NC STATE UNIVERSITY

Genomics-Assisted Sweetpotato Improvement: Strengthening the Speedbreeders CoP

Progress to date….

Craig Yencho[1] and Godwill Makunde[2]
North Carolina State University[1]
and CIP-Mozambique[2]

7th Annual SPHI Technical
and Steering Committee Meeting
ILRI Campus
Addis Ababa, Ethiopia
October 7-9, 2016

NC STATE UNIVERSITY

Craig Yencho, Lead PI
Ann Tomko, Program Coordinator
Bode Olukolu, Molecular Breeding
Ken Pecota, Breeder
Luis Duque, Breeder, Data capture
Sharon Williamson, Research Specialist
Bonny Oloka, PhD Student, Uganda
Victor Amankwaah, PhD Student, Ghana

Zhao-Bang Zeng (CoPI) Statistics/QTL/GS
Guilherme Da Silva Pereira (Postdoc)
Marcelo Mollinari (Postdoc)

Lina Quesada (coPI) Pathology

GT4SP
GENOMIC TOOLS FOR SWEETPOTATO IMPROVEMENT

MICHIGAN STATE UNIVERSITY

Robin Buell (coPI) Genome Browser
John Hamilton, Bioinformatics Engineer
Jeongwoon Kim, Postdoc
Grant Godden, Postdoc
Krystle Wiegert-Rininger, Research Associate

BTI BOYCE THOMPSON INSTITUTE

Zhangjun Fei (coPI) Sequencing
Lukas Mueller (coPI) Database Development
Alex Obgona, Postdoc
Shan Wu, Postdoc
Alex Ogbonna, Postdoc, Database Development
Bryan Ellerbrock, Database Development

CENTRO INTERNACIONAL DE LA PAPA CIP A member of the CGIAR Consortium
International Potato Center
biosciences eastern and central africa

CIP Lima
Awais Khan, (coPI) Geneticist, Molecular Breeding
Wolfgang Gruneberg, SP Breeding, Global Lead
Merideth Bonierbale, Progam Leader
Jan Kreuze, Virologist
Reinhard Simon, Database Development
Raul Ezaguiere, Statistician
Dorcas Gemenet, Postdoc, Molecular

CIP SSA
Marc Ghislain, (coPI) Global Biotechnology Lead
Ted Carey, SP Breeding, Ghana, West Africa
Robert Mwanga, SP Breeding Uganda, East Africa
Luka Wanjohi, Database Development, Website
Mercy Kitavi, Postdoc, Molecular Breeding, Capacity Development

NaCRRI NATIONAL CROPS RESOURCES RESEARCH INSTITUTE

Benard Yada, (coPI) Breeder
Gorrettie Ssemakula, Breeder
Milton Otema Anyanga, Entomologist

THE UNIVERSITY OF QUEENSLAND AUSTRALIA | IMB Institute for Molecular Bioscience

Lachlan Coin (coPI)
Chenxi Zhou, PhD Student, China

CSIR COUNCIL FOR SCIENTIFIC AND INDUSTRIAL RESEARCH GHANA

Marian Quain, Biotechnology, TC

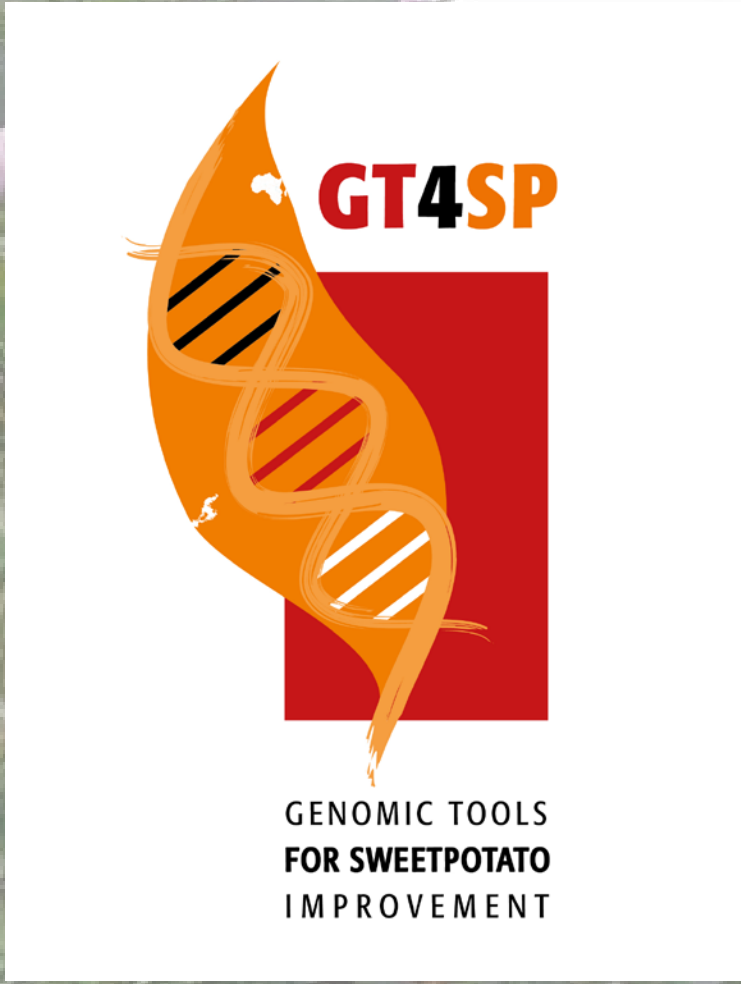
NC STATE UNIVERSITY

Sweetpotato Convening
A Vision for Sweetpotato Improvement in Africa:

Modern Breeding Tools
Increased Potential, Improved Genetic Gain,
Reduced Hunger and Poverty

