E

BMT-TWA/Maize/2/4 Add.
ORIGINAL: English
DATE: December 3, 2007

UPOV

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

AD HOC CROP SUBGROUP ON MOLECULAR TECHNIQUES FOR MAIZE

Second Session
Chicago, United States of America, December 3, 2007

ADDENDUM TO DOCUMENT BMT-TWA/MAIZE/2/4

SELECTION AND DEVELOPMENT OF REPRESENTATIVE SIMPLE SEQUENCE REPEAT PRIMERS AND MULTIPLEX SSR SETS FOR HIGH THROUGHPUT AUTOMATED GENOTYPING IN MAIZE

Document prepared by experts from China

This document is an addendum to document BMT-TWA/Maize/2/4 "Selection and Development of Representative Simple Sequence Repeat Primers and Multiplex SSR Sets for High Throughput Automated Genotyping in Maize" and contains a copy of the presentation made by experts from China at the second session of the *Ad Hoc* Crop Subgroup on Molecular Techniques for Maize.

Slide 1

Selection and development of representative simple sequence repeat primers and multiplex SSR sets for high throughput automated genotyping in maize

WANG Fengge, ZHAO Jiuran, DAI Jingrui, et al

Maize research center, Beijing academy
of agricultural & forestry science
China

Slide 2

Purpose of the study

1. **thoroughly screen and assess published maize SSR primers to establish a set of universal core SSR primers in maize**
   - More than 2000 SSR primers have been developed in maize
   - great variability among different markers in polymorphism, amplification quality, whether codominant inheritance, number of amplification sites, and location and distribution in the genome
   - A lack of systematic evaluations of SSR primers in maize has hampered their utilization in maize. only a few and unsystematic evaluation on the developed maize SSR primers

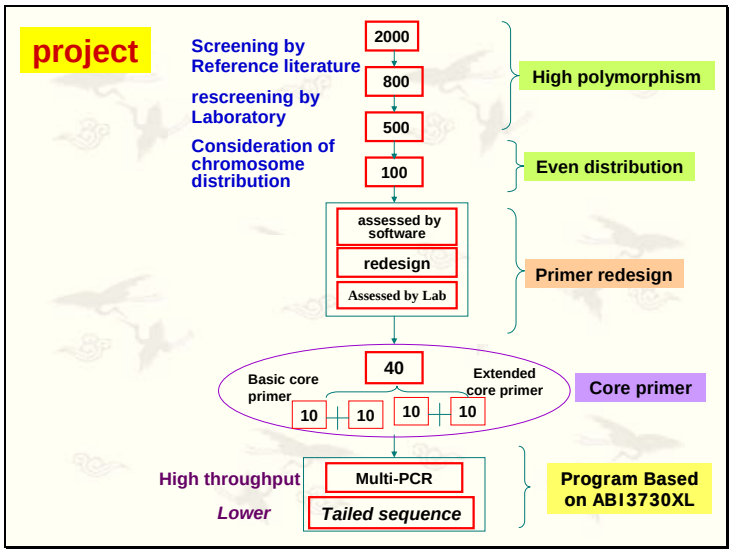
Slide 3

Purpose of the study

2. construct a high throughput multiplex PCR amplification system utilizing automated capillary electrophoresis fluorescence detection

- **The few recently published studies investigating multiplex PCR in maize were limited : research-specific, not universal ; typically ranging from two to four primers in a combination**
- **The work may significantly increase SSR analysis throughput and reduce costs associated with large-scale SSR genotyping**

Slide 4



Slide 5

method

- **1.Screening of primer polymorphism**
  - Reference literature screening
    - **MaizeGDB (http://www.maizegdb.org/ssr.php)；**
    - **CIMMYT（Xia et al, 2004; Marilyn et al, 2002）；**
    - **America(Smith et al, 1997；Matsuoka et al, 2002；Lu et al, 2001)；**
    - **France（Clerc et al, 2005）；**
    - **China（Li XH，2003；Teng WT，2004)；**
    - **Japan（Enoki et al, 2002）；**
    - **印度(Bantte et al, 2003)）**
  - Laboratory rescreening
    - **15 inbreds（for polymorphism）**
    - **15 hybrids（for heterozygosity ）**

