

Kevin J. McKernan, Yvonne Helbert, Liam T. Kane, Nathan Houde, Lei Zhang, Stephen McLaughlin

Abstract

We performed whole genome sequencing to assess the specificity of colony forming units derived from three different plating media's: Potato Dextrose Agar (PDA), PDA with chloramphenicol and Dichloran Rose Bengal with chloramphenicol (DRBC). Colonies were isolated from each media type and whole genome sequenced to identify the diversity of microbes presented with each media selection. ITS3 and 16S qPCR was performed to correlate these CFUs with fungi and bacterial specific qPCR. The plating mediums each displayed a log scale different colony forming unit count. PDA with chloramphenicol demonstrated the highest diversity and the highest concordance with whole genome sequencing. According to ITS3 and 16S qPCR DRBC under counts yeast and mold while PDA without chloramphenicol over counts CFUs due to bacterial growth without selection.

Introduction

Total yeast and mold testing is required in many states for safety testing cannabis prior to sale of cannabis flowers and cannabis infused products. Cannabis is an inhaled product and cases of cannabis acquired Aspergillosis exist in the clinical literature (Bal et al. 2010; Gargani et al. 2011; McKernan et al. 2015; Remington et al. 2015; Ruchlemer et al. 2015; McKernan et al. 2016). Cannabis is a unique matrix in that antibiotic cannabinoids can make up to 20% of the flowers weight and many fungi that infect cannabis are endophytes. Endophytes are not easily cultured from the plant without lysing open plant cell walls. Conditions which lyse open plant cells walls also lyse open fungal cell walls thus impacting the viability of the microbes in the lysis and homogenization process required for testing. Cannabis flowers contain both bacteria and fungi, further complicating fungal quantitation with colony forming units that lack speciation. Antibiotic selections are often utilized to reduce background bacteria but many of these antibiotics (chloramphenicol) inhibit the growth of the most pathogenic fungi found on cannabis (*Fusarium*, *Pythium* and *Aspergillus*)(Smith and Marchant 1968; Day et al. 2009; Joseph et al. 2015).

As part of an AOAC Emergency response validation (ERV) in the State of Michigan, we investigated the impact of media selection on surveying total yeast and mold on cannabis. Cured cannabis flowers were homogenized and tested on 3 different plating medias. These data were compared to ITS3 and 16S based qPCR and whole genome sequencing. To further complement these cannabis flower samplings, organisms were acquired from ATCC and plated as pure monocultures on different plating mediums to confirm the differential growth on each media.

Methods

Plating

Samples originated from Steadfast Analytical Laboratories (Hazel Park, MI) and were tested independently at a laboratory within the Michigan Coalition of Independent Cannabis Testing Laboratories. Briefly, 10 grams of dried cannabis flowers were sampled from three lots of homogenized cannabis containing high, medium and low quantities of CFUs as measured using

culture-based techniques with chloramphenicol selection. 10 grams of homogenized flower were soaked with 90ml of Tryptic Soy Broth (TSB) in a filtered Whirl-Pak bag. Samples were hand homogenized, and then 0.1 mL plated onto 3 plates (1:100 dilution). Two additional dilutions were prepared (10 mL into 90 mL) and same plating protocol was followed. All plates were incubated for 5 days at 25°C.

qPCR

ITS3 qPCR was performed as described in McKernan *et al.* with 2 modifications. Briefly, 1ml of homogenate from a WhirlPak bag if collected and briefly micro-centrifuged to enrich for live organisms. This pellet was resuspended in 200µl ddH2O and lysed with the addition of 12µl of Thaumatin-like protein (TLP) and incubated at 37°C for 30 minutes. This enzymatic lysis step (glucanase) ensures more complete lysis of fungal cell walls (MGC part #420206)(McKernan et al. 2015; McKernan et al. 2016). 12.5µl of MGC Lysis buffer was added, vortexed and incubated for 5 minutes at 25°C. Lysed sample were micro-centrifuged and 200µl of supernatant was aspirated and added to 250µl of MGC binding buffer (MGC part# 420001) for magnetic bead isolation. The samples were incubated with the bead mixture for 10 minutes, magnetically separated and washed 2 times with 70% ethanol. The beads were dried at 37°C for 5 minutes to remove excess ethanol and eluted with 25µl of ddH2O. qPCR was performed according to the manufacturers' instructions on a BioRad CFX instrument.

DNA isolation from colonies for whole genome sequencing.

Colonies were picked with a pipette tip and introduced into 200µl of ddH2O with 12.5µl of MGC TLP (MGC part #420206). TLP is a glucanase active at 37°C. Samples were digested for 30 minutes at 37°C and 12.5µl of MGC Lysis buffer was added, vortexed and incubated for 5 minutes at 25°C. Lysed sample were micro-centrifuged and 200µl of supernatant was aspirated and added to 250µl of MGC binding buffer (MGC part# 420001) for magnetic bead isolation. The samples were incubated with the bead mixture for 10 minutes, magnetically separated and washed 2 times with 70% ethanol. The beads were dried at 37°C for 5 minutes to remove excess ethanol and eluted with 25µl of ddH2O.

Library construction for whole genome sequencing.

Fragmentation

gDNA was quantitated with a Qubit (Thermo Fisher Scientific) and normalized to reflect 4-8ng/µl in 13µl of TE buffer. Libraries were generated using enzymatic fragmentation with the NEB Ultra II kits (NEB part # E7103). Briefly, 3.5µl of 5X NEB fragmentation buffer and 1µl of Ultra II fragmentation enzyme mix are added to 13µl of DNA. This reaction was tip mixed 10 times, vortexed, and quickly centrifuged. Fragmentation is performed in a thermocycler at 3.5 minutes at 37°C, 30 minutes at 65°C. The reaction was kept on ice until ready for adaptor ligation.

Adaptor ligation

1) Prepare master mix on ice

Component	Volume (µl)
SureSelect Adaptor Oligo Mix (brown cap)	0.75
ddH ₂ O	0.5
Ultra II Ligation Master Mix	15
Ligation enhancer	0.5
Total Volume	16.75

2) Add 16.75µl Master Mix to the 17.5µl Fragmentation/End Prep DNA reaction mixture, incubate for 15 minutes at 20°C.

3. Cleanup step to remove excess adaptors and adaptor dimers

- 1) Vortex room temp AMPure XP beads (Beckman Coulter) to resuspend
- 2) Add 16µl (approximately 0.45X) of resuspended AMPure XP beads to the ligation reactions. Mix well by pipetting up and down at least 10 times.
- 3) Incubate for 5 minutes at 25°C.
- 4) Put the PCR plate on an appropriate magnetic stand to separate beads from supernatant. After the solution is clear (about 5 minutes), carefully remove and discard the supernatant. Be careful not to disturb the beads that contain DNA targets.
- 5) Add 200µl of 70% ethanol to the PCR plate while in the magnetic stand. Incubate at room temperature for 30 seconds, and then carefully remove and discard the supernatant.
- 6) Repeat Step 5 once for a total of 2 washes. Remove trace amount of EtOH
- 7) Air dry beads for ~ 7 minutes while the PCR plate is on the magnetic stand with the lid open.
- 8) Remove PCR plate from magnet and elute DNA target from beads into 10µl of H₂O, transfer 9µl cleaned DNA to a fresh well.

4. PCR amplification

- 1) Add 12.5µl 2x NEBNext Q5 Hot Start Master Mix to 9µl ligated DNA, then add 3.5µl NEB 8bp index primer/universal primer, run the following program

STEP	TEMP	TIME	Cycle
Initial denaturation	98°C	30 sec	1
Denaturation Annealing/Extension	98°C 65°C	10 sec 75 sec	6
Final extension	65°C	5 min	1
Hold	4°C	forever	1

Please adjust total PCR volume to 25µl on Thermocycler

5. PCR reaction cleanup

- 1) Vortex room temp AMPure XP beads to resuspend
- 2) Add 15 μ l of resuspended AMPure XP beads to the PCR reactions (~ 25 μ l). Mix well by pipetting up and down at least 10 times.
- 3) Incubate for 5 minutes at room temperature.
- 4) Put the PCR plate on an appropriate magnetic stand to separate beads from supernatant. After the solution is clear (about 5 minutes), carefully remove and discard the supernatant. Be careful not to disturb the beads that contain DNA targets.
- 5) Add 200 μ l of 70% ethanol to the PCR plate while in the magnetic stand. Incubate at room temperature for 30 seconds, and then carefully remove and discard the supernatant.
- 6) Repeat Step 5) once.
- 7) Air dry beads for ~ 7 minutes while the PCR plate is on the magnetic stand with the lid open.
- 8) Elute DNA target from beads into 15 μ l of nuclease free H₂O, transfer 15 μ l into a fresh well.

Sample quality control

Libraries were evaluated on an Agilent Tape Station prior to pooling for Illumina sequencing. Sequencing was performed by GeneWiz, Cambridge MA. 473 million read pairs (2 x 150bp) were generated, averaging over 10 million read pairs per sample and 141Gb of total sequence.

Analysis

Fastq files were uploaded to OneCodex for Kmer analysis and Simpson's diversity index of each genome (Supplementary Table 1, sheet Summary). Reads were also assembled with MegaHit v.1.2.9 (Li et al. 2015; Li et al. 2016). The Nextflow mapping and assembly pipeline is published in GitHub (https://github.com/mclaugsf/mgc-public/tree/master/AOAC_TYM_ERV). Quast 5.0 was used to calculate the assembly quality statistics (Gurevich et al. 2013). Sequencing data is deposited in NCBI under Project ID PRJNA725256.

Results

Each colony that was imaged on plates and picked for whole genome sequencing and OneCodex analysis is displayed in Table 1 (DRBC), Table 2 (PDA-CAMP) and Table 3(PDA no CAMP). A link to each OneCodex analysis and its' respective NCBI submission ID is available in Supplementary Table 1- Sheet Summary. Some of the colonies from the plate merged with other colonies producing mixtures of genomes as evident in the OneCodex pie charts. These merged colonies were further evidenced with the display of bimodal sequence coverage (clusters of contigs at 1000X and 10X coverage) and compared with the plating images (Figure 1). A heatmap of sequencing read speciation and purity is seen in Figure 2. While merged colonies can be difficult to resolve visually, whole genome sequencing can resolve simple metagenomes and still extract additional diversity information for the samples. Colonies that were noticeably mixed according to sequence analysis and colony visual inspection were more prevalent with the PDA without selection colonies (Table 4).