Bill & Melinda Gates Foundation
Seattle, WA
June 3-5, 2013

The Genomic Tools for Sweetpotato Improvement Project – GT4SP



Scope

- Est. Sep. 2014
- 20 PI's
- 7 Institutions
- 6 Countries
- 4 Continents
- 15 Time Zones
- 1 Crop
- $12.4 M

BILL & MELINDA GATES *foundation*

NC STATE UNIVERSITY

A “Vision” for MAB Breeding in SSA

Breeding pipeline investments should include:

- **Genomic Resources –**
  - **A reference genome**
  - **Marker development – we are way behind the curve….**
  - **A robust set of SNP markers and a low-cost genotyping platform**
  - **Advanced laboratory sequencing linked with developing country phenotyping and breeding activities**
  - **2x and 6x mapping, training and test populations**

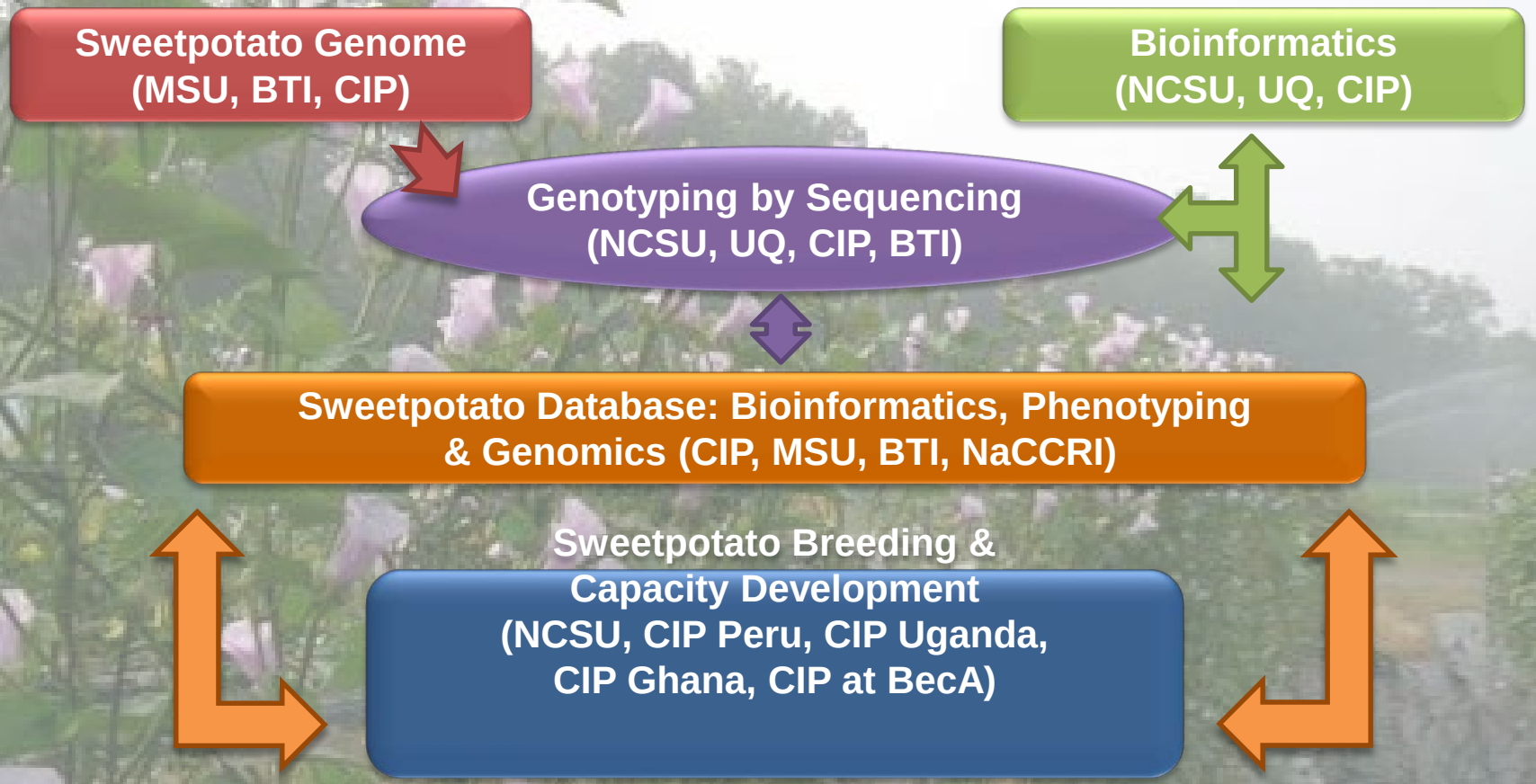
- **Phenotyping, analytics and database resources –**
  - **Improved phenotyping options**
  - **Web-based bioinformatic resources**
  - **New database, data collection and analysis resources**

- **Human Resources and Capacity Development –**
  - **Continue to develop a dynamic team of breeders and allied disciplines**
  - **Training in the use of traditional and genomic breeding methods**
  - **Effective communication and collaboration**
  - **Multi-institutional training and capacity development**

- **A common vision and continuity of effort.**

- **A stellar team to implement phase 1 !!**

The Genomic Tools for Sweetpotato Improvement Project – GT4SP



An ambitious project to sequence sweetpotato and develop modern breeding tools for a food crop that sustains million of people in SSA.