Slide 6

2. Candidate primer chromosome distribution and determination of a set of universal core primers

- Download IBM2 2004 neighbors frame , and identify the location of the highly polymorphic candidate primers ;
- Plot the framework with a Marco program compiled by CAU.
- evenly distributed 100 primers, 10 per chromosome, selected as a set of universal core primers.

Slide 7

3.Primer redesign and construction of a universal PCR amplification program

- retrieve the original genomic sequences corresponding to candidate primer loci from NCBI, MaizeGDB and PlantGDB
- Primers redesigned and assessed utilizing software Primer Premier 5.0 and Oligo 6.22
- a two-step amplification program was constructed fitting for the newly designed primers.
- Nomenclature of newly designed primers at the same polymorphic site : name of the original primer + code of the designer + serial number

Slide 8

4. Construction of fluorescence multiplex PCR combination

- Design requirements
  - **Each primer set was comprised of 10 primers, one per chromosome**
  - **the combination pattern was 3+3+2+2: divided into four groups based on fragment length, Primers of the same group were labeled with the same fluorescence (FAM, NED, PET, or VIC). No overlap of amplification products of primers labeled with the same fluorescence.**
  - **Interaction between sets of primers was as weak as possible, △G<13**
- Multiplex PCR assessment of primers:
  - **significant primer interactions excluded by software Primer Premier 5.0 and by lab experiment**
- Program:
  - **PCR amplification: two-step amplification program, same as single primer**
  - **Electrophoresis detection: a DNA analyzer (ABI3730XL)**

Slide 9

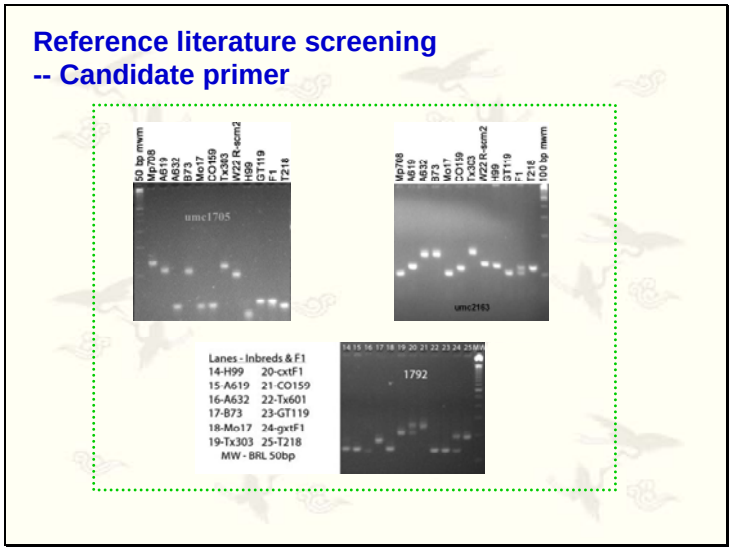
Result and analyses

Slide 10

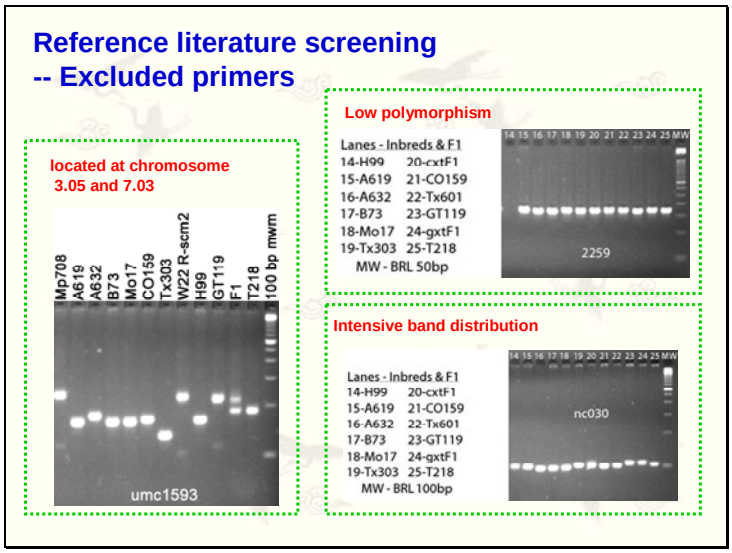
1. Determination of highly polymorphic primers

