Cannabis microbiome sequencing reveals several mycotoxic fungi native to dispensary grade cannabis flowers

Kevin McKernan, Jessica Spangler, Lei Zhang, Vasisht Tadigotla, Yvonne Helbert, Theodore Foss, Douglas Smith

Medicinal Genomics Corporation, Woburn MA, 01801, USA

Abstract
The Center for Disease Control estimates 128,000 people in the U.S. are hospitalized annually due to food borne illnesses. This has created a demand for food safety testing targeting the detection of pathogenic mold and bacteria on agricultural products. This risk extends to medical Cannabis and is of particular concern with inhaled, vaporized and even concentrated Cannabis products. As a result, third party microbial testing has become a regulatory requirement in the medical and recreational Cannabis markets, yet knowledge of the Cannabis microbiome is limited. Here we describe the first next generation sequencing survey of the microbial communities found in dispensary based Cannabis flowers and demonstrate the limitations in the culture based regulations that are being superimposed from the food industry.

Introduction
Many states in the U.S. are crafting regulations for microbial detection on cannabis in absence of any comprehensive survey of Cannabis microbiomes. A few of these regulations are inducing growers to “heat kill” or pasteurize Cannabis flowers to lower microbial content. While this is a harmless suggestion, we must remain aware of how these drying techniques often create false negatives in culture based safety tests used to monitor colony-forming units (CFU). Even though pasteurization may be effective at sterilizing some of the microbial content, it does not eliminate various pathogenic toxins or spores. Aspergillus spores and mycotoxins are known to resist pasteurization. Similar thermal resistance has been reported for E.coli produced Shiga Toxin. While pasteurization may reduce CFU’s used in petri-dish or plating based safety tests, it does not reduce the microbial toxins, spores or DNA encoding these toxins.

Mycotoxin monitoring in Cannabis preparations is important since aflatoxin produced by Aspergillus species is a carcinogen. The clearance of aflatoxin requires the human liver enzyme CYP3A4 and this liver enzyme is potently inhibited by cannabinoids. Modern day cannabis flowers can produce up to 25% (w/v) cannabinoids presenting potent inhibition of CYP3A4 and CYP2C19. Health compromised patients exposed to aflatoxin and clearance-inhibiting cannabinoids raise new questions in regards to the current safety tolerances to aflatoxin. Similarly, Fusarium species are known to produce fungal toxins and has proven to be difficult to selectively culture with tailored media. This is a common fault of culture-based systems as carbon sources are not exclusive to certain microbes and only 1% of microbial species are believed to be culturable.

While these risks have been well studied in the food markets, the presence of the microbial populations present on cannabis flowers has never been surveyed with next generation sequencing techniques. With the publication of the Cannabis genome and many other pathogenic microbial genomes, quantitative PCR assays have been developed that can accurately quantify fungal DNA present in Cannabis samples.
we analyze the yeast and mold species present in 10 real world, dispensary-derived Cannabis samples by quantitative PCR and sequencing, and demonstrate the presence of several mycotoxin producing fungal strains that are not detected by widely used culture-based assays.

Results

A commercially available Total Yeast and Mold qPCR assay (TYM-PathogINDICAtor, Medicinal Genomics, Woburn) was used to screen for fungal DNA in a background of host Cannabis DNA. The TYM qPCR assay targets the 18S rDNA ITS (Internal Transcribed Spacer) region using modified primers described previously. Fungal DNA amplified using these primers may also be subjected to next generation sequencing to identify the contributing yeast and mold species. ITS sequencing has been widely used to identify and enumerate fungal species present in a given sample.

We purified DNA from Cannabis samples obtained from two different geographic regions (Amsterdam and Massachusetts) several years apart (2011 and 2015). The majority of samples purified and screened with ITS qPCR were negative for amplification signal implying reagents clean of fungal contamination. Six of the 17 dispensary-derived Cannabis samples tested positive for yeast and mold in the TYM qPCR assay. These results were compared with the results derived from three commercially available culture based detection systems for each of the 17 samples (3M Petrifilm™, 3M Microbiology, St. Paul, MN, USA, SimPlates™, Biocontrol Systems, Bellevue, WA, USA, BioLumix™ Neogen, Lansing MI, USA). Of the 6 qPCR positive samples, two tested negative in all 3 culture-based assays and four tested negative in 1 or 2 of the culture-based assays (Table 1). None of the qPCR negative samples tested positive in any of the culture based assays. Each of the 6 discordant samples was subjected to ITS sequencing to precisely identify the collection of microbes present. Four additional samples from a different geographic origin (Amsterdam) were also subjected to ITS sequencing, for a total of 10 Cannabis samples.

