P449

Using Genotyping-By-Sequencing to Understand *Musa* Diversity

Agnes P. Chan[1], Xavier Perrier[2], Christophe Jenny[2], Jean-Pierre Jacquemoud-Collet[2], Mathieu Rouard[3], Julie Sardos[3], Nicolas Roux[3], and Christopher D. Town[1]

[1]J. Craig Venter Institute, Rockville, USA; [2]CIRAD, Montpellier, France
[3]Bioversity International, Montpellier, France;

ABSTRACT

Background: This project is part of a larger effort to apply genomics technologies to assess genetic diversity and to advance genetic improvement efforts in *Musa* (banana and plantain), a major staple food crop in the developing world. Most cultivated varieties of bananas result from intra- or inter-specific crosses of the wild diploid species, *Musa acuminata* (A genome) and *Musa balbisiana* (B genome). Somaclonal mutation and human selection has resulted in current day bananas with a wide morphological diversity. The Cavendish (AAA) subgroups are believed to have derived from an individual unique initial genotype, and similarly for the subgroup plantain (AAB). However, little or no genetic diversity can be detected within these groups using conventional molecular markers such as RFLP, SSR, DArT.

Methods: To assess genetic diversity with an improved resolution, we have selected 65 accessions with diploid and triploid combinations of the A and/or B genomes including AAB plantains and AAA Cavendish, and cultivated or wild *Musa* accessions from the core collection at the Global Musa Genomics Consortium (GMGC). A high-throughput reduced representation genome sequencing approach - genotyping-by-sequencing (GBS) is used to obtain high density sequence markers [1].

Results: Using GBS reads, genotypes were determined for each diploid and triploid accession, and dissimilarity computed across all accessions. Genetic diversity analysis was carried out using the DARwin software [2, 3].

Conclusions: GBS markers provides a high resolution approach to characterize the genetic diversity of individual Musa subgroups.

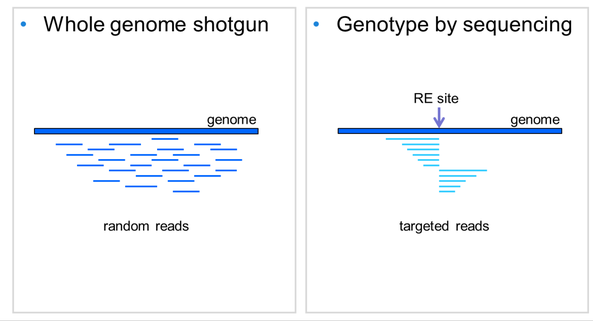
METHODS

(1) A total of 65 Musa accessions including diploid and triploid combinations of the A or B genomes, AAB plantains, AAA Cavendish, and cultivated or wild *Musa* accessions were selected for GBS sequencing.