A Simpson's diversity index analysis demonstrated PDA with CAMP provides the highest diversity score (Figure 3.) While the DRBC had 100-fold lower CFU counts than PDA without selection, it predominantly displayed fungal colonies (80%) while PDA without selection was biased toward bacteria (22%). PDA with chloramphenicol displayed more fungi (55%) than bacteria and also produced a half log more fungal colonies than DRBC with chloramphenicol (Table 5).

One fungal sample (*Cladosporium*) presented delayed Ct (31.79) with TYM qPCR primers. Scrutiny of the primer sequences against the *Cladosporium* genome demonstrates proper primer binding locations but missing probe sequence. The coverage of this genome is only 10X and the ITS regions are often in copy number flux resulting in assembly breaks in the ITS regions. This may explain the missing probe sequence in the low coverage fragmented assembly. Additionally, some significantly delayed TAC signal was observed in fungal colonies. This is the result of the use of a Thaumatin-like protein which is cloned and expressed in *E.coli* and is utilized to lyse the fungal cells. This background TLP expression vector *E.coli* signal can be seen in blank preps. In some cases this signal is elevated due to mixed colonies observed in the sequencing data.

The qPCR data represents an increased selectivity compared to DRBC delivering 80% fungal colonies. Even if we discount the mixed colonies and count only a single bacterial colony out of 13 on DRBC we obtain 92% fungal colonies on DBRC. Quantitative PCR demonstrated over 10 Cts (1024 fold) differences between the TYM and TAC primer sets and the majority of the residual TAC Ct can be discounted due to the background *E.coli* TLP DNA.

To confirm these observations several *Aspergillus* species and *Botrytis cinerea* were ordered from ATCC and plated on various plating medias in absence of background cannabis matrix (Table 6 and Figure 4). In all cases DRBC shows reduced CFU counts.

Whole genome sequencing of colonies derived from cannabis flowers and the impact of media selection on benchmarking total yeast and mold detection tools



Table 1. DRBC with Chloramphenicol:

Colony Image (Left), Assembly sequence coverage (Y) compared with contig Length (X) where the contigs are sorted largest to smallest from left to right (Mid-Left). Assembly statistics calculated with Quast 5.0 (Mid Right). OneCodex speciated kmer count (Right).

Colony Image	Assembly Coverage (Y) Large->Small Contigs (X)	Assembly Statistics (Quast 5.0)	OneCodex Kmer Classification	TAC Ct	TYM Ct
DRBC-D-2-1-B 001.fastq.gz 	DRBC-E-2-1-B 	Assembly contigs # contigs (>= 25000 bp) 408 # contigs (>= 50000 bp) 123 Total length (>= 25000 bp) 19142192 Total length (>= 50000 bp) 9227130 # contigs 2100 Largest contig 171,869 Total length 31,609,610 GC (%) 46.53 N50 31465	OneCodex Kmer Classification Name Estimated Abundance Epicoccum nigrum 96.67% Pseudomonas oryzae 0.83% Pantoea agglomerans 0.78% Pantoea ananatis 0.46% Bacillus velezensis 0.39% Pseudomonas coelestis 0.25% Penicillium citrinum 0.11% Fusarium fujikuroi 0.11% Enterobacteriaceae 0.11% Alternaria alternata 0.07% (Remaining) 0.13%	34.05	19.52
DRBC-D-2-1-C 001.fastq.gz 	DRBC_D-2-1-C 	Assembly contigs # contigs (>= 25000 bp) 402 # contigs (>= 50000 bp) 205 Total length (>= 25000 bp) 24470385 Total length (>= 50000 bp) 17528385 # contigs 1431 Largest contig 237,982 Total length 31,648,269 GC (%) 46.49 N50 55985	OneCodex Kmer Classification Name Estimated Abundance Penicillium citrinum 97.46% Pantoea ananatis 0.65% Pseudomonas oryzae 0.58% Pantoea agglomerans 0.38% Bacillus velezensis 0.34% Pseudomonas coelestis 0.20% Fusarium fujikuroi 0.09% Enterobacteriaceae 0.07% Alternaria alternata 0.06% Heligella maydis 0.04% (Remaining) 0.11%	32.87	15.71
DRBC-E-2-1-A 001.fastq.gz 	DRBC_E-2-1-A 	Assembly contigs # contigs (>= 25000 bp) 47 # contigs (>= 50000 bp) 35 Total length (>= 25000 bp) 6848402 Total length (>= 50000 bp) 5899910 Largest contig 540,523 Total length 6,101,403 GC (%) 61.70 N50 161275	OneCodex Kmer Classification Name Estimated Abundance Pseudomonas montellii 95.84% Pantoea ananatis 1.54% Pseudomonas sp. UB4652 0.71% Deftia acidovorans 0.51% Enterobacteriaceae 0.30% Bacillus velezensis 0.23% Enterobacteriaceae bacterium UB4398 0.18% Pantoea agglomerans 0.18% Pseudomonas syringae group genom. 7 0.15% Fusarium fujikuroi 0.07% (Remaining) 0.31%	17.90	ND
DRBC-E-2-1-B 001.fastq.gz 	DRBC-E-2-1-B 	Assembly contigs # contigs (>= 25000 bp) 408 # contigs (>= 50000 bp) 123 Total length (>= 25000 bp) 19142192 Total length (>= 50000 bp) 9227130 # contigs 2100 Largest contig 171,869 Total length 31,609,610 GC (%) 46.53 N50 31465	OneCodex Kmer Classification Name Estimated Abundance Penicillium citrinum 97.46% Pantoea agglomerans 0.57% Pseudomonas oryzae 0.57% Pantoea ananatis 0.46% Bacillus velezensis 0.39% Pseudomonas coelestis 0.21% Enterobacteriaceae 0.10% Fusarium fujikuroi 0.09% Alternaria alternata 0.07% Pseudomonas sp. UB4652 0.06% (Remaining) 0.11%	33.80	18.79
DRBC-D-2-2-A 001.fastq.gz 	DRBC-D-2-2-A 	Assembly contigs # contigs (>= 25000 bp) 188 # contigs (>= 50000 bp) 156 Total length (>= 25000 bp) 66304032 Total length (>= 50000 bp) 65220039 # contigs 1367 Largest contig 2,008,609 Total length 69,008,622 GC (%) 58.07 N50 633911	OneCodex Kmer Classification Name Estimated Abundance Alternaria alternata 39.72% Aspergillus sp. M4-0081 23.74% Alternaria arborescens 5.90% Pantoea ananatis 2.41% Pantoea agglomerans 1.96% Pseudomonas aeruginosa 1.56% Bacillus velezensis 1.15% Pseudomonas coelestis 0.60% Pseudomonas oryzae 0.55% Penicillium citrinum 0.38% (Remaining) 1.96%	31.48	20.48
DRBC-D-2-2-B 001.fastq.gz 	DRBC-D-2-2-B 	Assembly contigs # contigs (>= 25000 bp) 54 # contigs (>= 50000 bp) 42 Total length (>= 25000 bp) 33106235 Total length (>= 50000 bp) 32704184 # contigs 533 Largest contig 2,165,682 Total length 33,924,917 GC (%) 52.66 N50 1232928	OneCodex Kmer Classification Name Readcount (% of classified reads) Cladosporium sp. SI-16 4253135 (61.72%) Preussia sp. RSL19 2726374 (26.74%) Dikarya 821060 (8.09%) Bacteria 489176 (4.74%) Cladosporium cladosporioides 390740 (3.83%) Dothidiomycetes 365987 (3.59%) Neolamprospora 310477 (3.04%) Cladosporium sphaerosporum 98878 (0.97%) Cladosporium sphaerosporum UM 843 70186 (0.69%) Cladosporium 49862 (0.49%) (Remaining) 824683 (8.09%)	34.56	31.79

Whole genome sequencing of colonies derived from cannabis flowers and the impact of media selection on benchmarking total yeast and mold detection tools