Each discordant sample presented with an array of microbial species, as shown in figure 2. No sample presented with a single dominant species, and each sample displayed multiple species of interest. Of particular concern were the identified DNA sequences from toxin producing species: Aspergillus versicolor, Aspergillus terreus, Penicillium citrinum, Penicillium paxilli.

We further analyzed the ITS sequence alignments using the whole genome shotgun based microbiome classification software known as One Codex. Nine of the ten samples sequenced showed the presence of P. paxilli (Figure 3). To verify the accuracy of this ITS phylotyping, a gene involved in the Paxilline toxin Biosynthesis pathway of P. paxilli was amplified with PaxPss1 and PaxPss2 primers described by Saikia et al. The resulting 725bp amplicon (expected size) was sequenced to confirm the presence of the P. Paxilli biosynthesis gene in the cannabis sample KD8 (Figure 4). While there are some discrepancies between the two software platforms, our analysis used merged paired reads with MG-RAST and correlate better with PCR results. While One Codex predicted and confirmed KD8 as having the highest Paxilli content, the One Codex platform is optimized for whole genome shotgun data.

With the confirmed presence of P. paxilli, we are curious to find out whether the toxin, paxilline, is present in the samples. Development of monoclonal antibodies to paxilline...
has recently been described\textsuperscript{35}, but commercial ELISA assays with sensitivity under 50ppb do not appear to be available at this time. A >50ppb multiplexed ELISA assay is available from Randox Diagnostics (Crumlin, UK). Detection with LC-MS/MS has also been described\textsuperscript{36, 37}, however, and experiments are underway to determine whether Paxilline can be identified in the background of cannabinoids and terpenes present in Cannabis samples.

**Discussion**

Several potentially harmful fungal species were detected in dispensary-derived Cannabis samples by qPCR and subsequent sequencing in this study. Three different culture-based assays failed to detect all of the positive samples and one, BioLumix\textsuperscript{TM}, detected only one out of 7 positive samples. A review of the literature suggests that Penicillium microbes can be cultured on CYA media, but some may require colder temperatures (21-24C) and 7 day growth times\textsuperscript{38}. Of the Penicillium, only P. Citrinum has been previously reported to culture with 3M Petri-Film\textsuperscript{39}. In addition, several studies have demonstrated plant phytochemicals and terpenoids like eugenol can inhibit the growth of fungi\textsuperscript{40}. It is possible the different water activity of the culture assay compared to the natural terpene rich flower environment is contributing to the false negative test results.

Quantitative PCR is agnostic to water activity and can be performed in hours instead of days. The specificity and sensitivity provides important information on samples that present risks invisible to culture based systems. The draw back to qPCR is the method’s indifference to living or non-living DNA. While techniques exist to perform live-dead qPCR, the live status of the microbes is unrelated to toxin potentially produced while the microbes were alive. ELISA assays exists to screen for some toxins\textsuperscript{41}. Current State-recommended ELISA’s do not detect Citrinin or Paxilline, the toxins produced by P. citrinum and P. paxilli, respectively. The predominance of these Penicillium species in a majority of the samples tested is interesting. Several Penicillium species are known to be endophytes on various plant species, including P. citrinum\textsuperscript{10}, and this raises the question of whether they are also Cannabis endophytes.

Paxilline is a tremorgenic and ataxic potassium channel blocker and has been shown to attenuate the anti-seizure properties of cannabidiol in certain mouse models\textsuperscript{42-44}. Paxilline is reported to have tremorgenic effects at nanomolar concentrations and is responsible for Ryegrass-staggers disease\textsuperscript{45}. Cannabidiol is often used at micromolar concentrations for seizure reduction implying sub-percentage contamination of Paxilline could still be a concern. Citrinin is a mycotoxin that disrupts Ca\textsuperscript{2+} efflux in the mitochondrial permeability transition pore (mPTP)\textsuperscript{46-53}. Ryan et al. demonstrated that cannabidiol affects this pathway suggesting a potential concern for CBD-mycotoxin interaction\textsuperscript{54}. Considering the hydrophobicity of Paxilline and the recent interest in the use of cannabidiol derived from cannabis flower oils for drug resistant Epilepsy, more precise molecular screening of fungal toxins may be warranted\textsuperscript{55-60}.

