

Thousands of crops and very little money – how can marker technology be applied to underutilised species?



www.cffresearch.org



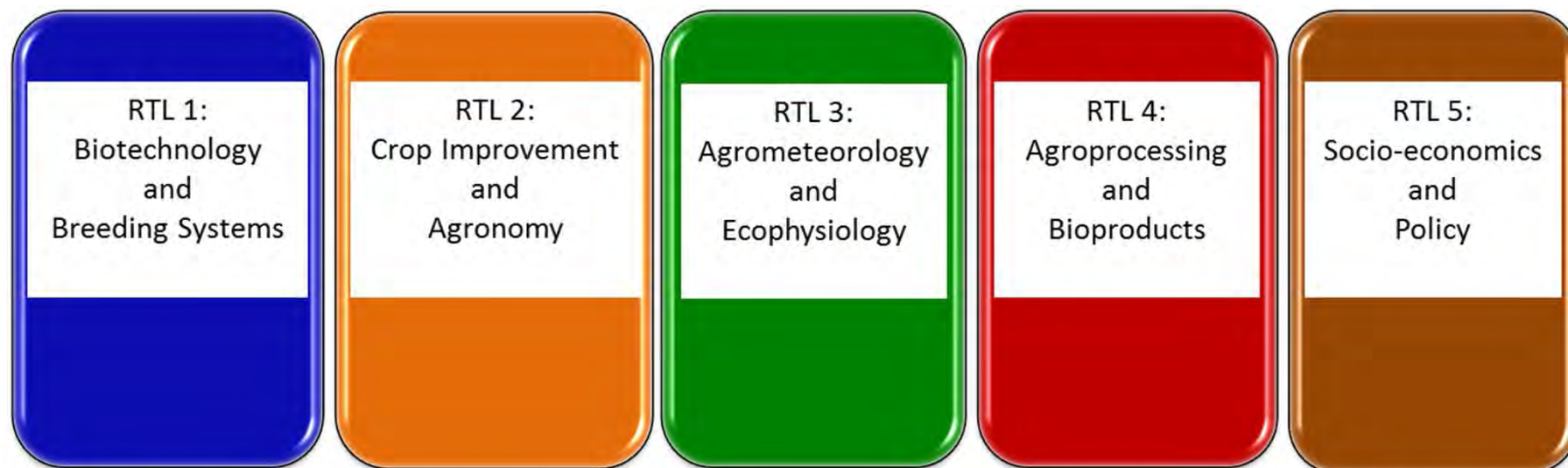
Sean Mayes, Festo Massawe, Martin Blythe, Joanne Morton, Andrzej Kilian, Sayed Azam-Ali

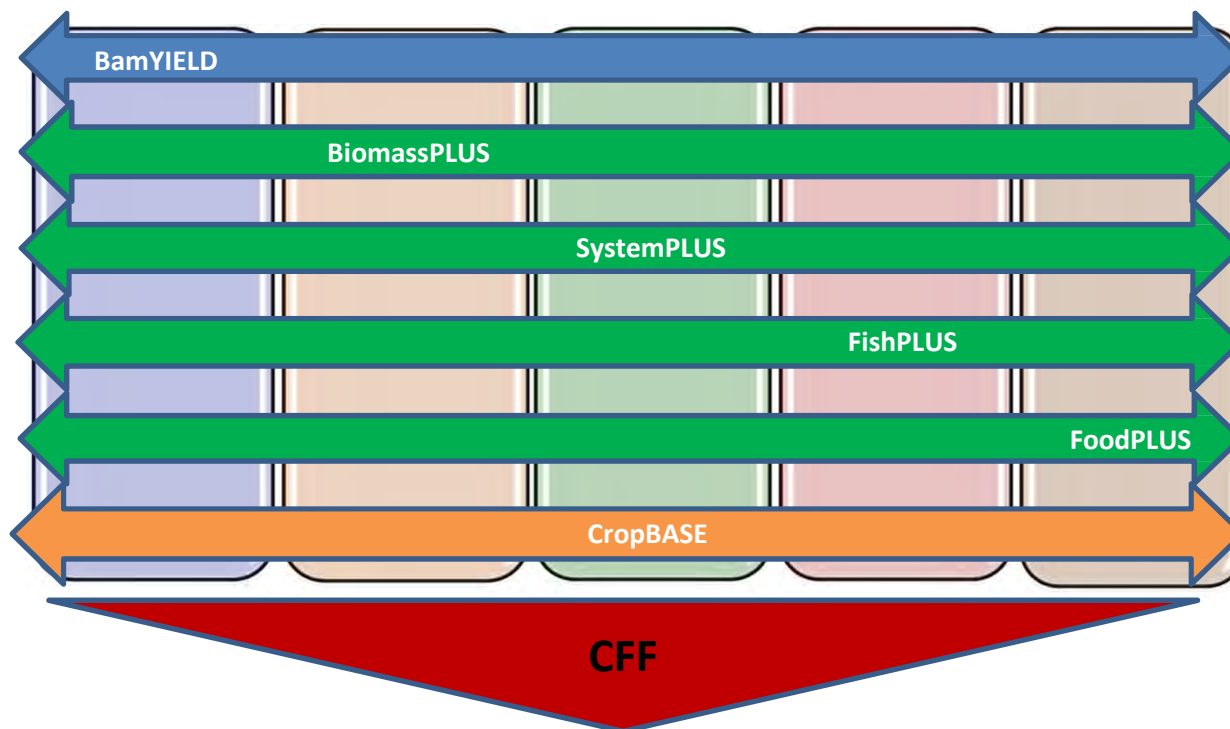
Crops for the Future Research Centre (CFFRC)

How do we define underutilised crop?

