E

BMT/11/10 Rev. Add.
ORIGINAL: English
DATE: October 9, 2008

UPOV

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

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

Eleventh Session
Madrid, September 16 to 18, 2008

ADDENDUM

CONSTRUCTION OF AN INTEGRATED MICROSATELLITE AND KEY MORPHOLOGICAL CHARACTERISTIC DATABASE OF POTATO VARIETIES ON THE EU COMMON CATALOGUE
PART 2: THE DATABASE (REVISED)

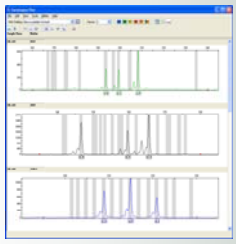
Document prepared by experts from the United Kingdom

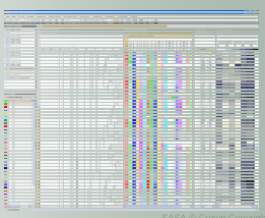
SASA

CONSTRUCTION OF AN INTEGRATED MICROSATELLITE AND KEY MORPHOLOGICAL CHARACTERISTIC DATABASE OF POTATO VARIETIES ON THE EU COMMON CATALOGUE

PART 2: THE DATABASE

Alex Reid & Lysbeth Hof





SASA

Introduction

During the course of the project a large amount of different types of data were generated

Binary data for the 9 SSR markers

Multi-state data for the morphological characteristics

Textual data for the information about each sample

Lightsprout photographs

Requirements for the software

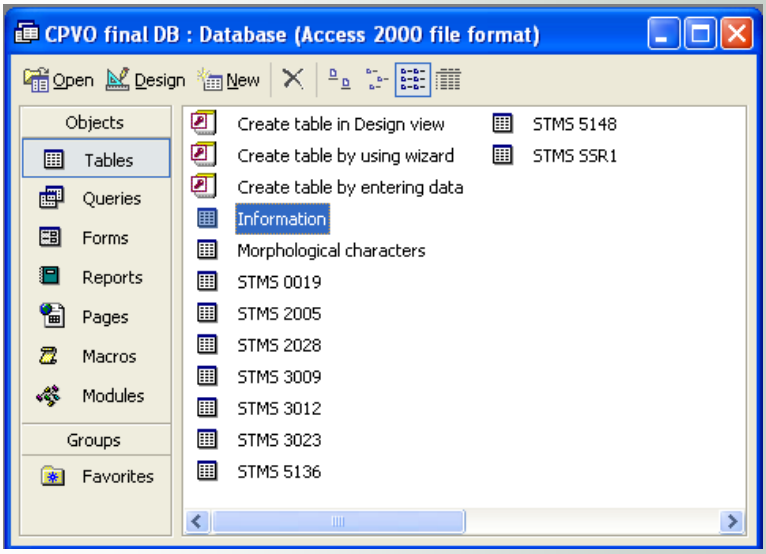
The software had to be able to store and analyse different types of data

It had to be able to analyse these data in different combinations

A highly desirable feature was the ability to use the database to identify unknown samples rapidly and accurately

The solution was to store the data in an Access database which was linked to BioNumerics via an ODBC (Open DataBase Connectivity) for analysis

Access database structure



The Information Table

The structure of the Access database was specifically designed so that it could serve BioNumerics

BioNumerics requires each entry to have a unique identifier or Key

The variety name could not be used for the DNA samples as some varieties had samples from numerous sources

For example cv. Ditta had samples submitted for SSR analysis from the Netherlands, Poland and the UK and therefore the DNA sample codes were used as the Key (in this case NL-089, PL-021 and UK-0598)

There is only one entry per country for the morphological data for a variety and in these cases the Key is the variety name followed by the country of origin (e.g. Ditta_DE, Ditta_NL, Ditta_PL and Ditta_UK)

The Information Table Fields

Key	Unique identifier for use in BioNumerics
Variety Denomination	The name of the variety
Origin of sample e.g. Breeder (B), office (O) or other (T)	Where the sample was obtained from either the breeder or from one of the partners collections or from another source
Submitting office (S SASA, B BSA, C COBORU, N Naktuinbouw)	The office which submitted the sample for DNA analysis or submitted the morphological description
Harvest year	The year the sample analyzed was harvested

The Information Table Fields

DNA extraction laboratory (S SASA, N Naktuinbouw)	The laboratory the DNA sample was extracted (either UK or Netherlands)
Extraction year	The year the DNA extraction was made
Place of storage of DNA sample (S SASA, N Naktuinbouw)	The place where the DNA sample is kept in long term storage
SSR analysis performed at (S SASA, N Naktuinbouw)	The laboratory where the SSR analysis was performed
SSR analysis year	The year the SSR analysis was performed

