

**Working Group on Biochemical and Molecular Techniques
and DNA-Profiling in Particular**

BMT/18/6

**Eighteenth Session
Hangzhou, China, October 16 to 18, 2019**

Original: English
Date: September 11, 2019

ADVANCES IN THE CONSTRUCTION AND APPLICATION OF DNA FINGERPRINT DATABASE IN MAIZE

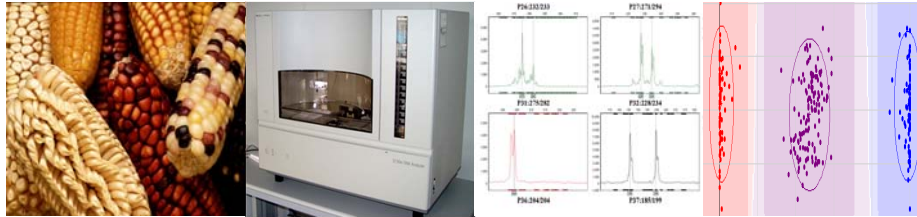
Document prepared by an expert from China

Disclaimer: this document does not represent UPOV policies or guidance

The annex to this document contains a copy of a presentation on “Advances in the construction and application of DNA fingerprint database in maize”, to be made at the eighteenth session of the BMT.

[Annex follows]

Advances in the construction and application of DNA fingerprint database in maize



Maize Research Center, Beijing Academy of
Agriculture and Forestry Sciences, China

BMT/18, Hangzhou, China

Oct 16 - 18, 2019

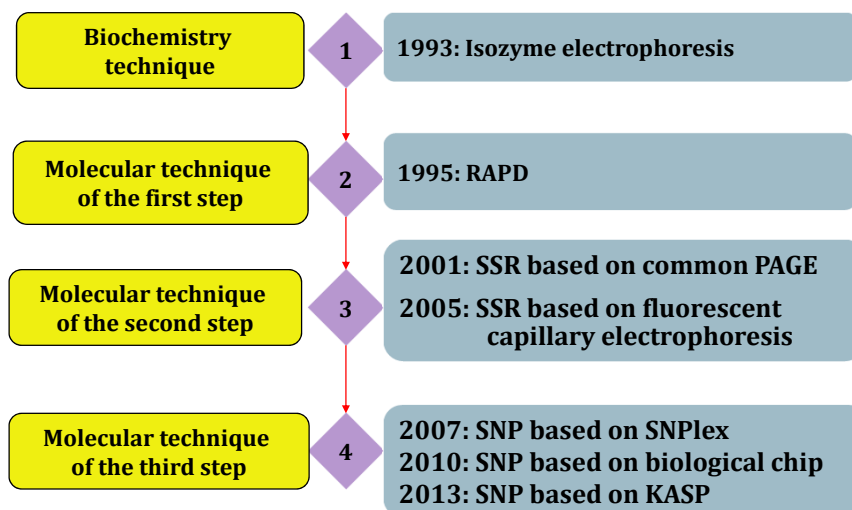


Outline

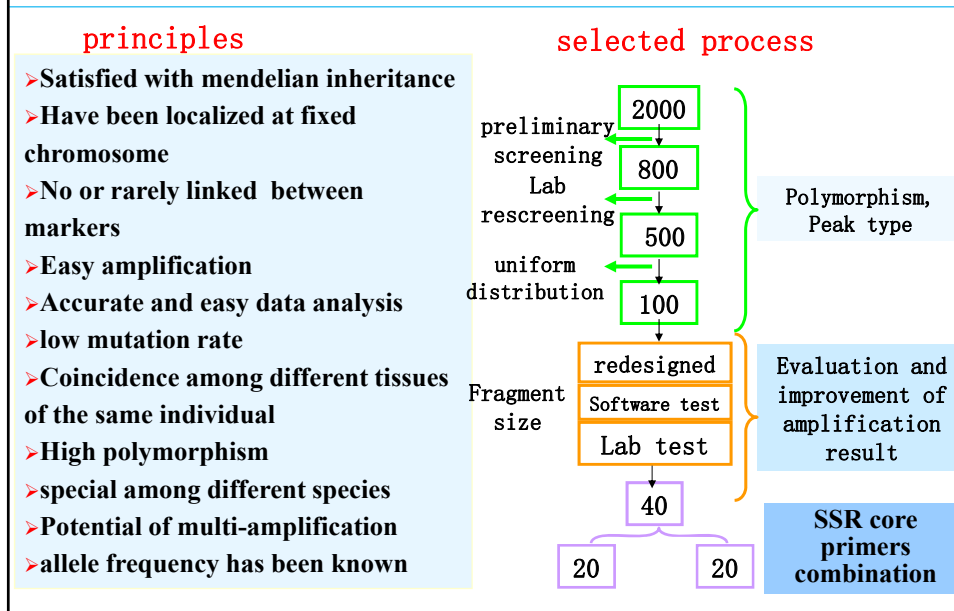
- The construction of the maize SSR-DNA fingerprint database and the development and application of the key technologies.
- The progresses in the SNP-based molecular fingerprint detection technology of maize seed.

