METHOD ARTICLE

Whole genome sequencing of colonies derived from cannabis flowers and the impact of media selection on benchmarking total yeast and mold detection tools [version 1; peer review: 1 approved, 1 approved with reservations]

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

Research and Development, Medicinal Genomics, Beverly, Mass, 01915, USA

V1 First published: 21 Jul 2021, 10:624
<https://doi.org/10.12688/f1000research.53467.1>
 Latest published: 12 Aug 2021, 10:624
<https://doi.org/10.12688/f1000research.53467.2>

Abstract

Background:

Cannabis products are subjected to microbial testing for pathogenic fungi and bacteria. These testing requirements often rely on non-specific colony forming unit (CFU/g) specifications without clarity on which medium, selection or growth times are required. We performed whole genome sequencing to assess the specificity of colony forming units (CFU) derived from three different plating media: Potato Dextrose Agar (PDA), PDA with chloramphenicol and Dichloran Rose Bengal with chloramphenicol (DRBC).

Methods:

Colonies were isolated from each medium type and their whole genomes sequenced to identify the diversity of microbes present on each medium selection. Fungal Internal Transcribed Spacer (ITS3) and Bacterial 16S RNA(16S) quantitative polymerase chain reactions (qPCR) were performed, to correlate these CFUs with fungi- and bacterial-specific qPCR.

Results:

Each plating medium displayed a ten-fold difference in CFU counts. PDA with chloramphenicol showed the highest diversity and the highest concordance with whole genome sequencing. According to ITS3 and 16S qPCR confirmed with whole genome sequencing, DRBC under counted yeast and mold while PDA without chloramphenicol over counted CFUs due to bacterial growth without selection.

Conclusions:

Colony Forming Unit regulations lack specificity. Each medium

Open Peer Review

Approval Status

	1	2
version 2 (revision) 12 Aug 2021		 view
version 1 21 Jul 2021	 view	 view

1. **Cindy Orser**, CLIP Laboratories, San Diego, USA

2. **Zamir K. Punja** , Simon Fraser University, Burnaby, Canada

Any reports and responses or comments on the article can be found at the end of the article.

produces significant differences in CFU counts. These are further dependent on subjective interpretation, failure to culture most microbes, and poor selection between bacteria and fungi. Given the most pathogenic microbes found on cannabis are endophytes which culture fails to detect, molecular methods offer a solution to this long-standing quantification problem in the cannabis testing field.

Keywords

Cannabis, Total Yeast and Mold, Microbiome, Whole Genome Sequencing, qPCR

This article is included in the **Cell & Molecular Biology** gateway.

Corresponding author: Kevin McKernan (Kevin.McKernan@medicinalgenomics.com)

Author roles: **McKernan K:** Conceptualization, Data Curation, Methodology, Software, Supervision, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Helbert Y:** Data Curation, Investigation, Methodology, Project Administration, Supervision; **Kane L:** Data Curation, Investigation, Methodology; **Houde N:** Data Curation, Formal Analysis, Investigation, Methodology; **Zhang L:** Data Curation, Formal Analysis, Investigation, Methodology; **McLaughlin S:** Data Curation, Formal Analysis, Investigation, Software, Validation, Visualization, Writing – Original Draft Preparation

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

Grant information: This work was privately funded by Medicinal Genomics. No public funding declared.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2021 McKernan K *et al.* This is an open access article distributed under the terms of the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: McKernan K, Helbert Y, Kane L *et al.* **Whole genome sequencing of colonies derived from cannabis flowers and the impact of media selection on benchmarking total yeast and mold detection tools [version 1; peer review: 1 approved, 1 approved with reservations]** F1000Research 2021, 10:624 <https://doi.org/10.12688/f1000research.53467.1>

First published: 21 Jul 2021, 10:624 <https://doi.org/10.12688/f1000research.53467.1>