Collaborators: Boyce Thompson Institute at Cornell, Michigan State University, University of Queensland, Australia; The International Potato Center, Peru; BioSciences East and Central Africa, Kenya; National Crops Resources Research Institute, Uganda; Crops Research Institute, Ghana

GT4SP Project Meetings

Start-Up, San Diego
Jan. 2015

1st Annual Mtg.
San Diego, Jan. 2016

NC STATE UNIVERSITY

SASHA
Sweetpotato Action for Security and Health in Africa

GT4SP
GENOMIC TOOLS FOR SWEETPOTATO IMPROVEMENT

1st Annual SASHA-GT4SP Joint Meeting, Mukono, Uganda

NC STATE UNIVERSITY

SASHA
Sweetpotato Action for Security and Health in Africa

GT4SP
GENOMIC TOOLS FOR SWEETPOTATO IMPROVEMENT

2nd Annual SASHA-GT4SP Joint Meeting, CIP-ILRI, Nairobi, Kenya

NC STATE UNIVERSITY

15th SPHI Sweetpotato *Speed*Breeders Annual Meeting 6-7 June 2016 Workshop

BecA-ILRI Research Platform, Nairobi, Kenya

Theme: An eye opener in planning the integration of molecular markers in sweetpotato breeding programs

Time	Activity/Session	Responsible
June 5	**Arrival**	Tassy Kariuki
June 6	***Day 1, Monday***	
7:30 am	***Registration***	Tassy Kariuki
	Chair: Dr. Edward Carey	***Note taker:*** *Dr. Bode Olukolu*
8:00 am	Welcome note from the BecA-ILRI hub Director, and SASHA Leader	Dr. Appolinaire Djikeng, Dr. Jan Low
8:30 am	**Keynote Address** Genomics-assisted sweetpotato improvement: Hope or hype?	Dr. Craig Yencho
9:30 am	Genetic diversity studies of the sweetpotato using DNA markers: what have we learnt?	Dr. Marc Ghislain
10:10 am	**Health Break**	Tassy Kariuki
10:30 am	Basics of marker-assisted selection/breeding and its potential use for sweetpotato improvement	Dr. Awais Khan
11.10 am	Accurate sweetpotato phenotyping of agronomic traits; why it is important for MAS	Dr. Awais Khan & Bramwel Wanjala
12:00 pm	Sweetpotato genotyping procedures for SSA (recorded video/ppt)	Dr. Mercy Kitavi
12.30 pm	**Lunch break**	Tassy Kariuki
2:00 pm	**BecA lab tour facility; An advantage in SSA** what is available and how the sweetpotato community can make use of it regarding the incorporation of molecular procedures in their breeding programs	Dr. Mercy Kitavi
3:00 pm	**Laboratory practical exercise** DNA extraction from sweetpotato roots and leaves Assess DNA quality & quantity PCR preparation (markers assays)	Dr. Dorcus Gemenet Dr. Mercy Kitavi Dorcah Ndege
4:00 pm	**Health Break**	Tassy Kariuki
4:30-6.00 pm	Laboratory practical exercise – continued	Dr. Dorcus Gemenet Dr. Mercy Kitavi Dorcah Ndege

NC STATE UNIVERSITY

Genomics-Assisted Sweetpotato Improvement Workshop (day 2)

June 7	*Day 2, Tuesday*	
8:00 -8.30 am	**Briefing session**	Dr. Dorcus Gemenet & Dr. Mercy Kitavi
8:30 am	**Practical exercise –** Clone identification, germplasm fingerprinting, parent selection & Genetic diversity analysis using SSR data	Dr. Mercy Kitavi
10:00 am	**Health Break**	Tassy Kariuki
10:20 am	**Practical exercise –** Sweetpotato QTL mapping	Dr. Dorcus Gemenet
12:30 pm	**Lunch Break**	Tassy Kariuki
2:00 pm	**Practical exercise –** Genome wide maps; SNP marker development in hexaploid sweetpotato using Genotyping by sequencing (GBS)	Dr. Bode Olukolu
4.00 pm	**Health Break**	Tassy Kariuki
4:20 pm	What do we know about the *Ipomoea* genome and how can this be used in breeding improved sweetpotato varieties?	Dr. Robin Buell
5:00 pm	Discussions/Consultations and AOB	Dr. Mercy Kitavi
5:30 pm	**End**	





NC STATE UNIVERSITY

On-Site Capacity Building

- Mercy Kitavi visited and trained staff in Uganda and Ghana – between July and August 2016
- Pending visit to Mozambique
- Several webinars conducted and pending
- Posted on SweetpotatoKnowledgePortal

NC STATE UNIVERSITY

SSA NARS Genomic infrastructure

- Basic molecular genetic facilities exist among most NARS partners, but not all.
- Sequencing capabilities generally not available.
- Computational and bioinformatics capabilities lacking.
- Biosciences Eastern and Central Africa (BecA-ILRI Hub) – a key partner.
  - Africa Biosciences Challenge Fund - 3 GT4SP fellows identified

NC STATE UNIVERSITY

Major Accomplishments

GT4SP

GENOMIC TOOLS FOR SWEETPOTATO IMPROVEMENT

Two Diploid Reference Genomes for Sweetpotato – A first…

Dr. Zhangjun Fei

I. trifida
NCNSP-0306

	Scaffold*		Contig	
	Size (bp)	Number	Size (bp)	Number
N100 (Shortest)	500	30,343	1	44,841
N90	23,196	984	4,204	10,947
N80	212,487	311	14,711	5,596
N70	518,084	179	28,599	3,494
N60	911,096	112	45,660	2,303
N50	**1,480,604**	**72**	**65,826**	**1,516**
N25	3,646,777	21	155,697	414
N00 (Longest)	16,580,315	1	1,067,799	1
Total	462,089,635	30,343	433,252,193	44,841

*Scaffolds >= 500bp are included

87.8%

82.3%

Estimated genome size: 526.5 Mb

I. triloba
NCNSP-0323

	Scaffold*		Contig	
	Size (bp)	Number	Size (bp)	Number
N100 (Shortest)	500	4,018	1	31,277
N90	2,012,206	76	8,536	12,672
N80	2,667,000	57	15,561	8,955
N70	3,686,450	42	22,163	6,606
N60	4,855,166	32	29,383	4,891
N50	**6,461,009**	**24**	**36,931**	**3,562**
N25	9,551,788	9	61,831	1,252
N00 (Longest)	19,833,707	1	313,171	1
Total	457,090,054	4,018	437,557,497	31,277