Part I. The construction of the maize SSR-DNA fingerprint database and the development and application of the key technologies.

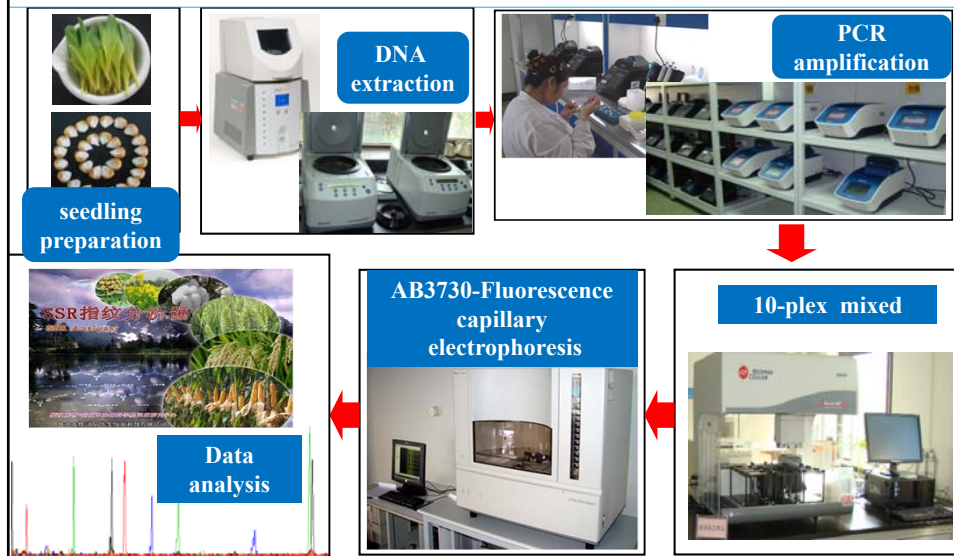
Four stages of maize variety identification



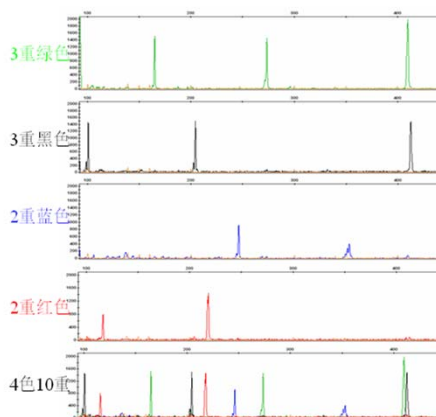
1. Based on 11 principles, 40 core primers were selected, determined and redesigned from 2000 SSRs.



2. Six technical steps were optimized and standardized to build a fast, accurate, economical, simple, stable and reliable detection technology system, and six DNA fingerprint testing standards were established.



3. A 10-plex electrophoresis testing system with 5 fluorescence was established, increasing the testing efficiency by 10 times.