- A total of 500 highly polymorphic candidate primers, presenting with high levels of polymorphism were identified
- Excluded primers:
  - low polymorphism
  - intensive band distribution making statistical analysis difficult
  - discrepancy in the chromosome location

Slide 11

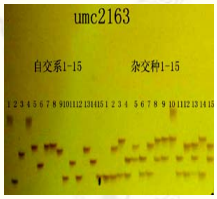


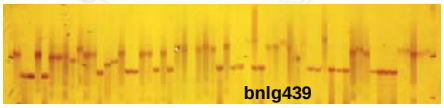
Slide 12

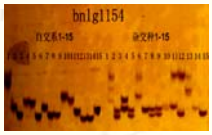


Slide 13

Laboratory rescreening
-- Candidate primer







Slide 14

2. Determination of a set of universal core primer

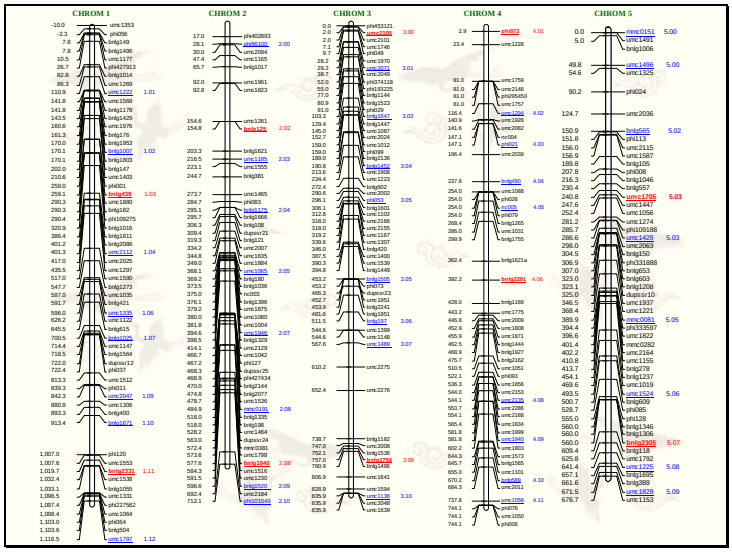
- A maize HP composite map plotted: used to demonstrate the chromosomal distribution of the 500 candidate primers and to provide a reference for core primer selection.
- A total of 100 evenly distributed primers, 10 from each chromosome, were selected as a set of universal core primers

Slide 15

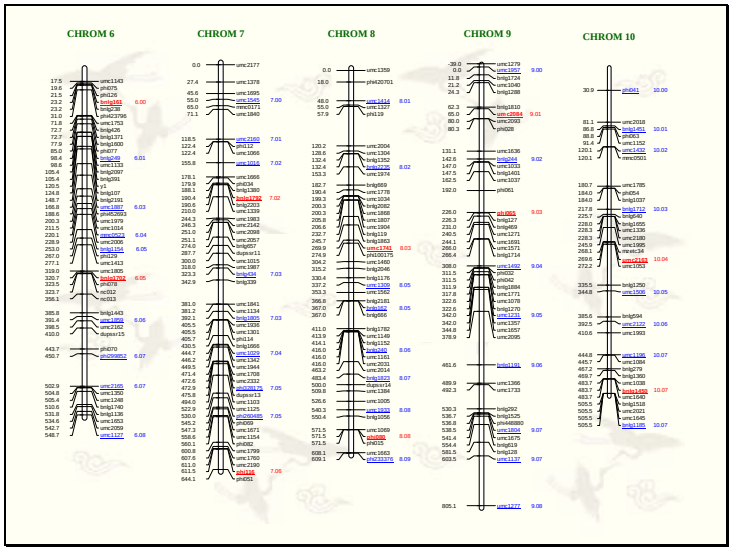
Information of 500 highly polymorphic candidate primers --based on IBM2 2004neighborsframe