Colony Image	Assembly Coverage (Y) Large->Small Contigs (X)	Assembly Statistics (Quast 5.0)	OneCodex Kmer Classification	TAC Ct	TYM Ct
DRBC-D-2-2-C_001.fastq.gz 	 Fusarium fujikuroi Fungi	https://app.onecodex.com/analysis/public/c09b904c5614d49 Assembly DRBC_D_2_2_C # contigs (>= 25000 bp) 112 # contigs (>= 50000 bp) 92 Total length (>= 25000 bp) 19094822 Total length (>= 50000 bp) 18347245 # contigs 616 Largest contig 942,560 Total length 19,895,240 GC (%) 53.84 N50 257012	 Name Estimated Abundance Fusarium fujikuroi 87.23% Fusarium proliferatum 6.52% Fusarium mangiferae 1.68% Fusarium oxysporum 0.83% Fusarium appaerthi 0.79% Fusarium commune 0.70% Fusarium rygmaei 0.60% Fusarium verticillioides 0.40% Pseudomonas oryzae 0.40% (Remaining) 0.88%	32.14	16.85
DRBC-D-2-2-D_001.fastq.gz 	 Verticillium/Pantoea Mix	https://app.onecodex.com/analysis/public/71e136d914129493 Assembly DRBC_D_2_2_D # contigs (>= 25000 bp) 94 # contigs (>= 50000 bp) 88 Total length (>= 25000 bp) 33607324 Total length (>= 50000 bp) 33392936 # contigs 483 Largest contig 1,606,949 Total length 34,091,419 GC (%) 57.03 N50 488399	 Name Readcount (% of classified reads) Verticillium 171271 (3.42%) Colletotrichum 86641 (6.71%) Fusarium 55063 (4.31%) Pseudomonas 47297 (3.71%) Pantoea 37622 (2.95%) Ustilago 33891 (2.61%) Trichoderma 23457 (1.78%) Penicillium 19410 (1.52%) Alternaria 14808 (1.16%) Aspergillus 9882 (0.77%) (Remaining) 123082 (9.64%)	32.31	19.92
DRBC-D-2-2-E_001.fastq.gz 	 Fusarium fujikuroi Fungi	https://app.onecodex.com/analysis/public/58d9f76e304980c Assembly DRBC_D_2_2_E # contigs (>= 25000 bp) 107 # contigs (>= 50000 bp) 95 Total length (>= 25000 bp) 43394942 Total length (>= 50000 bp) 42948681 # contigs 584 Largest contig 2,312,367 Total length 44,242,770 GC (%) 48.13 N50 714660	 Name Estimated Abundance Fusarium fujikuroi 88.67% Fusarium proliferatum 5.79% Fusarium mangiferae 1.40% Fusarium oxysporum 0.88% Fusarium appaerthi 0.82% Fusarium commune 0.73% Fusarium rygmaei 0.62% Fusarium verticillioides 0.61% Fusarium verticillioides 0.40% Fusarium solum 0.34% (Remaining) 1.42%	32.69	17.00
DRBC-D-2-2-F_001.fastq.gz 	 Alternaria alternata Fungi	https://app.onecodex.com/analysis/public/434efbcf564a04f36 Assembly contigs # contigs (>= 25000 bp) 289 # contigs (>= 50000 bp) 209 Total length (>= 25000 bp) 32113182 Total length (>= 50000 bp) 29147380 # contigs 645 Largest contig 682,840 Total length 33,720,522 GC (%) 50.96 N50 150893	 Name Estimated Abundance Alternaria alternata 57.93% Pantoea ananatis 0.59% Pseudomonas oryzae 0.49% Sacillus velosus 0.26% Pantoea agglomerans 0.23% Pseudomonas caryophyllorum 0.19% Fusarium fujikuroi 0.08% Enterobacter cloacae 0.07% Penicillium citrinum 0.07% Ustilago maydis 0.04% (Remaining) 0.06%	33.15	18.68
DRBC-E-2-2-A_001.fastq.gz 	 Fusarium fujikuroi Fungi	https://app.onecodex.com/analysis/public/637a9480297244d6 Assembly DRBC_E_2_2_A # contigs (>= 25000 bp) 107 # contigs (>= 50000 bp) 98 Total length (>= 25000 bp) 43700655 Total length (>= 50000 bp) 43384462 # contigs 2430 Largest contig 1,740,874 Total length 45,935,783 GC (%) 47.74 N50 627582	 Name Estimated Abundance Fusarium fujikuroi 84.87% Fusarium proliferatum 5.96% Fusarium mangiferae 1.48% Fusarium oxysporum 1.09% Fusarium appaerthi 1.01% Pantoea agglomerans 0.90% Fusarium commune 0.89% Fusarium rygmaei 0.68% Pseudomonas morsella 0.67% Pantoea ananatis 0.59% (Remaining) 1.91%	31.48	20.48
DRBC-E-2-2-B_001.fastq.gz 	 Pantoea/Fusarium Mix	https://app.onecodex.com/analysis/public/2b9786e853ca495c Assembly DRBC_E_2_2_B # contigs (>= 25000 bp) 111 # contigs (>= 50000 bp) 95 Total length (>= 25000 bp) 40404941 Total length (>= 50000 bp) 39815152 # contigs 353 Largest contig 2,203,702 Total length 40,669,557 GC (%) 46.87 N50 615580	 Name Readcount (% of classified reads) Fusarium fujikuroi 30689 (2.14%) Penicillium citrinum 24006 (1.68%) Ustilago maydis 11310 (0.79%) Epistocum nigrum 10249 (0.72%) Pseudomonas oryzae 9195 (0.64%) Trichoderma harzianum 9181 (0.64%) Pantoea agglomerans 9137 (0.64%) Pseudomonas caryophyllorum 8623 (0.62%) Pseudomonas sp. LB40562 4474 (0.31%) Alternaria alternata 4457 (0.31%) (Remaining) 36689 (2.56%)	31.03	21.17
DRBC-E-2-2-C_001.fastq.gz 	 Fusarium fujikuroi Fungi	https://app.onecodex.com/analysis/public/c09b904c5614d49 Assembly DRBC_E_2_2_C # contigs (>= 25000 bp) 102 # contigs (>= 50000 bp) 98 Total length (>= 25000 bp) 43604791 Total length (>= 50000 bp) 43437963 # contigs 410 Largest contig 1,738,365 Total length 44,311,822 GC (%) 48.14 N50 630437	 Name Estimated Abundance Fusarium fujikuroi 87.23% Fusarium proliferatum 6.52% Fusarium mangiferae 1.44% Fusarium oxysporum 0.83% Fusarium appaerthi 0.79% Fusarium commune 0.70% Fusarium rygmaei 0.60% Fusarium verticillioides 0.57% Fusarium verticillioides 0.40% Pseudomonas oryzae 0.40% (Remaining) 0.88%	32.14	16.85

Whole genome sequencing of colonies derived from cannabis flowers and the impact of media selection on benchmarking total yeast and mold detection tools



Table 2. PDA with Chloramphenicol:

Colony Image (Left), Assembly sequence coverage (Y) compared with contig Length (X) where the contigs are sorted largest to smallest from left to right (Mid-Left). Assembly statistics calculated with Quast 5.0 (Mid Right). OneCodex speciated kmer count (Right).

Colony Image	Assembly Coverage (Y) Large->Small Contigs (X)	Assembly Statistics (Quast 5.0)	OneCodex Kmer Classification	TAC Ct	TYM Ct
 PC-E-2-2-A_001.fastq.gz	 Pseudomonas oryzae Bacteria	https://app.onecodex.com/analysis/public/4215767e704f64341 Assembly PC_E_2_2_A # contigs (>= 25000 bp) 18 # contigs (>= 50000 bp) 17 Total length (>= 25000 bp) 4902243 Total length (>= 50000 bp) 4864257 # contigs 1892 Largest contig 676,208 Total length 6,184,687 GC (%) 62.85 NS0 331094	 Pseudomonas oryzae 52.65% Pseudomonas fluorescens 28.96% Pseudomonas putida 12.16% Bacillus subtilis 2.48% Pantoea ananatis 0.96% Pseudomonas colesporium 0.42% Pantoea agglomerans 0.29% Bacillus velezensis 0.28% Pseudomonas syringae 0.25% Fusarium fujikuroi 0.12% (Remaining) 0.43%	18.27	ND
 PC-E-2-2-B_001.fastq.gz	 Pseudomonas putida Bacteria	https://app.onecodex.com/analysis/public/a903e091a6466c Assembly PC_E_2_2_B # contigs (>= 25000 bp) 26 # contigs (>= 50000 bp) 22 Total length (>= 25000 bp) 5144265 Total length (>= 50000 bp) 4994836 # contigs 193 Largest contig 741,966 Total length 5,420,450 GC (%) 65.11 NS0 296382	 Pseudomonas putida 38.10% Pseudomonas psychrotolerans 37.20% Pseudomonas oryzae 19.93% Pseudomonas colesporium 3.19% Pantoea agglomerans 0.54% Pseudomonas sp. UBA6562 0.42% Bacillus velezensis 0.34% Enterobacteriaceae 0.11% Pantoea ananatis 0.09% Klebsiella pneumoniae 0.08% (Remaining) 0.20%	17.04	13.48
 PC-E-2-2-C_001.fastq.gz	 Pseudomonas colesporium Bacteria	https://app.onecodex.com/analysis/public/f49dca33c6d4887 Assembly PC_E_2_2_C # contigs (>= 25000 bp) 19 # contigs (>= 50000 bp) 17 Total length (>= 25000 bp) 4541770 Total length (>= 50000 bp) 4453469 # contigs 152 Largest contig 703,285 Total length 4,688,673 GC (%) 62.04 NS0 304475	 Pseudomonas colesporium 62.36% Pseudomonas oryzae 37.06% Pseudomonas cannabina 0.17% Pseudomonas sp. UBA6562 0.16% Pantoea agglomerans 0.15% Enterobacteriaceae 0.09% Bacillus velezensis 0.09% Fusarium fujikuroi 0.01% Penicillium oblongum 0.01% Alternaria alternata 0.01% (Remaining) 0.01%	14.41	ND
 PC-E-2-2-D_001.fastq.gz	 Pseudomonas fluorescens Bacteria	https://app.onecodex.com/analysis/public/8cfd9772d0848db Assembly PC_E_2_2_D # contigs (>= 25000 bp) 13 # contigs (>= 50000 bp) 12 Total length (>= 25000 bp) 4877180 Total length (>= 50000 bp) 4837473 # contigs 157 Largest contig 927,974 Total length 4,987,292 GC (%) 65.40 NS0 524142	 Pseudomonas fluorescens 97.58% Pseudomonas oryzae 7.88% Pseudomonas colesporium 1.44% Pseudomonas sp. UBA6562 0.29% Pantoea agglomerans 0.29% Enterobacteriaceae 0.08% Bacillus velezensis 0.09% Fusarium fujikuroi 0.01% Penicillium oblongum 0.01% (Remaining) 0.02%	15.68	ND
 PC-E-2-2-E_001.fastq.gz	 Pantoea agglomerans Bacteria	https://app.onecodex.com/analysis/public/8efc080a9b64fb Assembly PC_E_2_2_E # contigs (>= 25000 bp) 24 # contigs (>= 50000 bp) 21 Total length (>= 25000 bp) 4641789 Total length (>= 50000 bp) 4537931 # contigs 306 Largest contig 641,210 Total length 4,925,766 GC (%) 53.80 NS0 247445	 Pantoea agglomerans 97.83% Pantoea vagans 10.45% Pantoea ananatis 1.23% Pantoea sp. UBA3707 0.29% Pseudomonas oryzae 0.12% Bacillus velezensis 0.06% Pseudomonas colesporium 0.05% Singella sp. SF-2015 0.03% Fusarium fujikuroi 0.02% Penicillium oblongum 0.01% (Remaining) 0.04%	15.64	ND
 PC-D-2-2-B_001.fastq.gz	 Pseudomonas sp. UBA6562 Bacteria	https://app.onecodex.com/analysis/public/d6fc10db5916448f Assembly PC_D_2_2_B # contigs (>= 25000 bp) 12 # contigs (>= 50000 bp) 11 Total length (>= 25000 bp) 4529755 Total length (>= 50000 bp) 4498749 # contigs 1111 Largest contig 1,001,953 Total length 5,422,716 GC (%) 61.61 NS0 419050	 Pseudomonas sp. UBA6562 86.24% Pseudomonas putida 8.84% Bacillus subtilis 2.17% Pantoea ananatis 0.54% Pantoea agglomerans 0.50% Bacillus velezensis 0.32% Pseudomonas colesporium 0.28% Pantoea vagans 0.16% Pseudomonas oryzae 0.12% Fusarium fujikuroi 0.10% (Remaining) 0.40%	18.20	ND