5 colors -10 plex fluorescence electrophoresis

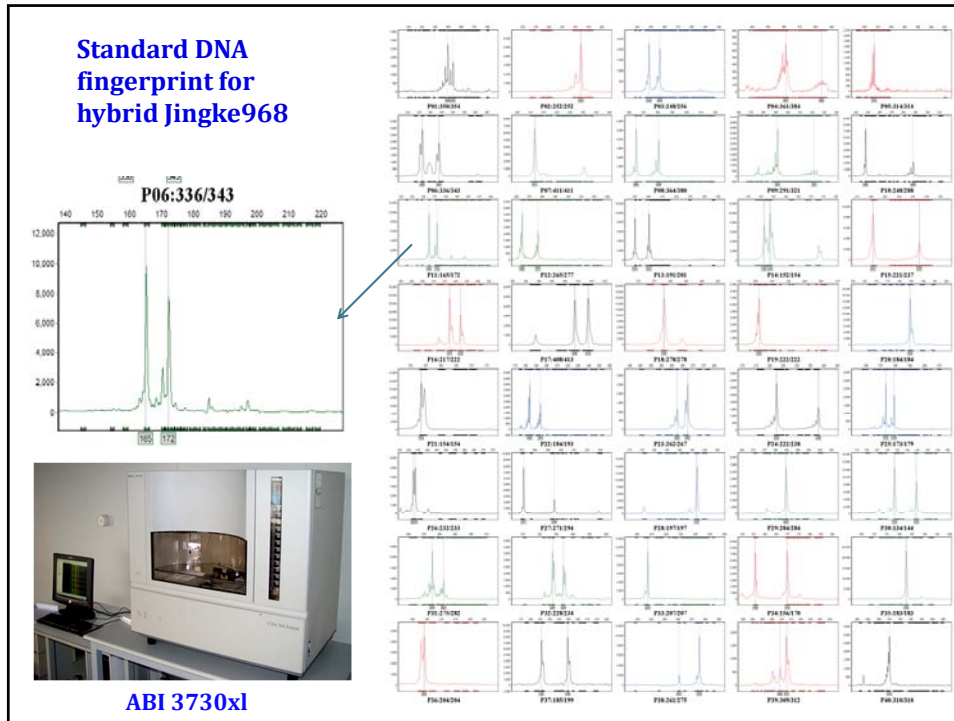
Set	Locus name	Chr	Bin range	Allele range	Fluorescence labeled	Set	Locus name	Chr	Bin range	Allele range	Fluorescence labeled
I-1	umc14326	10.02	211-229	VIC	VIC	II-1	umc14297	5.03	125-145	VIC	VIC
I-1	umc210543	3.00	280-320	VIC	VIC	II-1	umc19995	4.00	167-200	VIC	VIC
I-1	ph07234	4.01	405-432	VIC	VIC	II-1	umc14095	3.07	231-265	VIC	VIC
I-1	umc15452	7.00	180-240	NED	NED	II-1	umc11474	1.07	149-172	NED	NED
I-1	hdg17024	6.05	200-247	NED	NED	II-1	ph2299522	6.07	200-254	NED	NED
I-1	ph06289	9.03	391-415	NED	NED	II-1	umc216365	10.04	280-352	NED	NED
I-1	hdg14081	5.06	220-229	PET	PET	II-1	ph033337e1	8.00	180-222	PET	PET
I-1	hdg13961	1.03	310-360	PET	PET	II-1	hdg14094	4.04	245-331	PET	PET
I-1	hdg13494	5.07	240-312	FAM	FAM	II-1	umc20840c	9.01	184-214	FAM	FAM
I-1	hdg19407	2.08	324-338	FAM	FAM	II-1	ph09009e1	2.00	231-237	FAM	FAM
I-2	umc11259	7.04	149-175	VIC	VIC	II-2	umc19504	7.03	153-176	VIC	VIC
I-2	umc13355	1.06	233-257	VIC	VIC	II-2	umc216083	7.01	189-244	VIC	VIC
I-2	ph05382	3.05	333-363	VIC	VIC	II-2	ph04126	10.00	296-334	VIC	VIC
I-2	hdg16105	6.00	154-216	NED	NED	II-2	hdg12355	8.02	174-198	NED	NED
I-2	umc20074	2.04	233-300	NED	NED	II-2	umc15504b	2.07	216-238	NED	NED
I-2	hdg129104	4.06	342-421	NED	NED	II-2	hdg14912	6.01	259-313	NED	NED
I-2	umc159032	10.05	165-196	PET	PET	II-2	hdg152081	2.00	156-204	PET	PET
I-2	umc1705e1	5.03	254-340	PET	PET	II-2	umc123104	9.05	230-283	PET	PET
I-2	ph09045	8.08	202-238	FAM	FAM	II-2	hdg1673e7	1.00	173-255	FAM	FAM
I-2	umc149243	9.04	270-290	FAM	FAM	II-2	umc211953	5.02	265-295	FAM	FAM

The 10-plex mode for 40 core SSRs.

