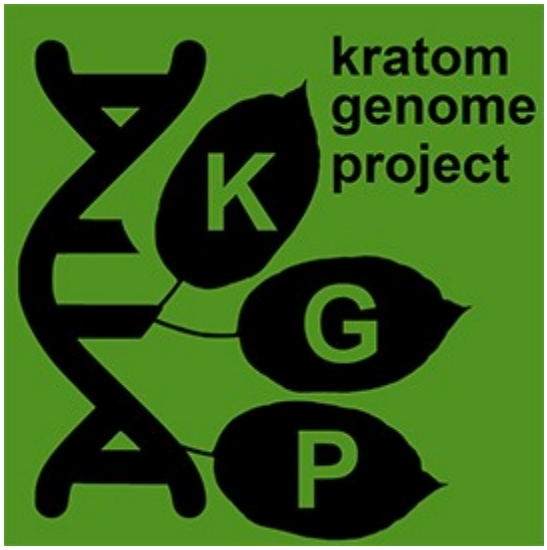
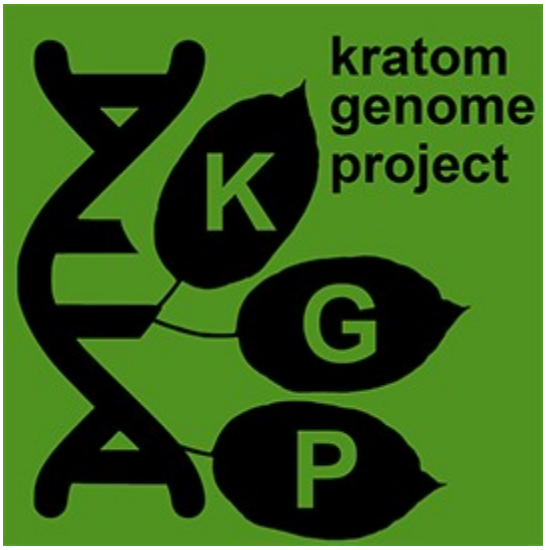




Welcome to the Kratom Genome Project

We have proof here that in 2017, we sequenced **10X the amount of the Kratom nucleotides or bases anyone else has released to** (NCBI / NIH) to date.

This torrent includes all the data.

This PDF contains the website that was previously at kratomdna.org

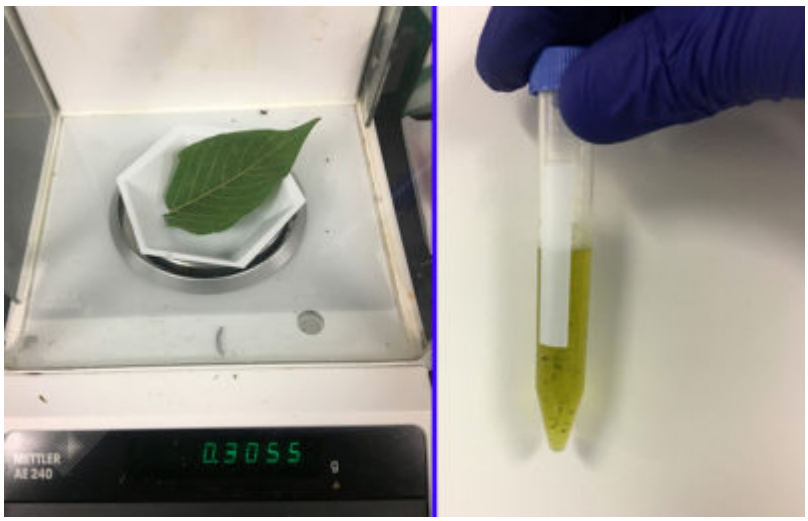
Kratom is under attack.

Long used as a helpful medicinal and recreational herb, *Mitragyna speciosa*, or Kratom, is now gaining the attention of those who don't want you to control your own body and mind.

Various governments have recently outlawed or regulated the plant, or have plans to do so.

This behavior is historically a prelude to large pharmaceutical companies patenting natural medicines to monopolize their derivatives and sell them at high prices….

...By sequencing the Kratom genome and releasing this information free to the public, we will establish "prior art" to render attempts to secure patents moot.



A lot of people think that plant DNA can't be patented

Plants can still can be patented, and recently have been. More here.

We're not trying to get a patent. We're trying to *block* patents on one very common kratom strain (Red Vein Thai). To that end, so far we've sequenced and published 10X the amount of the Kratom nucleotides or bases anyone else has released to (NCBI / NIH) to date.

What we've done so far

We've "shotgun sequenced" 10X the amount of the Kratom nucleotides or bases that anyone else has. (The FDA and a few others have released some into NCBI / NIH , but we increased this 10X in a few days.)

And we're sharing it free with the public.

What we plan and what we need

We plan to complete the sequencing of Red Vein Thai kratom, then move on to sequencing a white strain and a green kratom strain. All the while, sharing in public as soon as we have the data, in compliance with the suggestions of Bermuda Principles. And we will continue to engage in education and outreach to spread the word.

This will serve to cut patents off at the pass. But it is also valuable research that will help scientists around the world find out more about useful applications and safe consumption of Kratom for medicine and recreation.

This will cost money.

We also need people to torrent our data to keep it alive in the wild.

And we need people to share links to this project and talk it up. Public data few people know about isn't truly public.