Practically, an underutilised crop is one that does not have a major (CGIAR?) research Centre with a clear research mandate

We also need to generate **quantitative and comparative** data with the major crops where it does not exist





- BamYIELD:** Bambara groundnut as an exemplar crop for Africa and Asia
- BiomassPLUS:** Novel biomass crops for sustainable renewable energy and high value products
- SystemPLUS:** Diversifying agricultural systems using underutilised crops and cropping systems
- CropBASE:** World's first web-based knowledge and decision-support system for underutilised crops
- FishPLUS:** Novel plant products to increase the nutritional value of aquaculture feeds
- FoodPLUS:** Diversification of the food basket for enhanced community nutrition and health

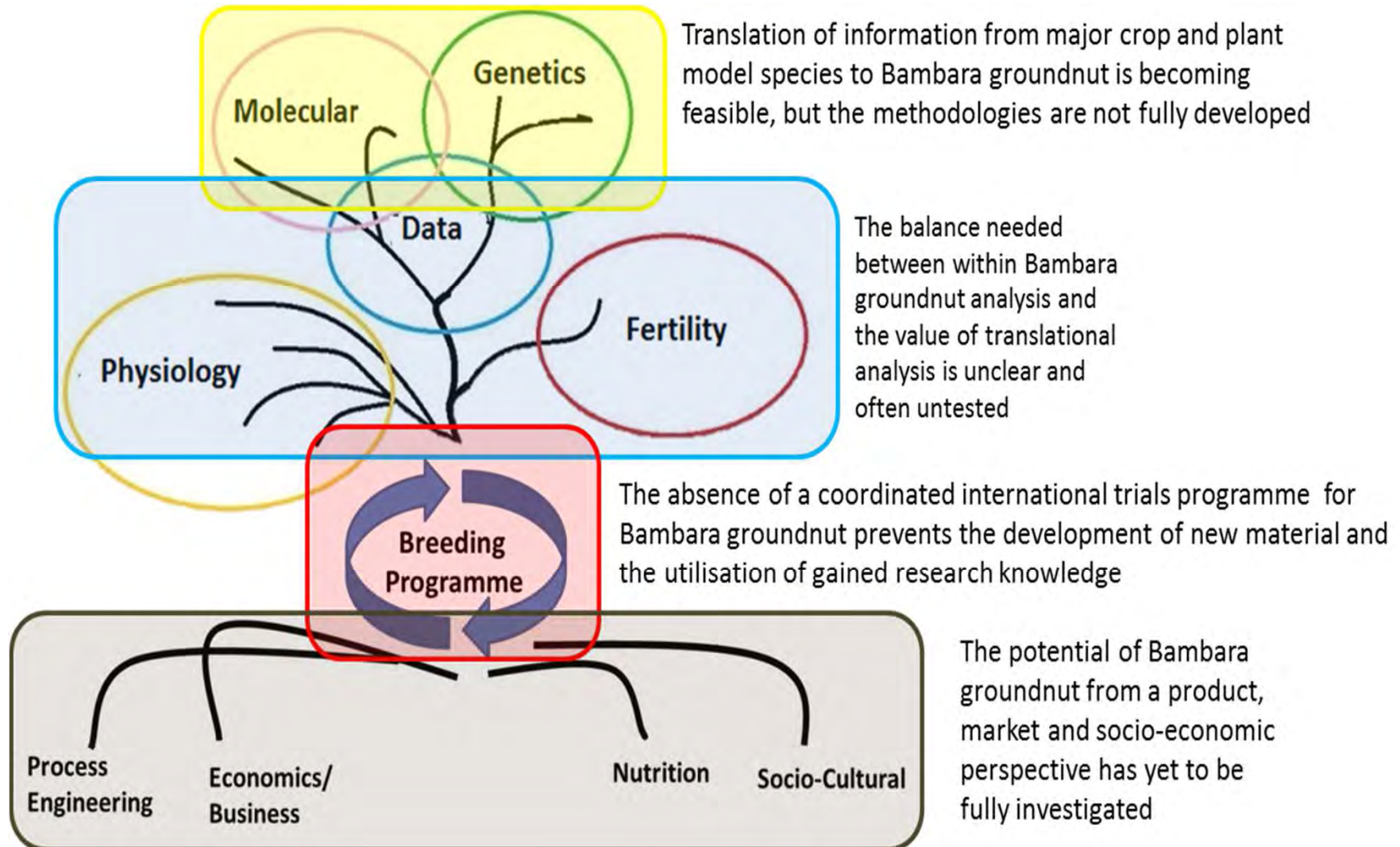
CFFRCPLUS/CFFRC DTC – University of Nottingham (UK/Malaysia/China)

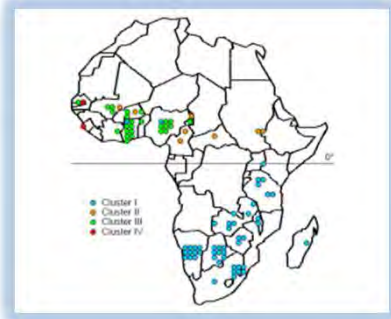
A post-graduate (MRes and PhD) research studentship programme for an initial 300 research years

Round 1 complete and 43 studentships approved (38 PhD and 5 MRes; 65 supervisors across five continents)

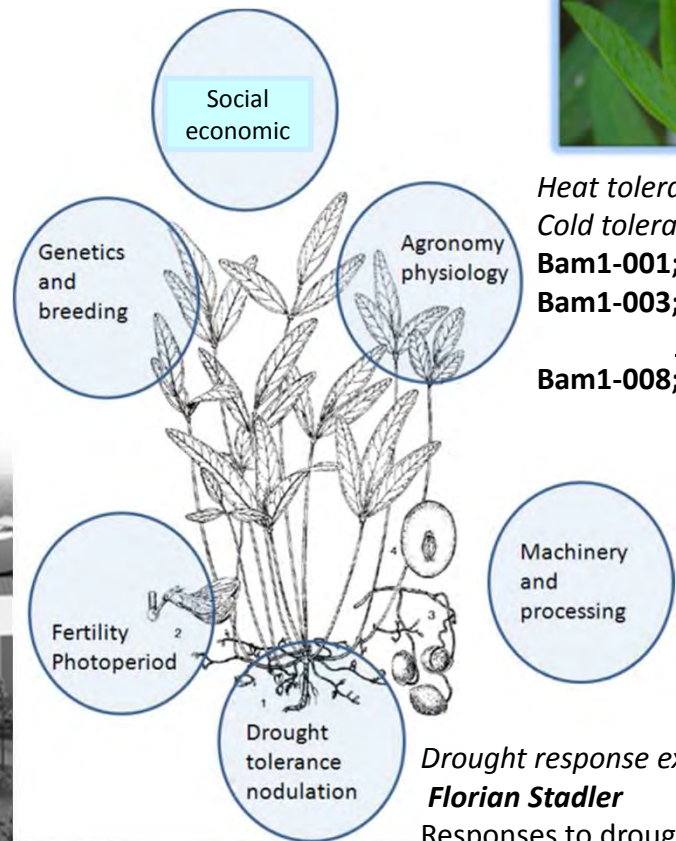
Round 2 Deadline in 2013, a further 40 PhD and 20 MRes expected