Whole genome sequencing of colonies derived from cannabis flowers and the impact of media selection on benchmarking total yeast and mold detection tools



Whole genome sequencing of colonies derived from cannabis flowers and the impact of media selection on benchmarking total yeast and mold detection tools



Colony Image	Assembly Coverage (Y) Large->Small Contigs (X)	Assembly Statistics (Quast 5.0)	OneCodex Kmer Classification	TAC Ct	TYM Ct
 PC-D-2-1-C1_001.fastq.gz	 Trichoderma harzianum Fungi	https://app.onecodex.com/analysis/public/312727a8313b4b68 Assembly PC_D_2_1_C1 # contigs (>= 25000 bp) 340 # contigs (>= 50000 bp) 37 Total length (>= 25000 bp) 12361853 Total length (>= 50000 bp) 2316186 # contigs 5131 Largest contig 131,469 Total length 41,502,375 GC (%) 46.82 N50 16662	OneCodex Kmer Classification Trichoderma harzianum 71.5% Trichoderma arbidonenseum 5.34% Trichoderma guihouense 5.19% Pantoea agglomerans 2.28% Pseudomonas ananatis 1.83% Pantoea ananatis 1.45% Bacillus velezensis 1.39% Pseudomonas coelesterrum 0.70% Pseudomonas cryohabitaris 0.37% Fusarium fujispori 0.31% (Remaining) 1.86%	32.17	26.16
 PC-D-2-1-C2_001.fastq.gz	 Trichoderma harzianum Fungi	https://app.onecodex.com/analysis/public/7a69f1b1132462e Assembly PC_D_2_1_C2 # contigs (>= 25000 bp) 12 # contigs (>= 50000 bp) 0 Total length (>= 25000 bp) 346836 Total length (>= 50000 bp) 0 # contigs 12671 Largest contig 34,730 Total length 41,758,812 GC (%) 47.09 N50 5262	OneCodex Kmer Classification Trichoderma harzianum 75.5% Trichoderma arbidonenseum 5.61% Trichoderma guihouense 5.48% Pantoea agglomerans 3.16% Pseudomonas cryohabitaris 1.85% Pantoea ananatis 1.26% Bacillus velezensis 1.28% Pseudomonas coelesterrum 1.09% Pseudomonas psophobolensis 1.02% Pseudomonas putida 0.62% (Remaining) 3.09%	30.47	24.95
 PC-D-2-1-C3_001.fastq.gz	 Trichoderma harzianum Fungi	https://app.onecodex.com/analysis/public/58726a182b94d85 Assembly PC_D_2_1_C3 # contigs (>= 25000 bp) 251 # contigs (>= 50000 bp) 19 Total length (>= 25000 bp) 8677054 Total length (>= 50000 bp) 1137009 # contigs 6487 Largest contig 94,369 Total length 41,771,792 GC (%) 46.85 N50 13238	OneCodex Kmer Classification Trichoderma harzianum 71.62% Trichoderma arbidonenseum 5.61% Trichoderma guihouense 5.49% Pantoea agglomerans 3.16% Pseudomonas cryohabitaris 2.18% Pantoea ananatis 1.52% Bacillus velezensis 1.39% Pseudomonas coelesterrum 1.24% Pseudomonas psophobolensis 0.71% Pseudomonas sp. UBA562 0.51% (Remaining) 2.71%	32.18	23.22
 PC-E-2-1-B_001.fastq.gz	 Pseudomonas oryzae Bacteria	https://app.onecodex.com/analysis/public/ba62d76d564298e Assembly PC_E_2_1_B # contigs (>= 25000 bp) 20 # contigs (>= 50000 bp) 19 Total length (>= 25000 bp) 4933410 Total length (>= 50000 bp) 4895124 # contigs 102 Largest contig 570,530 Total length 5,032,442 GC (%) 65.69 N50 356575	OneCodex Kmer Classification Pseudomonas oryzae 54.17% Pseudomonas fluorescens 40.34% Pseudomonas coelesterrum 4.81% Pseudomonas sp. UBA562 3.32% Pantoea agglomerans 3.26% Enterobacteriaceae 0.97% Bacillus velezensis 0.96% Enterobacteriaceae bacterium UBA473 0.95% Pantoea ananatis 0.93% Penicillium citrinum 0.92% (Remaining) 3.94%	16.23	ND
 PC-E-2-1-D_001.fastq.gz	 Pseudomonas coelesterrum Bacteria	https://app.onecodex.com/analysis/public/ba62d76d564298e Assembly PC_E_2_1_D # contigs (>= 25000 bp) 14 # contigs (>= 50000 bp) 14 Total length (>= 25000 bp) 4591239 Total length (>= 50000 bp) 4591239 # contigs 113 Largest contig 824,814 Total length 4,664,689 GC (%) 62.21 N50 328847	OneCodex Kmer Classification Pseudomonas coelesterrum 66.59% Pseudomonas fluorescens 29.16% Pseudomonas oryzae 3.83% Pantoea agglomerans 2.18% Pseudomonas caroliniana 0.07% Pseudomonas sp. UBA562 0.06% Enterobacteriaceae 0.03% Bacillus velezensis 0.03% Pantoea ananatis 0.01% Fusarium fujispori 0.01% (Remaining) 0.01%	15.31	ND
 PC-E-2-1-E_001.fastq.gz	 Penicillium citrinum Fungi	https://app.onecodex.com/analysis/public/8c4f1a8811cc44e Assembly PC_E_2_1_E # contigs (>= 25000 bp) 185 # contigs (>= 50000 bp) 142 Total length (>= 25000 bp) 30190784 Total length (>= 50000 bp) 28692835 # contigs 706 Largest contig 908,148 Total length 32,026,679 GC (%) 46.40 N50 226531	OneCodex Kmer Classification Penicillium citrinum 97.72% Pantoea agglomerans 0.52% Pantoea ananatis 0.38% Bacillus velezensis 0.29% Pseudomonas oryzae 0.25% Pseudomonas putida 0.24% Pseudomonas coelesterrum 0.22% Clostracter freundii 0.11% Fusarium fujispori 0.08% Alternaria alternata 0.06% (Remaining) 0.13%	31.91	15.45

Whole genome sequencing of colonies derived from cannabis flowers and the impact of media selection on benchmarking total yeast and mold detection tools



Table 3. PDA without Chloramphenicol: Colony Image (Left), Assembly sequence coverage (Y) compared with contig Length (X) where the contigs are sorted largest to smallest from left to right (Mid-Left). Assembly statistics calculated with Quast 5.0 (Mid Right). OneCodex speciated kmer count (Right).