Introduction

Total yeast and mold testing are required in many states to test the safety of cannabis, prior to the sale of cannabis flowers and cannabis-infused products. Cannabis is an inhaled product, and cases of cannabis-transmitted *Aspergillosis* have been reported in the clinical literature (Bal *et al.*, 2010; Gargani *et al.*, 2011; McKernan *et al.*, 2015, 2016; Remington *et al.*, 2015; Ruchlemer *et al.*, 2015). Cannabis is a unique matrix, in that antibiotic cannabinoids can make up to 20% of the flowers' weight, and many fungi infecting cannabis are endophytes. Endophytes are not easily cultured from the plant without lysing open plant cell walls. The 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 processes required for testing. Cannabis flowers contain both bacteria and fungi, further complicating fungal quantification for colony forming units (CFU) that lack speciation. Antibiotic selections are often utilized to reduce background bacteria, but many of these antibiotics (e.g. chloramphenicol) inhibit the growth of the most pathogenic fungi found on cannabis (*Fusarium*, *Pythium* and *Aspergillus*) (Smith & 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 medium selection on surveying total yeast and mold on cannabis. Cured cannabis flowers were homogenized and tested on 3 different plating media. These data were compared to ITS3- and 16S-based qPCR and whole genome sequencing. To further complement these cannabis flower samples, organisms were acquired from the American Tissue Culture Collection (ATCC) and plated as pure monocultures on different plating media to confirm the differential growth on each medium.

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 fungal and bacterial CFUs, as measured using culture-based techniques with chloramphenicol selection. 10 grams of homogenized flower were soaked with 90 ml of Tryptic Soy Broth (TSB, Medicinal Genomics #420205) in a filtered Nasco Whirl-Pak bag (#B01385). Samples were homogenized by hand, and then 0.1 mL of solution plated onto three media (DRBC, PDA with chloramphenicol, PDA, at 1:100 dilution). Two additional dilutions were prepared (10 mL into 90 mL) and the 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 two modifications. Briefly, 1ml of homogenate from a Whirl-Pak bag was collected and briefly micro-centrifuged to enrich for live organisms. This pellet was resuspended in 200 µl ddH₂O and lysed with the addition of 12 µl 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 (Medicinal Genomics part #420206, McKernan *et al.*, 2015, 2016). 12.5 µl of MGC Lysis buffer was added, vortexed and incubated for 5 minutes at 25°C. Lysed samples were micro-centrifuged and 200 µl of supernatant was aspirated and added to 250 µl of Medicinal Genomics binding buffer (MGC part# 420001) for magnetic bead isolation. The samples were incubated with the Medicinal Genomics magnetic bead mixture for 10 minutes, magnetically separated and washed two times with 70% ethanol. The beads were dried at 37°C for 5 minutes to remove excess ethanol and eluted with 25 µl of ddH₂O. Quantitative PCR was performed using Medicinal Genomics PathoSEEK Total Yeast and Mold detection assay (MGC# 420103) and Medicinal Genomics PathoSEEK Total Aerobic Count Assay (MGC# 420106) according to the manufacturers' instructions on a BioRad CFX96 thermocycler.

DNA isolation from colonies for whole genome sequencing

A total of 45 colonies were picked with a pipette tip and introduced into 200 µl of ddH₂O with 12.5 µl of MGC TLP (MGC part #420206). TLP is a glucanase active at 37°C. Samples were digested for 30 minutes at 37°C and 12.5 µl of MGC Lysis buffer was added, vortexed and incubated for 5 minutes at 25°C. Lysed sample were micro-centrifuged and 200 µl of supernatant was aspirated and added to 250 µl 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 25 µl of ddH₂O.

Library construction for whole genome sequencing.

Fragmentation

Genomic DNA (gDNA) was quantified with a Qubit (Thermo Fisher Scientific) and normalized to reflect 4-8 ng/µl in 13 µl of TE buffer. Libraries were generated using enzymatic fragmentation with the NEB Ultra II kits (NEB part # E7103).

Briefly, 3.5 µl of 5X NEB fragmentation buffer and 1 µl of Ultra II fragmentation enzyme mix are added to 13 µl of DNA. This reaction was tip-mixed 10 times, vortexed, and quickly centrifuged. Fragmentation was performed in a BioRad CFX96 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

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

The master mix for ligation was prepared on ice using 0.75 µl of Agilent SureSelect Adaptor Oligo Mix, 0.5 µl of ddH₂O, 15 µl of NEB Ultra II Ligation Master Mix, 0.5 µl of Ligation enhancer (New England Biolabs) for a total reaction volume of 16.75 µl.