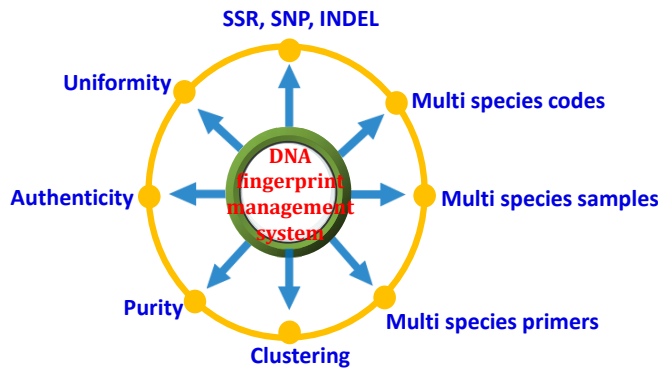
4. The maize DNA fingerprint database with more than 30,000 samples was created, providing an accurate molecular ID for each varieties (inbred lines, hybrids, landraces etc.).

containing > 30,000 maize varieties

- Expansion database: > 30,000
 - VCU varieties: > 24,000
 - PVP varieties: > 2,000
 - Registered varieties: > 6,000
 - Inbred lines: > 2,000
 - Core landraces: 124



5. The first plant DNA fingerprint management system with independent intellectual property rights was developed, which is compatible with many species, various markers and multiple functions.



The development of robust SSR-DNA database management system

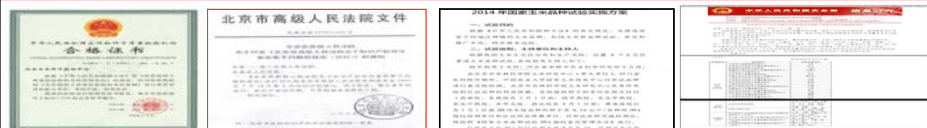


6、SSR analyzer imbedded in the DNA fingerprint database management system



7. Multiple qualifications and recognitions were obtained from MOA and Supreme People's Court

- First designated center for variety genuineness identification by MOA, China;
- Designated center for variety genuineness identification of forensic authentication by Beijing higher people's court;
- Leading center of DNA fingerprinting technology for national maize region trial;
- Authorized center for region trial by more than 20 provinces;
- Assigned center for database construction of standard samples by MOA, China;
- Uniqueness assigned center for variety cultivar genuineness identification of Supervise Year Action by MOA, China.



Part II . The progresses in the SNP-based molecular fingerprint detection technology of maize seed.

1. The screening process of SNP loci suitable for maize IP confirmation, identification and purity test

Pool of outstanding loci across the genome

High density chip (60K)

Expanded loci for confirmation ($\approx 3K$)

Core loci for identification (200)

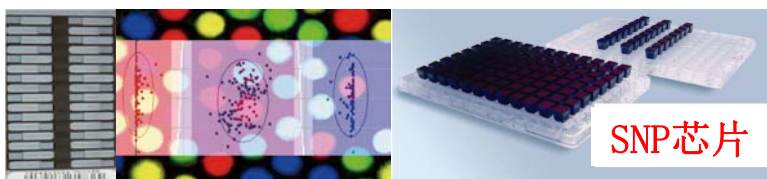
loci for purity test (20)

2. The development of the novel maize SNP chip Maize6H-60K

2.1 General info

➤ **Independent property rights and multi purpose:** Maize6H-60K is developed by the Maize Research Center, Beijing Academy of Agriculture and Forestry Sciences with independent IP right. It is the first maize SNP chip for maize variety identification, IP confirmation and molecular breeding in China.

➤ **Outstanding loci with multiple validation:** based on the deep sequencing of 400 Chinese and/or foreign inbred lines with broad representation, outstanding loci were revealed. Along with the previous 6 maize chips, the Maize6H-60K was finalized after multiple rounds of validation.



➤ **High coverage and high polymorphism:**

the chip contains more than 60,000 SNP loci, more than 6000 InDels, including loci from genome, chloroplast genome and core loci for maize variety identification. It presents six major merits of high density, high quality, high discrimination power, high compatibility, high flexibility and high cost efficiency. ◦

➤ **Ready to use:** Chips have been produced, and they are promoted with the Consortium mode to accomplish the patent licensing, price concessions, technical support, cooperation and mutual benefits.