Chrom	length (cM)	Number of candidate	Mean density	PIC(Min;Max,Aver)
1	1137	59	19.3	0.70;0.83;0.76
2	770	54	14.3	0.64;0.78;0.73
3	842	58	14.5	0.63;0.83;0.74
4	804	50	16.1	0.64;0.84;0.74
5	676	53	12.8	0.64;0.81;0.72
6	579	46	12.6	0.66;0.84;0.77
7	644	52	12.4	0.60;0.77;0.70
8	632	47	13.4	0.60;0.75;0.75
9	805	46	17.5	0.64;0.84;0.73
10	533	35	15.2	0.64;0.81;0.73
total	7422	500	14.8	0.64;0.81;0.74

Slide 16



Slide 17



Slide 18

3. Redesign of core primer sites and establishment of multiplex PCR combinations

Slide 19

original genomic sequences were downloaded on web

- 390 from NCBI
- 36 from Maize GDB
- 6 from Plant GDB

- total 432 of the 500 candidate primers downloaded. So primers could be redesigned for these sites
- remaining 68 were not found

Example:bnlg439 (see Annex I)

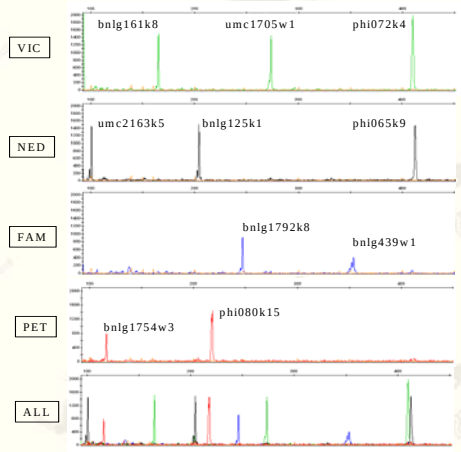
Slide 20

First set of ten-plex PCR combinations

- **ten-plex PCR combination was constructed in the 3+3+2+2 mode, the advantages is:**
  - **only one ten-plex PCR amplification and one electrophoresis to detect 10 primers on ABI3730XL**
  - **only utilized four PCR amplifications and four electrophoresis even if using a conventional denaturing PAGE system**
- redesigned primers showed significant improvements in every parameter compared with original primers: (see Annex II)
  - **primers had similar sequence features and were readily amplified under the same conditions**
  - **The amplification product with the expected size fitting for combination**
  - **weak primer interactions : Software and experimental assessments**

Slide 21

Profiling of the first ten-plex



Slide 22

Primers and combination modes of ten-plex PCR combination

First sets:

3: VIC			3: NED			2: FAM			2: PET		
bin	primer	range	bin	primer	range	bin	primer	range	bin	primer	range
4.01	phi072k4	413-433	9.03	phi065k9	399-419	1.03	bnlg439w1	319-385	9.03	phi080k15	203-233
5.03	umc1705w1	273-330	2.03	bnlg125k1	219-327	7.02	bnlg1792k8	192-250	2.03	bnlg1754w3	145
6.00	bnlg161k8	152-208	10.04	umc2163k5	154						

Second sets:

3: PET			3: NED			2: FAM			2: VIC		
bin	primer	range	bin	primer	range	bin	primer	range	bin	primer	range
1.11	bnlg2331k1	377-432	2.08	bnlg1940k9	407	10.07	umc2163k7	284-359	4.06	bnlg2291k4	384
3,00	umc21105k3	286-330	5.07	bnlg2305k4	292	8.04	phi080k15	203-233	6.05	bnlg1702k1	265
7.06	phi116k3	150-171	9.01	umc2084w2	188-211						

Slide 23

Slide 24

1. Polymorphism of SSR primer sites

polymorphism level of SSR primers is associated with the number and the type of repeat units, the genomic region, the database source for primer development and the materials used for detection

- positively correlated with the number of repeat units
- Uneven distribution of SSR primers throughout the genome
- generally higher in non-coding regions than in coding regions
- The evaluation greatly influenced by the material type and the range of materials selected

Slide 25

2. Primer selection criteria and multiplex PCR combination

Common Criteria

- high polymorphism;
- high amplification quality;
- even coverage of the entire genome;
- multiplex PCR amplification potential

- a stepwise screening strategy to obtain a set of universal SSR core primers
  - polymorphism screening
  - even distribution on the chrom
  - improve the amplification quality by redesign
  - ten-plex PCR primer combination

Slide 26

Other progress on the construction of DNA fingerprint database in maize

- Publish a standard "maize variety identification DNA fingerprint method"
- Construct a maize DNA fingerprint database with more than 2000 varieties.
- Apply the database in national regional trial to identify the uniformity and distinctness.
- Apply the database in more than 230 maize infringing cases entrusted by court. These DNA fingerprint was as the main basis for judgement.

Slide 27

Acknowledgements

- This work was supported by the Beijing Agricultural Innovative Platform-Beijing Natural Science Fund Program (No. YZPT02-06)

[Annex I follows]

ANNEX I

```
LOCUS       CC346513                968 bp    DNA     linear   GSS 16-MAY-2003
DEFINITION  OGIAN21TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0358C18,
            genomic survey sequence.
ACCESSION   CC346513
VERSION     CC346513.1  GI:30815920
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 968)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick
            ,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek
            ,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: methylation filtered.
FEATURES             Location/Qualifiers
     source          1..968
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0358C18"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN
        1 tcttcctggc aaacaataga aagttcagct tcgcccgtcc cggcctcggg ttctaaggct
       61 tccattgttc cgatcgcccg ggcaacgctg tactgctagc agtgcacatg agttagttgt
      121 tagtgttcct catccattac gccatttttc tgacaaaaaa aatgcagctt gtacagcaaa
      181 ctagtttcaa tcctgcatcg agccgccaag cctcgacctt gtctttagca aagatgtcaa
      241 gctgtcgaac aacaatttgg ctgacgtcgt gaaactgctg tttgcgagta gcaattctaa
      301 tcacagctgt taaggttaac cagcagaatt gagcatgtgt gtttttattc tgatttctga
      361 agcgcgtagc ggcttgcata gttactcact tactcgttgc tgatggtctt gtgcccttgt
```

421 acgaatcaga ccacggcctt tcaaactacc tctgcagtat gtaaccgtgg agacaggcac
481 ttccactctt tctctgaaat attccgacca ctcgatccag ctctccacag cagcagcata
541 gttgacatcg ccatcttggt gaccacaatt tacttttcag cgagcaaaga tcttggtact
601 ggtcgcaaca atggttaatt ccaggaagtt tcatctctgt ctgtctgcac ccgtttttc
661 tgactaatct ctctctctcc tctctctctc tctctctctc tctctaattt cattttcatc
721 accatgcttt ctcatttgaa tttccacaac ttcgtacgat cgcattaaga aacgctcacc
781 atggggcaca agcgctgcac aagaagaaga gccgtaagaa acgctcacga ccgacggggc
841 actcaagaga gacgggtccg gctgagctcg acaacccgct aagggcttgt tcgtcacact
901 ccaatccttc tgcgaccgac agaacccgaa atggagtcag ttatggaaat ggaaatggat
961 gtcgtatc
//

