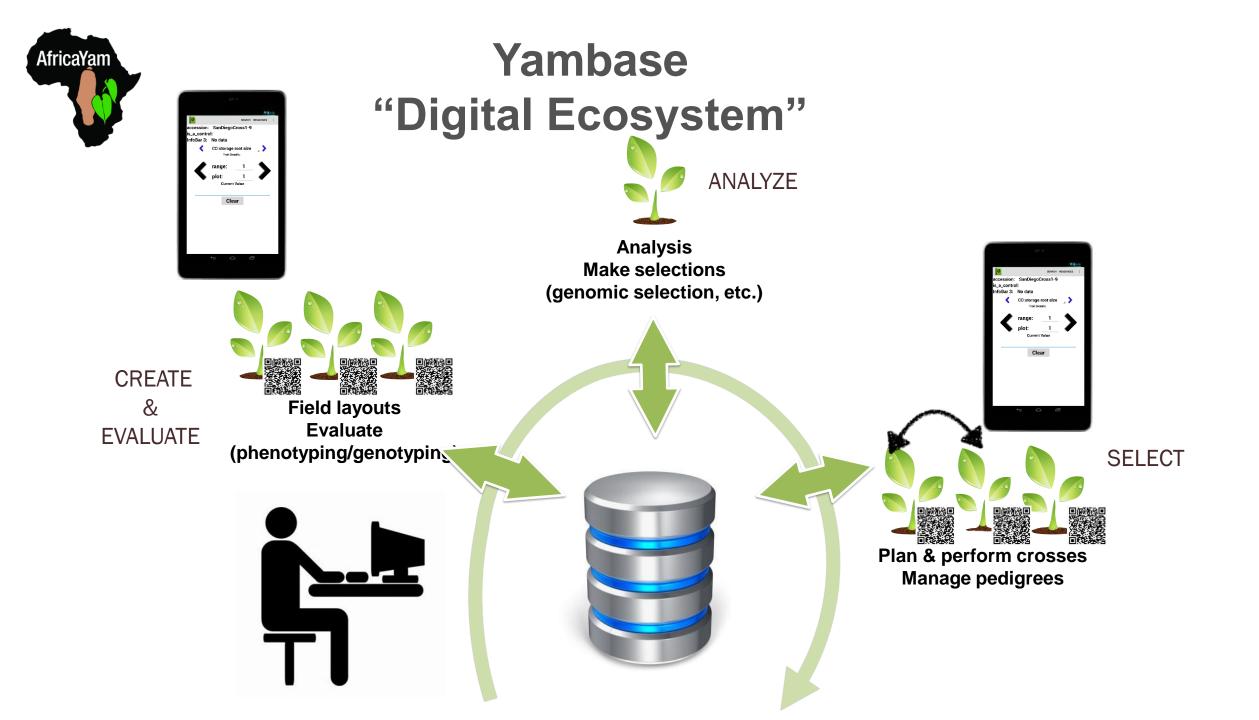
Yambase Update

Lukas Mueller and Isaak Tecle Boyce Thompson Institute



- Thanks to the entire Yam community
- And indeed the entire RTB community
- All the colleagues at BTI
- Thank you!!!







Yambase features

- Completely web-based
- Multi species, multi breeding program, multi user database
- Manage
 - accessions, fields, locations, traits, and measurements
 - genotypic information
 - Special high dimensionality phenotypes such as NIRS data
 - Product profiles
- Analyses
 - Mixed Models, Clustering, Stability, Heritability, SolGS

G3, 2022, **12(7)**, jkac078

GG3 Genes | Genetics

https://doi.org/10.1093/g3journal/jkac078 Advance Access Publication Date: 6 April 2022 Software and Data Resources Featured

