UPOV
International Union for the Protection of New Varieties of Plants

E

Technical Working Party on Testing Methods and Techniques **TWM/4/16**

Fourth Session
Cambridge, United Kingdom, June 2 to 5, 2026

Original: English
Date: May 11, 2026

DUS PHENO/GENO DATABASE

Document prepared by an expert from the International Seed Federation (ISF)

Disclaimer: this document does not represent UPOV policies or guidance

The annex to this document contains a copy of a presentation “DUS Pheno/Geno database”, to be made by an expert from the International Seed Federation (ISF), at the fourth session of the TWM.

[Annex follows]

1

Goal

Provide a **database framework** for UPOV and PBR offices to store and share data related with **DUS information** that covers simultaneously morphological characteristics (Phenotypes) and molecular information (Genotypes) for **plant varieties**.

2

Momentum from TWM/3

Many great proposals were presented in Beijing, including TWM/3/23

The atmosphere allowed for a spirit of collaboration

Profit from the momentum and turn into deliverables

Possibility to implement the “confidentiality of molecular information” (TWM/9/4) through server authentication

3

Concept

Simple:

- Minimalist design covering what’s necessary

Data Agnostic:

- Species, Authority, Marker Type etc

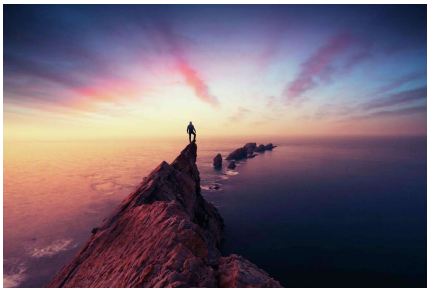
Robust:

- Well defined keys that contemplate the foundation

Flexible:

- Allows for changes in data without interruptions – version changes in TGs

Big Picture



Secure:

- Enhance data security by moving away of Excel and PDF files

4

Scope

Purely a database to store and manage molecular data (**genotypes**) and morphological characteristics (**phenotypes**) related with DUS ➔ no software analysis

Data integrity ➔ genotypic concordance

Reciprocal imputations between genotypes and phenotypes ➔ to serve the current requirement of association between markers and morphological characteristics

Robust ➔ following technical guidelines and already harmonized terminology under UPOV documentation and systems

5

Scope

Data integrity ➔ genotypic concordance

Variety	RST-24	Authority
Alpha	**A/C**	A
Beta	A/C	A
Gamma	C/C	A

Curation

Variety	RST-24	Authority
Alpha	**A/A**	B
Beta	A/C	B
Gamma	C/C	B

Variety	RST-24	Authority
Alpha	--	UPOV
Beta	A/C	UPOV
Gamma	C/C	UPOV

6

Scope

Reciprocal imputations between genotypes and phenotypes ➔ to serve the current requirement of association between markers and morphological characteristics

Variety	Marker 12	Petal Color
Alpha	A/A	White
Beta	--	Pink
Gamma	C/C	--

Imputation

Variety	Marker 12	Petal Color
Alpha	A/A	White
Beta	*A/C*	Pink
Gamma	C/C	*Red*

7

Scenarios and implications

Example Authority A

Databased built to fit specific requirements limited to vegetable crops

Example Authority B

Databased built to fit specific requirements limited to flowers and other asexually propagated crops

Example Authority C

Databased built to fit specific requirements limited to agricultural crops

8

Scenarios and implications

Example Authority A

Stores SNPs only as:
A/A
C/C
G/G
T/T
Uses **alpha characters** and backslash is the compulsory **delimiter**

Example Authority B

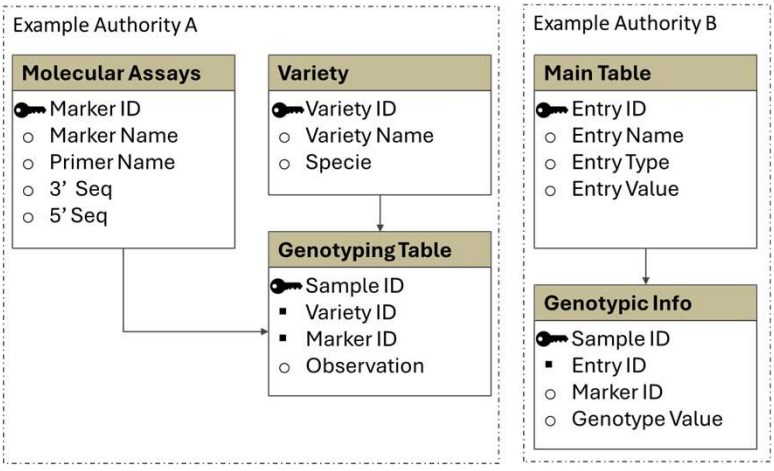
Stores SNPs only as:
AA
CC
GG
TT
Uses **alpha characters** and does not support delimiters