BamYIELD - using an exemplar species to develop generic approaches





Heat tolerance; Al-Shareef
Cold tolerance; Noah
Bam1-001; productivity and N₂ fixation; Mukhta Musa
Bam1-003; adaptation to different environments; Philip Cleasby
Bam1-008; Canopy modelling for ideotypes; Josie Dodd



[Proposals in development; Perlis and UNMC]

Drought response expression;
Florian Stadler
Responses to drought;
Yusuf Muhamamd
Genetical genomics of drought;
Hui Hui Chai



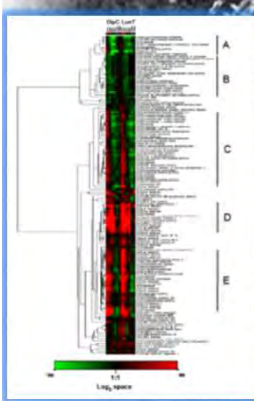
African genetic diversity; Stadler
African genetic diversity; Moloisiwa
Indonesian genetic diversity; Endah Redjeki
Genetic mapping and domestication;
Nariman Ahmad
Bam1-006; MAGIC breeding; Aliyu Siise

Photoperiod requirement;
Kendabie
Bam1-009 Fertility in bambara groundnut; Bhavya Dhanaraj

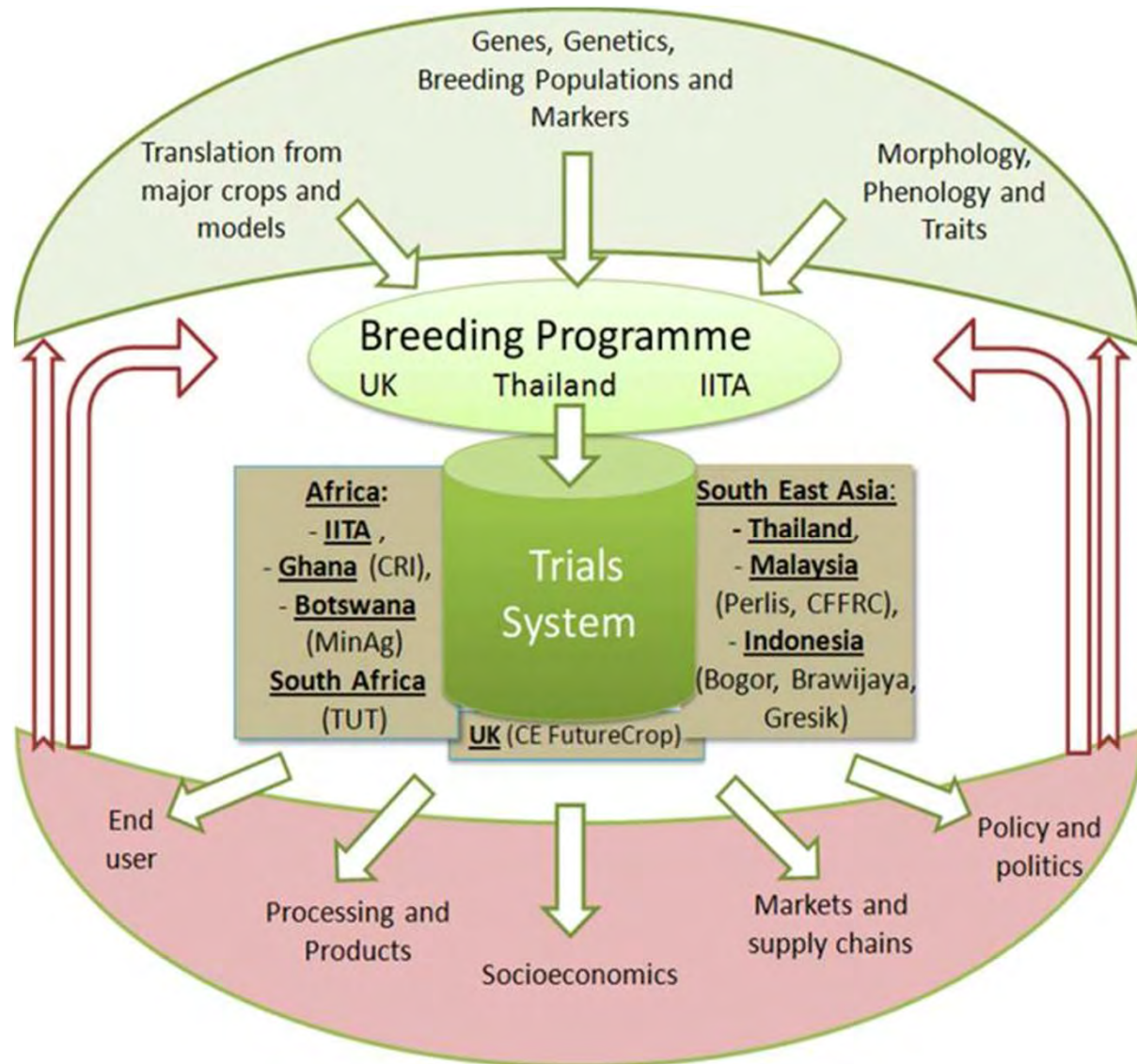


Translation projects

Overlaying the BG transcriptome onto Soybean (MSc); Faraz Khan
Linking genetic maps of BG with physical maps of Soybean; Hui Hui Chai
Bam1-004 Machine Learning to detect drought tolerance pathways; Venkata Suresh
Bam1-007 Combining multiple data types; Faraz Khan



Recent completed projects
Current on-going projects
BamYield agreed projects



International trial programme:

**Thailand
Malaysia
Indonesia**

**Ghana
Niger/Niger (IITA)
Botswana
South Africa**

**Crossing, controlled environment glasshouses
UK**

Applications of markers in NUS – an example focused around bambara groundnut

TRANSCRIPTOME SEQ; MISA.pl script results	Oil palm	Bambara groundnut	Lablab	Winged Bean
	454	454+SOLiD	454+Illumina	Illumina
Total number of sequences examined:	62,337	51,272	49,395	84,892
Total size of examined sequences (bp):	34,596,101	34,513,803	34,862,628	55,972,781
Total number of identified SSRs:	1,787	1,187	4,465	3,857
Number of SSR containing sequences:	1,641	1,110	3,900	3,561
Number of sequences containing more than 1 SSR:	99	61	489	269
Number of SSRs present in compound formation:	126	68	243	152
Developed polymorphic microsatellites	n/a	103	50+	to do
Number of gene models:	62,337	51,272	49,395	83,084
(Newbler assembly from tissue)	Mesocarp	Leaf	Leaf	Leaf/shoot/flowers