*Scaffolds >= 500bp are included

92.2%

88.2%

Estimated genome size: 495.9 Mb

Diploid *Ipomoea* v1 Assembly Annotation Summary

Annotation after final MAKER run and functional annotation assignment:

Structural Annotation Summary:

	I. trifida	*I. triloba*
Number of Gene Models:	34,277	34,122
Avg Gene Length:	3,527.3	3,835.22
Avg Transcript Length	1,462.7	1,530.4
Avg CDS Length:	1,319.8	1,391.4
Avg Exon Length:	263.8	259.9
Avg Intron Length:	456.2	473.4
Single Exon Transcripts:	11,412	9,798

Functional Annotation Summary:

	I. trifida	*I. triloba*
TAIR	27,549	26,751
PFAM	1,296	1,218
Swiss-Prot	318	328
Conserved Hypo.	1,439	1,457
Hypothetical	3,675	4,368

The Arabidopsis Information Resource (TAIR)
Protein Families Database
Swiss Protein Database

- Locus Name Assignment:
- *I. trifida* locus name: v1_itf_[GT]000000
- *I. triloba* locus name: v1_itb_[GT]000000

Sweetpotato Genomics Resource
http://sweetpotato.plantbiology.msu.edu/

MICHIGAN STATE UNIVERSITY
Dr. Robin Buell

NC STATE UNIVERSITY

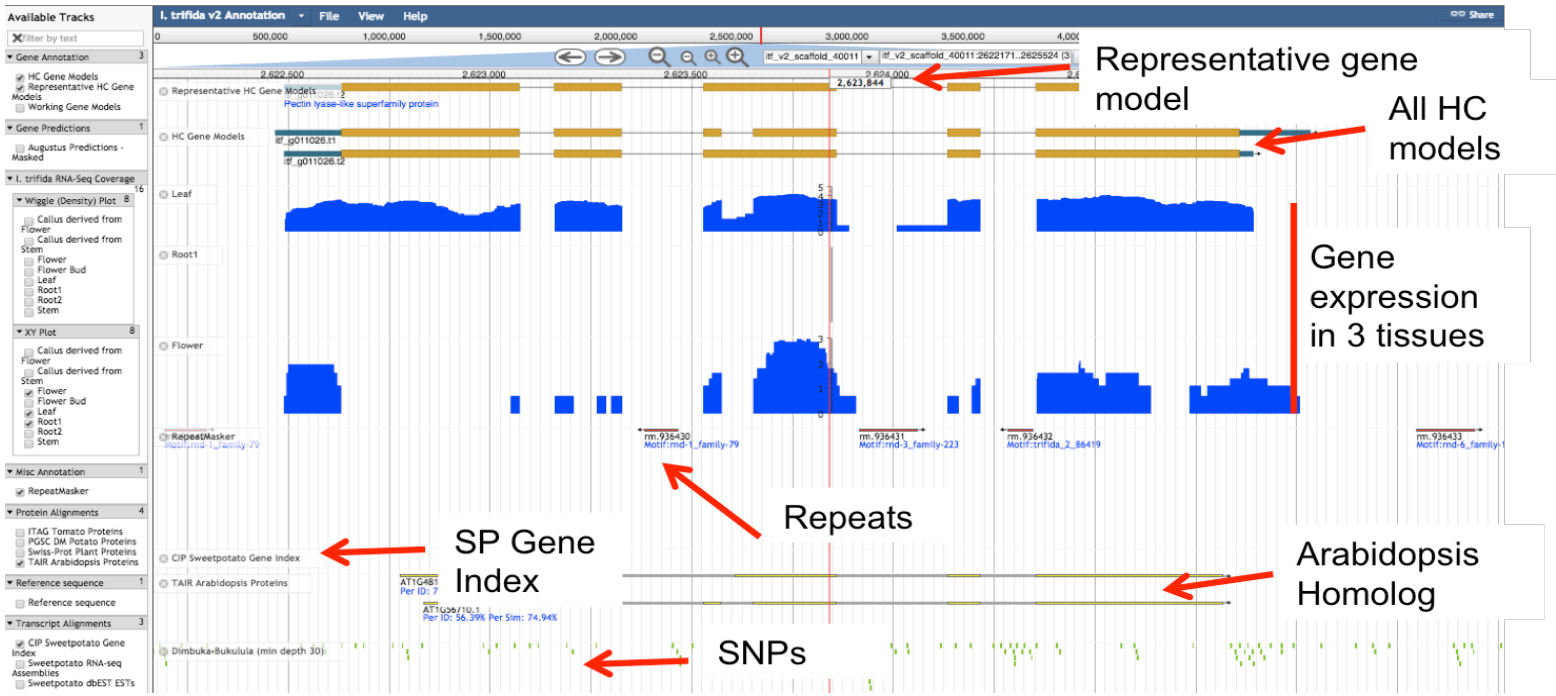
RNAseq Annotation

I. trifida (NCNSP-0306) and *I. triloba* (NCNSP-0323)

Accession	Library	Raw		Cleaned		mapped	
		No. pairs	length	No. pairs	length	No. mapped	%
I. trifida	Callus from flower	21,848,047	151	21,003,512	140	18,231,691	86.8
	Callus from stem	19,079,976	151	18,295,387	141	15,985,394	87.37
	Flower bud	9,983,745	151	9,466,255	142	8,028,621	84.81
	Flower	16,243,164	151	14,485,548	141	11,942,268	82.44
	Leaf	16,460,756	151	15,594,184	140	13,765,383	88.27
	Root1	10,371,929	151	8,620,931	145	6,890,570	79.93
	Root2	11,330,169	151	9,333,983	143	7,505,456	80.41
	Stem	13,732,963	151	13,213,076	136	11,521,863	87.2
	Total	**119,050,749**	**151**	**110,012,876**	**141**	**93,871,246**	**85.33**
I.triloba	Flower bud	14,805,985	151	14,179,939	140	13,665,688	96.37
	Flower	15,722,899	151	13,559,219	143	12,318,763	90.85
	Leaf	17,295,636	151	16,652,971	139	16,107,108	96.72
	Root1	12,626,254	151	10,903,081	142	7,681,265	70.45
	Root2	10,380,650	151	9,090,481	144	8,422,508	92.65
	Stem	22,110,610	151	21,160,167	140	19,950,834	94.28
	Total	**92,942,034**	**151**	**85,545,858**	**141**	**78,146,166**	**91.35**