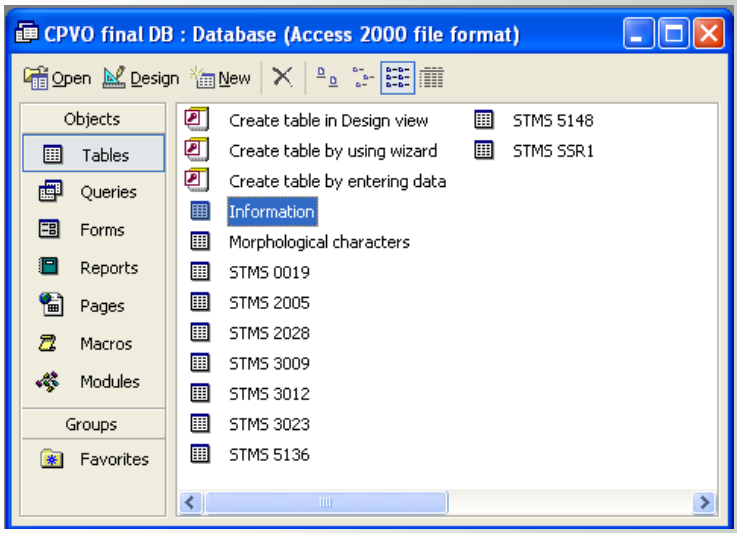
The Information Table Fields

Technical protocol used for morphological description	The technical protocol used for the morphological description (if known)
Description year (either before 1995 or actual year)	The year that the latest description was carried out, N.B. not necessarily from official description
Place description carried out	The office the description was carried out at
Photograph availability (and link to photograph)	The file name for the photograph (if available) and hyperlink

The Information Table Fields

Place photograph taken (office)	The office the photograph was taken at
Photograph year (if known)	The year the photograph was taken
National Reference	The national reference number of the sample (if there is one)
Comments	Any other comments (e.g. is the variety a mutant of another variety)

