Differentiating Marijuana from Hemp using an NGS Panel and Rapid Genotyping Assays Targeting Cannabinoid Synthase Genes
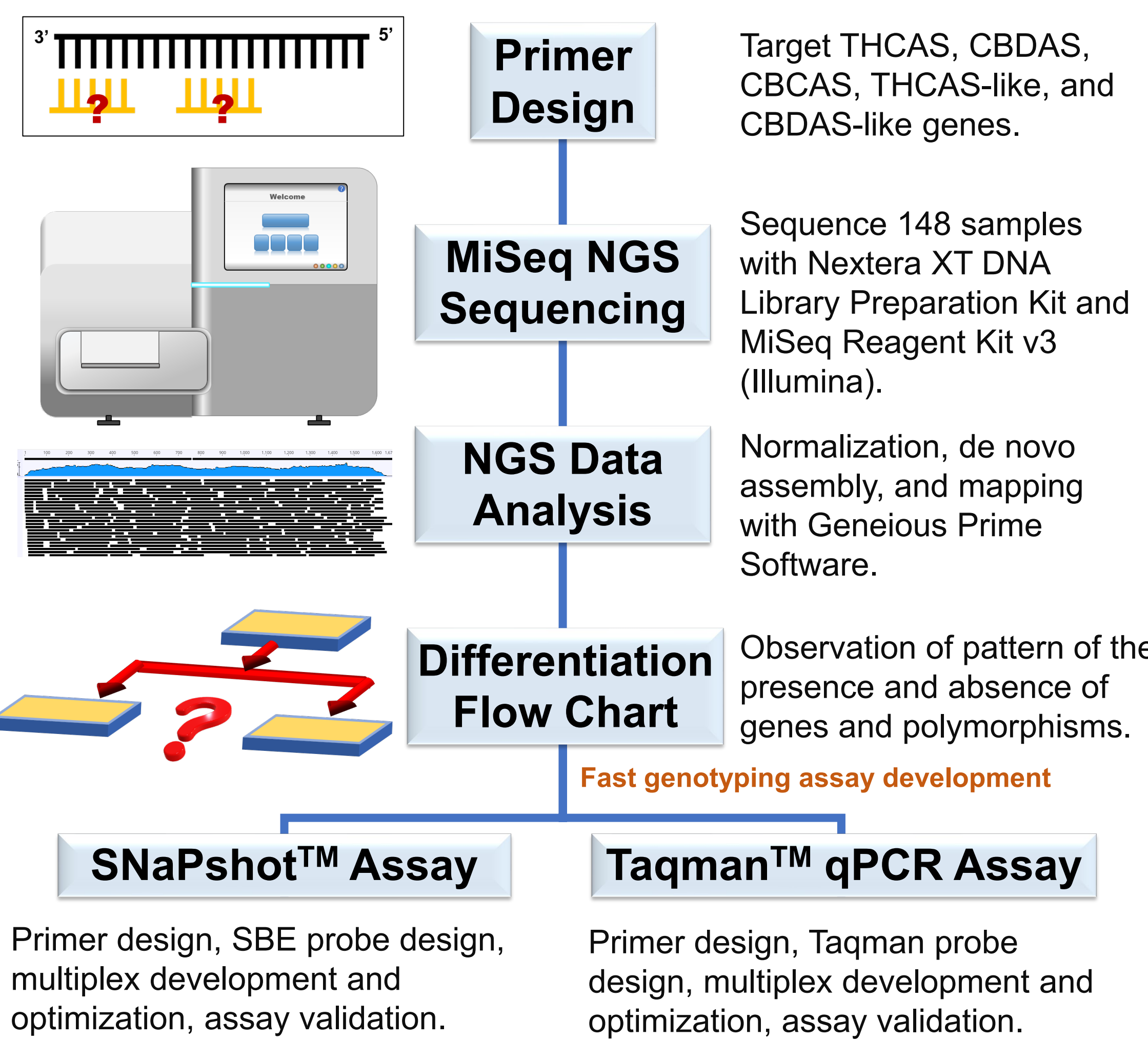
Ya-Chih (Jessica) Cheng, PhD and Rachel Houston, PhD