Our survey of cannabis flowers in this study was limited. Further screening will be required to define a set of tests that can adequately capture all risks. While ELISA assays are easy point of use tests that can be used to detect fungal toxins, they can suffer from lack of sensitivity and cross reactivity. ITS amplification and sequencing offers hypothesis-free testing that can complement the lack of specificity in ELISA assays. Appropriate primer design can survey a broad spectrum of microbial genomes while affording rapid iteration of design. Quantitative PCR has also demonstrated single
molecule sensitivity and linear dynamic range over 5 orders of magnitude offering a very robust approach for detection of microbial risks. This may be important for the detection of nanomolar potency mycotoxins. Further studies are required to validate better detection methods for these toxins and verify whether Paxilline or Citrinin are present on cannabis at concentrations that present a clinical risk.

Conclusions

These results demonstrate that culture based techniques superimposed from the food industry should be re-evaluated based on the known microbiome of actual Cannabis flowers in circulation at dispensaries. Several mycotoxin producing molds were detected that can potentially interfere with the medical use of cannabidiol. These microbes failed to grow on traditional culture based platforms but were rapidly detected with molecular based techniques. Further studies are required to quantitate the presence and concentration of mycotoxin production.

Methods

Culture based methods

3.55ml of Tryptic Soy Broth (TSB) was used to wet 250mg of homogenized flower in a whirlpack bag. TSB was aspirated from the reverse side of the 100um mesh filter and placed into a Biolumix growth vial and spread onto a 3M Petri Film™ and a SimPlate™ according to the respective manufacturers recommendations. Biolumix™ vials were grown and monitored for 48 hours while Petri-films™ and SimPlates™ were grown for 5 days. Petri-films™ and SimPlates™ were colony counted manually by three independent observers. Samples were tested on Total Coliform, Total Entero, Total Aerobic, and Total Yeast and Mold. Only Total Yeast and Mold discrepancies were graduated to sequencing.

DNA Purification

Plant DNA was extracted with SenSATIVAx according to manufacturers instructions (Medicinal Genomics part #420001). DNA is eluted with 50ul ddH2O.

Primers used for PCR and sequencing

PCR was performed using 5ul of DNA (3ng/ul) 12.5ul 2X LongAmp (NEB) with 1.25ul of each 10uM MGC-ITS3 and MGC-ITS3 primer (MGC-ITS3; TACACGACGTTGTAAGCAAGCACACATCGATGAAGAAGAAGCAGC) and (MGC-ITS3R; AGGATAAACAATTTCACACAGGATTTGAGCTCTTGCCGCTTCA) with 10ul ddH2O for a 25ul total reaction. An initial 95C 5 minute denaturization was performed followed by 40 cycles of 95C for 15s and 65C for 90s. Samples were purified with 75ul SenSATIVAx, washed twice with 100ul 70% EtOH and bench dried for 5minutes at room temperature. Samples were eluted in 25ul ddH2O.

Tailed PCR Cloning and Sequencing

DNA libraries were constructed with 250ng DNA using NEB's NEBNext Quick ligation module (NEB # E6056S). End Repair used 3ul of Enzyme Mix, 6.5ul of Reagent Mix, 55.5ul of DNA + ddH2O. Reaction was incubated at 30C for 20 minutes. After End Repair, Ligation was performed directly with 15ul of Blunt End TA Mix, 2.5ul of Illumina Adaptor (10uM) and 1ul of Ligation enhancer (assumed to be 20% PEG 6000). After 15-minute ligation at 25C, 3ul of USER enzyme was added to digest the hairpin adaptors and prepare for PCR. The USER enzyme was tip-mixed and
incubated at 37°C for 20 minutes. After USER digestion, 86.5µl of SenSATIVAx was
added and mixed. The samples were placed on a magnet for 15 minutes until the
beads cleared and the supernatant could be removed. Beads were washed twice
with 150µl of 70% EtOH. Beads were left for 10 minute to air dry and then eluted in
25µl of 10mM Tris-HCl.

Library PCR
25ul 2X Q5 Polymerase was added to 23ul of DNA with 1ul of i7 index primer (25uM)
and 1ul Universal primer (25um). After an initial 95°C for 10 secs, the library was
amplified for 15 cycles of 95°C 10sec, 65°C 90sec. Samples were purified by mixing 75ul
of SenSATIVAx into the PCR reaction. The samples were placed on a magnet for 15
minutes until the beads cleared and the supernatant could be removed. Beads were
washed twice with 150µl of 70% EtOH. Beads were left for 10 minute to air dry and
then eluted in 25ul of 10mM Tris-HCl. Samples were prepared for sequencing on the
MiSeq V2 chemistry according to the manufactures instructions. 2x250bp reads
were selected to obtain maximal ITS sequence information.