Access database structure

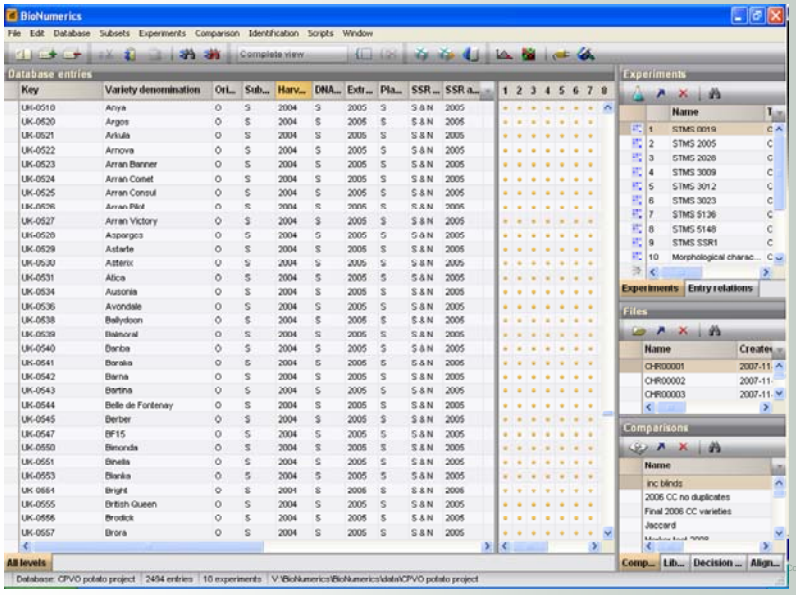


Examples of morphological and SSR data

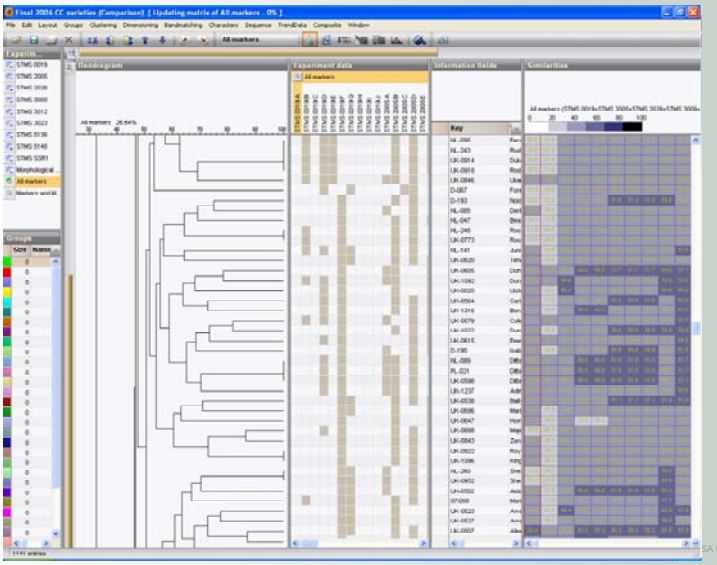
Key	Variety	1	2	3	4	5	6	7	8	9	10	11	24	28	29	31	32	34	35	36
Av_NL	Av	5	2	5	1	5	4	4	2	5	4	4	5	4	1	5	3	2	2	3
Ave_NL	Ave	7	2	7	3	4	6	3	3	5	4	5	4	1		8	2	2	2	1
Ax_DE	Ax	5	4	5	1	3	4	3	1	2	6	3								
Ba_DE	Ba	3	3	4	2	5	5	3	2	3	5	4								
Bal_NL	Bal	6	2	4	1	3	5	3	3	3	3	3	5	1		7	4	2	2	3

Key	Variety	A	B	C	D	E	F	G	H	I	J
UK-0523	Arran B	0	0	0	0	0	1	1	0	0	0
UK-0524	Arran C	0	1	0	0	0	1	1	0	0	0
UK-0525	Arran C	0	0	0	0	0	1	1	0	0	0
UK-0526	Arran P	0	1	0	1	0	1	0	0	0	0
UK-0527	Arran V	0	0	0	0	0	1	1	0	0	0