Department of Forensic Science, Sam Houston State University, Huntsville, TX 77340

INTRODUCTION

Differentiating between the two crop-type of *Cannabis sativa* (hemp and marijuana) has become an important task for law enforcement following the legalization of hemp through the 2018 Agriculture Improvement Act. Genetic markers, particularly cannabinoid synthase genes, have been identified as key elements in determining the chemical composition or chemotype of a cultivar. While previous studies proposed the use of tetrahydrocannabinolic acid synthase (THCAS) and cannabidiolic acid synthase (CBDAS) for predicting chemotypes, existing methods that rely on outdated theories might lack specificity and precision. The highly repetitive nature of the cannabis genome along with non-functional pseudogenes that share a significant nucleotide identity with functional synthase genes poses a challenge in accurate genetic testing [1-4]. To enhance precision, a custom Next Generation Sequencing (NGS) assay was developed, targeting THCAS, CBDAS, and their pseudogenes, along with another synthase gene (cannabichromenic acid synthase (CBCAS)). The resulting flowchart was utilized to create fast genotyping methods (SNaPshot™ and TaqMan™ SNP Genotyping Assays), promising efficient tools for law enforcement. The SNaPshot™ assay's effectiveness was confirmed using 15 *C. sativa* varieties (n=130), while further assessment is needed for the TaqMan™ SNP Genotyping assay.

MATERIALS & METHODS

Primer Design — Target THCAS, CBDAS, CBCAS, THCAS-like, and CBDAS-like genes.

MiSeq NGS Sequencing — Sequence 148 samples with Nextera XT DNA Library Preparation Kit and MiSeq Reagent Kit v3 (Illumina).

NGS Data Analysis — Normalization, de novo assembly, and mapping with Geneious Prime Software.

Differentiation Flow Chart — Observation of pattern of the presence and absence of genes and polymorphisms.

Fast genotyping assay development

SNaPshot™ Assay — Primer design, SBE probe design, multiplex development and optimization, assay validation.

Taqman™ qPCR Assay — Primer design, Taqman probe design, multiplex development and optimization, assay validation.