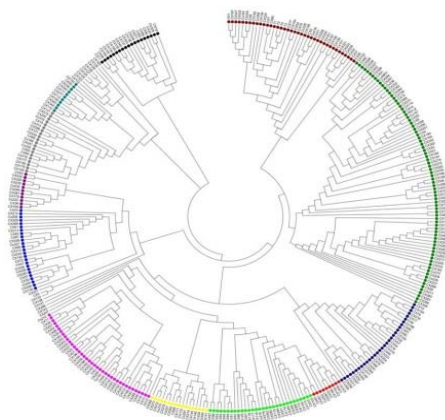
384 chip



Mini-96 chip

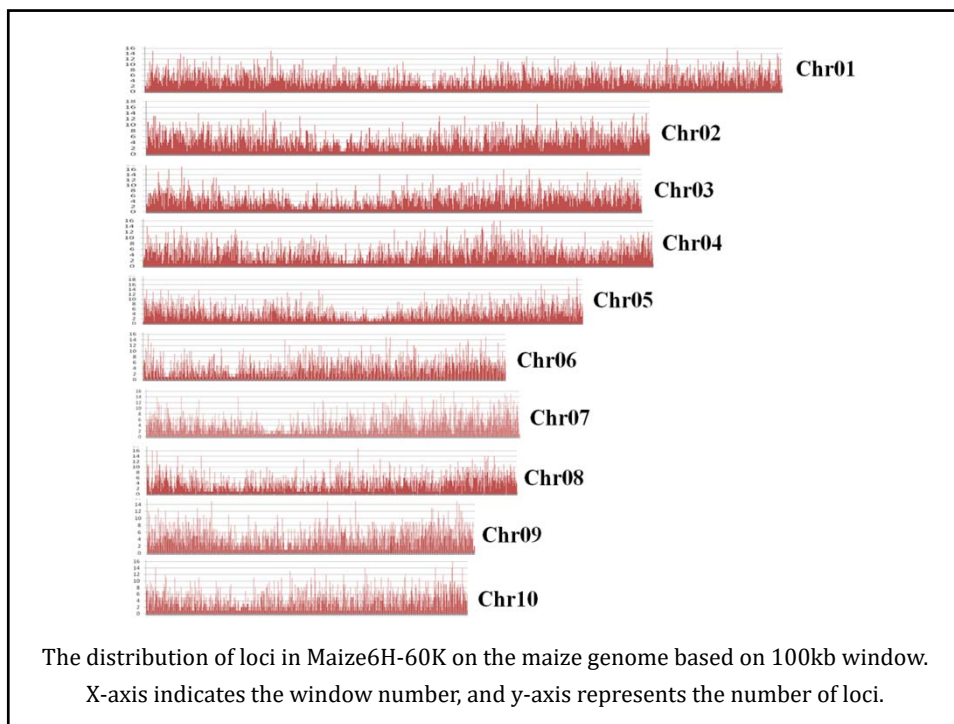
2.2 Materials for evaluation

The samples used for loci evaluation include domestic and foreign materials with wide genetic background. They covered nine germplasm groups, as well as the tropical, sweet and waxy, and local varieties.