Diploid *Ipomoea* Genome Annotation: Jbrowse Genome Browser – A first…

Robin Buell



Tracks available on the Jbrowse:

- Genome Annotation (Loci, Gene Models)
- Gene Predictions (Augustus, SNAP, FGENESH)
- MAKER Transcript Evidence
- MAKER Protein Evidence
- CIP Sweetpotato Gene Index Alignments
- RepeatMasker Repeats
- RNA-Seq Coverage – Wiggle
- RNA-Seq Coverage – XY
- I. trifida 0431-1 SNPs

Sweetpotato Genomics Resource
http://sweetpotato.plantbiology.msu.edu/

MICHIGAN STATE UNIVERSITY

NC STATE UNIVERSITY

Genotyping-by-Sequencing

Strategy: Optimized enzyme and fragment size selection for SP

Optimal # of fragments/SNPs:

- low density (genetic map)
- high density (GWAS/GS)

Cost effectiveness:

- exclude repeat regions
- unusable for SNP calling
- methylation sensitive enzyme

GT4SP

GENOMIC TOOLS FOR SWEETPOTATO IMPROVEMENT

DArTseq for hexaploid sweetpotato: B x T mapping population

33,086 SNPs identified in BT population

After quality analyses, 11,091 SNPs mapped on 95 linkage groups.

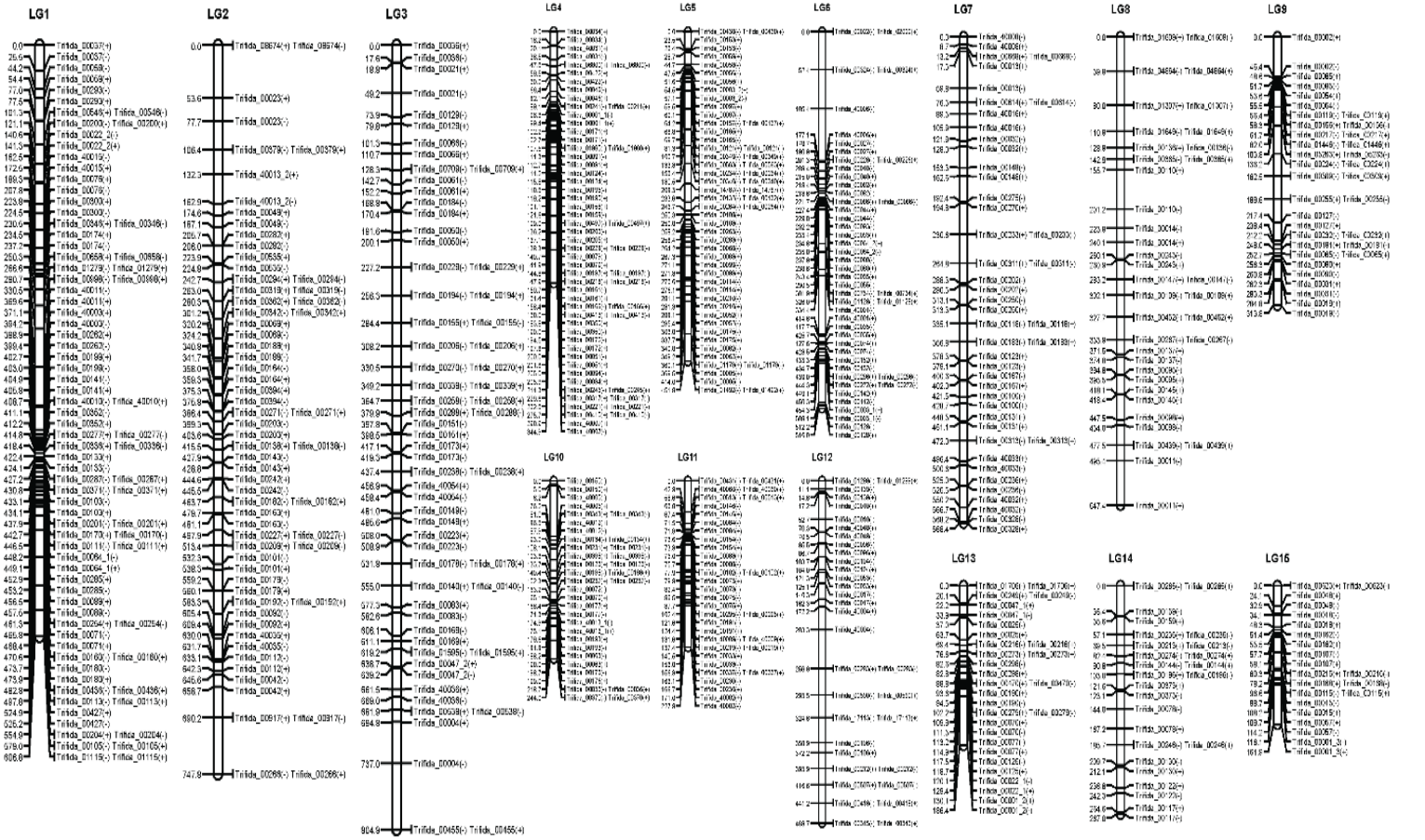
Linkage Groups	Number of Markers	Max Length (cM)	Average Interval (cM)	Max Interval (cM)
1	332	86.93	0.26	6.67
2	182	81.91	0.45	5.19
3	165	83.72	0.51	8.99
4	158	103.54	0.66	13.33
5	171	84.84	0.50	17.29
6	239	83.66	0.35	5.66
7	160	64.27	0.40	9.70
8	222	58.97	0.27	2.96
9	362	97.22	0.27	10.45
10	147	70.70	0.48	12.08
11	389	117.02	0.30	10.01
12	272	73.19	0.27	10.46
13	149	76.79	0.52	6.32
14	132	85.74	0.65	10.00
15	227	69.20	0.31	5.97
16	138	82.34	0.60	11.48
17	249	79.19	0.32	4.81
18	139	71.28	0.52	6.67
19	118	90.27	0.77	15.56
20	337	81.53	0.24	6.67
21	130	86.83	0.67	12.96
22	163	67.23	0.42	10.69
23	98	68.77	0.71	7.78
24	137	61.31	0.45	4.83
25	121	83.22	0.69	14.23
26	322	61.09	0.19	2.82
27	116	66.81	0.58	8.58
28	113	68.76	0.61	11.52
29	276	85.92	0.31	8.55
30	136	47.51	0.35	5.19
31	112	65.36	0.59	7.14
32	218	70.50	0.32	3.73
33	182	71.47	0.39	6.69
34	201	60.85	0.30	5.58
35	92	74.59	0.82	8.24
36	91	87.65	0.97	12.96
37	102	63.50	0.63	13.36
38	90	83.62	0.94	14.44
39	95	62.93	0.67	6.72
40	98	72.47	0.75	21.48
41	104	66.17	0.64	19.26
42	225	63.80	0.28	4.81
43	106	63.67	0.61	8.89
44	93	55.35	0.60	5.53
45	84	59.75	0.72	5.93
46	75	56.51	0.76	6.72
47	89	70.81	0.80	14.98
48	80	69.24	0.88	11.28
49	80	74.61	0.94	20.74
50	106	59.21	0.56	15.36

