E

BMT/13/19 Add.
ORIGINAL: English
DATE: December 8, 2011

UPOV

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

WORKING GROUP ON BIOCHEMICAL AND MOLECULAR TECHNIQUES, AND DNA-PROFILING IN PARTICULAR

Thirteenth Session
Brasilia, November 22 to 24, 2011

ADDENDUM

USE OF MOLECULAR MARKERS
FOR INFRINGEMENT DETECTION IN HYBRID CROPS

Document prepared by experts from the Monsanto Company

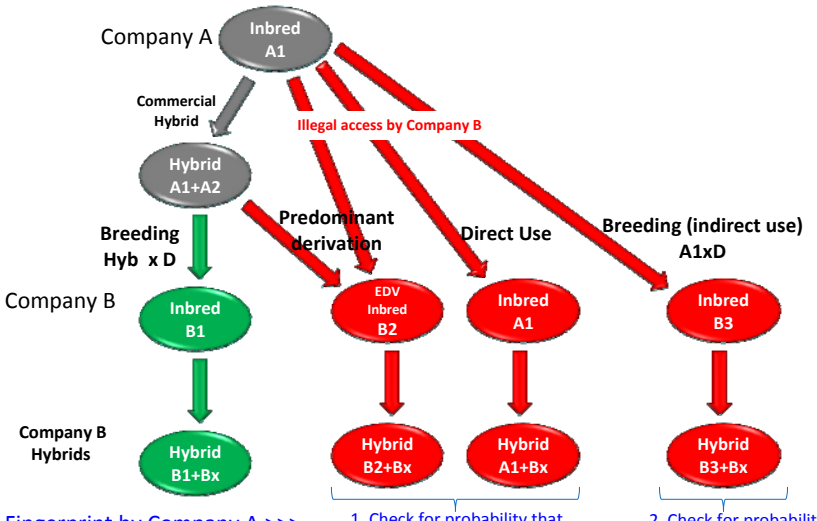
Use of Molecular Markers for Misuse Detection in Hybrid Crops

Rogerio Andrade
Monsanto Company
Brazil Corn Line Development Breeding Lead

Introduction

- Focus on hybrid crops,
  - Common misconception: *Misuse detection options are limited when considering hybrids.*
- Most common types of germplasm misuse:
  - **Direct use;** misappropriated inbred parent is used to produce a hybrid,
  - **Predominant derivation;** slightly different 'copy' of parental line is developed and used in hybrid combination,
  - **Derivation from illegaly accessed source;** parental line is illegally accessed to start new breeding population, resulting in a new inbred progeny.

Exemples of Misuse of germplasm in hybrid crops



UPOV-BMT meeting_Brasilia, Nov-22, 2011

Complexity of hybrid genotyping

- In hybrids, ~ 40-60% of the markers are heterozygotes, making fingerprint analysis and misuse detection more complex.
- Hybrids can be more complex than just single crosses (three-way cross and double-cross hybrids).
- In most countries, inbred parents are not available for DNA fingerprinting, and it would rely on fingerprinting hybrids.

UPOV-BMT meeting_Brasilia, Nov-22, 2011

1. Detection of Direct use or Predominant derivation:

Hybrid fingerprint is tested for suspected parentage relationship with proprietary inbreds.

Hybrid Fingerprint

Marker set

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
Hybrid X	AA	AT	GG	CG	AT	CG	CC	AA	CC	CG

Estimate probability of a particular inbred being present in the hybrid combination

Fingerprint database of proprietary inbreds

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
Inbred A	AA	AA	GG	GG	TT	CC	CC	AA	CC	GG
Inbred B	TT	TT	CC	CC	AA	GG	CC	TT	CC	CC
Inbred C	AA	AA	CC	CC	AA	CC	GG	TT	GG	GG
Inbred D	AA	TT	CC	GG	TT	CC	GG	AA	GG	GG
Inbred E	TT	AA	GG	GG	TT	GG	GG	TT	CC	GG

- **Test parentage hypothesis .** By using thousands of SNP markers, it is possible to establish parentage relationship with high probability (Infinium 50 K SNP markers chip is publically available)

- Algorithms are needed to screen large number of potential permutations (Multiple hybrids by multiple proprietary inbreds, at thousands of marker loci)

UPOV-BMT meeting_Brasilia, Nov-22, 2011

Use of haplotype sequences as a proof of derivation

DNA recombination «Cross-overs» occur during derivation-breeding.
Such Line Specific Recombination (LSR) events provide unique haplotypes that can be identified with confidence when comparing high-density fingerprinting profiles of progeny lines versus their parents.