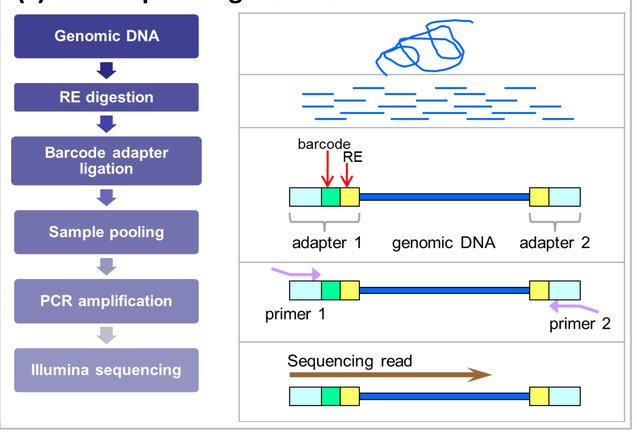
(a) 65 Musa accessions

	ITC	Name	ploidy	Group	Subgroup
1	ITC0557	Americani	3	AAA	Cavendish
2	ITC0002	Dwarf Cavendish	3	AAA	Cavendish
3	ITC1586	Grande Naine	3	AAA	Cavendish
4	ITC0654	Petite Naine	3	AAA	Cavendish
5	ITC0340	Pisang Masak Hijau	3	AAA	Cavendish
6	ITC0048	Valery	3	AAA	Cavendish
7	ITC1251	Vietnam No5	3	AAA	Cavendish
8	ITC0724	Cocos	3	AAA	GrosMichel
9	ITC0686	Pisang Umbuk	3	AAA	Orotava
10	ITC0671	Hom Thong Mokho	3	AAA	Ambon
11	ITC1305	Paji	3	AAA	EA AAA
12	ITC0184	One Hand Planty	3	AAB	Plantain
13	ITC1127	3 Vert	3	AAB	Plantain
14	ITC0111	Agbagba	3	AAB	Plantain
15	ITC0217	Akpakpak	3	AAB	Plantain
16	ITC0007	Asamiensa	3	AAB	Plantain
17	ITC1126	Batard	3	AAB	Plantain
18	ITC1129	Big Ebanga	3	AAB	Plantain
19	ITC0033	Bungaoisan	3	AAB	Plantain
20	ITC0034	Daluyao	3	AAB	Plantain
21	ITC0514	Eba Oboikpa	3	AAB	Plantain
22	ITC0121	Ihitisim	3	AAB	Plantain
23	ITC0631	Maqueno	3	AAB	Plantain
24	ITC0118	Mbang Okon	3	AAB	Plantain
25	ITC0489	Mimi Abue	3	AAB	Plantain
26	ITC0022	Mululou	3	AAB	Plantain
27	ITC0044	Niangafelo	3	AAB	Plantain
28	ITC1133	Njock Kon	3	AAB	Plantain
29	ITC0113	Ntanga 2	3	AAB	Plantain
30	ITC0512	Ntanga 5	3	AAB	Plantain
31	ITC0192	Ntie	3	AAB	Plantain
32	ITC0718	Nzumoigne	3	AAB	Plantain
33	ITC0109	Obino l'Ewai	3	AAB	Plantain
34	ITC0235	Obubit Ukom	3	AAB	Plantain
35	ITC1325	Orishele	3	AAB	Plantain
36	ITC0119	Osoaboaso	3	AAB	Plantain
37	ITC1140	Red Yade	3	AAB	Plantain
38	ITC0719	Mnalouki	3	AAB	plantain like Comoros
39	ITC0036	Muracho	3	AAB	plantain like Phillipino
40	ITC0945	Bira	3	AAB	Iholena
41	ITC0825	Uzakan	3	AAB	Iholena
42	ITC0332	Laknau	3	AAB	Laknau
43	ITC0733	Iho U Maohi	3	AAB	Maia Maoli Popoulu
44	ITC1327	Poingo	3	AAB	Maia Maoli Popoulu
45	ITC1135	Popoulou CMR	3	AAB	Maia Maoli Popoulu
46	ITC1232	Bagatow	3	AAB	PNG AAB
47	ITC0870	Horul	3	AAB	PNG AAB
48	ITC0909	Kupulik	3	AAB	PNG AAB
49	ITC0830	Rukumamb Tambey	3	AAB	Rukumamb
50	ITC0875	Tigua	3	AAB	Rukumamb
51	ITC0312	Pisang Jari Buaya	2	AA	Jari Buaya
52	ITC0653	Pisang Mas	2	AA	Sucrier
53	ITC1187	Tomolo	2	AA	AA PNG
54	ITC0245	Safet Velchi	2	AB	Ney Poovan
55	ITC0659	Namwa Khom	3	ABB	PisangAwak
56	ITC0623	Banksii	2	acuminata	banksii
57	ITC0283	Long Tavoy	2	acuminata	burmannica
58	ITC0249	Calcutta 4	2	acuminata	burmannicoides
59	ITC1511	DH Pahang	2	acuminata	malaccensis
60	ITC0609	Pahang	2	acuminata	malaccensis
61	ITC0253	Borneo	2	acuminata	microcarpa
62	ITC1177	Zebrina	2	acuminata	zebrina
63	ITC0247	Honduras	2	balbisiana	balbisiana
64	ITC1063	Pisang Klutuk Wulung	2	balbisiana	balbisiana
65	ITC0637	Musa Ornata	2	ornata	ornata