Example Authority C

Stores SNPs only as:
1,1
2,2
3,3
4,4
Uses **numeric characters** and comma is the compulsory **delimiter**

9

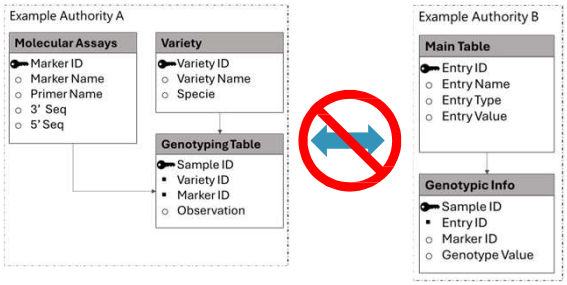
Scenarios and implications



10

Dataflow and communication

How do we store information?



What do we want to share?

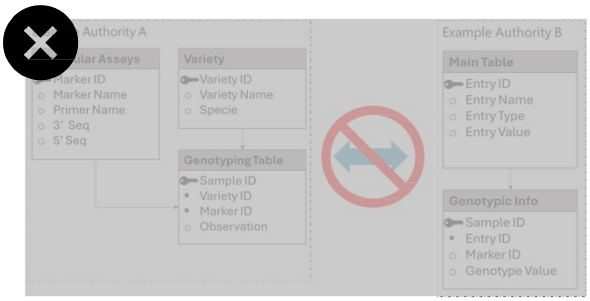
Variety	RST-24	RSTV Virus
Alpha	A/A	Resistant
Beta	A/C	Resistant
Gamma	C/C	Susceptible

Why are those not compatible?

11

Dataflow and communication

How do we store information?



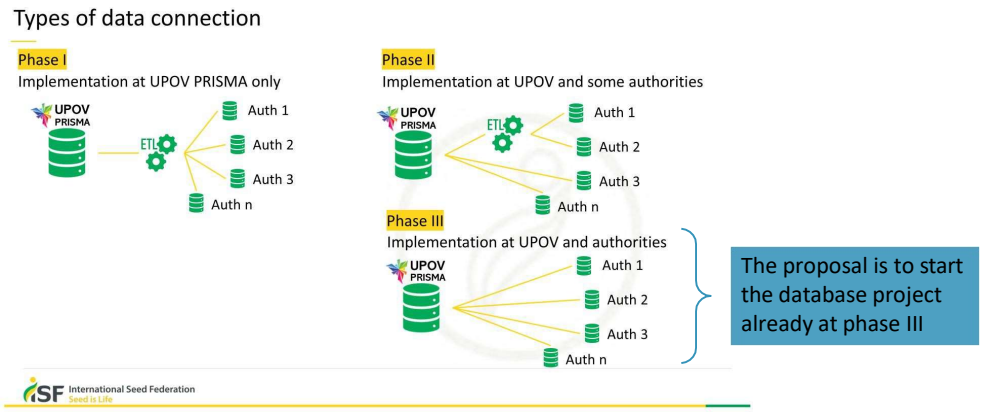
What do we want to share?

Variety	RST-24	RSTV Virus
Alpha	A/A	Resistant
Beta	A/C	Resistant
Gamma	C/C	Susceptible