Colony Image	Assembly Coverage (Y) Large->Small Contigs (X)	Assembly Statistics (Quast 5.0)	OneCodex Kmer Classification	TAC Ct	TYM Ct
		Assembly P_D_2_1_A # contigs (>= 25000 bp) 114 # contigs (>= 50000 bp) 49 Total length (>= 25000 bp) 7814145 Total length (>= 50000 bp) 5670116 # contigs 2129 Largest contig 400,518 Total length 15,498,447 GC (%) 53.80 N50 25286	OneCodex Kmer Classification Name Estimated Abundance Pantoea ananatis 51.59% Pantoea agglomerans 36.89% Pantoea vagans 5.66% Pantoea alli 4.42% Shigella dysenteriae 1.26% Pseudomonas oryzae 0.10% Bacillus velezensis 0.06% Pseudomonas coelesterrum 0.06% Pseudomonas psychrotolerans 0.03% Fusarium fujikuroi 0.03% (Remaining) 0.08%	15.99	ND
		Assembly P_D_2_1_B # contigs (>= 25000 bp) 26 # contigs (>= 50000 bp) 23 Total length (>= 25000 bp) 4810440 Total length (>= 50000 bp) 4722027 # contigs 886 Largest contig 514,181 Total length 5,456,028 GC (%) 54.22 N50 249921	OneCodex Kmer Classification Name Estimated Abundance Pantoea agglomerans 69.56% Pantoea ananatis 19.36% Pantoea vagans 5.68% Clostridium freundii 3.47% Pseudomonas oryzae 2.23% Enterobacter cloacae 0.23% Pseudomonas putida 0.19% Bacillus velezensis 0.18% Pseudomonas coelesterrum 0.17% Deffia aculeiformis 0.16% (Remaining) 0.47%	18.51	ND
		Assembly P_D_2_1_C # contigs (>= 25000 bp) 26 # contigs (>= 50000 bp) 20 Total length (>= 25000 bp) 4785028 Total length (>= 50000 bp) 4601583 # contigs 285 Largest contig 557,633 Total length 5,026,841 GC (%) 54.46 N50 306723	OneCodex Kmer Classification Name Estimated Abundance Pantoea agglomerans 75.20% Pantoea ananatis 15.31% Pantoea vagans 7.52% Pseudomonas oryzae 0.27% Enterobacter cloacae 0.23% Enterobacteriaceae bacterium UBA4753 0.22% Bacillus velezensis 0.21% Enterobacteriaceae bacterium UBA3398 0.17% Pseudomonas coelesterrum 0.16% Pseudomonas psychrotolerans 0.14% (Remaining) 0.56%	17.77	ND
		Assembly P_D_2_1_D # contigs (>= 25000 bp) 95 # contigs (>= 50000 bp) 28 Total length (>= 25000 bp) 4277137 Total length (>= 50000 bp) 2043450 # contigs 3474 Largest contig 169,384 Total length 11,841,942 GC (%) 53.74 N50 14877	OneCodex Kmer Classification Name Estimated Abundance Pantoea agglomerans 61.86% Pantoea ananatis 18.79% Pantoea vagans 7.14% Pantoea alli 2.93% Shigella sp. SF-2015 0.32% Pseudomonas montali 0.19% Bacillus velezensis 0.16% Enterobacter cloacae 0.14% Clostridium freundii 0.13% Pseudomonas coelesterrum 0.08% (Remaining) 0.25%	17.17	ND
		Assembly P_D_2_1_E # contigs (>= 25000 bp) 72 # contigs (>= 50000 bp) 28 Total length (>= 25000 bp) 3689171 Total length (>= 50000 bp) 2143023 # contigs 27681 Largest contig 177,453 Total length 29,272,617 GC (%) 51.77 N50 1032	OneCodex Kmer Classification Name Readcount (% of classified reads) Pantoea agglomerans 65481 (5.11%) Alternaria alternata 44258 (4.41%) Pantoea sp. all 27932 (2.7%) Pantoea sp. MB13 27854 (2.69%) Pantoea sp. UBA3398 20299 (2.02%) Alternaria sp. MCI 6993 (0.69%) Alternaria arborescens 5764 (0.57%) Pantoea sp. CF5ND3390 3165 (0.31%) Penicillium rotundum 2369 (0.23%) Fusarium fujikuroi 2343 (0.23%) (Remaining) 7868 (0.78%)	21.03	27.40
		Assembly P_D_2_2_A # contigs (>= 25000 bp) 94 # contigs (>= 50000 bp) 30 Total length (>= 25000 bp) 4422074 Total length (>= 50000 bp) 2253010 # contigs 5203 Largest contig 228,144 Total length 13,867,458 GC (%) 55.14 N50 10968	OneCodex Kmer Classification Name Estimated Abundance Pantoea agglomerans 72.20% Pantoea vagans 13.60% Pantoea alli 12.84% Pseudomonas coelesterrum 0.44% Pseudomonas putida 0.19% Shigella sp. SF-2015 0.18% Pseudomonas oryzae 0.14% Kosakonia coxeni 0.12% Bacillus velezensis 0.12% Pseudomonas psychrotolerans 0.05% (Remaining) 0.11%	17.18	ND

Whole genome sequencing of colonies derived from cannabis flowers and the impact of media selection on benchmarking total yeast and mold detection tools