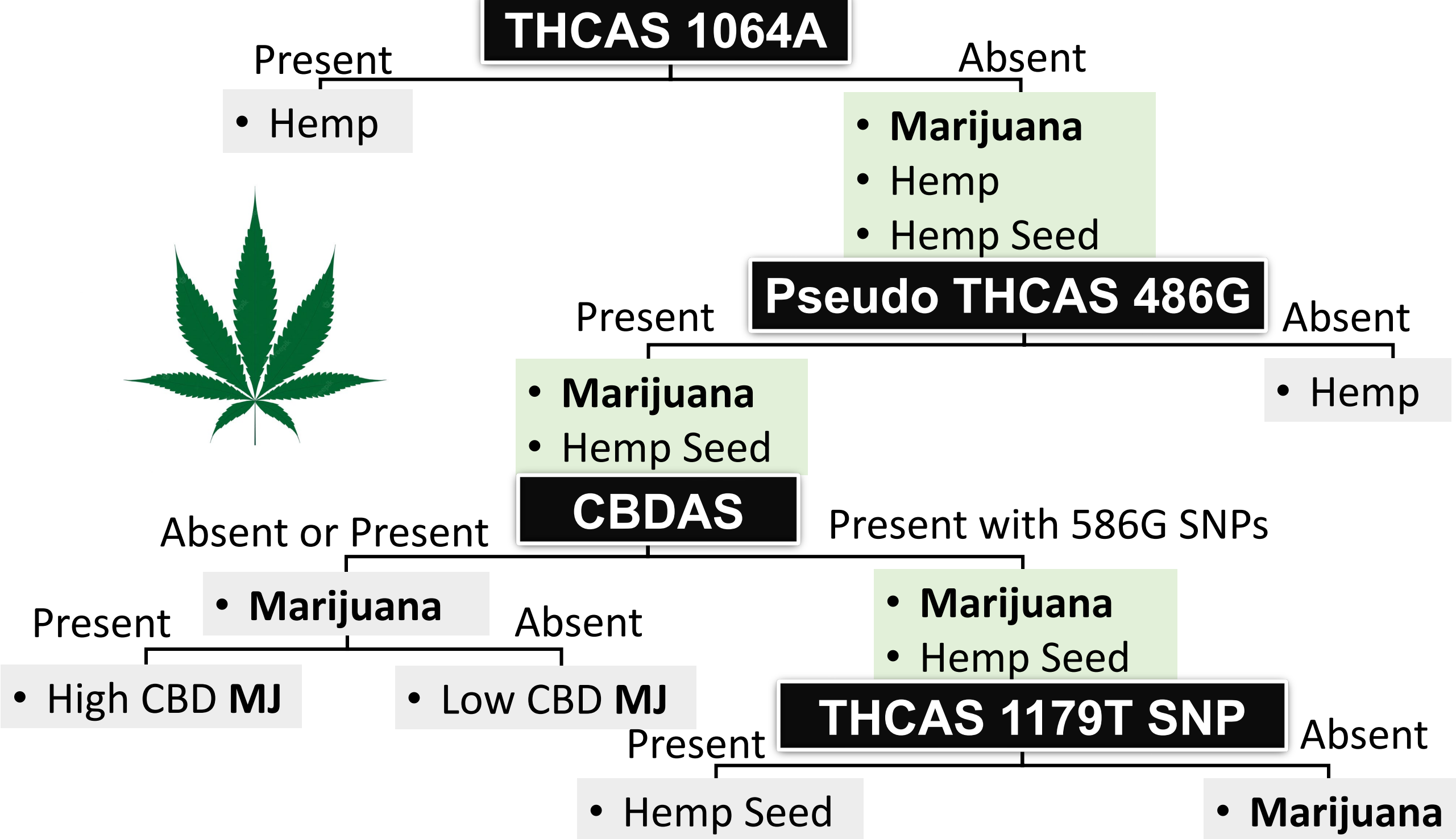## RESULTS & DISCUSSION

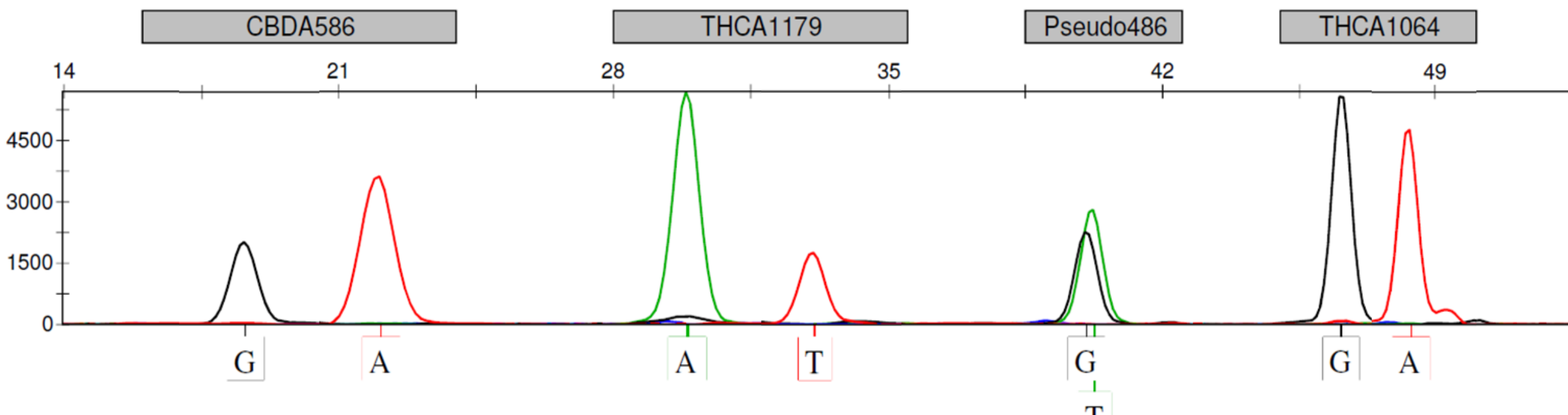Figure 1. Cannabis differentiation flow chart.

