UPOV
International Union for the Protection of New Varieties of Plants

E

Technical Working Party on Testing Methods and Techniques **TWM/1/17**

First Session
Virtual meeting, September 19 to 23, 2022

Original: English
Date: August 31, 2022

DEVELOPMENT OF A SNP MARKER SET IN *CANNABIS* TO SUPPORT DUS TESTING

Document prepared by an expert from the Netherlands

Disclaimer: this document does not represent UPOV policies or guidance

The annex to this document contains a copy of a presentation on "Development of a SNP marker set in *Cannabis* to support DUS testing", prepared by an expert from the Netherlands, to be made at the first session of the TWM.

[Annex follows]

Development of a SNP marker set in *Cannabis* to support DUS testing

UPOV-TWM/1 – September 19-23, 2022

Objective and Scope

A validated SNP marker-set suitable to genetically differentiate *Cannabis* varieties.

Validated SNP-set and SNP-genotypes are prerequisites for…

- DNA database for *Cannabis*
- Use the database in DUS procedure for the management of reference collection (UPOV-models)

Follow-up project

Deliverables

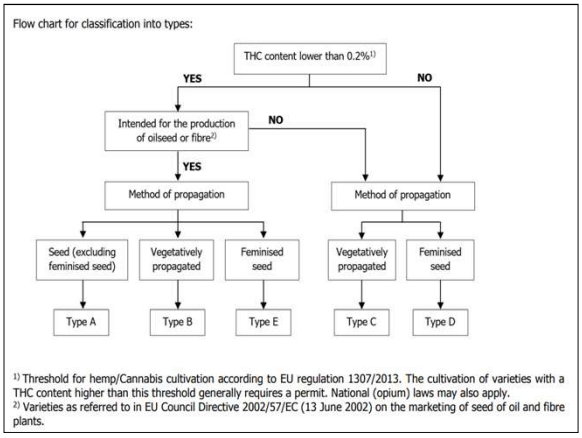
- A SNP set for *Cannabis* (circa 200-300 SNPs) that meet the quality criteria. (*reproducible, repeatable, robust, highly discriminative)*
- A genotyping method for *Cannabis* based on multiplex targeted amplicon sequencing (GT-Seq) to effectively and routinely detect the SNPs.

General Project information

- Project started January 2020 (grant agreement between CPVO and Naktuinbouw)
- Budget €53.000; financed by CPVO for 100%
- Duration 24 months (terminated at 1st January 2022)
- Project partners: Naktuinbouw & NÉBIH
- Stakeholders: GEVES & Bedrocan

Challenges in Cannabis

- Intended use:
  - Fibre/feed/seed/oil (low intoxicants)
  - Cannabinoids-THC (high intoxicants)
- Different ways of reproduction: Cuttings, normal seeds, feminised seeds
- Different Types (A to E)
- Common genetic pool (one SNP set and one database)



Plant Material

- Collection of plant material during 2019-2021 (still ongoing)
- Both low and high THC types
- In total 291 different varieties genotyped!!
  - 245 varieties 2019-2021, 46 varieties in 2022
- Individual plants per variety varied (depending on method of propogation) between #4 - #48
- Total number of entries in BN database = #2862
- Technical replica's and biological replica's are included (validation)

SNP discovery & selection

Discovery

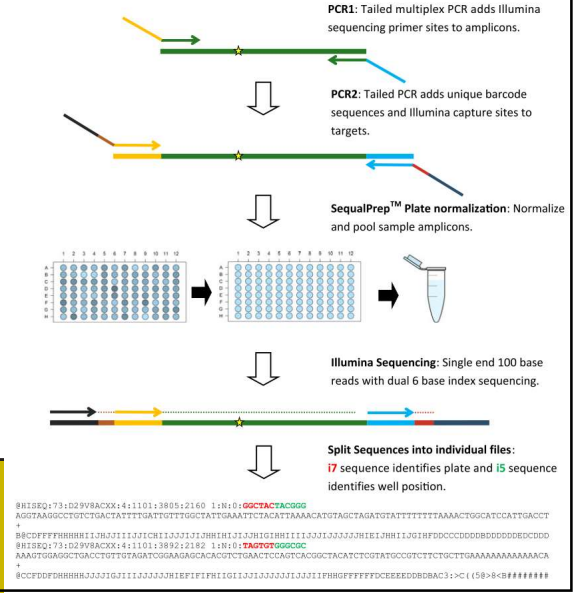
- Two GBS experiments (Bedrocan & Naktuinbouw)
- Naktuinbouw: 94 samples (both low and high THC types)
- Bedrocan: 150 samples (only 'coffeeshop'-types)
- After quality filtering and mapping: 125669 SNP positions identified