PaxP Verification PCR
Primers described by Shirazi-zand et al. were utilized to amplify a segment of the 725bp
PaxP gene. 25ul LongAmp (NEB) 4ul 10uM Primer, 1ul DNA (14ng/ul), 20ul ddH2O to
make a 50ul PCR reaction. Cycling conditions we slightly modified to accommodate a
different polymerase. 95°C for 30s followed by 28 cycles of 95°C 15s, 55°C for 30sec, 65°C
2.5 minutes. Samples were purified with 50ul of SenSATIVAx as described above. 1ul of
purified PCR product was sized on Agilent HS 2000 chip. Nextera libraries and
sequencing were performed according to instructions from Illumina using 2x75bp
sequencing on a version 2 MiSeq.

Analysis
Reads were demultiplex and trimmed with Casava 1.8.2 and trim_galore. FLASH was
used to merge the reads using max_overlap 150. The reads were aligned to microbial
references using MG-RAST. Alignments and classifications were confirmed with a
second software tool from One Codex and critical pathways identified for further
evaluation with PCR of toxin producing genes. Reads are deposited in NCBI under SRA
accession: SRP065410. Nextera 2x75bp sequencing of the PaxP gene was mapped to
accession number HM171111.1 with CLCbio Workstation V4 at 98% identity over
80% of the read. One Codex analysis was put into Public mode under the following
classification:

Australian Bastard:
https://app.onecodex.com/analysis/public/201e7f1642e04a3c
https://app.onecodex.com/analysis/public/58f1e03c10434bfa

KD4:
https://app.onecodex.com/analysis/public/2e86e262817246c4
https://app.onecodex.com/analysis/public/1abd5b6046140a0

KD6:
https://app.onecodex.com/analysis/public/a92d3df5485499d
https://app.onecodex.com/analysis/public/8d72e2514e564ecd
Author Contributions:

KJM designed the study and performed the OneCodex analysis
JS designed and ran the culture and qPCR laboratory experiments
LZ assisted in the figure generation and laboratory experiment
YH assisted in sequencing and PCR confirmation of Pax.
VT - read alignment, MG-RAST, primer design and analysis.
TF - Sample tracking software, figure generations, ITS software comparisons
DS - Manuscript construction and review

Acknowledgments: John McPartland, Cindy Orser, Brad Douglass, Joost Heeroma, Nick Greenfield and Kellie Dodd for thoughtful advice.


**Total Yeast & Mold**

*(10,000 CFU/g)*

<table>
<thead>
<tr>
<th>Samples</th>
<th>Yeast/Mold</th>
<th>Simplate(^\circ) (CFU/g)</th>
<th>3M(^TM) (CFU/g)</th>
<th>Biolumix</th>
<th>Cq</th>
</tr>
</thead>
<tbody>
<tr>
<td>KD4</td>
<td>0</td>
<td>0</td>
<td>pass</td>
<td>21.71</td>
<td>(f)</td>
</tr>
<tr>
<td>KD8</td>
<td>0</td>
<td>0</td>
<td>pass</td>
<td>22.5</td>
<td>(f)</td>
</tr>
<tr>
<td>PC3</td>
<td>0</td>
<td>0</td>
<td>pass</td>
<td>&gt;40</td>
<td>(p)</td>
</tr>
<tr>
<td>White Widow</td>
<td>0</td>
<td>0</td>
<td>pass</td>
<td>&gt;40</td>
<td>(p)</td>
</tr>
<tr>
<td>KD1</td>
<td>0</td>
<td>0</td>
<td>pass</td>
<td>29.33</td>
<td>(p)</td>
</tr>
<tr>
<td>KD2</td>
<td>0</td>
<td>0</td>
<td>pass</td>
<td>&gt;40</td>
<td>(p)</td>
</tr>
<tr>
<td>KD3</td>
<td>0</td>
<td>0</td>
<td>pass</td>
<td>30.16</td>
<td>(p)</td>
</tr>
<tr>
<td>KD5</td>
<td>1000</td>
<td>6000</td>
<td>pass</td>
<td>27.76</td>
<td>(p)</td>
</tr>
<tr>
<td>KD6</td>
<td>3000</td>
<td>19000</td>
<td>pass</td>
<td>24.72</td>
<td>(f)</td>
</tr>
<tr>
<td>KD7</td>
<td>0</td>
<td>0</td>
<td>pass</td>
<td>&gt;40</td>
<td>(p)</td>
</tr>
<tr>
<td>Liberty Haze</td>
<td>172000</td>
<td>89000</td>
<td>pass</td>
<td>24.02</td>
<td>(f)</td>
</tr>
<tr>
<td>Blueberry Kush</td>
<td>0</td>
<td>0</td>
<td>pass</td>
<td>37.99</td>
<td>(p)</td>
</tr>
<tr>
<td>Blueberry Kush (Exp)</td>
<td>&gt;738000</td>
<td>too high</td>
<td>fail</td>
<td>15.71</td>
<td>(f)</td>
</tr>
<tr>
<td>Girl Scout Cookies</td>
<td>&gt;738000</td>
<td>Too high</td>
<td>pass</td>
<td>19.66</td>
<td>(f)</td>
</tr>
<tr>
<td>Jake’s Grape</td>
<td>&gt;738000</td>
<td>Too high</td>
<td>pass</td>
<td>24.56</td>
<td>(f)</td>
</tr>
<tr>
<td>Serious Happiness</td>
<td>0</td>
<td>3000</td>
<td>pass</td>
<td>&gt;40</td>
<td>(p)</td>
</tr>
<tr>
<td>White Rhino</td>
<td>0</td>
<td>3000</td>
<td>pass</td>
<td>&gt;40</td>
<td>(p)</td>
</tr>
</tbody>
</table>