Figure 2. Four-plex SNaPshot™ assay with all possible alleles.

Table 1. Interpretation of the SNaPshot™ genotypes. Dark gray box indicates the SNP of that marker is not important in that genotype. CBDA marker in genotype 3 can provide CBD information of the marijuana sample. Genotype 5 needs peak height ratio (PHR) to determine the strains.

Genotype	CBDA 586	THCA 1179	Pseudo 486	THCA 1064	Identification
1				A	THC<1% cannabis or Hemp seed (n=56)
2				Abs.	
3	CBD info.		G	G	Marijuana (n=69)
4	G	T or A/T	G	G	Hemp Seed (n=2)
5			G or G/T	G or G/A	PHR (1064G/486G) (n=3)

- Cannabis differentiation flow chart based on NGS and PCR results (**Figure 1**).
- SNaPshot™ assay was developed using all four markers and validated with 130 samples (**Figure 2**).
- Interpretation of the SNaPshot™ data is explained in **Table 1**.
- Taqman™ qPCR assay was developed using two markers and preliminarily validated with 94 samples (**Figure 3**).
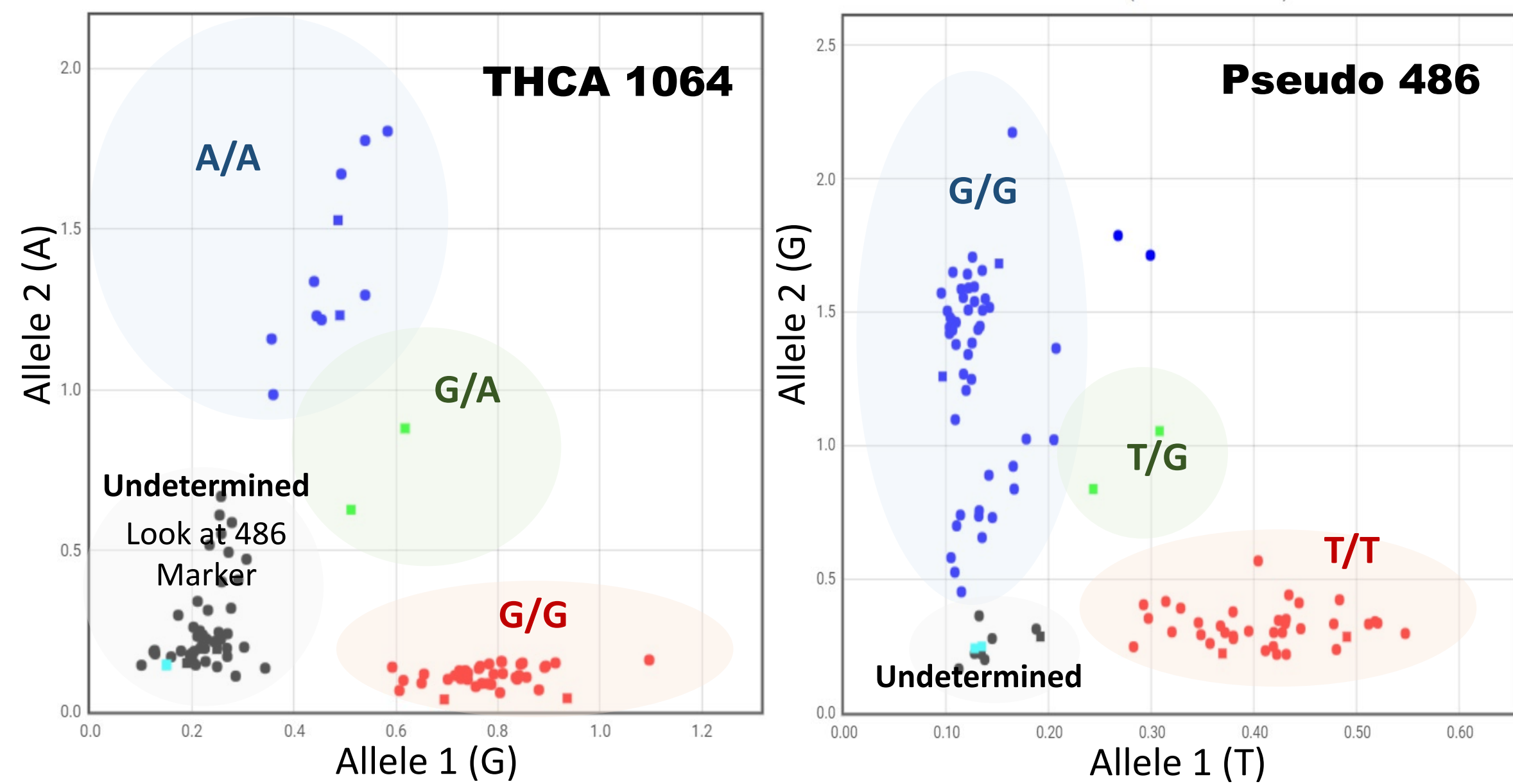- Interpretation of the Taqman™ data is proposed in **Table 2**.