Breedbase: a digital ecosystem for modern plant breeding

Nicolas Morales (b), ^{1,2,†} Alex C. Ogbonna (b), ^{1,2,†} Bryan J. Ellerbrock, ^{1,†} Guillaume J. Bauchet (b), ^{1,16,†} Titima Tantikanjana, ¹ Isaak Y. Tecle (b), ¹ Adrian F. Powell (b), ¹ David Lyon (b), ^{1,17} Naama Menda (b), ¹ Christiano C. Simoes, ¹ Surya Saha (b), ¹ Prashant Hosmani (b), ^{1,18} Mirella Flores (b), ¹ Naftali Panitz (b), ¹ Ryan S. Preble, ¹ Afolabi Agbona (b), ³ Ismail Rabbi (b), ³ Peter Kulakow (b), ³ Prasad Peteti (b), ³ Robert Kawuki (b), ⁴ Williams Esuma (b), ⁴ Micheal Kanaabi, ⁴ Doreen M. Chelangat (b), ⁴ Ezenwanyi Uba, ⁵ Adeyemi Olojede (b), ⁵ Joseph Onyeka, ⁵ Trushar Shah (b), ⁶ Margaret Karanja (b), ⁶ Chiedozie Egesi (b), ^{1,3,5} Hale Tufan (b), ² Agre Paterne (b), ³ Asrat Asfaw (b), ⁷ Jean-Luc Jannink (b), ^{2,8} Marnin Wolfe (b), ² Clay L. Birkett (b), ^{2,8} David J. Waring (b), ^{2,8} Jenna M. Hershberger (b), ² Michael A. Gore (b), ² Kelly R. Robbins (b), ² Trevor Rife (b), ⁹ Chaney Courtney, ⁹ Jesse Poland (b), ⁹ Elizabeth Arnaud (b), ¹⁰ Marie-Angélique Laporte (b), ¹⁰ Heneriko Kulembeka, ¹¹ Kasele Salum, ¹¹ Emmanuel Mrema, ¹¹ Allan Brown (b), ³ Stanley Bayo, ³ Brigitte Uwimana (b), ³ Violet Akech, ³ Craig Yencho (b), ¹² Bert de Boeck (b), ¹³ Hugo Campos (b), ¹³ Rony Swennen (b), ¹⁴ Jeremy D. Edwards (b), ¹⁵ Lukas A. Mueller (b), ^{1,*}

¹Boyce Thompson Institute, Ithaca, NY 14853, USA,
²Cornell University, Ithaca, NY 14853, USA,
³IITA Ibadan, 200001 Ibadan, Nigeria,
⁴NaCCRI, Namulonge, Uganda,
⁵National Root Crops Research Institute (NRCRI), 463109 Umudike, Nigeria,
⁶IITA Nairobi, 30709-00100 Nairobi, Kenya,
⁷IITA Abuja, 901101 Abuja, Nigeria,
⁸USDA-ARS, Ithaca, NY 14853, USA,
⁹Kansas State University, Manhattan, KS 66506, USA,
¹⁰Bioversity-CIAT Alliance, 34397 Montpellier, France,
¹¹TARI, 33518 Ukiriguru, Tanzania,
¹²North Carolina State University (NCSU), Raleigh, NC 27695, USA,
¹³CIP, 15000 Lima, Peru,
¹⁴KU Leuven, 3000 Leuven, Belgium,
¹⁵USDA-ARS, Stuttgart, AR 72160, USA

¹⁶Procent address: Terra da Lin, Saint Pierra la Vigar, Franco

OXFORD

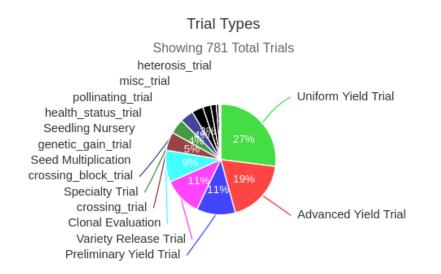


Data in Yambase

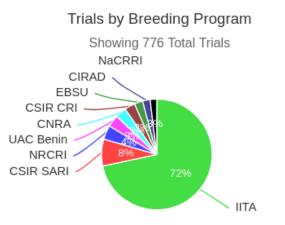
https://yambase.org/breeders/dbstats

Database Stats

Data categories



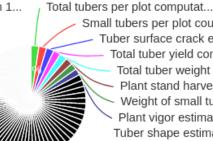
Traits





Total trait measurements: 1305371 Twining direction estimation 1... Tuber flesh color measurement Tuber dormancy period estimati... ig tuber yield adjusted compu... ybug incidence computation... knot nematode incidence c... iber of leaves per plant cou... uit set number estimation 1-... First leaf emergence time esti... Fruit index computation 0 to 1 Tuber flour starch content com...

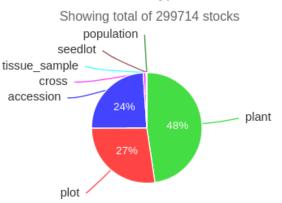
Tuber flour sugar content comp... Tuber flour protein content co...