Colony Image	Assembly Coverage (Y) Large->Small Contigs (X)	Assembly Statistics (Quast 5.0)	OneCodex Kmer Classification	TAC Ct	TYM Ct																								
		https://app.onecodex.com/analysis/public/c7adba5b63645be Assembly P_D_2_2_B # contigs (>= 25000 bp) 16 # contigs (>= 50000 bp) 15 Total length (>= 25000 bp) 4058977 Total length (>= 50000 bp) 4010903 # contigs 309 Largest contig 1,085,044 Total length 4,404,170 GC (%) 45.59 N50 462229	<table border="1"> <thead> <tr> <th>Name</th> <th>Estimated Abundance</th> </tr> </thead> <tbody> <tr><td>Bacillus velezensis</td><td>83.5%</td></tr> <tr><td>Bacillus amyloquelificans</td><td>14.05%</td></tr> <tr><td>Bacillus citreus</td><td>2.22%</td></tr> <tr><td>Bacillus subtilis</td><td>0.34%</td></tr> <tr><td>Pantoea agglomerans</td><td>0.21%</td></tr> <tr><td>Pseudomonas aeruginosa</td><td>0.19%</td></tr> <tr><td>Pantoea ananatis</td><td>0.13%</td></tr> <tr><td>Pseudomonas coelesterrum</td><td>0.04%</td></tr> <tr><td>Bacillus atrophaeus</td><td>0.04%</td></tr> <tr><td>Pseudomonas oryzae</td><td>0.03%</td></tr> <tr><td>(Remaining)</td><td>0.22%</td></tr> </tbody> </table>	Name	Estimated Abundance	Bacillus velezensis	83.5%	Bacillus amyloquelificans	14.05%	Bacillus citreus	2.22%	Bacillus subtilis	0.34%	Pantoea agglomerans	0.21%	Pseudomonas aeruginosa	0.19%	Pantoea ananatis	0.13%	Pseudomonas coelesterrum	0.04%	Bacillus atrophaeus	0.04%	Pseudomonas oryzae	0.03%	(Remaining)	0.22%	21.59	ND
Name	Estimated Abundance																												
Bacillus velezensis	83.5%																												
Bacillus amyloquelificans	14.05%																												
Bacillus citreus	2.22%																												
Bacillus subtilis	0.34%																												
Pantoea agglomerans	0.21%																												
Pseudomonas aeruginosa	0.19%																												
Pantoea ananatis	0.13%																												
Pseudomonas coelesterrum	0.04%																												
Bacillus atrophaeus	0.04%																												
Pseudomonas oryzae	0.03%																												
(Remaining)	0.22%																												
		https://app.onecodex.com/analysis/public/4b4c8477b184d400 Assembly P_D_2_2_C # contigs (>= 25000 bp) 62 # contigs (>= 50000 bp) 27 Total length (>= 25000 bp) 3905762 Total length (>= 50000 bp) 2679800 # contigs 20418 Largest contig 268,520 Total length 21,304,064 GC (%) 52.23	<table border="1"> <thead> <tr> <th>Name</th> <th>Readcount (% of classified reads)</th> </tr> </thead> <tbody> <tr><td>Pantoea agglomerans</td><td>3178930 (25.18%)</td></tr> <tr><td>Alternaria alternata</td><td>306673 (2.27%)</td></tr> <tr><td>Pantoea sp. CFAN032000</td><td>237901 (1.99%)</td></tr> <tr><td>Pseudomonas oryzae</td><td>42383 (0.32%)</td></tr> <tr><td>Pantoea sp. aff.</td><td>37064 (0.31%)</td></tr> <tr><td>Pantoea sp. MML3</td><td>36178 (0.3%)</td></tr> <tr><td>Alternaria sp. MG1</td><td>34273 (0.29%)</td></tr> <tr><td>Pantoea sp. UBA3896</td><td>33750 (0.28%)</td></tr> <tr><td>Alternaria arborescens</td><td>24987 (0.21%)</td></tr> <tr><td>Pantoea vagans</td><td>19289 (0.16%)</td></tr> <tr><td>(Remaining)</td><td>37818 (0.32%)</td></tr> </tbody> </table>	Name	Readcount (% of classified reads)	Pantoea agglomerans	3178930 (25.18%)	Alternaria alternata	306673 (2.27%)	Pantoea sp. CFAN032000	237901 (1.99%)	Pseudomonas oryzae	42383 (0.32%)	Pantoea sp. aff.	37064 (0.31%)	Pantoea sp. MML3	36178 (0.3%)	Alternaria sp. MG1	34273 (0.29%)	Pantoea sp. UBA3896	33750 (0.28%)	Alternaria arborescens	24987 (0.21%)	Pantoea vagans	19289 (0.16%)	(Remaining)	37818 (0.32%)	20.70	25.41
Name	Readcount (% of classified reads)																												
Pantoea agglomerans	3178930 (25.18%)																												
Alternaria alternata	306673 (2.27%)																												
Pantoea sp. CFAN032000	237901 (1.99%)																												
Pseudomonas oryzae	42383 (0.32%)																												
Pantoea sp. aff.	37064 (0.31%)																												
Pantoea sp. MML3	36178 (0.3%)																												
Alternaria sp. MG1	34273 (0.29%)																												
Pantoea sp. UBA3896	33750 (0.28%)																												
Alternaria arborescens	24987 (0.21%)																												
Pantoea vagans	19289 (0.16%)																												
(Remaining)	37818 (0.32%)																												
		https://app.onecodex.com/analysis/public/0044a7c2b05d4295 Assembly P_E_2_1_A # contigs (>= 25000 bp) 23 # contigs (>= 50000 bp) 19 Total length (>= 25000 bp) 4758256 Total length (>= 50000 bp) 4623610 # contigs 3340 Largest contig 1,009,934 Total length 7,358,440 GC (%) 53.86 N50 180723	<table border="1"> <thead> <tr> <th>Name</th> <th>Estimated Abundance</th> </tr> </thead> <tbody> <tr><td>Pantoea ananatis</td><td>88.74%</td></tr> <tr><td>Pantoea agglomerans</td><td>7.80%</td></tr> <tr><td>Pantoea vagans</td><td>2.34%</td></tr> <tr><td>Enterobacteriaceae bacterium UBA5138</td><td>0.26%</td></tr> <tr><td>Enterobacteriaceae bacterium UBA3398</td><td>0.17%</td></tr> <tr><td>Pseudomonas oryzae</td><td>0.13%</td></tr> <tr><td>Enterobacteriaceae bacterium UBA4753</td><td>0.13%</td></tr> <tr><td>Kosakonia cowanii</td><td>0.13%</td></tr> <tr><td>Bacillus velezensis</td><td>0.07%</td></tr> <tr><td>Pseudomonas coelesterrum</td><td>0.06%</td></tr> <tr><td>(Remaining)</td><td>0.14%</td></tr> </tbody> </table>	Name	Estimated Abundance	Pantoea ananatis	88.74%	Pantoea agglomerans	7.80%	Pantoea vagans	2.34%	Enterobacteriaceae bacterium UBA5138	0.26%	Enterobacteriaceae bacterium UBA3398	0.17%	Pseudomonas oryzae	0.13%	Enterobacteriaceae bacterium UBA4753	0.13%	Kosakonia cowanii	0.13%	Bacillus velezensis	0.07%	Pseudomonas coelesterrum	0.06%	(Remaining)	0.14%	15.81	ND
Name	Estimated Abundance																												
Pantoea ananatis	88.74%																												
Pantoea agglomerans	7.80%																												
Pantoea vagans	2.34%																												
Enterobacteriaceae bacterium UBA5138	0.26%																												
Enterobacteriaceae bacterium UBA3398	0.17%																												
Pseudomonas oryzae	0.13%																												
Enterobacteriaceae bacterium UBA4753	0.13%																												
Kosakonia cowanii	0.13%																												
Bacillus velezensis	0.07%																												
Pseudomonas coelesterrum	0.06%																												
(Remaining)	0.14%																												
		https://app.onecodex.com/analysis/public/15a94f72d73142ce Assembly PC_E_2_1_B # contigs (>= 25000 bp) 20 # contigs (>= 50000 bp) 19 Total length (>= 25000 bp) 4933410 Total length (>= 50000 bp) 4895124 # contigs 102 Largest contig 570,530 Total length 5,032,442 GC (%) 65.69 N50 356575	<table border="1"> <thead> <tr> <th>Name</th> <th>Estimated Abundance</th> </tr> </thead> <tbody> <tr><td>Bacillus velezensis</td><td>80.22%</td></tr> <tr><td>Bacillus amyloquelificans</td><td>15.29%</td></tr> <tr><td>Bacillus citreus</td><td>3.58%</td></tr> <tr><td>Pantoea agglomerans</td><td>0.23%</td></tr> <tr><td>Pantoea ananatis</td><td>0.13%</td></tr> <tr><td>Pseudomonas aeruginosa</td><td>0.13%</td></tr> <tr><td>Pseudomonas stutzeri</td><td>0.06%</td></tr> <tr><td>Bacillus atrophaeus</td><td>0.06%</td></tr> <tr><td>Pseudomonas coelesterrum</td><td>0.06%</td></tr> <tr><td>Fusarium fujikuroi</td><td>0.03%</td></tr> <tr><td>(Remaining)</td><td>0.20%</td></tr> </tbody> </table>	Name	Estimated Abundance	Bacillus velezensis	80.22%	Bacillus amyloquelificans	15.29%	Bacillus citreus	3.58%	Pantoea agglomerans	0.23%	Pantoea ananatis	0.13%	Pseudomonas aeruginosa	0.13%	Pseudomonas stutzeri	0.06%	Bacillus atrophaeus	0.06%	Pseudomonas coelesterrum	0.06%	Fusarium fujikuroi	0.03%	(Remaining)	0.20%	21.15	ND
Name	Estimated Abundance																												
Bacillus velezensis	80.22%																												
Bacillus amyloquelificans	15.29%																												
Bacillus citreus	3.58%																												
Pantoea agglomerans	0.23%																												
Pantoea ananatis	0.13%																												
Pseudomonas aeruginosa	0.13%																												
Pseudomonas stutzeri	0.06%																												
Bacillus atrophaeus	0.06%																												
Pseudomonas coelesterrum	0.06%																												
Fusarium fujikuroi	0.03%																												
(Remaining)	0.20%																												
		https://app.onecodex.com/analysis/public/364959307b14446 Assembly P_E_2_1_C # contigs (>= 25000 bp) 98 # contigs (>= 50000 bp) 52 Total length (>= 25000 bp) 7299781 Total length (>= 50000 bp) 5828972 # contigs 7153 Largest contig 306,404 Total length 14,692,225 GC (%) 54.09 N50 24878	<table border="1"> <thead> <tr> <th>Name</th> <th>Estimated Abundance</th> </tr> </thead> <tbody> <tr><td>Pantoea agglomerans</td><td>28.17%</td></tr> <tr><td>Enterobacteriaceae bacterium UBA4753</td><td>19.52%</td></tr> <tr><td>Enterobacteriaceae bacterium UBA3398</td><td>19.38%</td></tr> <tr><td>Kosakonia cowanii</td><td>13.33%</td></tr> <tr><td>Pantoea ananatis</td><td>10.71%</td></tr> <tr><td>Pantoea vagans</td><td>5.74%</td></tr> <tr><td>Bacillus velezensis</td><td>1.73%</td></tr> <tr><td>Alternaria alternata</td><td>0.57%</td></tr> <tr><td>Alternaria arborescens</td><td>0.52%</td></tr> <tr><td>Pseudomonas oryzae</td><td>0.12%</td></tr> <tr><td>(Remaining)</td><td>0.22%</td></tr> </tbody> </table>	Name	Estimated Abundance	Pantoea agglomerans	28.17%	Enterobacteriaceae bacterium UBA4753	19.52%	Enterobacteriaceae bacterium UBA3398	19.38%	Kosakonia cowanii	13.33%	Pantoea ananatis	10.71%	Pantoea vagans	5.74%	Bacillus velezensis	1.73%	Alternaria alternata	0.57%	Alternaria arborescens	0.52%	Pseudomonas oryzae	0.12%	(Remaining)	0.22%	21.83	27.03
Name	Estimated Abundance																												
Pantoea agglomerans	28.17%																												
Enterobacteriaceae bacterium UBA4753	19.52%																												
Enterobacteriaceae bacterium UBA3398	19.38%																												
Kosakonia cowanii	13.33%																												
Pantoea ananatis	10.71%																												
Pantoea vagans	5.74%																												
Bacillus velezensis	1.73%																												
Alternaria alternata	0.57%																												
Alternaria arborescens	0.52%																												
Pseudomonas oryzae	0.12%																												
(Remaining)	0.22%																												
		https://app.onecodex.com/analysis/public/ed70a63714574eaf Assembly P_E_2_1_D # contigs (>= 25000 bp) 86 # contigs (>= 50000 bp) 57 Total length (>= 25000 bp) 9215888 Total length (>= 50000 bp) 8188817 # contigs 3291 Largest contig 530,633 Total length 12,137,385 GC (%) 53.49 N50 105,803	<table border="1"> <thead> <tr> <th>Name</th> <th>Readcount (% of classified reads)</th> </tr> </thead> <tbody> <tr><td>Enterobacteriaceae bacterium UBA4753</td><td>1466340 (15.8%)</td></tr> <tr><td>Enterobacteriaceae bacterium UBA3398</td><td>1309108 (14.04%)</td></tr> <tr><td>Kosakonia cowanii</td><td>1060939 (11.4%)</td></tr> <tr><td>Pantoea agglomerans</td><td>1008425 (10.8%)</td></tr> <tr><td>Pantoea sp. CFAN032000</td><td>83510 (0.9%)</td></tr> <tr><td>Mucor circinelliae</td><td>52200 (0.56%)</td></tr> <tr><td>Mucor antarcticus</td><td>46765 (0.5%)</td></tr> <tr><td>Pantoea aff.</td><td>37504 (0.4%)</td></tr> <tr><td>Penicillium citrinum</td><td>20800 (0.22%)</td></tr> <tr><td>Fusarium fujikuroi</td><td>20740 (0.22%)</td></tr> <tr><td>(Remaining)</td><td>118469 (12.7%)</td></tr> </tbody> </table>	Name	Readcount (% of classified reads)	Enterobacteriaceae bacterium UBA4753	1466340 (15.8%)	Enterobacteriaceae bacterium UBA3398	1309108 (14.04%)	Kosakonia cowanii	1060939 (11.4%)	Pantoea agglomerans	1008425 (10.8%)	Pantoea sp. CFAN032000	83510 (0.9%)	Mucor circinelliae	52200 (0.56%)	Mucor antarcticus	46765 (0.5%)	Pantoea aff.	37504 (0.4%)	Penicillium citrinum	20800 (0.22%)	Fusarium fujikuroi	20740 (0.22%)	(Remaining)	118469 (12.7%)	14.10	ND
Name	Readcount (% of classified reads)																												
Enterobacteriaceae bacterium UBA4753	1466340 (15.8%)																												
Enterobacteriaceae bacterium UBA3398	1309108 (14.04%)																												
Kosakonia cowanii	1060939 (11.4%)																												
Pantoea agglomerans	1008425 (10.8%)																												
Pantoea sp. CFAN032000	83510 (0.9%)																												
Mucor circinelliae	52200 (0.56%)																												
Mucor antarcticus	46765 (0.5%)																												
Pantoea aff.	37504 (0.4%)																												
Penicillium citrinum	20800 (0.22%)																												
Fusarium fujikuroi	20740 (0.22%)																												
(Remaining)	118469 (12.7%)																												
		https://app.onecodex.com/analysis/public/b283af48bc0d4e6f Assembly P_E_2_2_A # contigs (>= 25000 bp) 24 # contigs (>= 50000 bp) 19 Total length (>= 25000 bp) 4739071 Total length (>= 50000 bp) 4550105 # contigs 120 Largest contig 721,879 Total length 5,007,137 GC (%) 53.36 N50 347,627	<table border="1"> <thead> <tr> <th>Name</th> <th>Estimated Abundance</th> </tr> </thead> <tbody> <tr><td>Pantoea ananatis</td><td>91.91%</td></tr> <tr><td>Pantoea agglomerans</td><td>7.93%</td></tr> <tr><td>Pseudomonas oryzae</td><td>0.06%</td></tr> <tr><td>Pseudomonas coelesterrum</td><td>0.02%</td></tr> <tr><td>Bacillus velezensis</td><td>0.02%</td></tr> <tr><td>Citrobacter freundii</td><td>0.02%</td></tr> <tr><td>Fusarium fujikuroi</td><td>0.01%</td></tr> <tr><td>Penicillium citrinum</td><td>0.01%</td></tr> <tr><td>Pseudomonas sp. UBA5062</td><td>0.01%</td></tr> </tbody> </table>	Name	Estimated Abundance	Pantoea ananatis	91.91%	Pantoea agglomerans	7.93%	Pseudomonas oryzae	0.06%	Pseudomonas coelesterrum	0.02%	Bacillus velezensis	0.02%	Citrobacter freundii	0.02%	Fusarium fujikuroi	0.01%	Penicillium citrinum	0.01%	Pseudomonas sp. UBA5062	0.01%	13.72	ND				
Name	Estimated Abundance																												
Pantoea ananatis	91.91%																												
Pantoea agglomerans	7.93%																												
Pseudomonas oryzae	0.06%																												
Pseudomonas coelesterrum	0.02%																												
Bacillus velezensis	0.02%																												
Citrobacter freundii	0.02%																												
Fusarium fujikuroi	0.01%																												
Penicillium citrinum	0.01%																												
Pseudomonas sp. UBA5062	0.01%																												
		https://app.onecodex.com/analysis/public/e8578bf12aa4092 Assembly P_E_2_2_B # contigs (>= 25000 bp) 17 # contigs (>= 50000 bp) 15 Total length (>= 25000 bp) 4022231 Total length (>= 50000 bp) 3958214 # contigs 93 Largest contig 1,085,017 Total length 4,234,940 GC (%) 45.80 N50 310014	<table border="1"> <thead> <tr> <th>Name</th> <th>Estimated Abundance</th> </tr> </thead> <tbody> <tr><td>Bacillus velezensis</td><td>83.75%</td></tr> <tr><td>Bacillus amyloquelificans</td><td>13.39%</td></tr> <tr><td>Bacillus citreus</td><td>2.22%</td></tr> <tr><td>Bacillus subtilis</td><td>0.31%</td></tr> <tr><td>Pantoea ananatis</td><td>0.08%</td></tr> <tr><td>Pantoea agglomerans</td><td>0.06%</td></tr> <tr><td>Pseudomonas subtilis</td><td>0.04%</td></tr> <tr><td>Bacillus atrophaeus</td><td>0.04%</td></tr> <tr><td>Pseudomonas oryzae</td><td>0.03%</td></tr> <tr><td>Pseudomonas coelesterrum</td><td>0.03%</td></tr> <tr><td>(Remaining)</td><td>0.09%</td></tr> </tbody> </table>	Name	Estimated Abundance	Bacillus velezensis	83.75%	Bacillus amyloquelificans	13.39%	Bacillus citreus	2.22%	Bacillus subtilis	0.31%	Pantoea ananatis	0.08%	Pantoea agglomerans	0.06%	Pseudomonas subtilis	0.04%	Bacillus atrophaeus	0.04%	Pseudomonas oryzae	0.03%	Pseudomonas coelesterrum	0.03%	(Remaining)	0.09%	18.20	ND
Name	Estimated Abundance																												
Bacillus velezensis	83.75%																												
Bacillus amyloquelificans	13.39%																												
Bacillus citreus	2.22%																												
Bacillus subtilis	0.31%																												
Pantoea ananatis	0.08%																												
Pantoea agglomerans	0.06%																												
Pseudomonas subtilis	0.04%																												
Bacillus atrophaeus	0.04%																												
Pseudomonas oryzae	0.03%																												
Pseudomonas coelesterrum	0.03%																												
(Remaining)	0.09%																												