Ligation was performed by the addition of 16.75 µl of ligation master mix to the 17.5 µl Fragmentation/End Prep DNA reaction mixture, incubate for 15 minutes at 20°C. To purify excess adaptors and adaptor dimers, AMPure XP beads (Beckman Coulter #A63881) were vortexed at room temperature for resuspension and 16 µl (approximately 0.45X) of resuspended AMPure XP beads were added to the ligation reactions. This was well-mixed by pipetting up and down at least 10 times. The mixture was incubated for 5 minutes at 25°C. The PCR plate was placed on an appropriate magnetic stand (Medicinal Genomics #420202) to separate the beads from the supernatant. After the solution was clear (about 5 minutes), the supernatant was carefully removed and discarded. We were careful not to disturb the beads containing target DNA molecules. The magnetic beads were washed by adding 200 µl of 70% ethanol to the PCR plate while on the magnetic stand. Followed incubation at room temperature for 30 seconds, and then careful removal and discarding of the supernatant. The ethanol wash was repeated once for a total of 2 washes. Trace amounts of ethanol were removed. The beads were air dried for ~ 7 minutes while the PCR plate was on the magnetic stand with the lid open. The PCR plate was then removed from the magnet and target DNA eluted from the beads into 10 µl of H₂O, then 9 µl of cleaned DNA was transferred to a fresh well.

Figure 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).

PCR amplification

A volume of 12.5 µl 2x NEBNext Q5 Hot Start Master Mix (New England Biolabs #M0492S) was added to 9 µl ligated DNA, then 3.5 µl of NEB 8bp index primer/universal primer were added to the mix. The reaction ran in a cycling program set at 98°C for 30 seconds as an initial denaturation step; six cycles of denaturation, annealing and extension were performed, cycling between 98°C for 10 seconds and 65°C for 75 seconds. A final 5-minute step at 65°C was performed, with a final 4°C forever step.

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

Figure 1. (continued)

Figure 1. (continued)

PCR reaction cleanup

AMPure XP beads were resuspended at room temperature with a brief vortex. A volume of 15 µl of resuspended AMPure XP beads was added to the PCR reactions (~25 µl). To mix well, we pipetted up and down at least 10 times. The mixture was incubated for 5 minutes at room temperature. The PCR plate was put on an appropriate magnetic stand to separate the beads from the supernatant. After the solution was clear (about 5 minutes), the supernatant was carefully removed and discarded. We were careful not to disturb the beads containing the target DNA. A volume of 200 µl of 70% ethanol was added to the PCR plate while on the magnetic stand. The mix was incubated at room temperature for 30 seconds, and then the supernatant was carefully removed and discarded. The ethanol wash was repeated once more. The beads were air dried for 7 minutes while the PCR plate was on the magnetic stand with the lid open. The target DNA molecules were eluted from the beads into 15 µl of nuclease-free H₂O, and 15 µl were transferred 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. A total of 473 million paired reads (2 × 150bp) were generated, averaging over 10 million read pairs per sample and a total sequence of 141Gb.

Analysis

Fastq files were uploaded to OneCodex for Kmer analysis and Simpson's diversity index analysis for each genome (Extended data: Supplementary Table 1, sheet Summary <https://doi.org/10.5281/zenodo.4759883>). Reads were also assembled with MegaHit v.1.2.9 (Li *et al.*, 2015, 2016). The Nextflow mapping and assembly pipeline is published on GitHub. 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 which was imaged on plates and chosen for whole genome sequencing and OneCodex analysis is displayed in Figure 1 (DRBC), Figure 2 (PDA-chloramphenicol) and Figure 3 (PDA no chloramphenicol). A link to

Figure 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).

each OneCodex analysis and its respective NCBI submission ID is available in Supplementary Table 1 - Sheet Summary (Extended data, McKernan *et al.*, 2021). 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 by the display of bimodal sequence coverage (clusters of contigs at 1000X and 10X coverage) and compared with the plating images (Figure 4). A heatmap of sequencing read speciation and purity is seen in Figure 5. While merged colonies can be difficult to resolve visually, whole genome sequencing can resolve simple metagenomes and still extract additional diversity information from 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 1).

A Simpson's diversity index analysis demonstrated PDA with CAMP provides the highest diversity score (Figure 6) 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 2).

One fungal sample (*Cladosporium*) presented delayed Ct (31.79) with PathoSEEK Total Yeast and Mold (ITS3-TYM) qPCR primers. Scrutiny of the primer sequences against the *Cladosporium* genome shows proper primer binding locations but missing probe sequences. This genome has low coverage (10X) and the repetitive ITS qPCR target regions are often

Figure 2. (continued)

poorly assembled in low coverage-genomes. This may explain the missing probe sequence in the low coverage fragmented assembly. Additionally, some significantly delayed PathoSEEK Total Aerobic Count (TAC) signal was observed in fungal colonies. This is the result of the use of the lytic enzyme (TLP) which is cloned and expressed in *E. coli* and contains some background *E. coli* DNA. This background TLP expression in *E. coli* produces signals that can be seen in blank preparations. In some cases, this signal is elevated due to mixed colonies observed in the sequencing data.