Small tubers per plot counting ... Tuber surface crack estimation ... Total tuber yield computation ... Total tuber weight per plot co.. Plant stand harvested count Weight of small tubers per plo Plant vigor estimation 1-3 Tuber shape estimation 1-4 Appearance of tuber surface es. Tuber thorniness intensity est...

Tuber surface texture estimati... Big tubers per plot counting n...

Stock Types



Analyze

Maps Ab

Analyses & Models

Breeder Tools

Selection Index Genomic Selection Population Structure Clustering Kinship & Inbreeding Stability AMMI/GGE Heritability

Accession Usage Mixed Models Compare Trials Graphical Filtering BoxPlotter Tool GWAS Image Analysis

Sequence Analysis

Analysis Tools

- Many routine analyses available on Cassavabase
- "Analysis framework" most analyses work in a similar way
- Review Cassavabase statistics capability





Titima Tantikanjana



Isaak Tecle

Bryan Ellerbrock



David Lyon





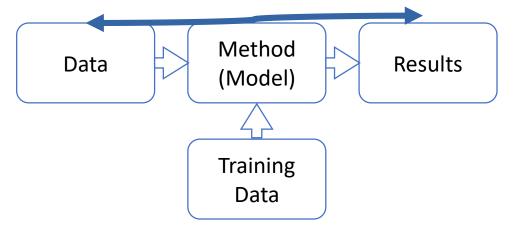
Adrian Powell

Chris Simoes

Christine Nyaga

AfricaYam Working with Analyses

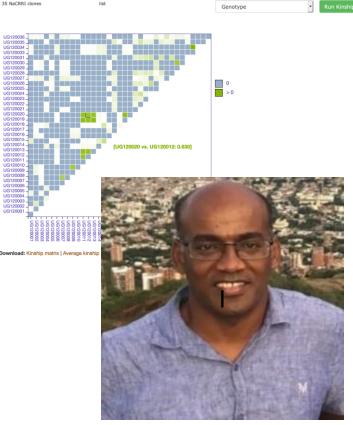
- Analyses consist of input data, methods, models, training data, and results
- All these metadata can be stored for an analysis
- Analysis results, for example, predictions from solGS, can be stored in the database to inform downstream breeding decisions (selection index etc)
- Stored analysis results can be used in other analyses





Genomic Selection (solGS tool)

- Integrates genotypic and phenotypic data
- Training and prediction
- New features:
 - Input based on Wizard datasets
 - Explore data, estimate and visualize kinship and inbreeding coefficients
 - K-means clustering tool
 - Improved visualization of results
 - Store GEBVs in the database for further decision support
 - Performance improvements



Isaak Tecle



Mate selection prediction

- Started with Africa Yam Project
- Predict best parent pairs
- Based on genomic predictions
- Implemented by Christine Nyaga, based on code by Marlee Labroo, Marnin Wolfe, and Isaak Tecle
- Re-implementation of Mixed Model tool using Rpackage "sommer" as backend to allow inclusion of more co-variates