Reposition ourselves to a generic and common perspective

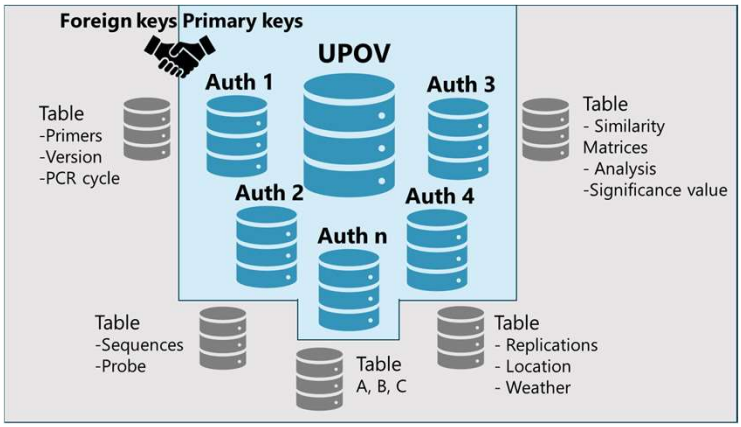
12

UPOV PRISMA – similar concept



13

Central database and satellite databases



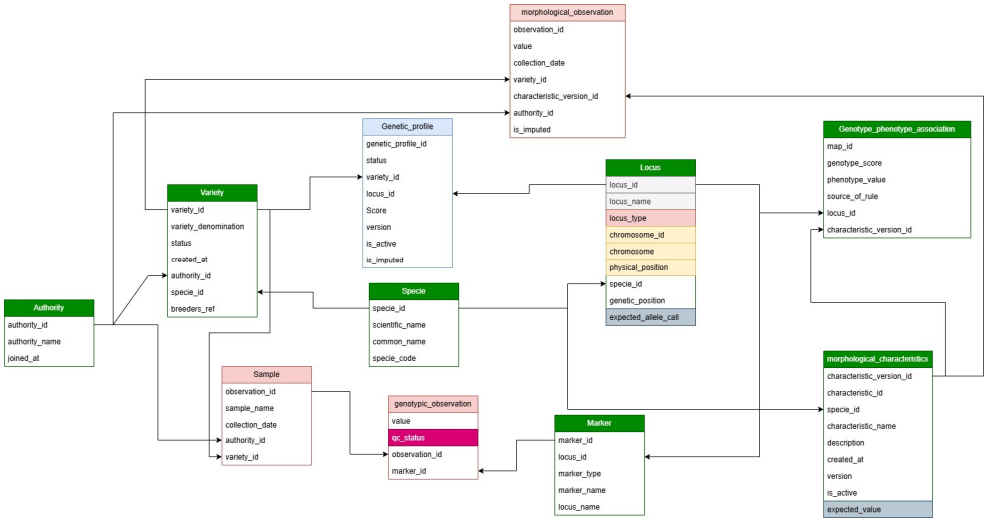
The design allows authorities customized tables and attributes without detriment to the core data and data connection.

Mutual tables, attributes, values and types.

Independent tables, attributes, values and types to accommodate specific needs that are not shared.

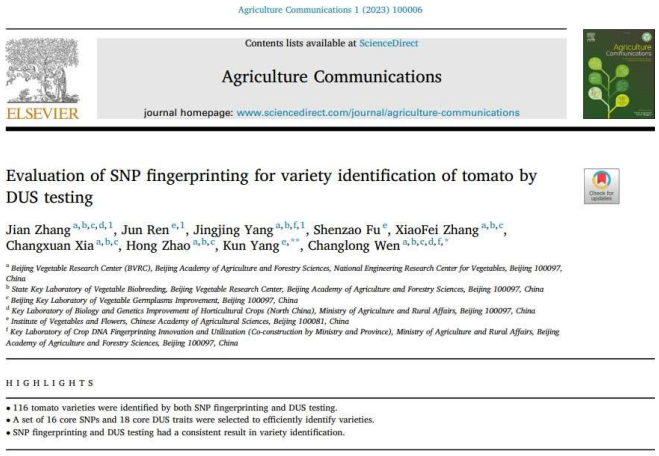
14

Architecture



15

Demonstration



Agriculture Communications 1 (2023) 100006

Contents lists available at ScienceDirect

Agriculture Communications

journal homepage: www.sciencedirect.com/journal/agriculture-communications

Evaluation of SNP fingerprinting for variety identification of tomato by DUS testing

Jian Zhang [a,b,c,d,1], Jun Ren [e,1], Jingjing Yang [a,b,f,1], Shenzao Fu [e], XiaoFei Zhang [a,b,c], Changxuan Xia [a,b,c], Hong Zhao [a,b,c], Kun Yang [e,**], Changlong Wen [a,b,c,d,f,*]

[a] Beijing Vegetable Research Center (BVRC), Beijing Academy of Agriculture and Forestry Sciences, National Engineering Research Center for Vegetables, Beijing 100097, China
[b] State Key Laboratory of Vegetable Biobreeding, Beijing Vegetable Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China
[c] Beijing Key Laboratory of Vegetable Germplasms Improvement, Beijing 100097, China
[d] Key Laboratory of Biology and Genetics Improvement of Horticultural Crops (North China), Ministry of Agriculture and Rural Affairs, Beijing 100097, China
[e] Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing 100081, China
[f] Key Laboratory of Crop DNA Fingerprinting Innovation and Utilization (Co-construction by Ministry and Province), Ministry of Agriculture and Rural Affairs, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China

HIGHLIGHTS

- 116 tomato varieties were identified by both SNP fingerprinting and DUS testing.
- A set of 16 core SNPs and 18 core DUS traits were selected to efficiently identify varieties.
- SNP fingerprinting and DUS testing had a consistent result in variety identification.