Parents P1 X P2

Breeding process

Progeny: A1

3 Cross-overs = LSR-haplotypes

If A1 is illegally accessed & used for breeding purpose

A1 X D1

Breeding

B3

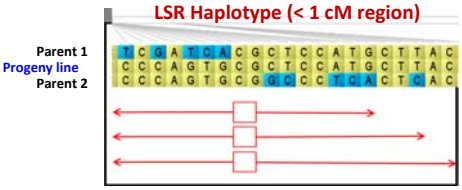
A1 is crossed with D1 to develop a segregating pop.
For breeding purpose.
New cross-overs occur during breeding process…

After derivation, a subset of the LSR-haplotypes from parent A1 will still remain in new progeny B3.
Even if B3 has low Sim% with A1, derivation act could be confirmed by looking at presence of LSR haplotypes.

LSR-haplotypes
as inventor's encrypted «signature»

UPOV-BMT meeting_Brasilia, Nov-22, 2011

LSR-haplotypes as efficient IP protection tool



- Generally, LSR haplotypes have 3 components:
1. Monomorphic region
2. Left flanking set of markers
3. Right flanking set of markers
- Length of these flanking regions is optimized for strong specificity, trying to keep the region small (<1 cM)

- Some 20-25 LSR haplotypes could be identified for an individual inbred line.

Inbred	Pedigree	LSR-ID#	Chr.	Pos.	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16
INBRED A1	P1 x P2	LSR-A01	1	118.8 - 119.2	X	0	0	0	0	0	X	0	0	0	X	0	X	0	0	0
		LSR-A02	2	119.6 - 119.7	X	0	0	0	0	0	X	0	0	0	X	0	X	0	0	0
		LSR-A03	2	120.0 - 120.1	X	0	0	0	0	0	X	0	0	0	X	0	X	0	0	0
		LSR-A04	3	156.6 - 158.8	X	0	0	0	0	0	X	0	0	X	X	0	X	0	0	0
		LSR-A05	5	134.5 - 135.4	X	0	0	0	0	0	0	0	X	X	0	X	0	X	X	X
		LSR-A06	7	58.2 - 59.3	X	0	0	0	0	0	X	X	0	0	0	X	0	0	X	X
		LSR-A07	7	117.9 - 118.8	X	0	0	0	0	0	X	0	X	0	X	X	X	0	X	0
		LSR-A08	9	133.2 - 134.1	X	0	0	0	0	0	0	0	X	0	0	X	0	X	0	0
						Non related lines					Cycle 1 derivation						Cycle 2 derivation			

- A set of 15-20 LSR haplotypes, evenly distributed accross the genome, is sufficient for accurate tracking of parental origin.

2. Detection of indirect use (unauthorized access + derivation):

Objective: Check hybrids for presence of particular proprietary LSR haplotypes

Proprietary Fingerprinting and
LSR haplotypes Database

INBRED A	LSRA1	LSRA2	LSRA3	LSRA4	LSRA5	LSRA6	LSRA7	LSRA8	LSRA9	LSRA10	LSRA11	LSRA12	LSRA13	LSRA14						
INBRED B	LSRB1	LSRB2	LSRB3	LSRB4	LSRB5	LSRB6	LSRB7	LSRB8	LSRB9	LSRB10	LSRB11	LSRB12	LSRB13	LSRB14	LSRB15	LSRB16				
INBRED C	LSRC1	LSRC2	LSRC3	LSRC4	LSRC5	LSRC6	LSRC7	LSRC8	LSRC9	LSRC10	LSRC11	LSRC12	LSRC13	LSRC14	LSRC15	LSRC16	LSRC17	LSRC18		
INBRED D	LSRD1	LSRD2	LSRD3	LSRD4	LSRD5	LSRD6	LSRD7	LSRD8	LSRD9	LSRD10	LSRD11	LSRD12	LSRD13	LSRD14	LSRD15					