Christine Nyaga



Marlee Labroo

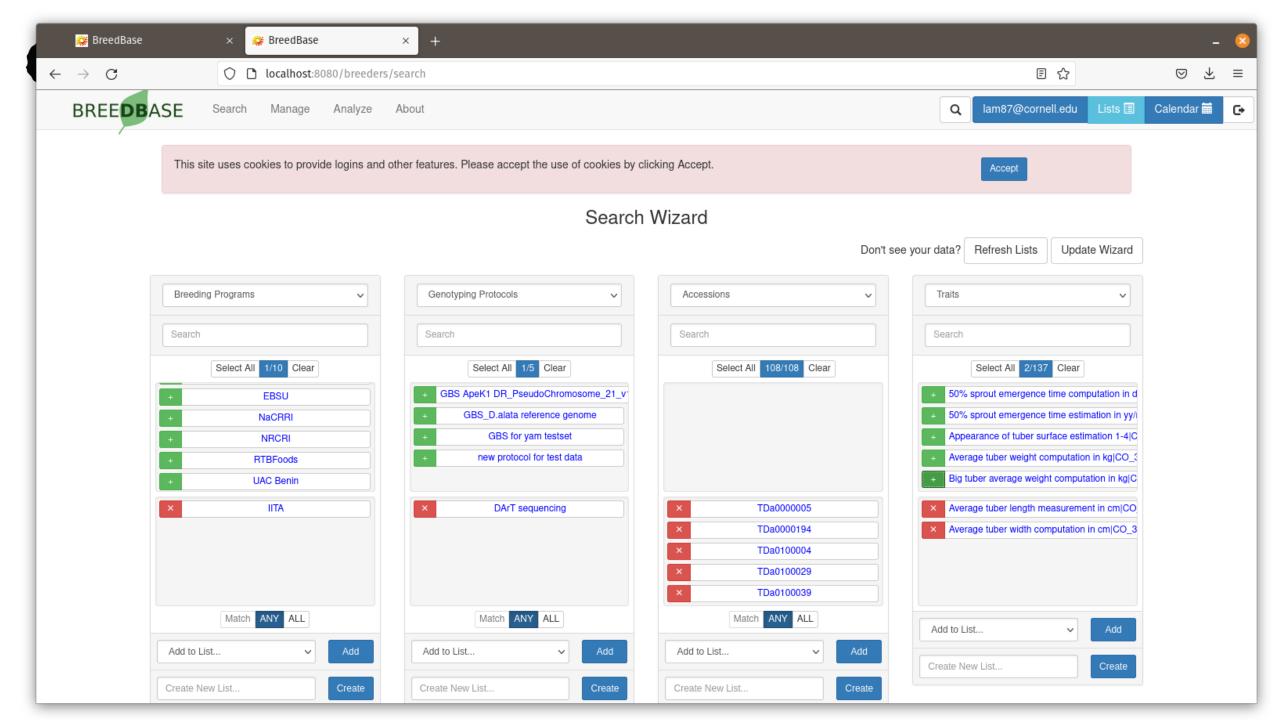
Isaak Tecle

With Agre Paterne and Asrat Amele



GCPC Workflow

- Create Dataset with phenotypic and genotypic data using the Wizard
- Create Selection Index using Selection Index Tool
- Open GCPC tool, choose the Dataset and Selection Index
 Choose the
- Choose the
- Hit submit!



	🔅 BreedBase		× 🔅 BreedB	ase	× +							- 😣
÷	\rightarrow G		🔿 🗅 localho	st :8080/selectio	on/index				Ē	120% 公	⊠ 2	⊻ ≡
	BREED	BASE	Search	Manage	Analyze Abo	ut			Q lam87@cornell	.edu Lists 🗐	Calendar 🚞	G
	,	This si	te uses cookie	s to provide lo	ogins and other feat	tures. Please accept	t the use of cookies by click	ing Accept.	Acce	pt		
						Build a S	election Index 🤅					
		Parame	ters							~	•	
		Trial:										
		01a	ytTDrAB(breed	dersline)						~		
		Traits	and coefficie	nts for <mark>01ayt</mark>	TDrAB(breederslir	ne):						
		Trait r	name		Trait synony	m Index coeffi	cient	Normalize	with a control?	Remove?		
		Plants	stand harvested	count	None	1		No cont	trols found 🗸	×		

