Cannabis sativa target capture kit for exon sequencing, genotyping, and marker discovery

Jacob Enk[1], Alison Devault[1], Biao Liu[2], Kevin McKernan[2]
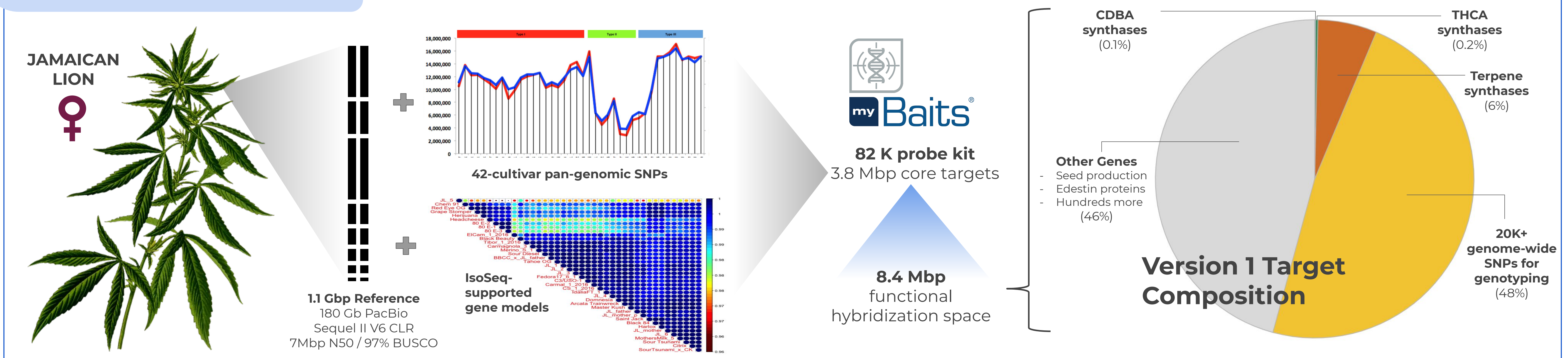
1 Arbor Biosciences, Ann Arbor, MI, USA | 2 Medicinal Genomics Corporation, Beverly, MA, USA
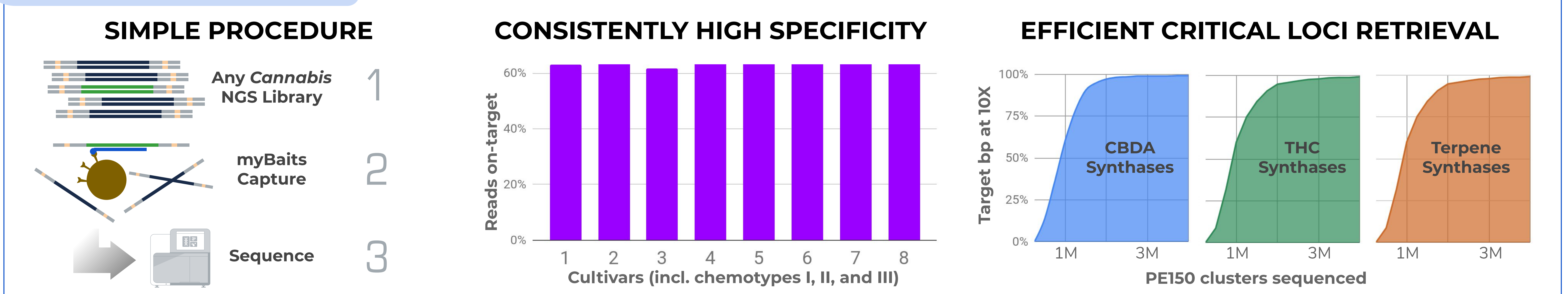
arbor biosciences