	Linkage Groups	Number of Markers	Max Length (cM)	Average Interval (cM)	Max Interval (cM)
	51	81	68.63	0.86	11.32
	52	82	57.55	0.71	12.22
	53	85	70.77	0.84	16.23
	54	80	49.01	0.62	7.84
	55	153	56.28	0.37	5.80
	56	197	94.69	0.48	7.40
	57	94	77.16	0.83	17.67
	58	208	98.44	0.48	9.63
	59	163	66.03	0.41	5.22
	60	198	74.01	0.38	9.12
	61	90	51.53	0.58	9.51
	62	72	46.00	0.65	9.85
	63	74	76.09	1.04	20.74
	64	70	54.15	0.78	13.49
	65	67	51.29	0.78	13.70
	66	143	61.32	0.43	15.25
	67	70	70.64	1.02	18.15
	68	53	55.00	1.06	18.96
	69	63	55.34	0.89	13.15
	70	58	43.96	0.77	10.49
	71	52	32.38	0.63	11.11
	72	49	45.37	0.95	10.07
	73	64	36.87	0.59	5.93
	74	45	23.05	0.52	4.07
	75	49	32.40	0.67	9.29
	76	95	42.08	0.45	4.51
	77	40	28.53	0.73	6.67
	78	30	52.00	1.79	16.60
	79	40	49.53	1.27	11.28
	80	30	15.19	0.52	3.33
	81	31	9.32	0.31	1.85
	82	25	27.99	1.17	16.00
	83	21	20.13	1.01	3.70
	84	24	15.67	0.68	5.24
	85	27	14.20	0.55	3.72
	86	26	11.52	0.46	3.70
	87	15	17.88	1.28	5.19
	88	15	9.48	0.68	5.73
	89	12	8.28	0.75	1.90
	90	11	19.96	2.00	6.74
	91	17	4.12	0.26	1.49
	92	13	3.41	0.28	1.91
	93	16	4.08	0.27	1.48
	94	8	5.31	0.76	2.33
	95	12	5.59	0.51	2.96
Total	95	11091	5504.06		
Max		389	117.02		21.48
Min		8	3.41		
Average		117	57.94	0.63	

CENTRO INTERNACIONAL DE LA PAPA
CIP
A member of the CGIAR Consortium

International Potato Center

High-density genetic map of the *I. trifida* CIP *M9 x M19* mapping population.



Map was made using GBS data generated by Bode Okulu at NCSU on 212 progeny plus 2 parents in a F1 cross of M9 X M19. This map indicates the position of the scaffolds obtained from the Trifida assembly. Only the starting and ending position of each scaffold is shown (indicated with + and – respectively). As a result some apparently large gaps greater than 50cM are long scaffolds. Where scaffolds have been split, this is indicated using _1 and _2 to indicate the split scaffolds.