Quality control

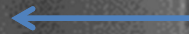
Hybrid F1



Parent 1&2

Simple agarose system :
Confirmation of hybrids

Hybrid F1



Horizontal acrylamide system:
Fingerprinting and quality control, anchoring map markers,

Hybrid F1



Parent 1&2

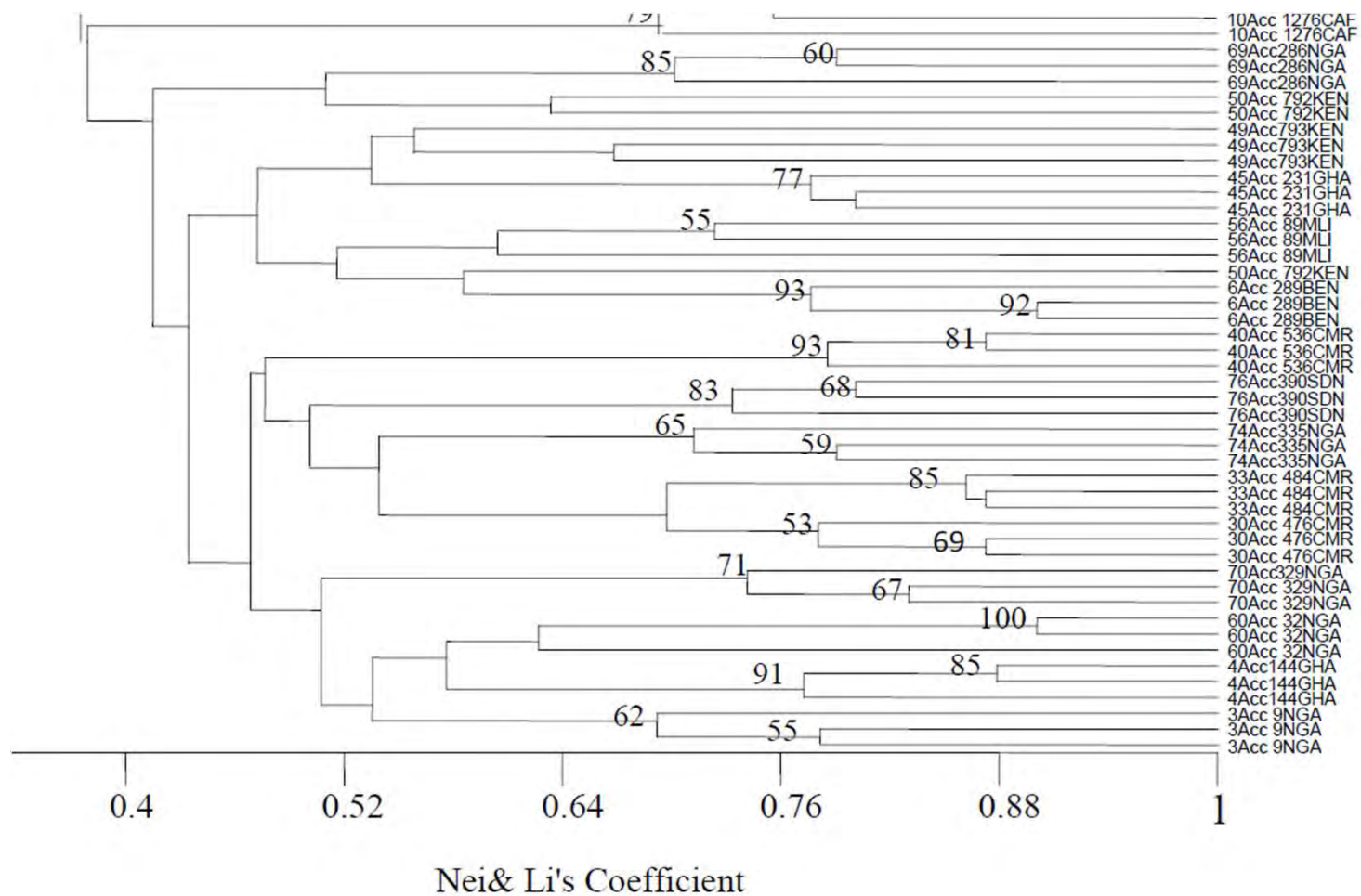
Hybrid F1



Breeding programme based QC needed

Genetic diversity

20 SSRs amplified from 35 landraces (3 accessions per landrace); UPGMA tree



3 accession from single
landraces generally cluster

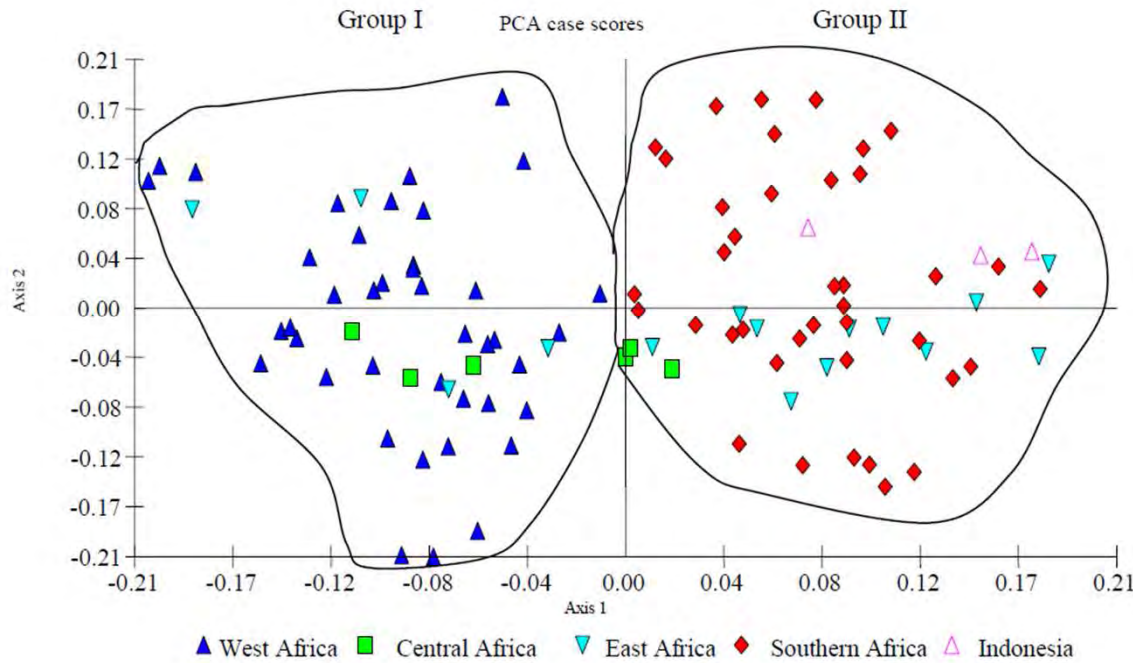
Average heterozygosity for single genotype accessions from 121 landraces = 0.01