Figure 1. Plating images. DRBC (Top). PDA with CAMP (Middle). PDA no CAMP (Bottom).

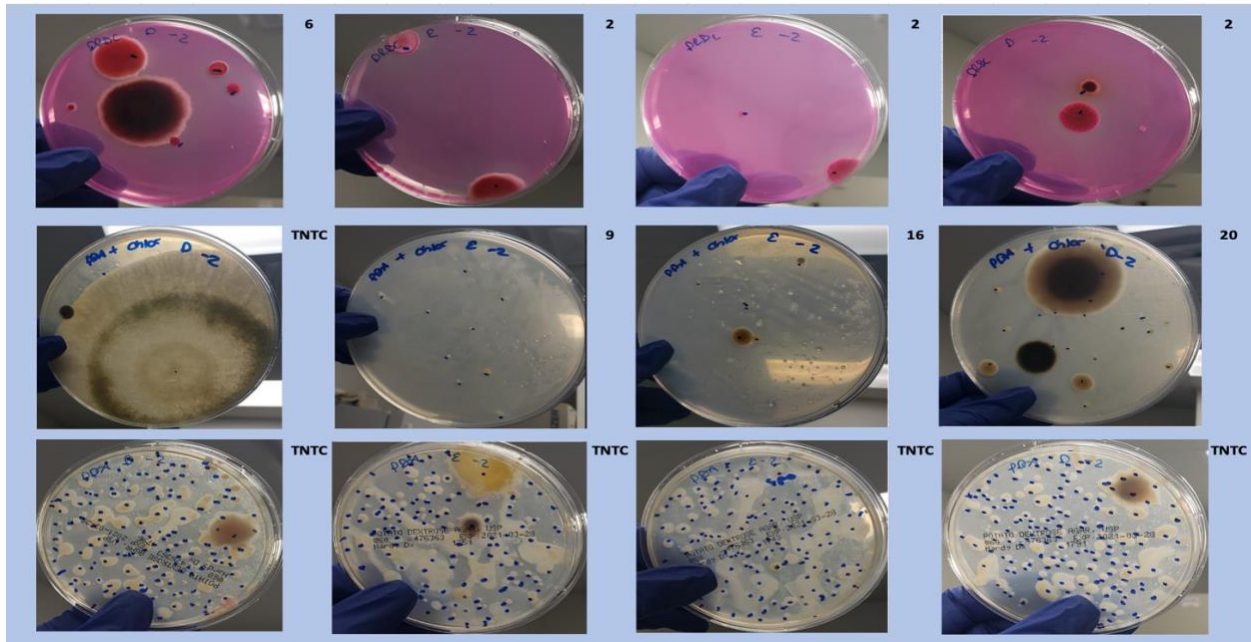


Figure 2. Summary heatmap of colony classification by whole genome sequencing.

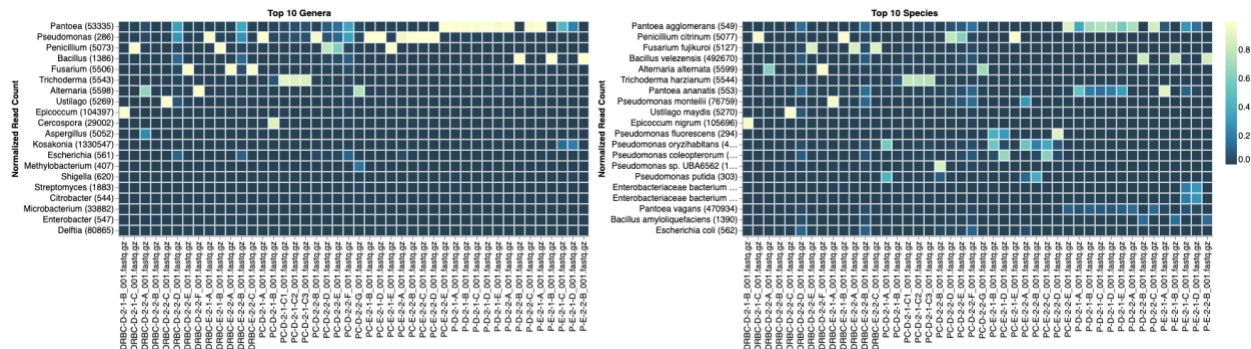


Table 4. Summary of colony forming unit classification:

	Bacteria	Fungi	Mixed
DRBC CAMP	3	12	2
PDA CAMP	10	9	1
PDA No CAMP	14	4	4

Figure 3: Simpson's Diversity Index

Simpson's diversity index is used to quantify the biodiversity of a habitat on a 0 to 1 scale. It takes into account the number of species present, as well as the relative abundance of each species. A diversity index of 1 represent infinite diversity where 0 reflects no diversity. DRBC plating demonstrates the lowest diversity. This is not surprising given DRBC contains 3 different selection agents (Dichloran, Rose Bengal and Chloramphenicol). While this limits bacterial contamination it also limits yeast and mold growth.

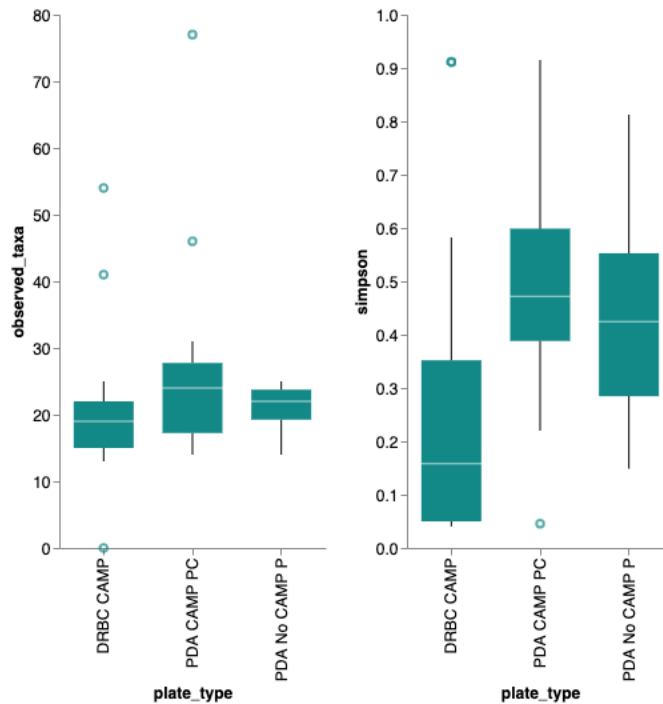


Table 5: cannabis samples plated on 3 different mediums.

Plating on different media's demonstrates a LOG scale difference in CFU with each plating medium. Sequencing can attribute only half of the colonies as bacteria on PDA with Chloramphenicol. This implies a 5-fold under counting of yeast and mold on DRBC.