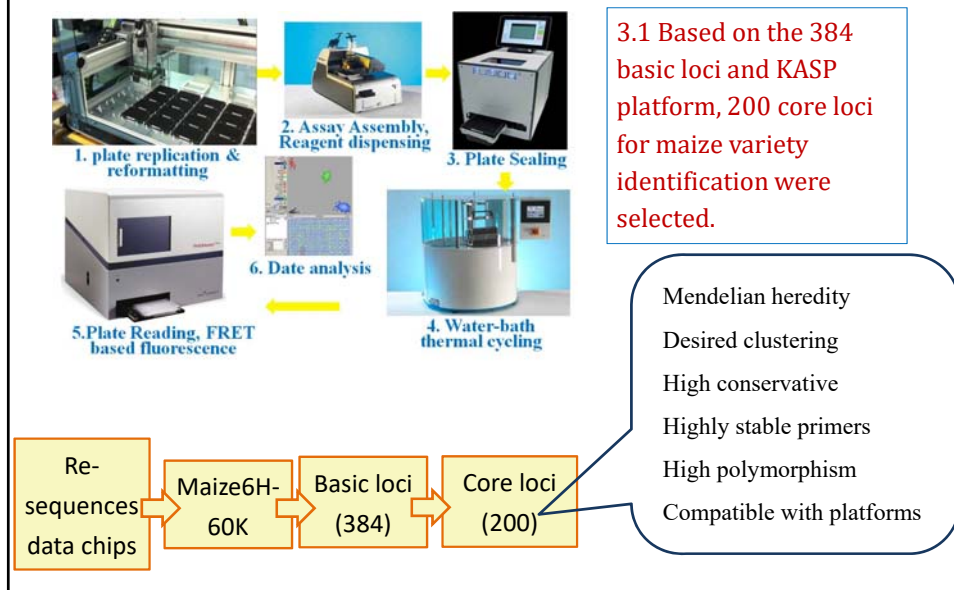


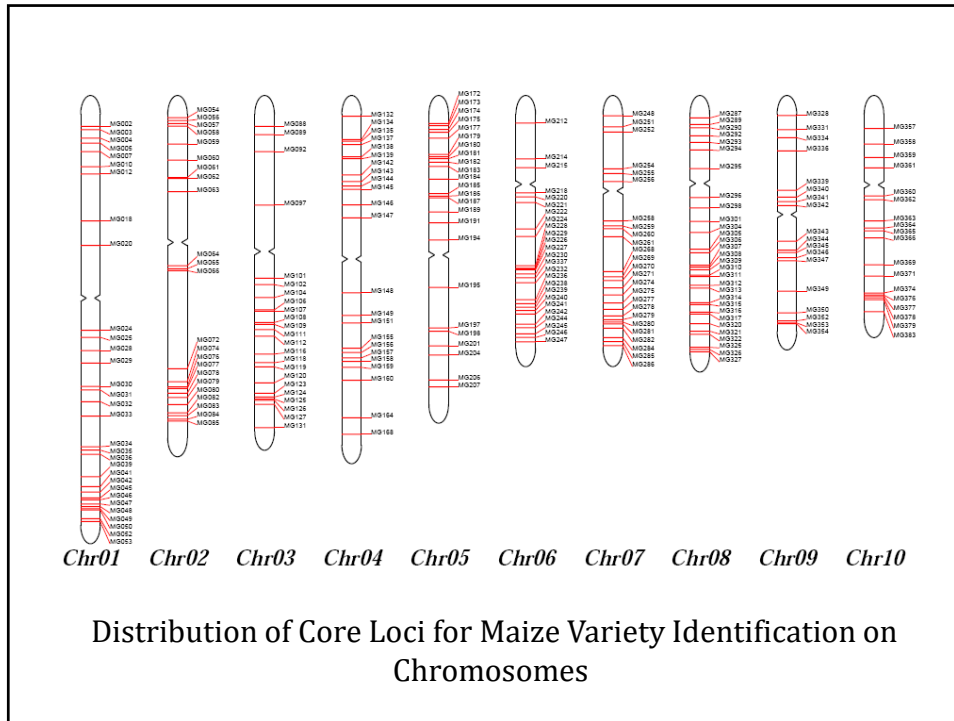
2.3 Characteristics-6H

- High Density: more than 60,000 loci, evenly distributed on the genome;
- High Quality : over 90% are effective loci, and the genotyping data can be analyzed automatically ;
- High Discrimination Power : Loci MAF ≥ 0.2 ;
- High Compatibility: compatible with both Affymetrix and Illumina chip platform;
- High Flexibility: formats of both 384 and 96 samples are available to accommodate various
- High Cost Efficiency: the best cost performance among chips with mid to high density.

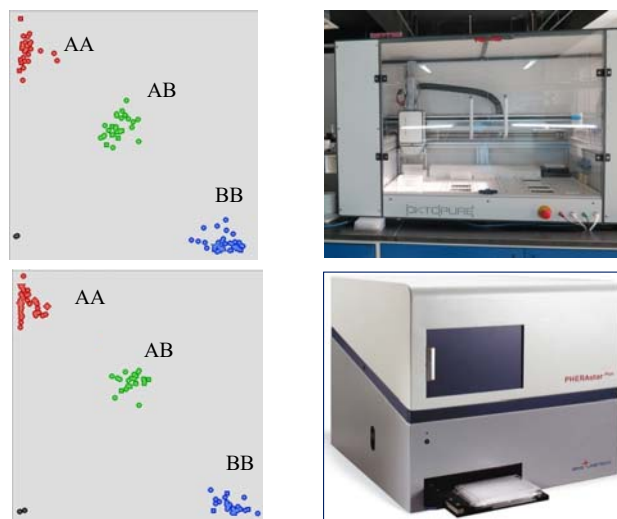


3. Determination of core SNP loci for maize variety identification and the establishment of technology system for authenticity identification