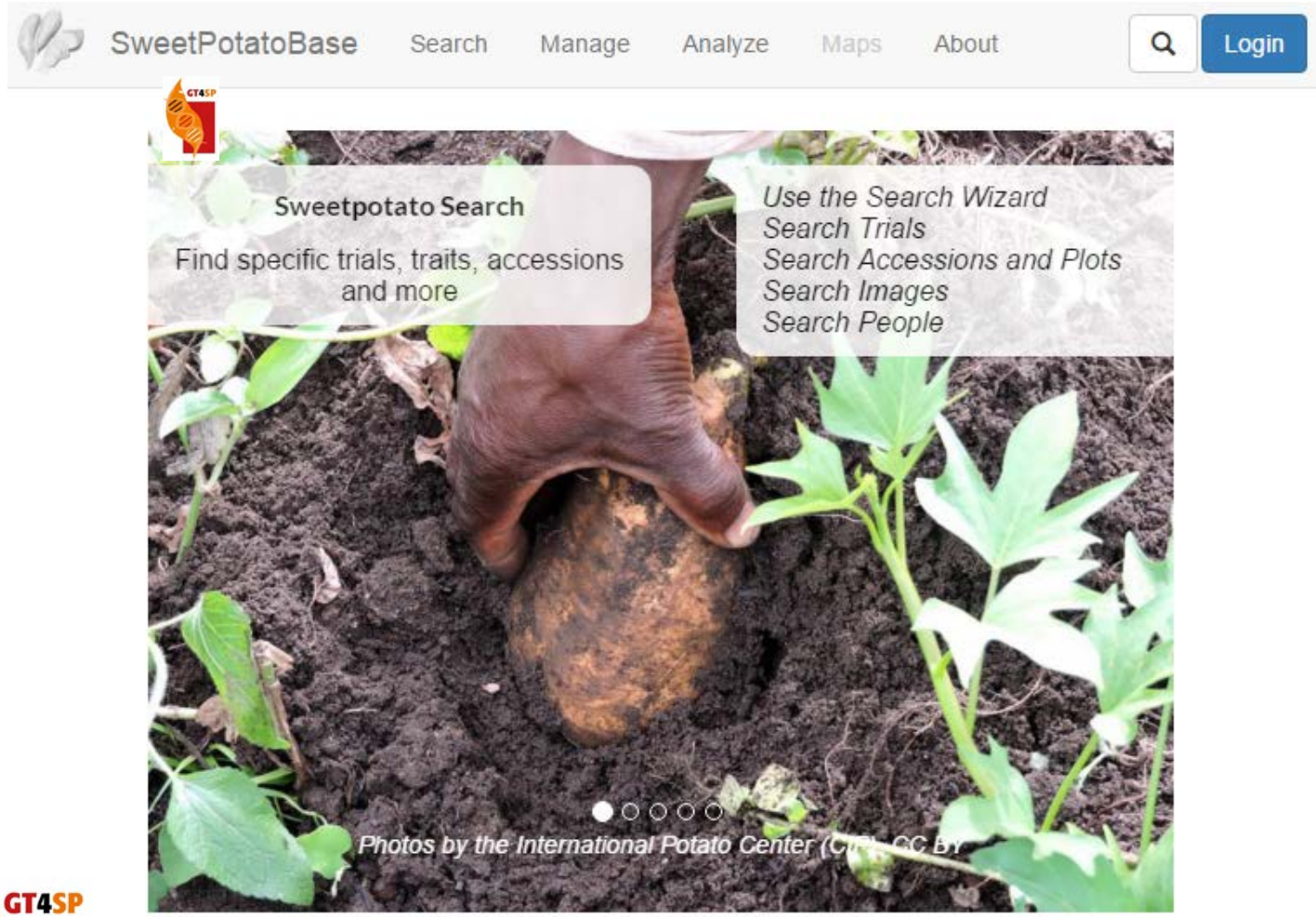
NC STATE UNIVERSITY

GT4SP Mapping Populations

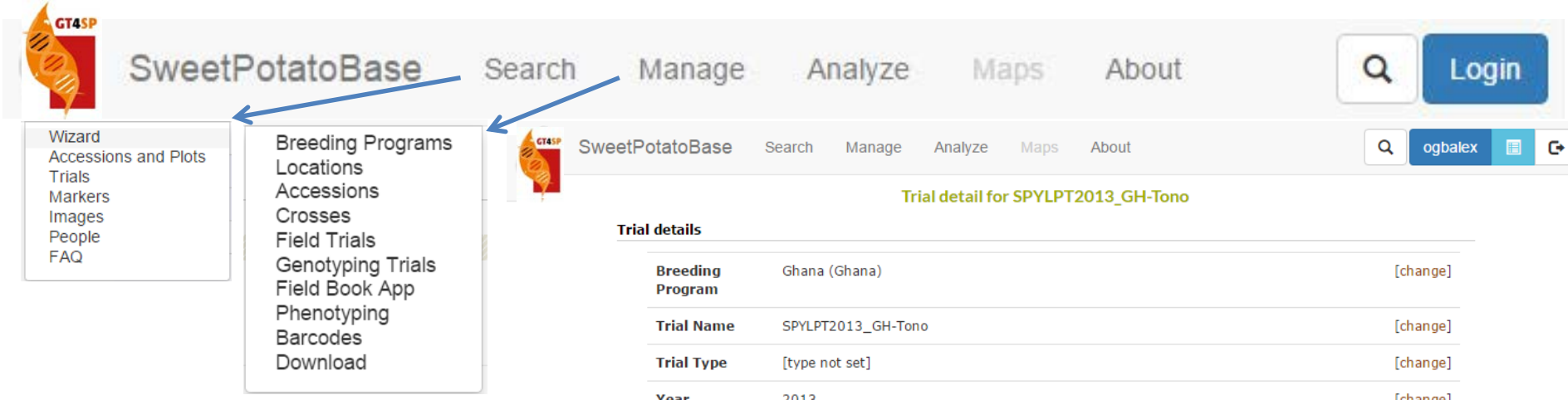
- M9 x M19 – Diploid *I. trifida*
  - 212 clones
- Beauregard x Tanzania (BT)
  - 316 clones
- Tanzania x Beauregard (TB)
  - 247 clones
- New Kawogo x Beauregard (NKB),
  - 287 clones
- Mwanga Diversity Population,
  - 8 x 8 mating design,
  - 1,920 clones
- Multi-location, multi-year phenotyping in Peru, USA, Ghana and Uganda

Managing Phenotypic Data
SweetPotatoBase and FieldBook App.