Develop 'single genotype' accession lines (aim 400-500)

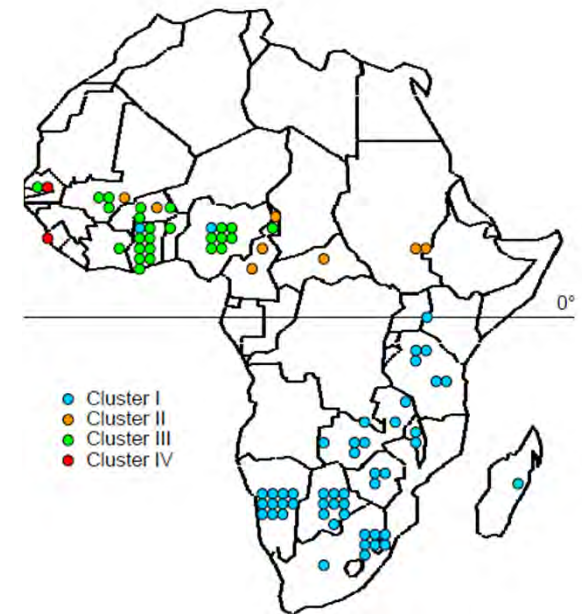
Genetic diversity



Clear differentiation between West and South+East African accessions



SSR-based genetic diversity analysis (PCA)

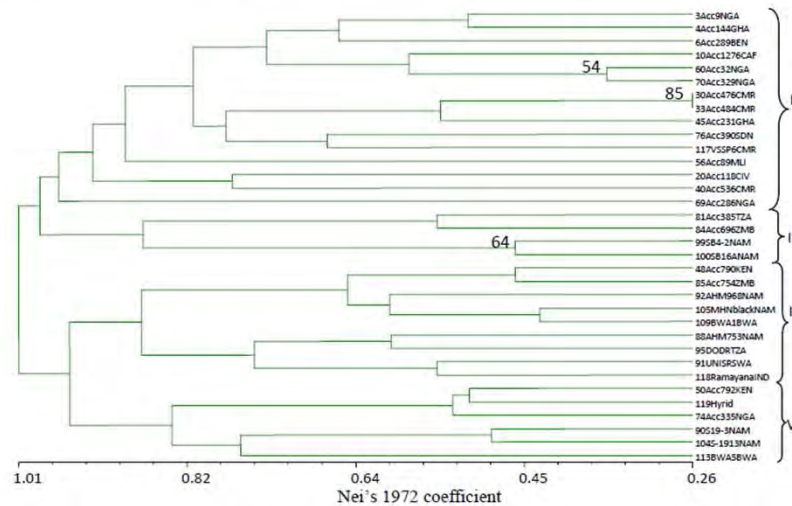


DArT microarray analysis

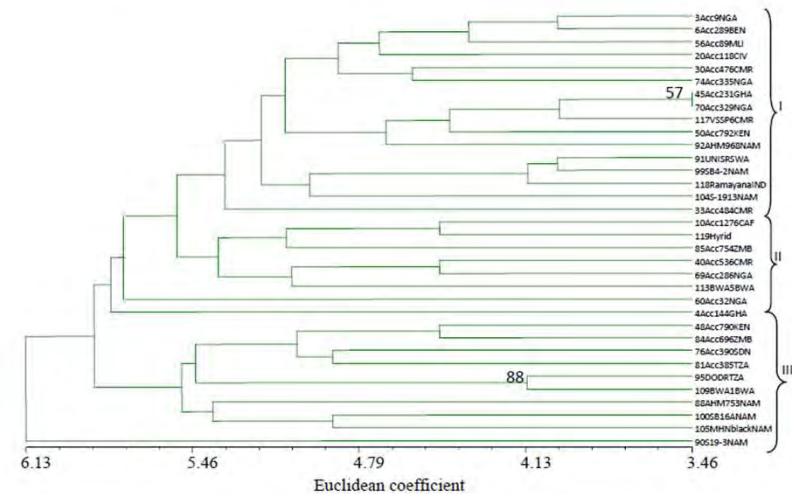
Crosses to introduce new variation could be made between agro-ecologically matched accessions but from different breeding groups

Genetic diversity

a) SSR marker dendrogram



b) Morphological marker dendrogram



What is the relationship between genetic and morphological/trait diversity estimates for selection of breeding parent?

Can we sample variation in genes known to be involved in morphology/traits – a 'meaningful' genetic diversity index?