Selection (500 SNPs)

- High discriminative power (0,4 < h > 0,6)
- Unique SNP location when mapped on the reference genome
- Flanking sequence suitable for primer design
- SNPs at least 1000 bp apart to avoid linkage

GT-Seq

- Genotyping-in-Thousands by sequencing (GT-seq)
- Multiplexed targeted amplicon sequencing
- Paired-end 150 bp sequencing



Campbell *et al.*, 2015

GT-Seq Primer Design

Design criteria:

- No primer dimers
- No hairpins
- Same annealing temperature for all primers (C/G content)
- No repeats
- In silico prediction of unique and on-target vs off-target amplification of primers and primer pairs according to the reference genome

500 SNPs selected (based on GBS) → 240 SNPs with suitable primers (in silico predicted)

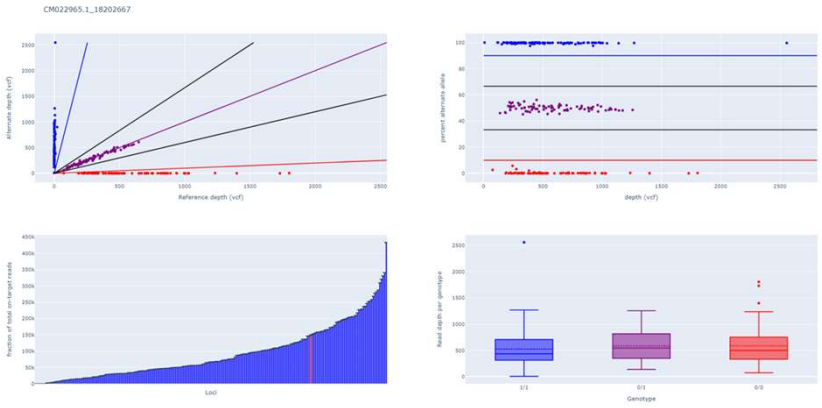
GT-Seq Results

From 240 SNPs in GT-Seq, 211 SNPs performed well.

Amplification issues for 29 SNPs (removed)

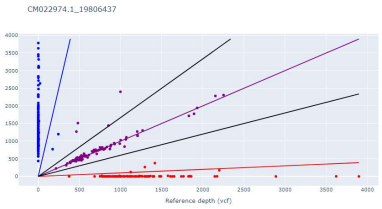
Distribution of reads over the SNPs

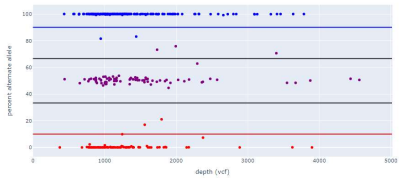
Classification of samples in the 3 genotypes (RR; RA; AA)



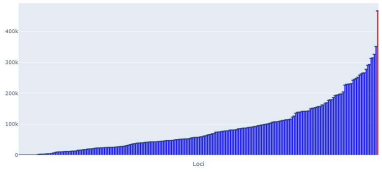
GT-Seq odd-Results

SNP with the highest fraction of on-target reads.

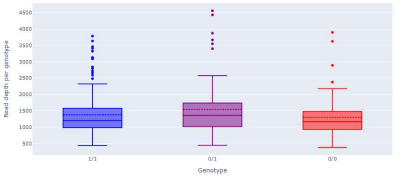




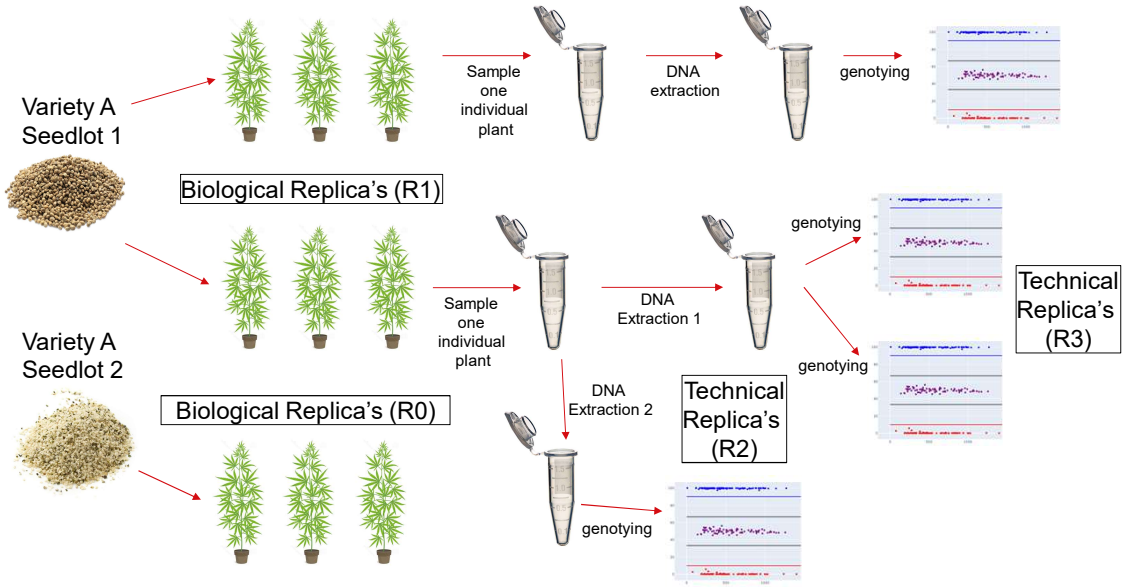
Some samples with divergent genotype patterns