Dr. Lukas Mueller



Meeting the Data Management Needs



Genomic Tools for Sweetp... | Sweetpotato Genomics Res... | http://hidap.sgn.cornell.edu/hd/

hidap.sgn.cornell.edu/hd/

Search

Most Visited | Apple | News | Amazon | SECU | NCSU | MyPack Portal | HS | GT4SP | CALS GC | SPBRD | ACCTS | SPARCS | CALS: eFAR | Expedia | WRAL | WCPSS

HIDAP4RTB

International Potato Center

A member of the CGIAR Consortium

- Phenotype
- Environment
- About new

Highly Interactive Data Analysis Platform for Clonal Plant Breeding

HIDAP v1.0 [07/06/2016]

HIDAP is a tool designed to help breeders of clonal plants (likw potato and sweetpotato) carry out field trial planning, documentation, analysis and reporting

International Potato Center (CIP)

GT4SP

What's needed and how do we get there?
Lesson's Learned -

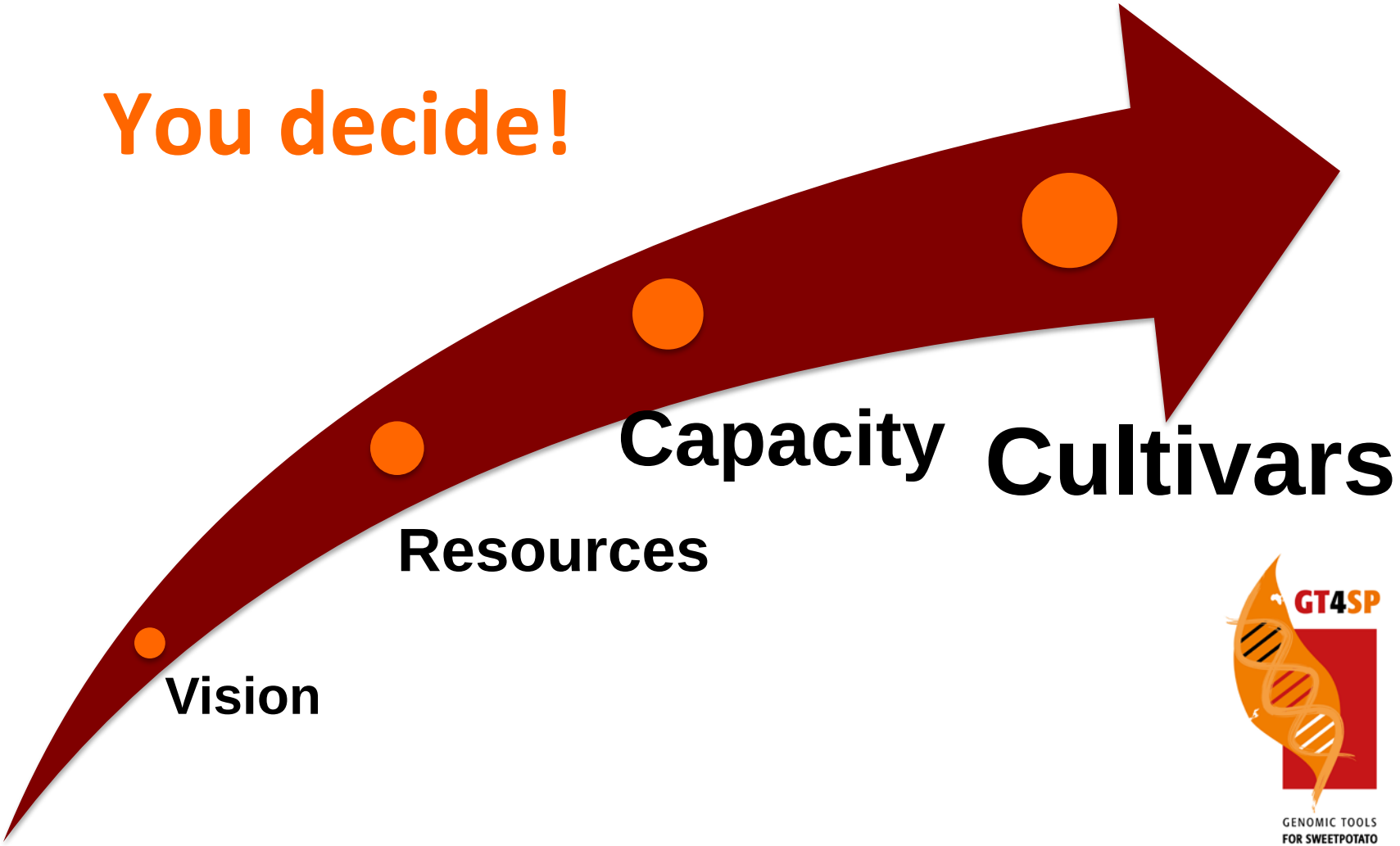
- Reference genome(s)
- A sequence-based genotyping platform
- Breeder friendly bioinformatics and analytical environments
  - This is easier said than done! What will this look like?
  - Special needs for clonally propagated, polyploid crops?
  - Assured technical support to breeders required
- Populations (QTL mapping and GWAS)– We're breeders….no worries
- Improved phenotyping and data collection capabilities including mechanization equipment…
- Sustained funding – multiple sources (NGO, Gov., private industry)
- Sustained capacity development – people, institutional, programs, farmers, industry – all are critical.

For more information see:
SweetpotatoGenomics.CALS.NCSU.edu

GT4SP
GENOMIC TOOLS FOR SWEETPOTATO IMPROVEMENT

Genomic-Assisted Breeding in Sweetpotato – Hope or Hype?

You decide!



GT4SP
GENOMIC TOOLS FOR SWEETPOTATO IMPROVEMENT

Acknowledgements

BILL & MELINDA GATES foundation

http://www.blakeleypritchard.com/ncsu-chancellors-office-e-card/