×

×

 \sim

 \sim

 \sim

No controls found

No controls found

Total tubers per plot computation

virus severity estimation in 1-5

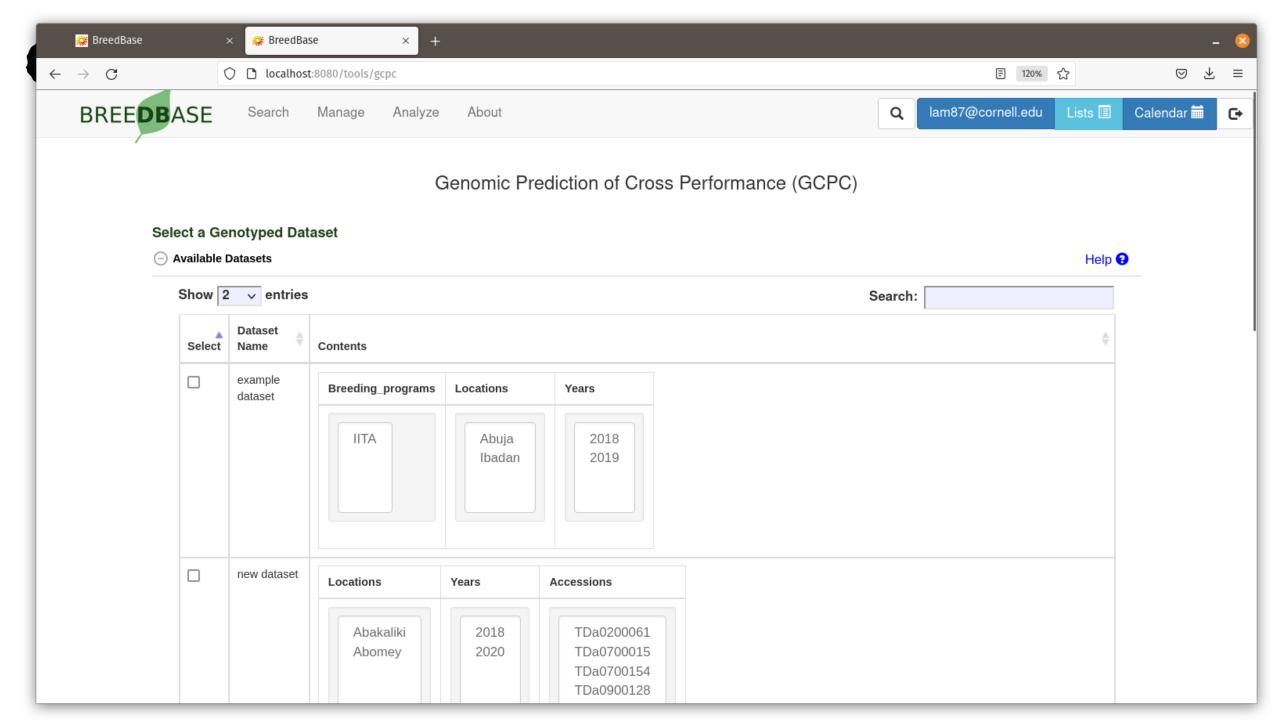
virus severity estimation in 1-5

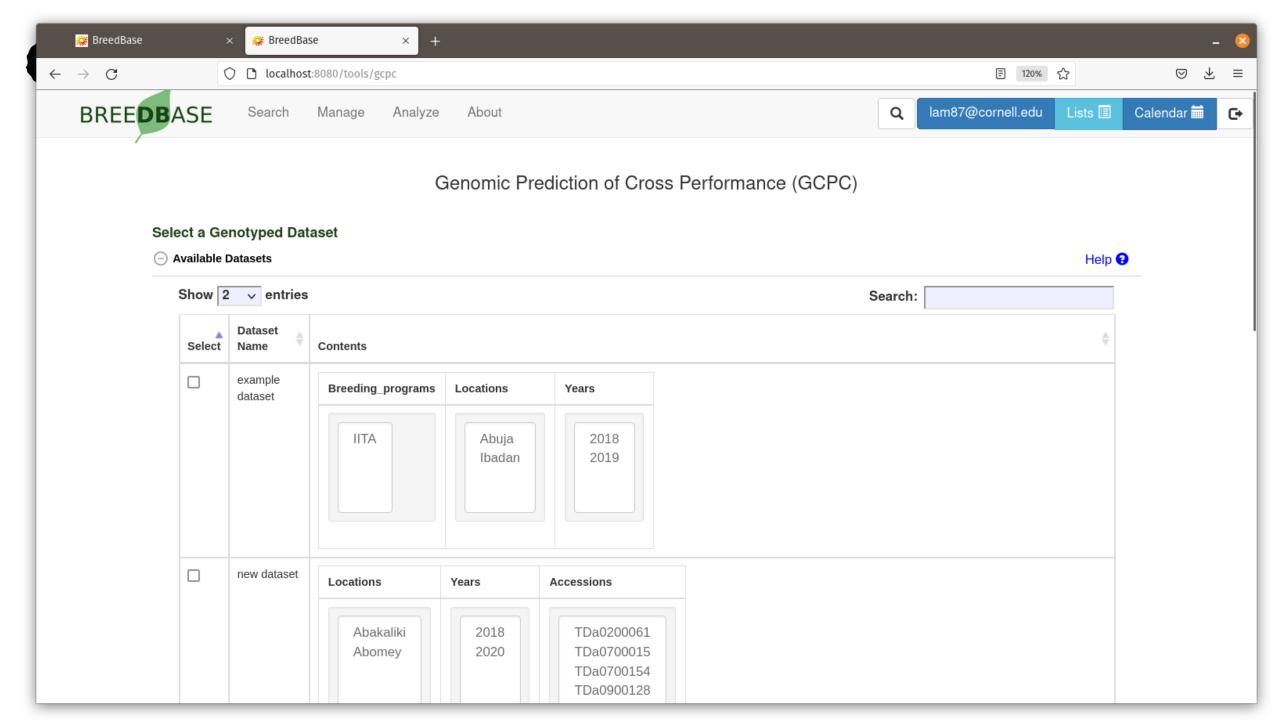
None

VRSS

3

2





	👺 BreedBase	× 🗱 BreedBase × +		- 🙁
←	\rightarrow G	localhost:8080/tools/gcpc	틥 120% ☆	\boxtimes $\overline{+}$ \equiv
	BREEDBA	SE Search Manage Analyze About	Q lam87@cornell.edu Lists 🗏	Calendar 🛗 🕞
	, (Proceed to Factor Selection		
		Select fixed and random factors to be included in the model		
		locationName		
		O fixed ○ random ○ None		
		replicate		
		○ fixed ○ random ○ None		
		germplasmName		
		⊙ random		

This is a drop-down menu showing your pre-constructed selection indices. Select one and proceed to run the model.	

ultimate selection index



~ S



Download Results

Output

Plant Sex info: 1= Male, 2=Female, 3= Monoecious male (m>f) and 4= Monoecious female(f>m)

Show 10 v e	ntries			Search	h:	
Cross #	Parent 1	Parent 2	Cross Predicted Merit	▼ Parent1 Sex	Parent2 Sex	\$
10325	TDr1615003	TDr1619204	1.81256644943338	2	1	
10854	TDr1615003	TDr1774001	1.78515821688158	2	1	
806	TDr1621012	TDr1615003	1.78142414874577	1	2	
10727	TDr1615003	TDr1685017AB	1.74883184716264	2	1	
10294	TDr1615003	TDr1612901	1.71457601416463	2	1	
68691	TDr1618715	TDr1774001	1.70089942571952	2	1	
62328	TDr1619306	TDr1774001	1.69940200793324	2	1	
68563	TDr1618715	TDr1689033AB	1.68707850758771	2	4	