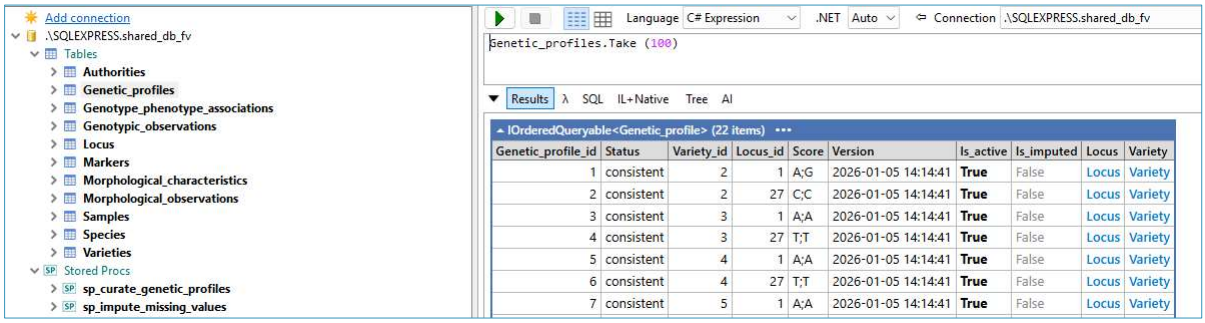
The data used is from the publication from Zhang et al., 2023

Genotype and phenotypes were randomly assigned and do not represent the real state of the varieties, because that data was not shared.

16

Demonstration

Consolidated genetic profile

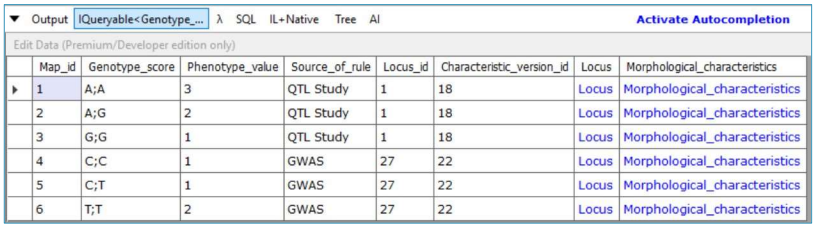


Genetic_profile_id	Status	Variety_id	Locus_id	Score	Version	Is_active	Is_imputed	Locus	Variety
1	consistent	2	1	A;G	2026-01-05 14:14:41	True	False	Locus	Variety
2	consistent	2	27	C;C	2026-01-05 14:14:41	True	False	Locus	Variety
3	consistent	3	1	A;A	2026-01-05 14:14:41	True	False	Locus	Variety
4	consistent	3	27	T;T	2026-01-05 14:14:41	True	False	Locus	Variety
5	consistent	4	1	A;A	2026-01-05 14:14:41	True	False	Locus	Variety
6	consistent	4	27	T;T	2026-01-05 14:14:41	True	False	Locus	Variety
7	consistent	5	1	A;A	2026-01-05 14:14:41	True	False	Locus	Variety

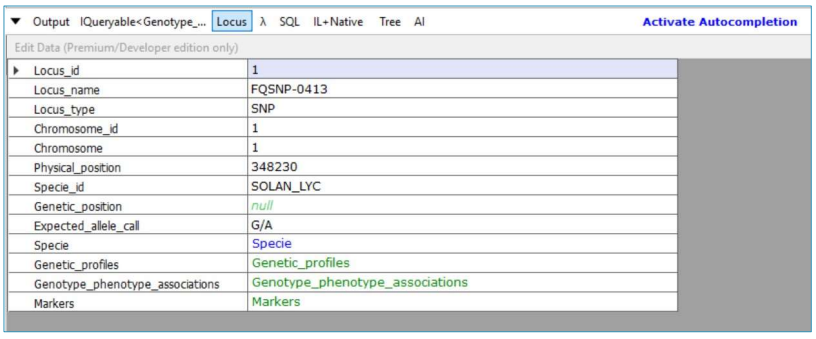
17

Demonstration

Phenotypic and genotypic annotation



Map_id	Genotype_score	Phenotype_value	Source_of_rule	Locus_id	Characteristic_version_id	Locus	Morphological_characteristics
1	A;A	3	QTL Study	1	18	Locus	Morphological_characteristics
2	A;G	2	QTL Study	1	18	Locus	Morphological_characteristics
3	G;G	1	QTL Study	1	18	Locus	Morphological_characteristics
4	C;C	1	GWAS	27	22	Locus	Morphological_characteristics
5	C;T	1	GWAS	27	22	Locus	Morphological_characteristics
6	T;T	2	GWAS	27	22	Locus	Morphological_characteristics