MEDICINAL GENOMICS

Second- and third-generation sequencing technologies have recently been leveraged to resolve genomes of a variety of cannabis cultivars. Using this tremendous resource, Medicinal Genomics Corporation (MGC) and Arbor Biosciences have partnered to develop a new hybridization capture panel for comprehensive, routine sequencing of key cannabis genomic loci. Comprising **8.4 Mbp total target space**, the kit consistently retrieves not only **tens of thousands of SNP markers** for high resolution genotyping, but also regions involved in cannabinoid and terpene synthesis, as well as edestin and seed production. The co-branded panel will be available as a kit from MGC for the global cannabis breeding market, and from Arbor Bio for the global hemp research market. This cost-effective sequencing tool enables researchers to better understand genetic pathways responsible for favored traits, and empowers growers, dispensaries and testing laboratories to ensure that patients and consumers have access to safe, quality cannabis.
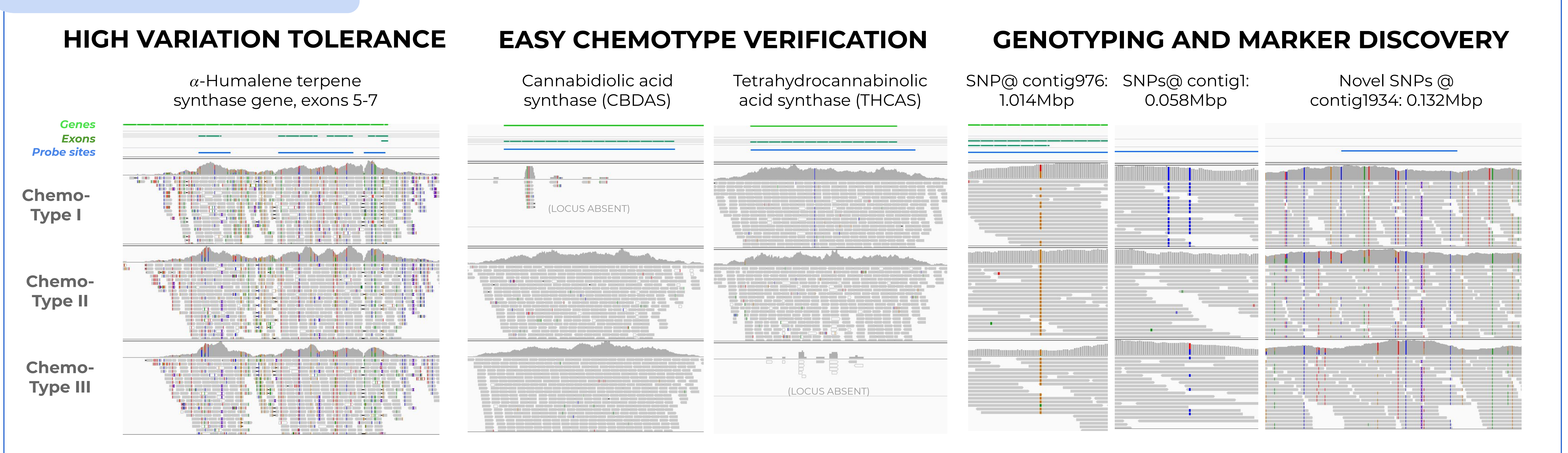
COMPOSITION

JAMAICAN LION ♀

42-cultivar pan-genomic SNPs

IsoSeq-supported gene models

1.1 Gbp Reference
180 Gb PacBio
Sequel II V6 CLR
7Mbp N50 / 97% BUSCO

myBaits®

82 K probe kit
3.8 Mbp core targets

8.4 Mbp functional hybridization space

Version 1 Target Composition

CDBA synthases (0.1%)

THCA synthases (0.2%)

Terpene synthases (6%)

Other Genes
- Seed production
- Edestin proteins
- Hundreds more
(46%)

20K+ genome-wide SNPs for genotyping (48%)

PERFORMANCE

SIMPLE PROCEDURE

1 — Any *Cannabis* NGS Library

2 — myBaits Capture

3 — Sequence

CONSISTENTLY HIGH SPECIFICITY

Reads on-target — Cultivars (incl. chemotypes I, II, and III)

EFFICIENT CRITICAL LOCI RETRIEVAL

CBDA Synthases | THC Synthases | Terpene Synthases

Target bp at 10X — PE150 clusters sequenced

VERSATILITY

HIGH VARIATION TOLERANCE

α-Humalene terpene synthase gene, exons 5-7

Genes
Exons
Probe sites

Chemo-Type I
Chemo-Type II
Chemo-Type III

EASY CHEMOTYPE VERIFICATION

Cannabidiolic acid synthase (CBDAS)

Tetrahydrocannabinolic acid synthase (THCAS)

(LOCUS ABSENT)

(LOCUS ABSENT)

GENOTYPING AND MARKER DISCOVERY

SNP@ contig976: 1.014Mbp

SNPs@ contig1: 0.058Mbp

Novel SNPs @ contig1934: 0.132Mbp

myBaits®

Targeted NGS via Hyb Capture
www.arborbiosci.com/genomics/targeted-sequencing/mybaits