10288	TDr1615003	TDr1619803	1.68598559238079	2	1	
68162	TDr1618715	TDr1619204	1.67799070197507	2	1	
Showing 1 to 10	of 100 entries			Previous 1 2	3 4 5 10	Next

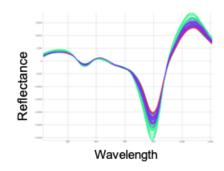


Multi-dimensional Phenotypes

- Datatypes that don't consist of one value per measurement, but a series of values or more complex structures
 - Expression data
 - Proteomics data
 - Metabolomics data
 - Ground Penetrating Radar (GPR) Data
- Collaboration with Jean-Luc Jannink's group
- Data exchange standard development with Peter Selby (BrAPI)







- NIRS spectra can be uploaded to Breedbase databases
- NIRS calibration can be performed in Cassavabase (based on standardized lab measurements)
- Collaboration with Jenna Hershberger (Michael Gore Lab, Cornell University).







Chris Simoes

Nick Morales

Jenna Hershberger



- Non destructive method of root visualization
- Current on-going work at IITA Ibadan by IDL (Delgado et al.)
- Providing support the GPR group with database infrastructure
- Storage of raw (currently) and analyzed data (in the future)
- Thanks for the dinner! :-)



Ground penetrating radar

Quality data and vocabularies

. . .



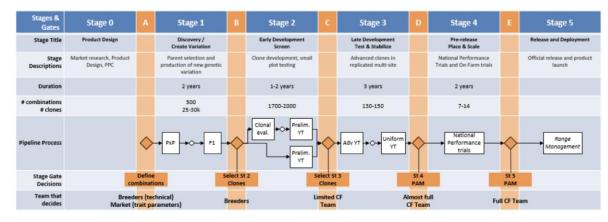
AfricaYam

- Add food quality and processing terms, ontologies and data
- Creation of post-composed ontologies
- RTBfoods major driver for NIRS integration
- Dominique Dufour and Eglantine Fauvelle (Cirad), Thierry Tran, Karima Meghar (CIAT),
- Elizabeth Arnaud, Marie Angélique Laporte (Bioversity)