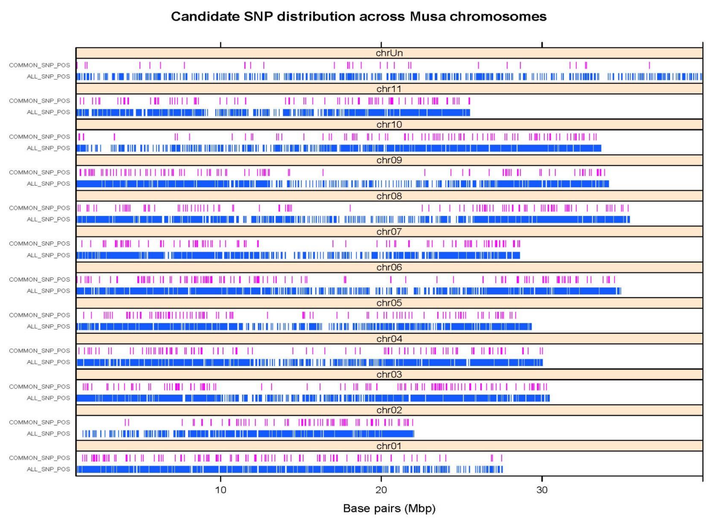
(b) Whole genome shotgun vs GBS sequencing



(c) GBS sequencing workflow

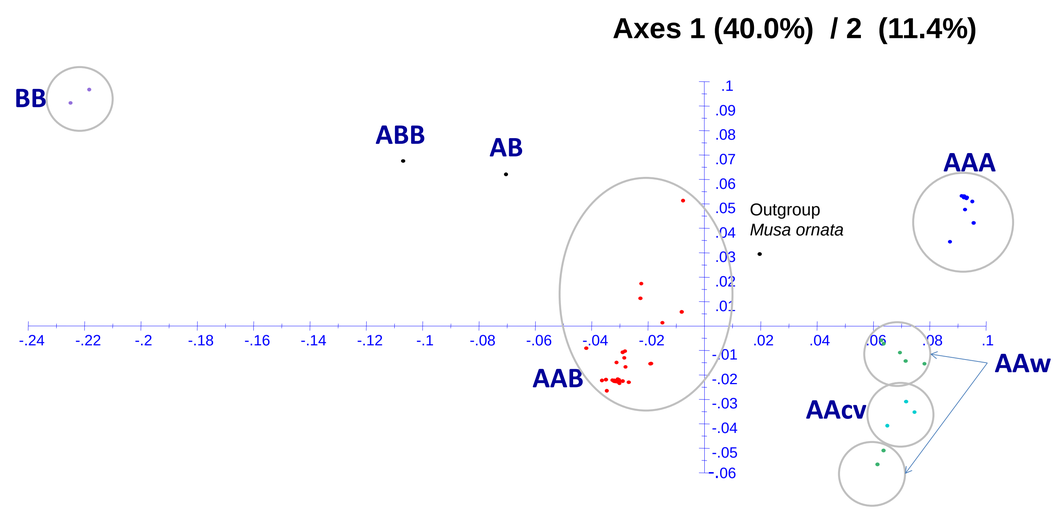


(2) About 1 to 4 million Illumina reads flanking PstI sites were generated for each Musa accession. Approximately 200,000 candidate SNP locations with reference to the doubled-haploid Pahang genome were obtained from all 65 accessions combined (labeled as "ALL_SNP_POS"). Of which, approximately one-tenth were common locations and shared obtained across all 65 accessions (labeled as "COMMON_SNP_POS"). GBS markers obtained were distributed throughout the entire lengths of the Musa chromosomes (as expected) as shown in the chromosome position plot below.