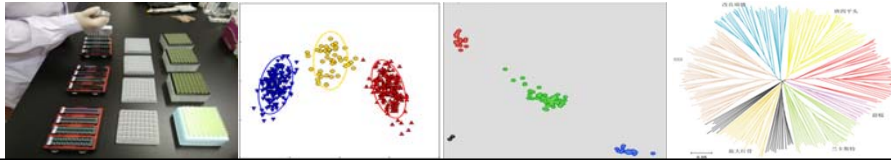
3.2 Based on KASP platform and the work station for DNA prep, a high throughput and high efficiency detection system was established for the core 200 SNP loci.



SNP genotyping on KASP

4. With these core loci, the standard SNP-DNA fingerprint database with 10,000 maize samples was established.

- Samples: registered varieties, varieties with protection rights and representative lines;
- Loci: Core loci, 200 SNP;
- Platform: Chip platform along with the KASP platform;
- Data: genotype shown as ATCG; Data Integration of various platforms through the SNP database management system;
- Reference samples: 2 homozygous DH lines;
- Function: Provides data support for DUS, variety Rights and authenticity identification.

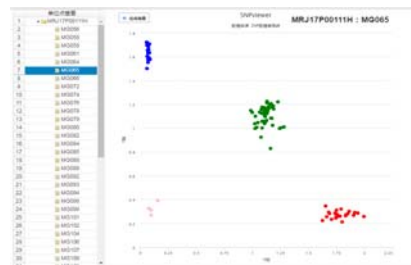


➤ Demonstration of the SNP-DNA fingerprint input management system and relevant data

Sample query

序号	样品名称	样品编号	样品来源	样品类型	负责人	所属	创建时间	操作
1	0502	0102122	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
2	0503	0102171	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
3	0507	0102043	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
4	0508	0102073	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
5	0509	0102170	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
6	0510	0102019	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
7	05011	0102015	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
8	05012	0102014	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
9	05013	0102010	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
10	05014	0102011	山东省农科院	杂种	王强	玉米所	2010-08-09	删除

Genotype query (KASP)



Fingerprint query

序号	样品名称	样品编号	样品来源	样品类型	负责人	所属	创建时间	操作
1	0501001	0102004	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
2	0501002	0102004	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
3	0501003	0102004	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
4	0501004	0102004	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
5	0501005	0102004	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
6	0501006	0102004	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
7	0501007	0102004	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
8	0501008	0102004	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
9	0501009	0102004	山东省农科院	杂种	王强	玉米所	2010-08-09	删除

Comparison outcomes

序号	样品名称	样品编号	样品来源	样品类型	负责人	所属	创建时间	操作
7	0501001	0102004	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
24	0501004	0102004	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
25	0501005	0102004	山东省农科院	杂种	王强	玉米所	2010-08-09	删除
210	0501009	0102004	山东省农科院	杂种	王强	玉米所	2010-08-09	删除

5. The application of each SNP sets

- Expanded loci, about 3,000 SNPs (selected from 60K chip), mainly used for the identification of maize EDV, inbred lines and germplasm resources;
- Core loci of 200 SNPs, mainly used for the ID of hybrids registration;
- 20 SNPs, mainly used for the purity testing of maize hybrids;
- Based on these 200 SNPs, the DNA-SNP fingerprint database of known hybrids and inbred lines was established to facilitate the confirmation of maize IP, DUS and authentication.

```
....ATGAC....ACACGCCA....TCGGGGTC....GTCGACCG....TCGT....  
....GTGAC....ACACGCCA....TCGAGGTC....GTCACCG....TCGC....  
....GTGAC....ACATGCCA....TCGGGGTC....GTCACCG....TCGT....  
....GTGAC....ACAAGCCA....TCGGGGTC....GTCGACCG....TCGT....
```

Thanks for your
attention!