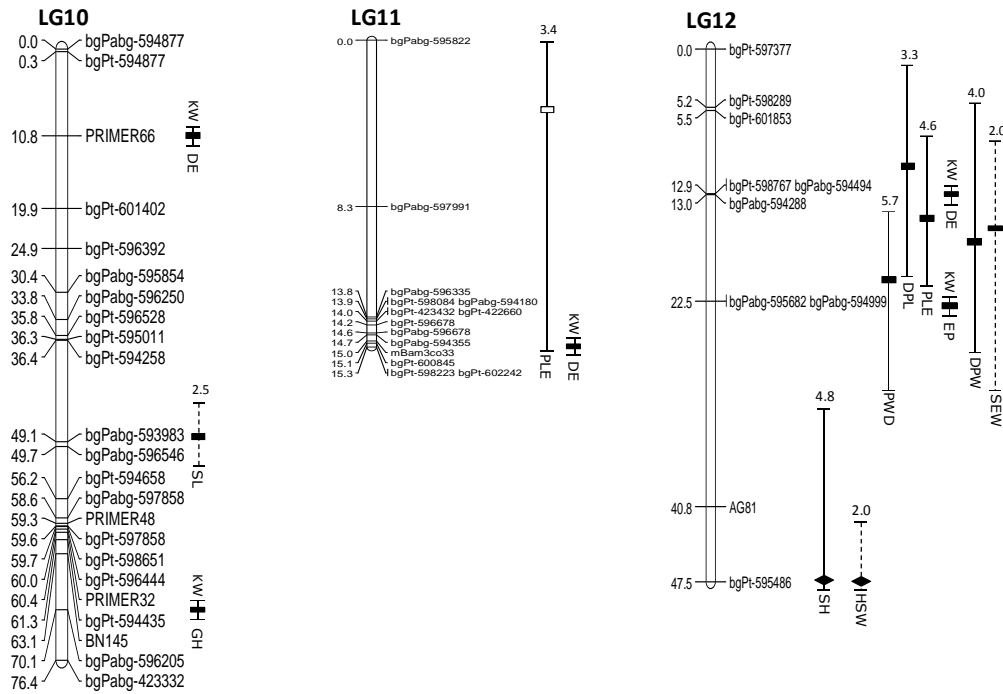
Trait diversity

100018070 F 0--48:A>T	TGCAGCATCATGCTCCTGAACAATTTTTCCAATTATTATTTTTATTTATTTTACAGATCGGAAG	0	1	0	1	1	1	1	1	0	1	1	0	1	0	1	0	0	1	0	1	0	1	1	0	0	1	1	1	0	0	1	0	1	0		
100018070 F 0-48:A>T-48:A>T	TGCAGTATCATGTCTGAACAATTTTTCCAATTATTATTTTTATTTATTTTACAGATCGGAAG	1	0	1	1	0	1	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
100024578 F 0-13:T>C	TGCAGCGCTGACCTAAGCATAAAGTGATCCGAATCCTTATTTAGATTTCCAGAAAAATAAAGTTA	1	1	1	0	1	1	1	0	1	1	0	1	0	1	0	0	1	0	0	1	0	1	1	1	1	0	0	1	0	1	1	0	1	1	1	
100024578 F 0-13:T>C-13:T>C	TGCAGCGCTGACCTAAGCATAAAGTGATCCGAATCCTTATTTAGATTTCCAGAAAAATAAAGTTA	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	
100026770 F 0-49:A>G	TGCAGCTAAAGGTTACAGTATCAATCATCATATGAATGAATGTCAGAAATTTCTAGAATTTGGG	1	1	1	0	0	1	1	1	1	1	1	1	0	0	0	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
100026770 F 0-49:A>G-49:A>G	TGCAGCTAAAGGTTACAGTATCAATCATCATATGAATGAATGTCAGAAATTTCTAGAATTTGGG	1	0	1	1	1	0	1	1	0	0	1	0	1	1	1	1	1	1	1	1	0	1	0	1	0	1	0	1	0	1	0	1	1	0	1	
100030646 F 0-20:G>C	TGCAGGATCTCTTGACTTGGGGTCCATCGCATTGATACTGAAAAAATGAACAATAACAACATTA	1	1	1	1	0	0	0	0	1	1	1	1	0	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
100030646 F 0-20:G>C-20:G>C	TGCAGGATCTCTTGACTTGGGGTCCATCGCATTGATACTGAAAAAATGAACAATAACAACATTA	1	0	1	0	1	1	1	1	0	0	0	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
100031990 F 0-45:T>C	TGCAGAAAGAAAAACCTGGCATCCGAAGCGTTGCGCAACAATCTTCTGCTGAAACGCATTCTTT	0	0	1	0	0	1	1	1	0	0	1	0	1	1	0	1	0	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	
100031990 F 0-45:T>C-45:T>C	TGCAGAAAGAAAAACCTGGCATCCGAAGCGTTGCGCAACAATCTTCTGCTGAAACGCATTCTTT	1	1	0	1	1	0	1	0	1	1	1	1	1	0	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	
100013573 F 0-53:A>G	TGCAGCACTGTGCGAACATTGGCACCATGTAGAGCTTGTGGATTAGGTTGAGACCTGTGAGATGTTA	1	0	1	1	1	1	0	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
100013573 F 0-53:A>G-53:A>G	TGCAGCACTGTGCGAACATTGGCACCATGTAGAGCTTGTGGATTAGGTTGAGACCTGTGAGATGTTA	0	1	0	1	1	0	1	1	1	0	1	1	1	1	1	0	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
100027885 F 0-43:G>C	TGCAGTCCATCTCTTGTCTCCCTCTGCTCCTCTCGAAGCCTCCCAATATATCCATTGAGTCA	1	0	1	1	1	0	1	1	0	0	1	0	1	1	1	1	0	1	1	1	0	1	1	1	0	0	1	0	1	0	1	1	1	1	1	
100027885 F 0-43:G>C-43:G>C	TGCAGTCCATCTCTTGTCTCCCTCTGCTCCTCTCGAAGCCTCCCAATATATCCATTGAGTCA	1	1	1	0	1	1	1	1	1	1	0	0	1	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
100028343 F 0-22:G>A	TGCAGCACTGTGCGAACATTGGCACCATGTAGAGCTTGTGGATTAGGTTGAGACCTGTGAGATGTTA	1	1	1	0	0	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
100028343 F 0-22:G>A-22:G>A	TGCAGCACTGTGCGAACATTGGCACCATGTAGAGCTTGTGGATTAGGTTGAGACCTGTGAGATGTTA	0	0	0	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
100028877 F 0-46:C>A	TGCAGGGGAAAAACCAAGCCTCCATTTGAGCAGGTAATGTCTCTTTACAGATCGGAAGAGCGG	0	0	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
100028877 F 0-46:C>A-46:C>A	TGCAGGGGAAAAACCAAGCCTCCATTTGAGCAGGTAATGTCTCTTTACAGATCGGAAGAGCGG	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
100038155 F 0-30:A>G	TGCAGAACTGAGCCACTGACTGTGCTCCCTACTCTGAAGTTCGCAATTGAGCGCAATCCATACAT	1	0	1	1	1	1	1	0	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
100038155 F 0-30:A>G-30:A>G	TGCAGAACTGAGCCACTGACTGTGCTCCCTACTCTGAAGTTCGCAATTGAGCGCAATCCATACAT	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
100015159 F 0-42:C>A	TGCAGCATCTACTATGTCTTTATCTTGTTCATCAAAACCGAGAACATCGTATTCAAGCGTGG	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
100015159 F 0-42:C>A-42:C>A	TGCAGCATCTACTATGTCTTTATCTTGTTCATCAAAACCGAGAACATCGTATTCAAGCGTGG	0	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
100020562 F 0-26:T>C	TGCAGTCAAATTAGTATTGGAATTGATGTAATTTTCATATGATGAATAGTTGAACGTGATAATGT	0	-	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
100020562 F 0-26:T>C-26:T>C	TGCAGTCAAATTAGTATTGGAATTGATGTAATTTTCATATGATGAATAGTTGAACGTGATAATGT	1	-	-	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
100028314 F 0-45:A>G	TGCAGACATGAGGAAGATTGCGATGAAATGAAGATGATGTTAGGATAGGAGAAACCAAGATT	1	1	0	1	1	0	0	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Within species marker data – genotyping by sequence – DArT Seq