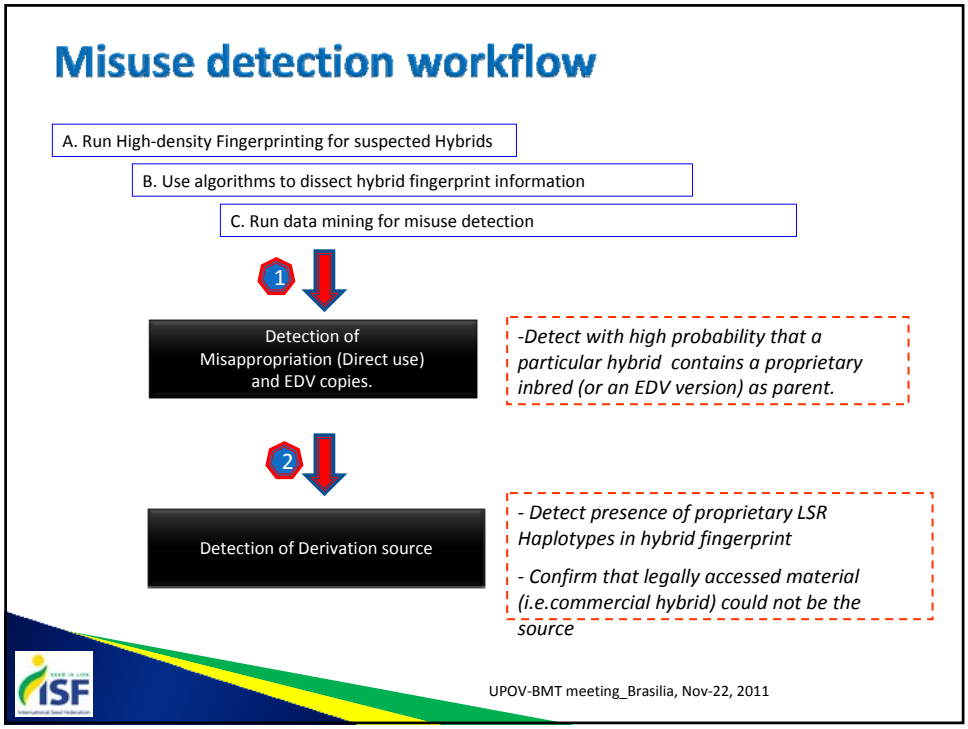
Fingerprint of LSR-A1 haplotype

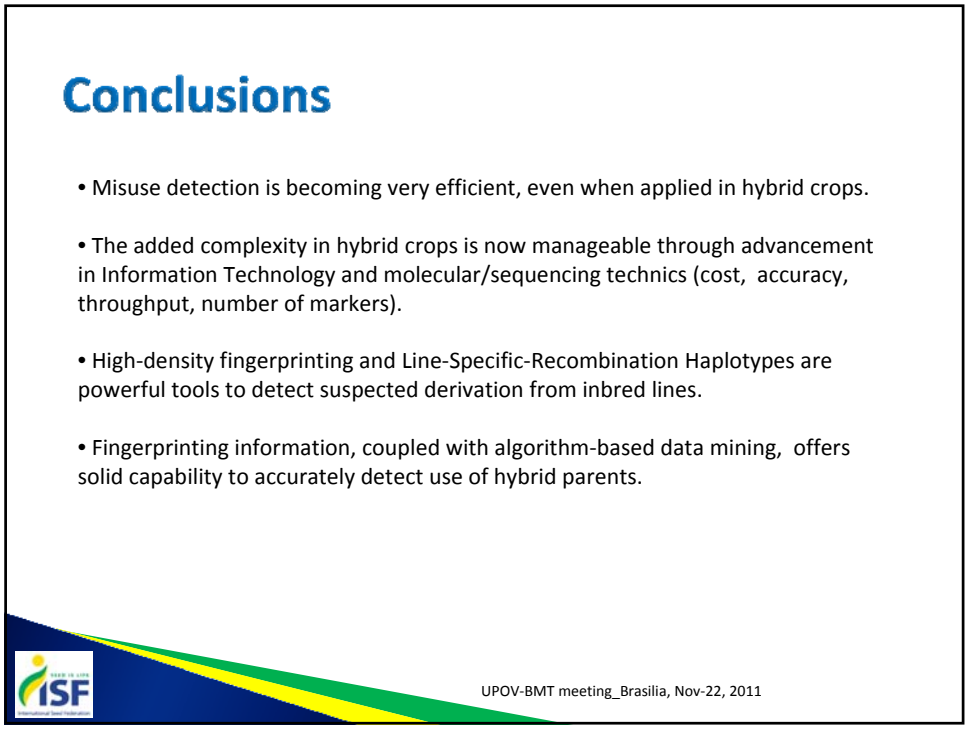
		M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20
Inbred A	LSR-A1	AA	AA	TT	GG	GG	AA	CC	CC	AA	TT	TT	AA	CC	CC	GG	CC	TT	TT	AA	GG

High-density Fingerprinting of Hybrids (50 K SNP markers):
>> Algorithm to asses fit between hybrid sequence and a particular LSR haplotype.

Hybrid Name	Versus	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20
HYBRID X	LSR-A1	AT	TT	TT	GG	GC	TT	GC	CC	AA	AT	AT	AA	GC	CC	GC	GG	TT	AA	TT	GC
HYBRID Y	LSR-A1	AT	AA	TT	GC	GC	AA	GC	CC	AA	AT	AT	AA	GC	CC	GC	CC	AT	TT	AA	GC
HYBRID Z	LSR-A1	TT	AA	AA	CC	GC	AA	GC	GG	TT	AT	AT	TT	GG	CC	GC	GG	AA	AA	AA	CC

>> Detect suspected presence of some LSR haplotypes in hybrid fingerprint
- with probability assigned to each LSR window
- with probability assigned to a derivation act, when fit is ascertained at multiple LSR windows.





[End of document]