The qPCR method represents an increased selectivity in assessing fungal and bacterial CFU compared to DRBC, where only ~92% of the colonies were fungal colonies. Quantitative PCR identified all fungi and never mistook one for bacteria. In a minority of cases we had visually mixed colonies. Even if we discount the mixed colonies and count, only the single bacterial colony out of 13 on DRBC, we obtain 92% (1/13) fungal colonies on DRBC where qPCR delivered perfect results. As a comparison, quantitative PCR demonstrated over 10 Cts (1024 fold) differences between the TYM and TAC signals on fungal colonies. The majority of the residual TAC signal being observed in fungi can be normalized and discounted with the background *E. coli* TLP DNA signal measured in blank preparations.

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 3 and Figure 7). In all cases DRBC showed reduced CFU counts.

Figure 2. (continued)

Discussion

Microbial media and their selection have a significant impact on the Simpson's diversity index of microbes observed with whole genome sequencing. This has been noted in prior microbiome surveys in cannabis, in which culturing the microbes changes the representation of the microbiome as measured by qPCR and sequencing performed directly off of the flower (McKernan *et al.*, 2015, 2016). In this study the DRBC selection reduced bacterial growth more than PDA with chloramphenicol, but also reduced 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. States exclusively considering DRBC for ease of colony visualization should be aware of the species-specific sensitivities of using a single medium type, and consider species-specific testing for such pathogenic organisms, to complement a partial yeast and mold test offered from a single selection-based medium. PCR-based techniques can identify 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 which is attempting to quantify 10,000 CFUs/gram. This results in multiplying diluted CFUs 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

Figure 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).

a 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) (*Extended data*: 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 three different forms of selection (DRBC), had bacterial contamination and each level of selection reduced fungal CFU counts.

Figure 3. (continued)

Figure 6. Simpson's Diversity Index. Simpson's diversity index (<https://geographyfieldwork.com/Simpson%27sDiversityIndex.htm>) 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. Dichloran Rose Bengal (DRBC) plating demonstrates the lowest diversity. This is not surprising given DRBC contains 3 different selection agents. While this limits bacterial contamination it also limits yeast and mold growth.

Table 2. Cannabis samples plated on 3 different media. Plating on different media demonstrates a LOG scale difference in Colony Forming Units (CFU) with each plating medium. Sequencing can attribute only half of the colonies as bacteria on Potato Dextrose Agar (PDA) with chloramphenicol. This implies a 5-fold under counting of yeast and mold on Dichloran Rose Bengal (DRBC).

Sample	DRBC					
	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
Low B	0	3	1	1	1	0
Low C	2	1	0	0	0	0
Low D	2	5	2	0	2	0
Low E	2	0	2	0	0	0
			Average CFU/g			
Sample	PDA with Chloramphenicol					
	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
Low B	8	12	8	1	0	3
Low C	13	19	13	1	2	1
Low D	3	12	21	2	0	1
Low E	9	7	4	0	3	3
			Average CFU/g			

Table 2. Continued

Sample	PDA with Chloramphenicol					
	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
Low B	151	157	101	26	20	28
Low C	TNTC	TNTC	TNC	41	45	37
Low D	147	141	123	32	26	26
Low E	138	102	119	23	15	24
			Average CFU/g			

Table 3. Mono-culture evaluations. Fungi species were ordered from American Tissue Culture Collection (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

A. niger ATCC#1015 plated on PDA, PDA-TA, DRBC

100ul plated and grown at 25C for 7 days

A. niger on PDA, PDA-CAMP & DRBC 100ul grown 25C
3 days***Botrytis cinerea*** ATCC#204446 plated on PDA, PDA-CAMP, DRBC (red)

100ul plated and grown at 25C for 5 days

Figure 7. *Aspergillus niger* and *Botrytis cinerea* monocultures plated on 3 different medias. Cultures were plated on Potato Dextrose Agar (PDA), PDA with selection (chloramphenicol or Tartaric acid), and Dichloran Rose Bengal (DRBC). Fewer colonies are consistently found on DRBC.

Data availability

Underlying data

NCBI Bioproject: Under Counting of Total Yeast and Mold on Cannabis using DRBC, Accession number PRJNA725256, <https://www.ncbi.nlm.nih.gov/bioproject?term=PRJNA725256>.

Extended data

Zenodo: Whole genome sequencing of colonies derived from cannabis flowers and the impact of media selection on benchmarking total yeast and mold detection tools, <https://doi.org/10.5281/zenodo.4759883> (McKernan *et al.*, 2021).