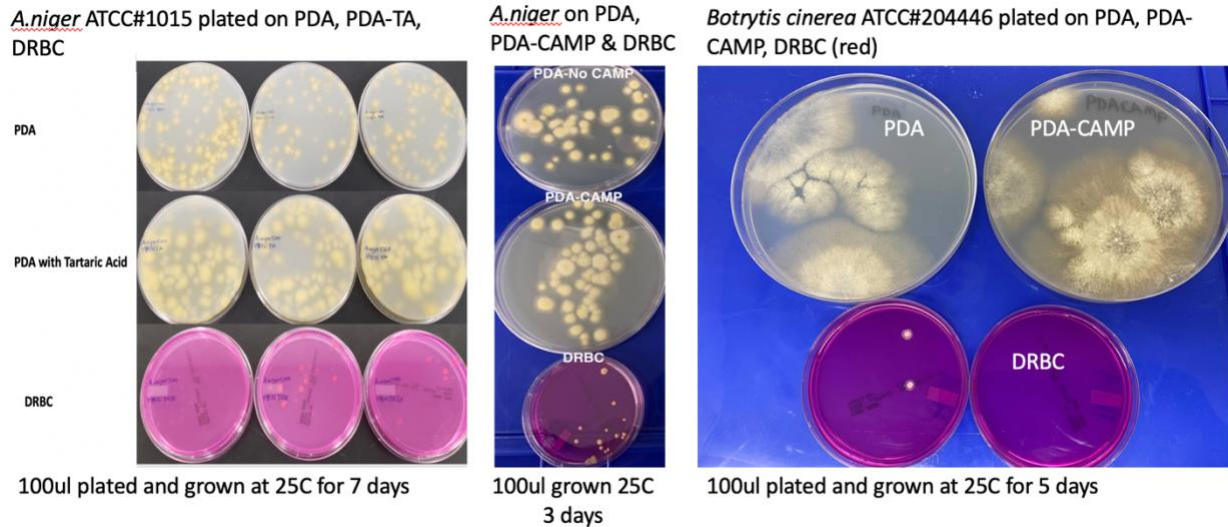
Sample	DRBC						Final Result CFU/g
	10 ⁻² CFU/g	10 ⁻² CFU/g	10 ⁻² CFU/g	10 ⁻³ CFU/g	10 ⁻³ CFU/g	10 ⁻³ CFU/g	
Low A	0	4	1	1	0	0	170
Low B	0	3	1	1	1	0	130
Low C	2	1	0	0	0	0	100
Low D	2	5	2	0	2	0	300
Low E	2	0	2	0	0	0	130
Average CFU/g							170
Sample	PDA with Chloramphenicol						Final Result CFU/g
	10 ⁻² CFU/g	10 ⁻² CFU/g	10 ⁻² CFU/g	10 ⁻³ CFU/g	10 ⁻³ CFU/g	10 ⁻³ CFU/g	
Low A	8	16	12	1	3	2	1200
Low B	8	12	8	1	0	3	930
Low C	13	19	13	1	2	1	1500
Low D	3	12	21	2	0	1	1200
Low E	9	7	4	0	3	3	670
Average CFU/g							1100
Sample	PDA						Final Result CFU/g
	10 ⁻² CFU/g	10 ⁻² CFU/g	10 ⁻² CFU/g	10 ⁻³ CFU/g	10 ⁻³ CFU/g	10 ⁻³ CFU/g	
Low A	127	133	124	32	32	21	14000
Low B	151	157	101	26	20	28	15000
Low C	TNTC	TNTC	TNTC	41	45	37	41000
Low D	147	141	123	32	26	26	15000
Low E	138	102	119	23	15	24	13000
Average CFU/g							20000

Table 6. Mono-culture evaluations

Organisms were ordered from ATCC and plated on 2 different medias (PDA and DRBC) to assess growth performance of the organisms in absence of cannabis background bacteria and matrix.

Species	ATCC Number	qPCR	DRBC Plating	PDA
<i>Aspergillus brasiliensis</i>	16404	Amp	5 Day Growth	5 Day Growth
<i>Aspergillus flavus</i>	9643	Amp	Less growth in 7 Days	5 Day Growth
<i>Aspergillus fumigatus</i>	204305	Amp	Less growth in 7 Days	5 Day Growth
<i>Aspergillus niger</i>	16888	Amp	Less growth in 7 Days	5 Day Growth
<i>Aspergillus terreus</i>	1012	Amp	Less growth in 7 Days	5 Day Growth
<i>Aspergillus tubigensis</i>	1004	Amp	Less growth in 7 Days	5 Day Growth
<i>Candida tropicalis</i>	13803	Amp	5 Day Growth	5 Day Growth
<i>Penicillium brevicompactum</i>	9056	Amp	Less growth than PDA	5 Day Growth
<i>Purpureocillium lilacinum</i>	10114	Partial Amp	5 Day Growth	5 Day Growth
<i>Rhizopus oryzae</i>	52748	Partial Amp	Less growth than PDA	5 Day Growth

Figure 4. *Aspergillus niger* and *Botrytis cinerea* monocultures plated on 3 different medias: Cultures were plated on PDA, PDA with selection (CAMP or Tartaric acid), and DRBC. Fewer colonies are consistently found on DRBC.



Conclusions

Microbial selection has a significant impact on the diversity index of microbes observed with whole genome sequencing. This has been noted in prior microbiome surveys in cannabis, where the act of culturing the microbes changes the representation of the microbiome as measured by qPCR and sequencing performed directly off of the flower and after culturing (McKernan et al. 2015; McKernan et al. 2016). In this study the DRBC selection reduces bacterial growth more than PDA with chloramphenicol but also reduces the fungal CFU 5-fold in the process. This has important implications for chloramphenicol sensitive cannabis endophytes like *Aspergillus*, *Pythium* and *Fusarium*. Both media types (PDA and DRBC) are referenced in the FDA Bacteriological Analytical Manual (FDA-BAM). States exclusively considering DRBC for ease of colony visualization should be aware of the species-specific sensitivities of using a single media type and consider species specific testing for such pathogenic organisms to complement a partial yeast and mold test offered from a single selection-based media. PCR based techniques can capture more organisms than DRBC alone as no selection is occurring given thorough cell lysis is achieved for qPCR analysis.

Plating also suffers from having a very limited dynamic range. Since it is difficult to count colonies when more than 100 colonies are present on a plate, multiple dilutions are often required to understand the full range of CFU counts one may encounter with a test that is attempting to quantitate 10,000 CFUs/gram. This results in multiplying diluted CFUs by 10, 100 and even a 1,000 fold to back estimate the total CFU count. In this scenario a single colony can swing the CFU count from passing to failing (9 colonies x 1,000 fold dilution vs 10 colonies at 1,000 fold dilution). Quantitative PCR has linear dynamic range over 5-6 orders of magnitude and no such multiplication is required. Thus, qPCR provides a more accurate itemization of actual CFUs counts.

In-vitro inclusion and exclusion testing with ITS3 qPCR on ATCC sourced organisms demonstrated over 96% inclusion (50 yeast and mold) and zero bacterial cross reactivity (30 bacteria) (Supplementary Table 1- Sheet TYM Inclusion & TYM Exclusion). *In-silico* analysis of ITS3 primer sequences, predicts over 1400 yeast and mold should amplify with the described ITS3 primer sequences. All plating media, even with 3 different forms of selection (DRBC), had bacterial contamination and each level of selection reduced fungal CFU counts.

Funding

This work was privately funding by Medicinal Genomics. No public funding declared.

Conflict Declaration

The authors are employees of Medicinal Genomics which manufacture qPCR reagents for Cannabis testing.

Author contributions

KJM- Wrote the manuscript, assisted in experimental design and sequencing analysis
YH, LTK, LZ, NH- isolated colonies, constructed sequencing libraries, performed plating and qPCR.

SM- Assembled genomes and performed sequencing analysis.

- Bal A, Agarwal AN, Das A, Suri V, Varma SC. 2010. Chronic necrotising pulmonary aspergillosis in a marijuana addict: a new cause of amyloidosis. *Pathology* **42**: 197-200.
- Day S, Lalitha P, Haug S, Fothergill AW, Cevallos V, Vijayakumar R, Prajna NV, Acharya NR, McLeod SD, Lietman TM. 2009. Activity of antibiotics against *Fusarium* and *Aspergillus*. *Br J Ophthalmol* **93**: 116-119.
- Gargani Y, Bishop P, Denning DW. 2011. Too many mouldy joints - marijuana and chronic pulmonary aspergillosis. *Mediterr J Hematol Infect Dis* **3**: e2011005.
- Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**: 1072-1075.
- Joseph MR, Al-Hakami AM, Assiry MM, Jamil AS, Assiry AM, Shaker MA, Hamid ME. 2015. In vitro anti-yeast activity of chloramphenicol: A preliminary report. *J Mycol Med* **25**: 17-22.
- Li D, Liu CM, Luo R, Sadakane K, Lam TW. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**: 1674-1676.
- Li D, Luo R, Liu CM, Leung CM, Ting HF, Sadakane K, Yamashita H, Lam TW. 2016. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* **102**: 3-11.
- McKernan K, Spangler J, Helbert Y, Lynch RC, Devitt-Lee A, Zhang L, Orphe W, Warner J, Foss T, Hudalla CJ et al. 2016. Metagenomic analysis of medicinal Cannabis samples; pathogenic bacteria, toxigenic fungi, and beneficial microbes grow in culture-based yeast and mold tests. *F1000Research* **5**: 2471.

- McKernan K, Spangler J, Zhang L, Tadigotla V, Helbert Y, Foss T, Smith D. 2015. Cannabis microbiome sequencing reveals several mycotoxic fungi native to dispensary grade Cannabis flowers. *F1000Research* **4**: 1422.
- Remington TL, Fuller J, Chiu I. 2015. Chronic necrotizing pulmonary aspergillosis in a patient with diabetes and marijuana use. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne* **187**: 1305-1308.
- Ruchlemer R, Amit-Kohn M, Raveh D, Hanus L. 2015. Inhaled medicinal cannabis and the immunocompromised patient. *Support Care Cancer* **23**: 819-822.
- Smith DG, Marchant R. 1968. Chloramphenicol inhibition of *Pythium ultimum* and *Rhodotorula glutinis*. *Arch Mikrobiol* **60**: 262-274.