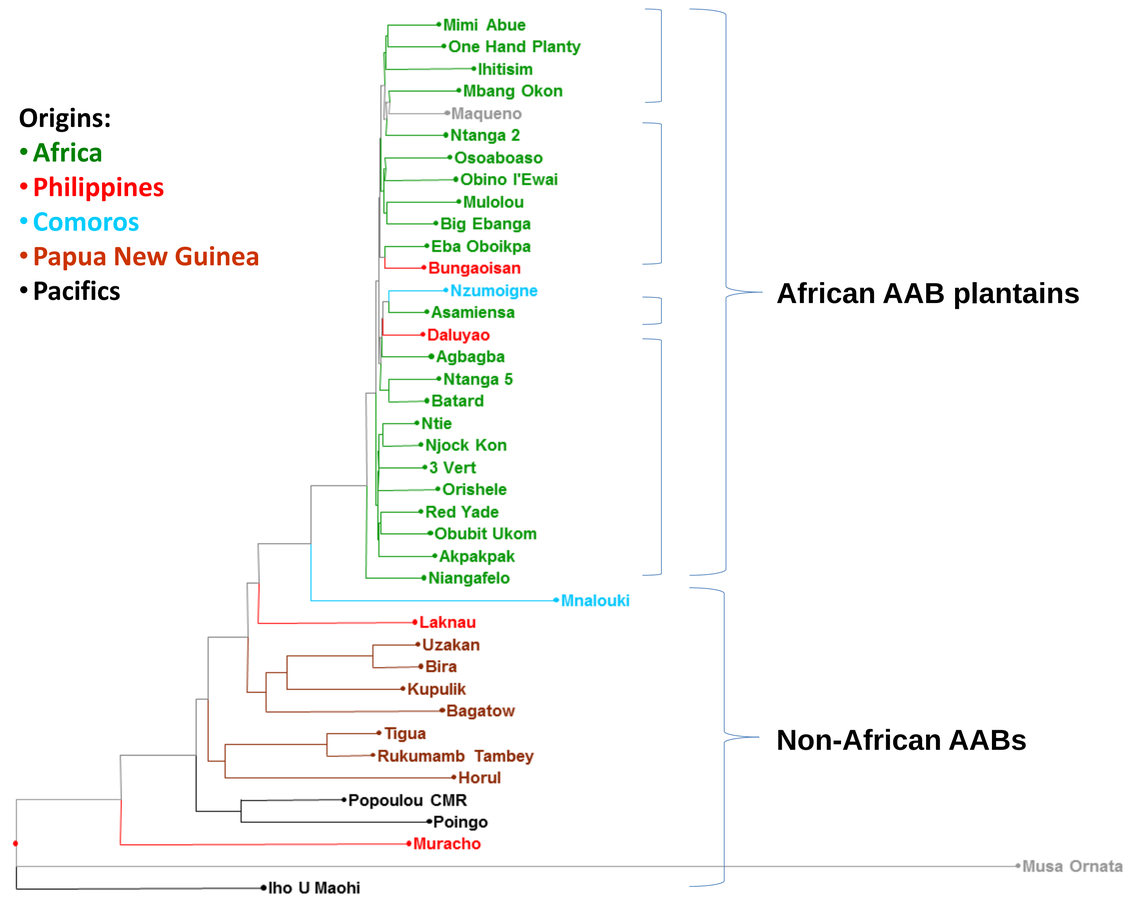


RESULTS

(1) Factorial analysis was performed using 25,115 sites shared across all 65 Musa accessions. The accessions were clustered according to the expected genome composition. The A-only and B-only containing genomes were seen as the extremes of the first axis as expected.



(2) A neighbor-joining tree was constructed using 75,981 sites shared across 39 AAB accessions. The node for *Musa ornata* was grafted afterwards for tree rooting. The majority of the African AAB plantains (green labels) were clearly separated from other AAB accessions.



CONCLUSIONS

(1) Genetic diversity between the A and B versions of the Musa genome is higher than inter A-genome variations, as expected.
(2) GBS markers provide a high resolution sequence-based approach to study genetic diversity for Musa subgroups such as the African plantains (AAB).
(3) Further analysis will be needed to compare the resolution of the GBS method with existing genotyping methods for studying Musa diversity.

REFERENCES

[1] Elshire R.J., Glaubitz J.C., Sun Q., Poland J.A., Kawamoto K., Buckler E.S., Mitchell S.E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 4;6(5):e19379. [PMID: 21573248]

[2] Perrier, X., Flori, A., Bonnot, F. (2003). Data analysis methods. In: Hamon, P., Seguin, M., Perrier, X.,Glaszmann, J. C. Ed., Genetic diversity of cultivated tropical plants. Enfield, Science Publishers. Montpellier. pp 43 - 76.

[3] Perrier, X., Jacquemoud-Collet, J.P. (2006). **DARwin** software http://darwin.cirad.fr

ACKNOWLEGEMENTS

We thank Ronny Swennen (KU Leuven, Belgium) for the selection of the Musa accessions, Eva Hřibová and Jaroslav Doležel (Institute of Experimental Botany, Czech Republic) for providing Musa genomic DNA, and Ed Buckler and Rob Elshire (Cornell University) for advices on the GBS method. This work is funded by the USAID linkage funds.