This project contains the following extended data:

Summary **Table 1**: OneCodex URLs and NCBI BioSample IDs for every sample.

TYM Inclusion: ATCC organisms tested for inclusion criteria

TYM Exclusion: ATCC organisms tested for exclusion criteria

Sequencing: Number of reads, Read Pairs and Total Gigabases sequenced for each sample.

Assembly: Complete Assembly statistics for each sample generated by QUAST

Data are available under the terms of the [Creative Commons Attribution 4.0 International license](#) (CC-BY 4.0).

References

- Bal A, Agarwal AN, Das A, *et al.*: **Chronic necrotising pulmonary aspergillosis in a marijuana addict: a new cause of amyloidosis.** *Pathology*. 2010; **42**: 197–200.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Day S, Lalitha P, Haug S, *et al.*: **Activity of antibiotics against *Fusarium* and *Aspergillus*.** *Br J Ophthalmol*. 2009; **93**: 116–119.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Gargani Y, Bishop P, Denning DW: **Too many mouldy joints - marijuana and chronic pulmonary aspergillosis.** *Mediterr J Hematol Infect Dis*. 2011; **3**: e2011005.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Gurevich A, Saveliev V, Vyahhi N, *et al.*: **QUAST: quality assessment tool for genome assemblies.** *Bioinformatics*. 2013; **29**: 1072–1075.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Joseph MR, Al-Hakami AM, Assiry MM, *et al.*: **In vitro anti-yeast activity of chloramphenicol: A preliminary report.** *J Mycol Med*. 2015; **25**: 17–22.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Li D, Liu CM, Luo R, *et al.*: **MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph.** *Bioinformatics*. 2015; **31**: 1674–1676.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Li D, Luo R, Liu CM, *et al.*: **MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices.** *Methods*. 2016; **102**: 3–11.
[PubMed Abstract](#) | [Publisher Full Text](#)
- McKernan K, Spangler J, Helbert Y, *et al.*: **Metagenomic analysis of medicinal Cannabis samples; pathogenic bacteria, toxigenic fungi, and beneficial microbes grow in culture-based yeast and mold tests.** *F1000Res*. 2016; **5**: 2471.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- McKernan K, Spangler J, Zhang L, *et al.*: **Cannabis microbiome sequencing reveals several mycotoxinogenic fungi native to dispensary grade Cannabis flowers.** *F1000Res*. 2015; **4**: 1422.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- McKernan KJ, Helbert Y, Kane LT, *et al.*: **Whole genome sequencing of colonies derived from cannabis flowers and the impact of media selection on benchmarking total yeast and mold detection tools (Version 1).** *Zenodo*. 2021, May 14.
[Publisher Full Text](#)
- Remington TL, Fuller J, Chiu I: **Chronic necrotizing pulmonary aspergillosis in a patient with diabetes and marijuana use.** *CMAJ*. 2015; **187**: 1305–1308.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Ruchlemer R, Amit-Kohn M, Raveh D, *et al.*: **Inhaled medicinal cannabis and the immunocompromised patient.** *Support Care Cancer*. 2015; **23**: 819–822.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Smith DG, Marchant R: **Chloramphenicol inhibition of *Pythium ultimum* and *Rhodotorula glutinis*.** *Arch Mikrobiol*. 1968; **60**: 262–274.
[PubMed Abstract](#) | [Publisher Full Text](#)

Open Peer Review

Current Peer Review Status:

Version 1

Reviewer Report 04 August 2021

<https://doi.org/10.5256/f1000research.56852.r90093>

© 2021 Punja Z. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Zamir K. Punja

Department of Biological Sciences, Simon Fraser University, Burnaby, BC, Canada

- Add “human pathogenic” where pathogenic is mentioned as the concern is with these particular microbes and not those potentially that could be affecting the plant as plant pathogens.
- The inclusion of the 3 selection media should be elaborated on. Why were these particular 3 media selected? Provide references to show where, or in what capacity, they may have been used in previously published work. There is abundant published work on the use of PDA with antibiotics to isolate fungi in the plant pathology literature. In fact, it is a standard medium used for isolation in labs worldwide. The addition of dichloran and rose bengal have also been used to restrict the growth of certain groups of fungi and bacteria as a semi-selective medium for isolation in particular from soil samples. Therefore, it would not be expected to provide a broad spectrum of recovery of fungi and yeasts. It is surprising that this medium would be used to assess total yeast and mold counts in cannabis.
- The comparison of the 3 media in this study sheds light on the differences in levels of recovery of fungi and yeasts. This is an important finding – not all media behave in the same manner. To observe a 10-fold difference in recovery between these media is quite significant as it illustrates the potential for under-representation in the recovery process.
- The use of whole genome sequencing to apply to the identification of colony-forming units is a definite plus for this work. It shows the ability to rapidly identify what is present on the culture media with regard to molds that originated from the samples.
- There are several prior reports of authors having recovered a range of fungi from cannabis buds and identified them using the ITS region. Please include these as a reference by which to compare the fungi and yeasts identified in the present study. It is important to build a body of knowledge on the exact identity of the general and species found on cannabis and how prevalent they are.