Locus_id	1
Locus_name	FQSNP-0413
Locus_type	SNP
Chromosome_id	1
Chromosome	1
Physical_position	348230
Specie_id	SOLAN_LYC
Genetic_position	null
Expected_allele_call	G/A
Specie	Specie
Genetic_profiles	Genetic_profiles
Genotype_phenotype_associations	Genotype_phenotype_associations
Markers	Markers

18

Demonstration

Phenotypic or genotypic imputations

IOrderedQueryable<Morphological_observation> (22 items)

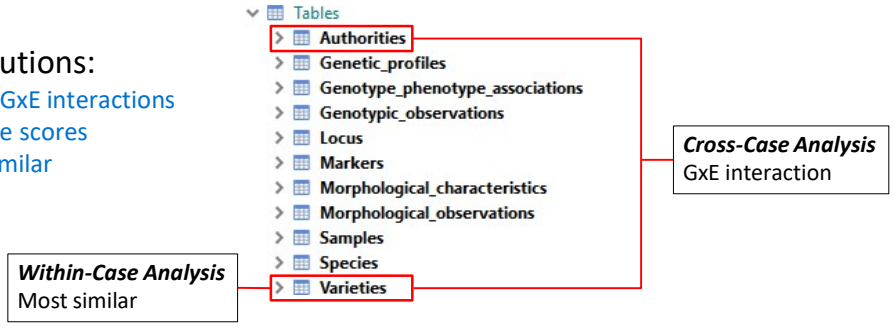
Observation_id	Value	Collection_date	Variety_id	Characteristic_version_id	Authority_id	Is_imputed	Authority	Morphological_characteristics	Variety
1	1	6/15/2025	1	18	CHN	False	Authority	Morphological_characteristics	Variety
2	1	6/15/2025	1	22	CHN	False	Authority	Morphological_characteristics	Variety
3	2	6/15/2025	2	18	CHN	False	Authority	Morphological_characteristics	Variety
4	1	6/15/2025	2	22	CHN	False	Authority	Morphological_characteristics	Variety
5	3	6/18/2025	3	18	CHN	False	Authority	Morphological_characteristics	Variety
6	2	6/18/2025	3	22	CHN	False	Authority	Morphological_characteristics	Variety
7	3	6/18/2025	4	18	CHN	False	Authority	Morphological_characteristics	Variety
8	2	6/18/2025	4	22	CHN	False	Authority	Morphological_characteristics	Variety
9	3	7/10/2025	5	18	CHN	False	Authority	Morphological_characteristics	Variety
10	2	7/10/2025	5	22	CHN	False	Authority	Morphological_characteristics	Variety
23	3	7/10/2025	7	18	CHN	True	Authority	Morphological_characteristics	Variety
24	1	7/10/2025	7	22	CHN	True	Authority	Morphological_characteristics	Variety
25	3	7/10/2025	8	18	CHN	True	Authority	Morphological_characteristics	Variety
26	2	7/10/2025	8	22	CHN	True	Authority	Morphological_characteristics	Variety
27	3	7/10/2025	9	18	CHN	True	Authority	Morphological_characteristics	Variety
28	2	7/10/2025	9	22	CHN	True	Authority	Morphological_characteristics	Variety
29	3	7/10/2025	10	18	CHN	True	Authority	Morphological_characteristics	Variety
30	2	7/10/2025	10	22	CHN	True	Authority	Morphological_characteristics	Variety
31	1	7/10/2025	11	18	CHN	True	Authority	Morphological_characteristics	Variety
32	1	7/10/2025	11	22	CHN	True	Authority	Morphological_characteristics	Variety
33	null	7/10/2025	12	18	CHN	False	Authority	Morphological_characteristics	Variety
34	null	7/10/2025	12	22	CHN	False	Authority	Morphological_characteristics	Variety

19

Demonstration

Code based solutions:

- Accounting for GxE interactions
- Fitting reference scores
- Finding most similar



20

Next steps

Continue sharing the concept and help defining a generic framework

Test and validation by interested PBR offices and UPOV

Connection with existing repositories and services:

- PLUTO
- ePVP DUS exchange module

PROPOSAL: Creation of an official working group with authorities and breeders

21

Thank you!

Contact: Emerson.Limberger@corteva.com

22

[End of Annex and of document