Africa Yam Product Profiles & Stage Gates

- Product Profiles management integrated in Cassavabase
- Product Profiles are described in terms of traits and trait values
- Future: Product profiles can be used to search matching/exceeding accessions
- Stage gate and roles integration

We mapped the Cassava Pipeline process to the Stage Gate Decision process



Weber & Fritz

Consulting

The content and concepts in this presentation are copyright of Weber & Fritz Consulting A8 unless marked otherwise



- Allows searching specific allele states in the database
- Based on MarkerSets
 - A markerset is a group of markers and specifies which allele is being searched
- List of accessions
 - The accessions that will be searched
- For example, a MarkerSet could correspond to a Intertek marker set



Titima Tantikanjana



- Request by Banana Breeding Project
- Users can order stocks from other users in Breedbase
- Users with the Vendor role can elevate stocks to be listed in the Catalog
- Vendor can set status of the order etc.
- Coming to all RTBbases soon



Titima Tantikanjana





- Important technology underpinning Breedbase functions
- Breeding Application Programming Interface different systems can exchange data programmatically, such as Cassavabase and Fieldbook
- Evolving standard, need to adapt from time to time
- Hackathon 2022 in October



Mirella Flores



Yambase FTP site

- FTP protocol is inherently insecure has been phased out
- FTP site contents is now available through the website, at:
 - <u>https://yambase.org/ftp</u>

$\bullet \bullet \bullet \blacksquare < >$	0	
Index of /ftp/		
/		
<u>3primeRNAseqRawData/</u>	09-Jun-2021 22:09	-
CASS/	18-Feb-2020 22:15	-
CIAT/	18-Jul-2016 13:56	-
<u>Cassava3primeRNAseqRawReads/</u>	08-Jun-2021 21:09	-
<u>GenotypeStoragePaper/</u>	11-May-2020 13:26	-
HapMapII/	15-Jan-2018 20:29	-
Manihot_v6.1/	14-Dec-2015 16:40	-
<u>Qchamps/</u>	01-Mar-2019 12:44	-
documents/	09-Jun-2021 00:29	-
<u>filtered_gbs_data/</u>	20-Mar-2015 02:54	-
<u>ftp/</u>	05-Sep-2021 19:01	-
<u>genotypes/</u>	04-Apr-2014 15:04	-
<u>images/</u>	17-Jul-2018 20:43	-
<u>intertek/</u>	15-Mar-2019 19:59	-
<u>jbrowse/</u>	05-Nov-2015 21:11	-
<u>manuscripts/</u>	03-Jul-2021 20:18	-
maps/	13-Jan-2014 08:03	-
<u>marnin_datasets/</u>	11-Aug-2021 13:33	-
<pre>static_content/</pre>	03-Nov-2020 18:04	-
<pre>static_content_test/</pre>	04-Oct-2019 13:47	-
ugo/	21-Apr-2015 17:27	-
user_requests/	12-Nov-2020 21:24	-
workshops/	26-Apr-2018 15:06	-
df <u>STAR HTSeq counts cassava DESeq2 v</u>	stNormaliz> 10-Jul-2018 17:15	397070315
news.txt	04-Sep-2015 18:31	763



- Simplifies development and deployment of software
- Generated 49 releases in 2 years
- about 1 release every two weeks
- Docker images were downloaded about 4000 times

We've up	dated our Terms of Service. <u>Le</u>	arn more.			
docker hub Q Search for great content (e.g., mysql)	Explore	e Repositorio	es Organizations	Get Help 👻 breedbase 👻	
Q Search by repository name		Creat	te Repository	0	
breedbase / breedbase_pg12.4 Updated 2 days ago	耸 0	<u>↓</u> 10 🔇) Public	Create an Organization Manage Docker Hub repositories with your team	
breedbase / breedbase Updated 3 days ago	☆ 1	¥ 474 🔇) Public		
breedbase / breedbase_pg12	耸 0	±0 &) Public		
breedbase / pg Updated a month ago	☆ 1	¥ 217 🔇) Public	Download Docker Desktop	
breedbase / jscview Updated a year ago	☆ 0	¥ 23 🔇) Public	Secure	
Tip: Not finding your repository? Try switching namespace via the top le	ft dropdown.			Private Repu	



Thank you!