- The report of endophytes in cannabis should be accompanied by a reference citation. These particular microbes are more difficult to recover in culture media and therefore a molecular approach has merit.
- The inclusion of confirmed ATCC culture specimens to demonstrate differences in growth on the 3 media is a good confirmatory experiment.
- The cannabis samples that originated from Steadfast Analytical Laboratories would have had an analysis of total yeast and mold conducted on them. Is it possible to have these results compared to those from the present study to show how the commercial lab testing may differ from the current study? Or was that not an objective of the current study?
- During the preparation of samples for the ITS3 qPCR procedure, was there a subset of samples included that did not contain the TLP lysis step to show that it made a difference? Or is that included in prior published work?
- For the 45 colonies that were selected for whole genome sequencing, could the identified genus and species be presented in a separate table? Perhaps in accordance with the media from which they were derived from? These would be a summary of what is shown in Figures 1, 2, 3. This is in addition to the OneCodex analysis and the NCBI submission ID available in Supplementary Table 1 It also helps clarify the data shown in Figure 5.
- The Simpson's diversity index analysis shown in Figure 6 is extremely helpful to show the differences between the 3 media types in recovery.
- The results from qPCR of the homogenate that was collected from the Whirl-Pak bags and subjected to PathoSEEK. How did this compare with the colony identification of the same sample plated on the 3 different media with regards to the identification of the genus and species present? Can this be shown in a Table?
- The use of DRBC, if conducted by testing laboratories, is worrisome. It is known that the addition of dichloran and rose bengal is specifically used to discourage certain types of microbes from growing when used for recovery of specific types of fungi from soil samples. The inclusion of DRBC in a testing laboratory for cannabis TYM counts should be discouraged, as shown in the present work where total CFU's recovered were significantly lower compared to PDA with chloramphenicol. DRBC would significantly under-estimate the TYM counts as shown in Figure 4.
- In Table 2, the headings seem incorrect as there are two with "PDA with chloramphenicol" and one should be "PDA w/o chloramphenicol"
- Table 3 and Figure 7 clearly show how DRBC provides reduced growth compared to PDA.

Overall, this is an informative study and the results merit publication. Once the items identified by the reviewer are addressed, this study will be a good addition to the slowly expanding studies showing how complex the assessment of total yeast and mold levels in cannabis is. The information from these types of studies should guide government agencies on the pitfalls of certain methods used to assess TYMC.

Is the rationale for developing the new method (or application) clearly explained?

Yes

Is the description of the method technically sound?

Yes

Are sufficient details provided to allow replication of the method development and its use by others?

Yes

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: microbiology, plant pathology, plant biology, cannabis pathogens, post-harvest quality

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 05 Aug 2021

Kevin McKernan

Thank you for this valuable feedback. I have interweaved our responses below and will update the manuscript accordingly:

- *Add "human pathogenic" where pathogenic is mentioned as the concern is with these particular microbes and not those potentially that could be affecting the plant as plant pathogens.*
- **Response:** Agree. Done

- *The inclusion of the 3 selection media should be elaborated on. Why were these particular 3 media selected? Provide references to show where, or in what capacity, they may have been used in previously published work. There is abundant published work on the use of PDA with antibiotics to isolate fungi in the plant pathology literature. In fact, it is a standard medium used for isolation in labs worldwide. The addition of dichloran and rose bengal have also been used to restrict the growth of certain groups of fungi and bacteria as a semi-selective medium for isolation in particular from soil samples. Therefore, it would not be expected to provide a broad spectrum of recovery of fungi and yeasts. It is*

surprising that this medium would be used to assess total yeast and mold counts in cannabis.