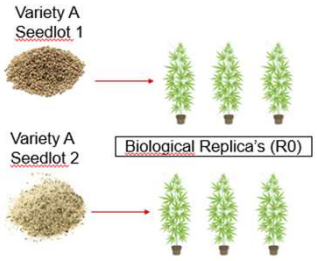
Possible explanation:
Tri- or tetraploid genomes

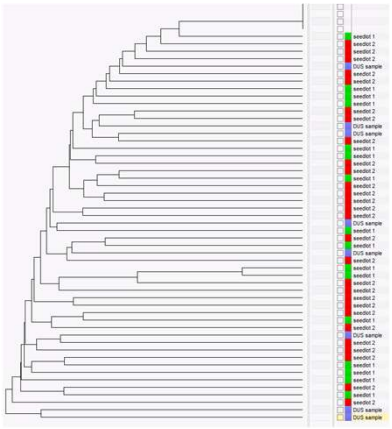


Controls for validation

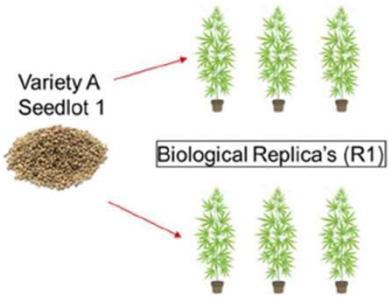


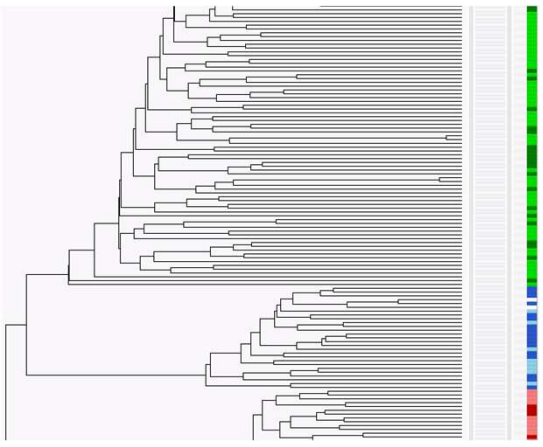
Validation R0 samples



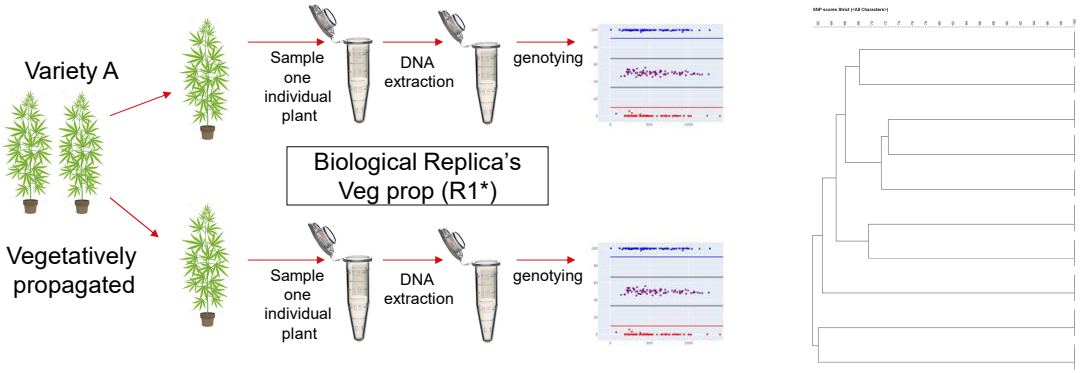


Validation R1 samples



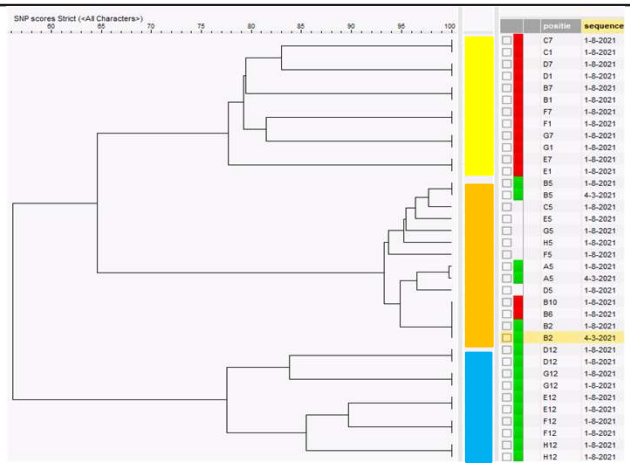


Validation R2 samples



Different individual plants from veg prop variety. Sampled, DNA extracted and Genotypes at two different moments in time. Identical genotypes (100% consistent). Distinct from all other varieties

Validation R1 and R2/R3 samples

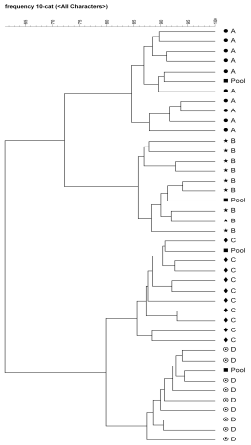


Variety A; several individual plants (R1) sampled in duplo (R2)

Variety B; several individual plants (R1) some sampled in duplo (R2) For others the same DNA extract was genotyped in duplo (R3)

Variety C; several individual plants (R1) for which the same DNA extract was genotyped in different moments in time

Individuals vs pools



For seed samples: pools in stead of individuals?

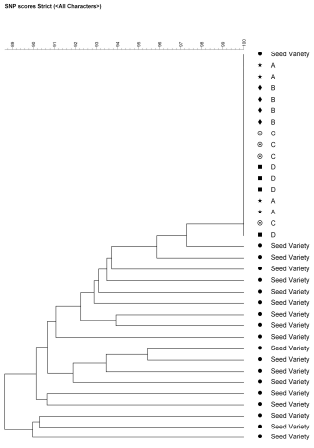
Advantages
- Representative for identity of the sample
- Cost reduction (genetic distance based on allele frequency)

Disadvantages
- Loss of overview on genetic relationships between seed samples and vegetatively propagated samples

For the time being: pooled sampled ***in addition to*** individuals in one database

Similarity based on allelic frequency – not on numeric genotype state (0, 1, 2)

Added value for DUS



4 vegetatively propagated applications currently in DUS trial (A, B, C and D)

100% match with one individual from seed variety.

All 4 veg propagated samples are genetically identical. Mutants?

Phenotype is leading

Added value for DUS

Cluster	Matching varieties	Conclusion based on morphology	Action
1	HNP221	Positive decision on D	PBR protection
1	HNP222	Positive decision on D in comparison with other reference variety in 2020 and before DNA matches were available.	Nullification is considered