Figure 3. Taqman™ qPCR multiplex assay. Two markers including THCA 1064 (left) and Pseudo THCA 486 (right).

Table 2. Interpretation of Taqman™ qPCR genotypes. Dark gray box indicates the SNP of that marker is not important.

Genotype	THCA1064	Pseudo486	Identification
1	A/A		THC<1% Cannabis
2	G/A		THC<1% Cannabis
3	G/G		Marijuana
4	Undetermined	T/T	THC<1% Cannabis
5	Undetermined	T/G	THC<1% Cannabis
6	Undetermined	Undetermined	THC<1% Cannabis
7	Undetermined	G/G	Marijuana

CONCLUSION

- A custom NGS panel was successfully utilized to analyze three cannabinoid synthase genes and their pseudogenes.
- A novel identification flow chart for the differentiation of hemp and marijuana is proposed.
- Two fast genotyping assays, SNaPshot™ assay and TaqMan™ qPCR assay, were successfully developed and optimized.
- The advantages and disadvantages of the two assays were compared and summarized (see table to the right).

Genotype	SNaPshot™	Taqman™
Targets	4	2
Instrumentation	PCR, CE	qPCR
Time	~3 hrs	~1 hrs
Differentiate cannabis with THC >1% and <1%	Yes	Yes
Differentiate hemp seed from marijuana	Yes	No
Marijuana CBD Info.	Yes	No

REFERENCES

[1] Kojoma, M., et al., DNA polymorphisms in the tetrahydrocannabinolic acid (THCA) synthase gene in "drug-type" and "fiber-type" Cannabis sativa L. Forensic Sci Int, 2006. 159(2-3): p. 132-40.
[2] Rotherham, D. and S.A. Harbison, Differentiation of drug and non-drug Cannabis using a single nucleotide polymorphism (SNP) assay. Forensic Sci Int, 2011. 207(1-3): p. 193-7.
[3] Cascini, F., et al., Highly Predictive Genetic Markers Distinguish Drug-Type from Fiber-Type Cannabis sativa L. Plants (Basel, Switzerland), 2019. 8(11): p. 496.
[4] Hurgobin, B., et al., Recent advances in Cannabis sativa genomics research. New Phytol, 2021. 230(1): p. 73-89.