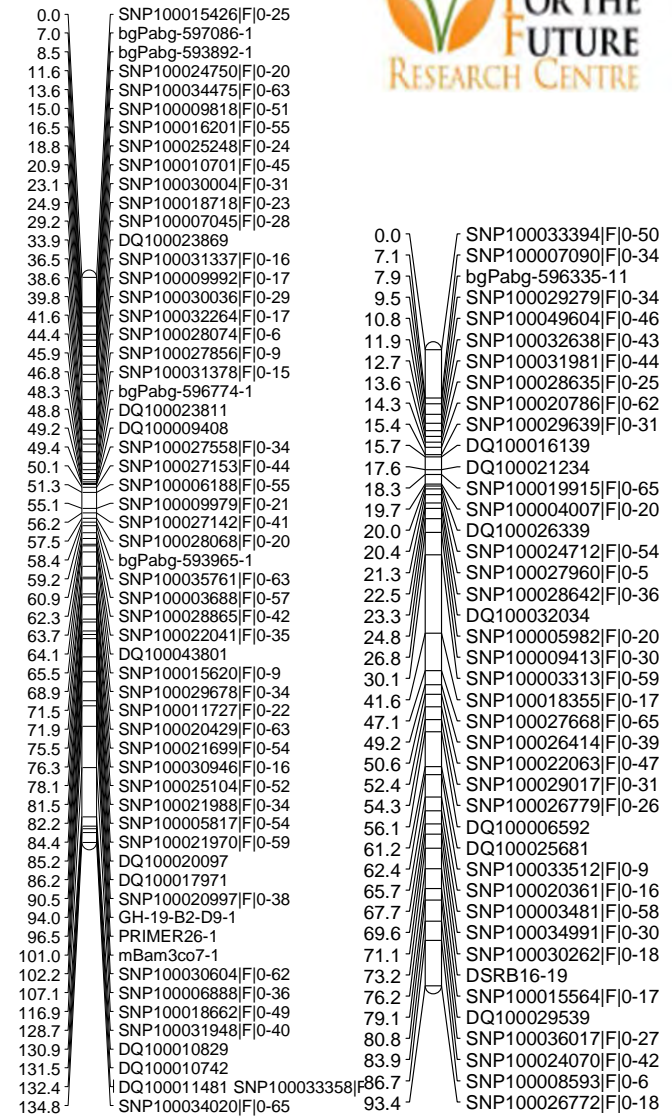
	probe ID	affyids	L2	L4	L5	L6	L7	L8	L9	L10	L13	L17	L19	L20	L21	L22	L23	L24	L26	L27	L35	L36	L37	L38	L39	L41	L43	L44	L45	L47	L48	L49	L51	L52	L53	L56	L57	L58	L60		
PM1	(a,b)	758237	GmaAffx.74553.1.S1_at	a	b	a	b	b	b	a	b	b	b	b	b	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	a	a	b	b	b		
PM2	(a,b)	53127	HgAffx.22497.1.S1_at	a	a	b	b	a	a	a	b	b	b	b	b	a	a	b	b	b	b	b	b	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	
PM3	(a,b)	924184	PsAffx.C30000035_at	b	b	b	b	b	a	b	b	b	b	a	b	a	b	b	b	a	b	a	b	a	a	b	b	a	b	a	b	b	b	b	b	b	a	a	b		
PM4	(a,b)	697325	GmaAffx.13860.1.S1_at	b	b	b	b	b	b	a	a	a	b	a	a	b	b	b	b	b	b	b	a	b	b	a	a	a	a	a	b	a	b	b	b	b	b	b	a	b	
PM5	(a,b)	1046961	Gma.4444.2.S1_a_at	a	b	b	b	b	b	b	b	b	b	b	b	a	b	b	b	a	a	b	b	b	b	b	b	b	a	a	a	b	a	b	b	b	b	b	a	b	
PM6	(a,b)	855943	GmaAffx.79256.1.S1_at	b	b	b	b	a	a	b	b	a	b	a	b	a	a	b	a	b	a	b	b	b	b	b	b	b	b	a	a	a	a	a	a	a	a	a	a	b	
PM7	(a,b)	1196466	Gma.1677.1.S1_a_at	a	b	b	a	a	b	b	b	b	a	b	a	a	a	b	b	b	b	b	a	b	b	b	b	b	a	b	b	b	a	b	b	b	b	b	a	b	
PM8	(a,b)	579583	GmaAffx.67791.1.S1_at	b	b	a	b	a	a	b	b	b	a	b	b	b	b	b	b	a	b	b	b	a	b	b	a	b	a	b	b	b	b	b	b	b	b	b	b	a	b
PM9	(a,b)	422914	HgAffx.2532.2.A1_at	b	a	a	b	a	a	a	b	b	b	a	a	b	a	b	b	b	b	b	b	b	b	b	a	a	b	a	b	b	b	b	b	b	b	b	b	b	a
PM10	(a,b)	725602	GmaAffx.89772.2.S1_at	a	a	b	a	a	a	a	b	a	a	a	a	a	a	a	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
PM11	(a,b)	825413	GmaAffx.58051.1.S1_s_at	b	a	b	b	b	a	b	b	a	b	b	a	a	b	a	b	b	b	b	b	b	b	b	b	a	b	b	b	b	b	b	b	a	b	a	a	b	a
PM12	(a,b)	101657	GmaAffx.83159.1.S1_at	b	a	b	a	b	b	b	a	b	b	a	b	b	b	b	b	b	b	b	a	a	b	b	a	a	a	b	a	a	a	a	b	b	b	b	a	b	
PM13	(a,b)	592511	GmaAffx.36292.1.S1_at	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	a	b	a	a	b	b	a	a	b	a	b	b	b	b	b	a	a	a	b
PM14	(a,b)	875044	Gma.17207.1.S1_at	b	a	a	b	b	b	a	b	b	a	a	b	b	b	b	b	a	b	a	b	b	b	b	b	a	b	a	b	b	b	a	b	b	a	b	b	a	b
PM15	(a,b)	82657	Gma.8070.1.S1_at	b	b	a	a	b	b	a	b	b	b	b	b	b	b	b	b	b	b	b	b	a	b	a	a	b	b	a	b	b	b	b	b	b	b	b	b	b	b
PM16	(a,b)	1093338	GmaAffx.82121.1.S1_at	b	b	b	a	a	b	b	a	a	b	a	a	b	a	b	b	b	b	b	b	b	b	b	a	a	b	b	b	b	b	b	b	b	b	a	a	b	a
PM17	(a,b)	311855	GmaAffx.58051.1.S1_s_at	b	a	b	b	a	a	b	b	a	a	b	a	a	b	a	b	b	b	b	b	b	b	b	b	a	b	b	b	b	b	b	b	b	a	b	b	b	b