2	HNP326	Positive decision on D	PBR protection
2	HNP328	D not yet clear	Extra year of testing - now side-by-side with HNP326
3	Ref A	Reference variety-PBR protected	
3	HNP252	D not yet clear	Extra year of testing - now side-by-side with Ref A
3	HNP411	Negative decision on D	Rejected based on lack of D
4	Ref B	Reference variety-PBR protected	
4	HNP138	Positive decision on D based on comparison with other reference variety in 2019. At that time DNA matches were not available	Nullification is considered
4	HNP416	D not yet clear	Extra year of testing - now side-by-side with HNP138 and Ref B
5	Ref D	Reference variety-PBR protected	
5	HNP139	Positive decision on D based on comparison with other reference variety in 2019. At that time DNA matches were not available	Nullification is considered
5	HNP159	Positive decision on D based on comparison with other reference variety in 2019. At that time DNA matches were not available	Nullification is considered
5	HNP291	Negative decision on D	Rejected based on lack of D
5	HNP412	Negative decision on D	Rejected based on lack of D
5	HNP415	Negative decision on D	Rejected based on lack of D
5	HNP318	Negative decision on D	Rejected based on lack of D
6	HNP231	Positive decision on D	PBR protection
6	HNP244	Positive decision on D in comparison with other reference variety in 2020 and before DNA matches were available.	Nullification is considered

- 6 clusters of matching varieties
- Nullifications are considered: D was based on comparison with other ref varieties.
- Rejected based on lack of D
- Extra year of testing, now side-by-side the DNA matching ref variety

Work in progress

'Improving DUS research for Cannabis through the use of DNA database'

- Collection of CK samples
- Add to database
- Develop a (new) DUS procedure with DNA and test the efficiency
- Develop DNA marker to predict male/female plants

[End of Annex and of document]