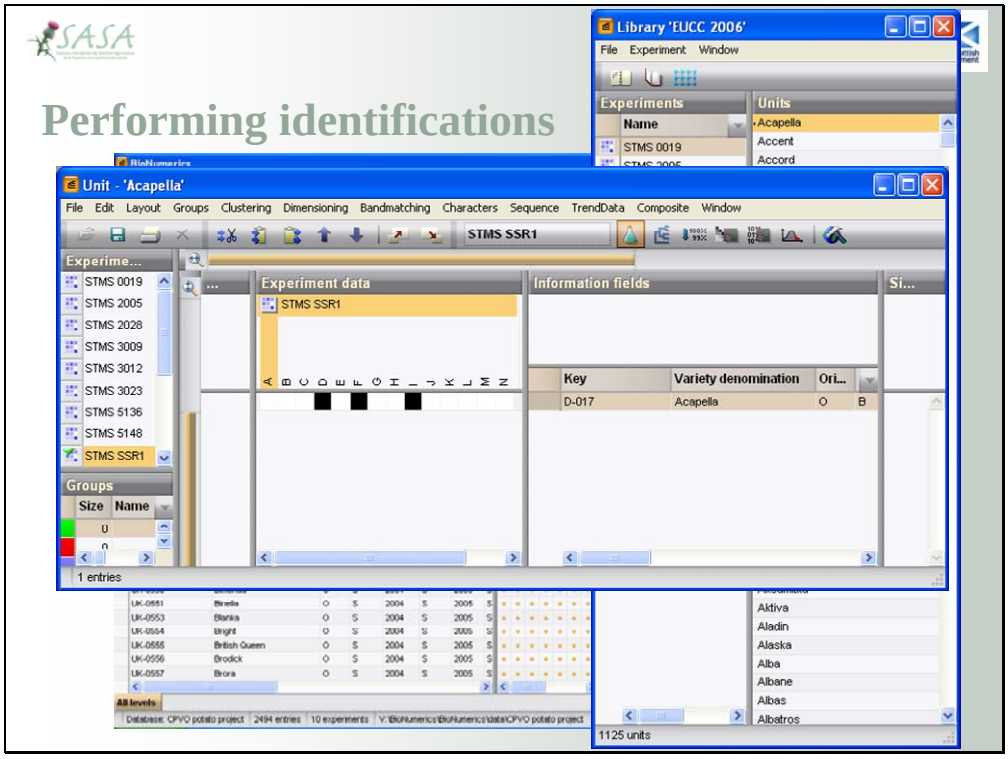
BioNumerics database

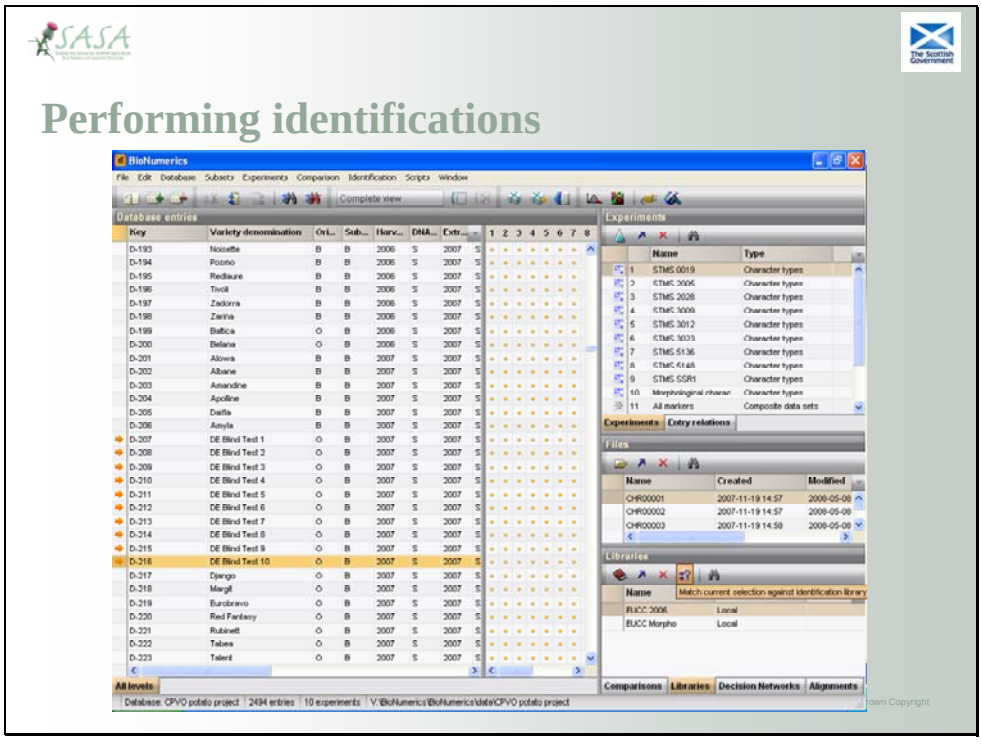


SSR data tree

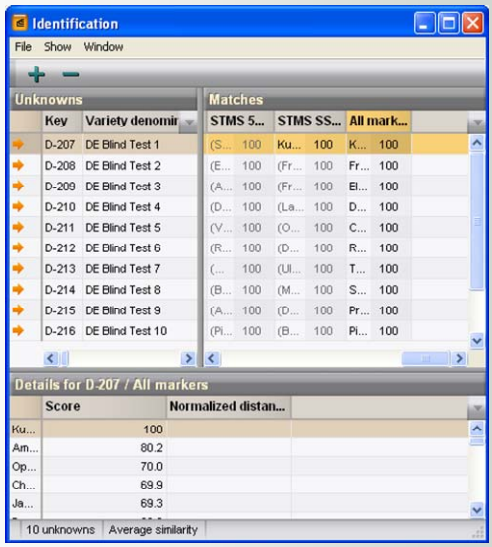


Morphological data tree


Performing identifications
Library 'EUCC 2006'
File Experiment Window
Experiments
Name
STMS 0019
Units
Acapella
Accent
Accord
Unit - 'Acapella'
File Edit Layout Groups Clustering Dimensioning Bandmatching Characters Sequence TrendData Composite Window
STMS SSR1
Experiment data
Information fields
Key
Variety denomination
D-017
Acapella
Groups
Size
Name
1 entries
Aktiva
Aladin
Alaska
Alba
Albane
Albas
Albatros
1125 units


Performing identifications
BioNumerics
Database entries
Experiments
Files
Libraries
Comparisons
Libraries
Decision Networks
Alignments

Library screen identification report



Practical uses

At SASA we now perform 100s of variety tests every year for

- Official field inspectors
- Breeders (national and international)
- Regulatory bodies

Cases have included

- Verification of varieties for export (when there have been queries)
- Labelling infringements in supermarkets
- Illegal dumping of potatoes in river systems
- Cannabis smuggling!

Summary

During the course of the project a large amount of different types of data were generated

These data are held in an Access database

Which is linked to BioNumerics via ODBC

Clustering analysis can be performed on single SSR markers or the morphological data or any combination of these

A library was created against which unknown samples can be screened resulting in rapid and accurate identifications

Without the help of...

Project partners

BSA - Beate Rücker and SwenjaTams

COBORU - Ewa Milczynska and Julia Borys

CPVO - Anne Weitz

SASA – Heather Campbell, Arlene Cameron and Sandra Goodfellow

None of this would have been possible!

[End of document]