注:新设计引物用方框标注。上游 20bp(70-89):tacggtttcgtttcctttgg,下游 20bp(272-291):ctaacagcatcccttgagcc; PCR 产物长度：在 240-340 间，比原引物短约 80bp。

原引物用黄色底纹标注。上游 24bp（542-565）：TTGACATCGCCATCTTGGTGACCA，下游 29bp(742-770)：TCTTAATGCGATCGTACGAAGTTGTGGAA; PCR 产物长度：在 320-420 间。

重复序列(668-704)：TC

[Annex II follows]

ANNEX II

Comparison of original and newly designed primers

Primer name	Product length	TM value (U;L;P) [a)]	GC content (U ;L ;P)	initiation efficiency (U;P)	False initiation (U;P)	immer (U;L;UL)	hair pin(U;L)	3'-termina l ΔG
bnlg439	201-253	79.2;77.9;82.2	50;41.4;41.5	492;483	91;131	4;8;5	3;3	7.9;8.5
bnlg439w1	319-385	75;75.4;85.6	50;56.5;47.9	469;522	117;99	3;3;4	3;3	6.1;6.7
bnlg125	323-422	66.6;67.9;83.7	47.6;43.5;43.7	377;388	81;146	2;2;2	0;0	6.7;6.6
bnlg125k1	219-327	74.9;75.1;83.9	52.2;50.0;45.7	465;445	31;64	3;4;3	3;3	8.8;8.2
bnlg1754	215	68.5;68.6;87.9	50.0;55.0;55.8	450;401	86;136	4;6;3	0;3	8.8;8.2
bnlg1754w3	145	75.4;75.9;88.0	61.9;63.6;56.6	447;431	88;58	6;3;3	0;3	8.4;8.2
phi072	142-162	78.6;78;78	50;42.9;34.5	533;484	103;113	6;6;4	3;0	6.4;9.7
phi072k4	413-433	74.3;;75.2;83.7	66.7;54.5;42.9	432;463	69;90	3;4;3	3;3	8.2;9.4
umc1705	57-114	70.7;76.4;82.9	50;50;49.5	411;480	193;81	4;6;3	0;3	6.4;6.4
umc1705w1	273-330	73.9;74.4;87.4	59.1;54.5;52.6	425;464	98;92	2;6;3	0;0	8.4;6.7
bnlg161	129-185	70.6;70.4;82.7	41.7;39.1;45.1	379;410	115;82	2;4;4	0;3	6.9;6.3
bnlg161k8	152-208	75.2;74.8;83.3	48.0;54.5;45.4	462;428	38;111	4;4;5	3;3	9.0;8.5
bnlg1792	113-171	68.0;68.4;81.9	50.0;45.0;44.4	431;444	120;97	4;2;4	0;0	6.4;7.0
bnlg1792k8	192-250	73.1;72.8;84.4	57.9;56.5;47.2	471;426	0;65	4;2;4	0;3	9.4;6.9
phi080	243-273	75.9;75.1;87.9	54.5;54.5;58	464;407	78;69	4;4;3	3;4	5.5;6.3
phi080k15	203-233	74.9;74.0;89.4	52.4;52.2;59.4	458;449	71;71	4;4;3	0;3	6.7;5.5
phi065	132-153	70.8;72; 83.6	50;50;47	379;415	33;76	4;4;3	3;3	5.5;6.7
phi065k9	399-419	75.2;75.7;88.4	50.0;66.7;54.2	497;436	94;88	4;3;3	4;3	9.4;9.4
umc2163	145	72.2;72.1;81.7	41.7;41.7;43.4	496;462	86;83	3;5;3	3;5	6.7;7.0
umc2163k5	154	73.6;73.4;82.5	50.0;55.0;44.8	495;462	83;63	4;4;4	3;3	6.6;7.9

a) U: forward primer sequence; L: reverse primer sequence; P: amplification products.

[End of document]