Marker Expression data generated using the Soybean Affymetrix microarray

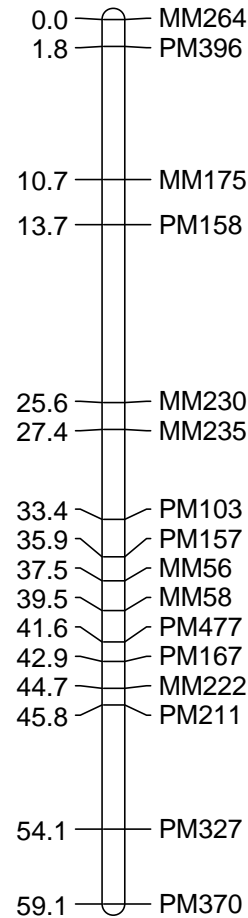


3 linkage groups based on DArT and SSR;
 QTLs shown; Dip C (Botswana) x Tiga
 necaru (Mali); 67 lines

**Generate dense genetic maps through an
 outsourced service provider/collaborator**

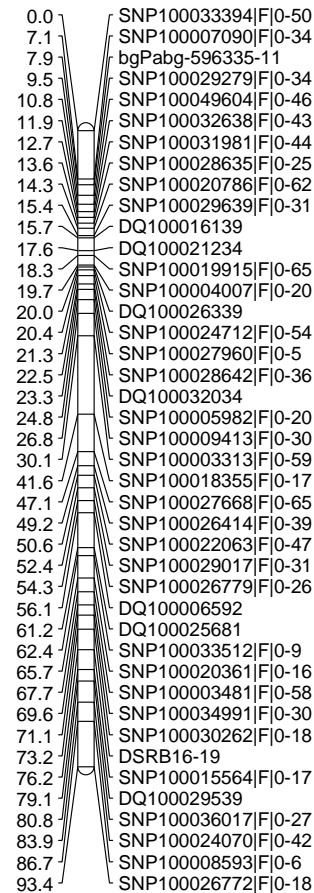


Composite linkage groups,
 DArT, SSR and DArT Seq.;
 raw 2996 SNP and 8500 +/- markers



Cross species
'soybean'
expression map

Same population - integration



Within species
DNA marker
map

Drawing links through cross-species array maps

Drawing links through the DArT seq. markers which link to sequence via 64bp

oil palm – 22 of 30 markers close to trait could be placed uniquely on the same scaffold

Date palm – 49% of DArT Seq Tags had unique hits in genome sequence

Bambara groundnut – 17% of Tags uniquely identified a gene model from the transcriptome

Relate the NUS genetic map to close relatives and model for 'pseudo'physical map

**Within species tool generation
'Within house' tool use**

**Within species tool generation
'Service' tool use**

**Cross-species species tool generation
'Within house' tool use**

**Cross-species species tool generation
'Service' tool use**

4. Development of SSRs for F1/ breeding system evaluation/breeding programme QC

5. Conversion of markers (via gene models/DArT tags?) to single locus for MAB

6. Linkage of in-species genetic maps to model and major crop species

7. Parental selection and ideotypes

**Within species tool generation
'Within house' tool use**

**Within species tool generation
'Service' tool use**

1. Leaf transcriptome sequencing
2. 'Genotype by sequencing' diversity and mapping
3. Other genetic analysis –
GWAS, Genomic selection
10. Next gen sequencing genetical genomics,

8. Genetically located gene candidates

**Cross-species species tool generation
'Within house' tool use**

**Cross-species species tool generation
'Service' tool use**

9. Cross species, genetical genomics (e.g. Affymetrix)

Overall conclusions:

1. Biotech has the potential to make a major contribution to NUS, but it should only be used to where there is a clear application and the same cannot be achieved by a good Breeding programme.
2. Some tools and information need to be developed in species, to be robust and allow accurate assessment and QC of breeding and trials programmes
3. Some cross-species tools may be directly valuable, but an information framework that allows translation of data from major crops and model species is critical for NUS

The current approach of :

- transcriptome (SSRs-20+?; Linkage across species using gene models)
- DArT Seq Markers (Diversity analysis (400?), mapping (200), links to gene models.)

Current costs around \$15K per species and is becoming cheaper.

**Funding: Government of Malaysia, University of Nottingham (UNiM)
EU, National government scholarships scheme**

SSR development:	Katie Mayes, Ozie Molosiwa, Nariman Ahmad
DArT Array:	Florian Stadler, Andrzej Kilian
DArT Seq:	Nariman Ahmad, Andrzej Kilian
DArT Array/SSR diversity:	Florian Stadler, Ozie Molosiwa
Mapping:	Rakhi Basu, Nariman Ahmad, Hui Hui Chai
QTL analysis:	Nariman Ahmad
Genetical Genomics:	Hui Hui Chai
Photoperiod and crossing:	Presidor Kendabie
LabLab:	Eliezah Kamau, Pavithravani, B.V. (Kirkhouse Trust)
Winged Bean:	Quin Nee Wong