- **Response:** We were not involved in the selection of these media types. The study was initiated on PDA in Michigan. After we completed the PDA study, the organizers informed us they were switching to DRBC based on Steadfast having used chloramphenicol based culture platforms to categorize the samples before shipping them to other labs in Michigan. We suspect the 3 media types were chosen due to their presence in the FDA BAM. We have added a sentence to clarify this.
- *The comparison of the 3 media in this study sheds light on the differences in levels of recovery of fungi and yeasts. This is an important finding – not all media behave in the same manner. To observe a 10-fold difference in recovery between these media is quite significant as it illustrates the potential for under-representation in the recovery process. The use of whole genome sequencing to apply to the identification of colony-forming units is a definite plus for this work. It shows the ability to rapidly identify what is present on the culture media with regard to molds that originated from the samples. There are several prior reports of authors having recovered a range of fungi from cannabis buds and identified them using the ITS region. Please include these as a reference by which to compare the fungi and yeasts identified in the present study. It is important to build a body of knowledge on the exact identity of the general and species found on cannabis and how prevalent they are.*
- **Response:** Very good point. We have added a paragraph to describe the substantial prior art here.
- *The report of endophytes in cannabis should be accompanied by a reference citation. These particular microbes are more difficult to recover in culture media and therefore a molecular approach has merit.*
- **Response:** Very good point. We have added references to emphasize Dichlorans inhibitory nature.
- *The inclusion of confirmed ATCC culture specimens to demonstrate differences in growth on the 3 media is a good confirmatory experiment.*
- **Response:** This is an important control but it should be known by people in the field that AOAC certification doesn't require the inclusion or exclusion testing to be performed in the presence of matrix and this inclusion and exclusion criteria can be obtained using a different media than DRBC. We believe this provides a false sense of safety. This is challenging to perform in the presence of matrix as there is no supplier of sterilized cannabis that ship across state lines and inclusion and exclusion testing can be impacted by background microbial content. Further legalization will improve this.
- *The cannabis samples that originated from Steadfast Analytical Laboratories would have had an analysis of total yeast and mold conducted on them. Is it possible to have these results compared to those from the present study to show how the commercial lab testing may differ from the current study? Or was that not an objective of the current study?*
- **Response:** As part of the AOAC study, we were blinded from these data. We do know that the high, medium and low categories were allocated according to culture-based

methods that used CAMP and that quantitative PCR over estimated microbial burden on the low DRBC samples.

- *During the preparation of samples for the ITS3 qPCR procedure, was there a subset of samples included that did not contain the TLP lysis step to show that it made a difference? Or is that included in prior published work?*
- **Response:** We did not include TLP lysis versus no TLP as the study required we settle on a single method for evaluation. We are in the process of writing up that comparison for another publication.
- *For the 45 colonies that were selected for whole genome sequencing, could the identified genus and species be presented in a separate table? Perhaps in accordance with the media from which they were derived from? These would be a summary of what is shown in Figures 1, 2, 3. This is in addition to the OneCodex analysis and the NCBI submission ID available in Supplementary Table 1 It also helps clarify the data shown in Figure 5.*
- **Response:** This is an important point. This does exist in Figure 5 but we failed to clarify the sample nomenclature that clarifies this. We have added a sample key to describe which samples are DRBC, PDA-CAMP, PDA-no-CAMP.
- *The Simpson's diversity index analysis shown in Figure 6 is extremely helpful to show the differences between the 3 media types in recovery. The results from qPCR of the homogenate that was collected from the Whirl-Pak bags and subjected to PathoSEEK. How did this compare with the colony identification of the same sample plated on the 3 different media with regards to the identification of the genus and species present? Can this be shown in a Table? The use of DRBC, if conducted by testing laboratories, is worrisome. It is known that the addition of dichloran and rose bengal is specifically used to discourage certain types of microbes from growing when used for recovery of specific types of fungi from soil samples. The inclusion of DRBC in a testing laboratory for cannabis TYM counts should be discouraged, as shown in the present work where total CFUs recovered were significantly lower compared to PDA with chloramphenicol. DRBC would significantly under-estimate the TYM counts as shown in Figure 4.*
- **Response:** We were only allowed to ship cultures on plates across state lines. As a result, we have Cq scores for the colonies that were picked and isolated in Figures 1,2,3 under the TYM and TAC Cq columns on the right. This only informs on inclusion and exclusion capabilities of the primers for the colonies harvested but loses quantitative information. We were not allowed to ship homogenized matrix in the mail to assess the Cq prior to plating. Labs local to Michigan have performed this comparison and are free to publish those results. The summary of the results communicated to us were that the qPCR had better concordance with PDA with CAMP and over estimated CFUs on the DRBC Low samples. This significantly differs from Michigans stated intentions with the ERV where they voiced concerns about molecular methods undercounting risk (<https://help.medicinalgenomics.com/hubfs/Regulatory%20Info%20for%20Sales/Michigan%20MRA%20FAQ.pdf>). The opposite turned out to be true. DRBC is undercounting risk compared to qPCR. The MRA was also led to believe that Klebsiella was not an appropriate validation organism as it was not commonly found on cannabis despite it having been published by Thompson et al previously. Candida albicans (which we have never seen

documented on Cannabis) was prioritized as a CRM. We have added some language to address this to the best of our ability.