Table 1. Samples were cultured with 3 different techniques and compared to quantitative PCR (qPCR). Biolumix had the lower sensitivity failing to pick up 4/17 samples detected with other culture-based platforms. qPCR identified 2 samples that were not picked by any other method. Positive qPCR samples were sequenced to identify the contributing signal. Highlighted samples fail the 10,000 CFU/g cutoffs which equates to a Cq of 26 on the qPCR assay according to the manufacturers instructions. (f) is Fail or over 10,000 CFU/g. (p) is Pass or under 10,000 CFU/g.
Figure 1A. qPCR signal from TYM (red line) test run concurrently (multiplexed) with a plant internal control marker (green line). This marker targets a conserved region in the cannabis genome and should show up in every assay (Upper left). SimPlates count the number of discolored wells (Purple to pink) as a proxy for CFU/gram. Only total aerobic show growth (Upper Right). Petrifilm only demonstrate colonies on Total Aerobic platings (lower left). Biolumix demonstrate no signal across all 4 test (lower right).
Figure 1B. Sample KD8 fails to culture any Total Yeast and mold yet demonstrates significant TYM qPCR signal. Sample was graduated to ITS based next generation sequencing.
Figure 1C. Sample Liberty Haze was tested with 3 culture based methods and compared to qPCR. Sample was graduated to ITS based next generation sequencing.
Figure 2- DNA sequencing of ITS3 amplicons from culture negative samples that are qPCR positive for Total Yeast and Mold tests. Penicillium and Aspergillus are commonly found (Y axis) but at different read counts in each sample (X axis). Read counts are more a reflection of sample normalization for sequencing than inter sample quantitation provided by qPCR.
Figure 3- One Codex classification of ITS reads. P.Paxilli is the most frequently found contaminant in Cannabis flowers. P.Citrinum is not in the One Codex database at this time. One Codex utilizes a fast k-mer based approach for whole genome shotgun classification and can be influenced by read trimming and database content. The reads provided to MG-RAST were trimmed and FLASH’d (Paired end reads merged when overlapping) prior to classification. K-mer based approaches can significantly differ from longer word size methods and this underscores the importance of confirmatory PCR in microbiome analysis.
Figure 4. PaxP PCR demonstrates amplification of a 725bp band in sample KD8(A). PCR products were made into a shotgun library with Nextera and sequenced on an Illumina MiSeq with 2x75bp reads to over 10,000X coverage. Reads were mapped with CLCbio 4 to NCBI Accession number HM171111.1. Paired reads are displayed as blue lines, Green and Red lines are unpaired reads. Read Coverage over the amplicon is depicted in a blue histogram over the cluster while paired end read distance is measured in a red histogram over the region. Off target read mapping is limited (B). Alignment of PCR primers to P.Paxilli reference shows a 5 prime mismatch that is a result of the primers being designed to target spliced RNA according to Saikia et al.