- *In Table 2, the headings seem incorrect as there are two with "PDA with chloramphenicol" and one should be "PDA w/o chloramphenicol" Table 3 and Figure 7 clearly show how DRBC provides reduced growth compared to PDA.*
- **Response:** Thank you!. Good catch. We have updated that.

Competing Interests: No competing interests were disclosed.

Reviewer Report 28 July 2021

<https://doi.org/10.5256/f1000research.56852.r90094>

© 2021 Orser C. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Cindy Orser

CLIP Laboratories, San Diego, CA, USA

A well-executed study of considerable breadth and size to further solidify the argument against State regulatory bodies requiring agar plating to evaluate microbial load on cannabis flower in lieu of proven superiority of molecular assays. In addition, this study demonstrates confirmation of chloramphenicol's ability to knock down *Fusarium* and *Aspergillus* growth when plating on DRBC at relatively low chloramphenicol levels [~ 0.1 mg/mL], not sure of the concentration in PDA agar. Nonetheless, this study should finally put to rest the issue of which method is most accurate at representing microbial load on cured cannabis flower. With qPCR assays shown to be fundamentally superior to culturing and plating, the discussion should now move over to "sampling" and how inadequate the normal sample size per batch size is to give a glimpse into the microbiome of the cannabis flower and the intrigue of how cultivation methods influence the cannabis microbiome.

Is the rationale for developing the new method (or application) clearly explained?

Yes

Is the description of the method technically sound?

Yes

Are sufficient details provided to allow replication of the method development and its use by others?

Yes

If any results are presented, are all the source data underlying the results available to

ensure full reproducibility?

Yes

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: plant microbiology; cannabis analytical testing; diagnostics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 10 Aug 2021

Kevin McKernan

Thank you for this feedback. The sampling comment is very pertinent. We were not in control of this aspect of the study so opted to save this topic for another manuscript. AOAC communicated to us that many different cannabis samples were mixed to normalize chemotype affects on culture. This is a very good idea as we have seen chemotype specific effects on cannabis microbiology and its published to occur in *Trema orientalis* (<https://pubmed.ncbi.nlm.nih.gov/34035994/>).

One of the concerns with plating, is that the antibiotic cannabinoids and terpenes may get liberated from trichomes in the aggressive lab homogenization and media saturation process. This may influence the viability of some of the microbes. This aggressive homogenization and fluid saturation is not what a consumer experiences.

Many publications demonstrate the antibiotic nature of cannabinoids and how different cannabinoids exhibit different antibiotic properties thus we should expect different chemotypes to plate differently given there is no purification step prior to plating (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7105690/>). Molecular methods lyse open cells and purify the DNA away from such potential growth inhibitors with Ethanol extractions.

To confirm these samples were indeed a mixture of different cannabis samples, we were sent DNA from these mixtures and performed 10Mb SureSelect capture and deep Illumina Sequencing on these samples to under stand how well they were mixed and how diverse they where. The read genotypes indeed suggested more than a single cannabis sample was present and likely more than 4 in each High, Medium and Low Categories.

We put these data public for anyone who is interested but felt it would bloat this manuscript with confirmatory data and distract from the core focus of the study.

<https://www.kannapedia.net/strains/rsp11748/>

<https://www.kannapedia.net/strains/rsp11749/>

<https://www.kannapedia.net/strains/rsp11750/>
<https://www.kannapedia.net/strains/rsp11752/>
<https://www.kannapedia.net/strains/rsp11753/>
<https://www.kannapedia.net/strains/rsp11754/>
<https://www.kannapedia.net/strains/rsp11755/>
<https://www.kannapedia.net/strains/rsp11756/>
<https://www.kannapedia.net/strains/rsp11757/>
<https://www.kannapedia.net/strains/rsp11758/>
<https://www.kannapedia.net/strains/rsp11759/>
<https://www.kannapedia.net/strains/rsp11760/>
<https://www.kannapedia.net/strains/rsp11761/>
<https://www.kannapedia.net/strains/rsp11762/>

Competing Interests: No additional conflicts than those already described in the manuscript.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research