Specifications of sequencing

DNA isolation: Plant DNA was isolated from 300mg of homogenized leaf from Red Vein Thai cultivar utilizing a 45 minute heated shaker (37C) with steel ball bearings. Lysate was purified utilizing SenSATIVAx and eluted in ddH20 at 2ng/ul.

Library Construction: Nextera transposition was performed on 20ng of DNA at 55C for 10 minutes with 1ul of Nextera Enzyme. 12 cycles of PCR were performed and the library size selected to 600-800 bases on Blue Pippin Prep SAGE station.

Sequencing: 2x151bp reads were run on an Illumina MiSeq with Version 2 chemistry. 42 Million reads were generated.

Analysis: Reads were assembled with SPades and CLCbio workbench version 9. Quast and Bondage were used to generate assembly statistics and assembly graphs. All 40 DNA sequences in NCBI were compared to the reference genome and 100% of these had alignments to the assembled reference with a 100% match to previous ITS sequence of Mitrogyna speciosa.

Timeline

Leaf samples were overnight shipped on a Friday and Sequence assembly was public 6 days later. 2 days were lost due to optimizing DNA clustering for high AT content DNA and low yield DNA isolations.

What you can do to help

1. Tell two friends (or two hundred)

Tweet, blog, post our website. Tell as many friends as possible. Here's a cut and paste that will work anywhere. Tailor as needed to your voice and your audience, but please keep the web address in there:

Protect Kratom from Big Pharma!

We just sequenced 10X more Kratom genome than can be found in public databases.

And we're sharing it to everyone via torrents as prior art. Help us sequence and share the Kratom Genome to keep it open source and not a corporate monopoly.
The Kratom Genome Project

https://KratomDNA.org

---

2. Torrent and seed the actual data

This makes the "prior art" to prevent patents universal and unstoppable. Learn more and torrent from here.

3. Talk to others

Become educated on the advantages of kratom and other botanical medicines. Explain patiently online in person with family, friends and strangers why it's false to say "There oughta be a law", remembering that all laws are backed by the threat of a gun and the threat of a cage.

Embrace this era of new growth, and spread the word.

Thank you for helping us help keep nature's gifts open to all!

=-=--=

Why We're Sequencing the Kratom Genome

…and why we're releasing it all free to the public as a protection against monopolies from big pharma and governments.

Long answer of why we're doing this and why it's important for people to torrent and donate is below.

Layman version – tl/dr is: While an entire plant genome can no longer be patented in the USA, specific strains can be, which can be bad for kratom users in general.
We beat big pharma to a public release and that's good for keeping kratom from being locked down. Sequencing multiple strains with collected chemotype information is the best path forward.

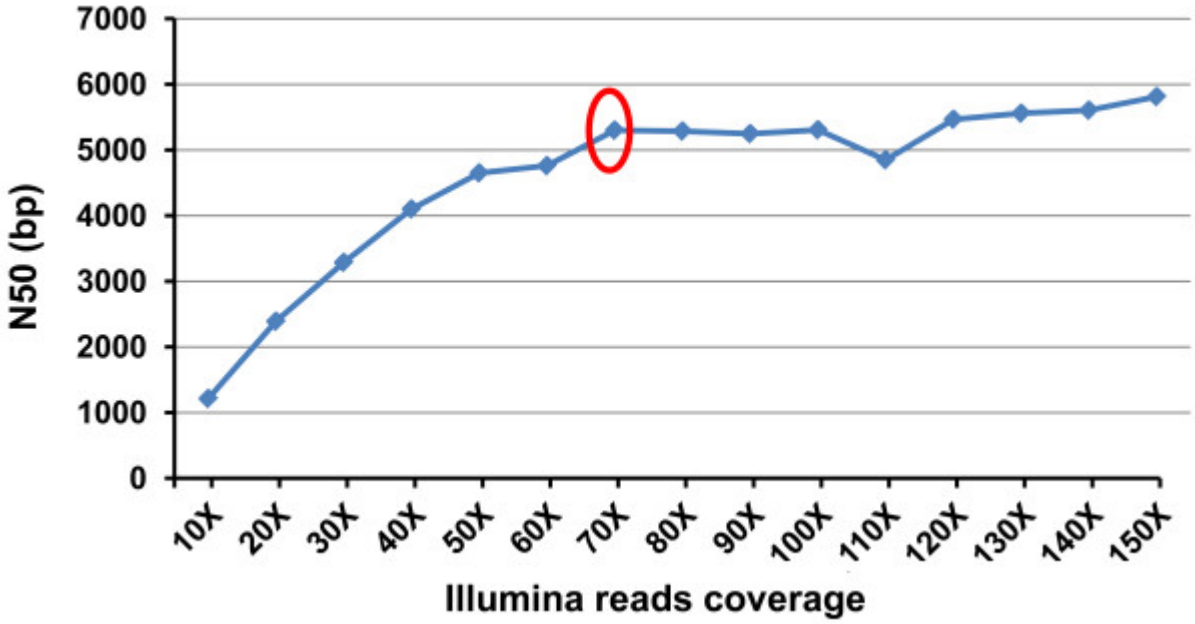
We're not trying to get a patent. We're trying to *block* patents on one very common common strain (Red Vein Thai). To that end, so far we've sequenced and published 10X the number of nucleotides (bases) of the Kratom genome than the largest amount anyone else has released into (NCBI / NIH) to date.

With enough resources, we could sequence more of red, and also move on to white, green and other strains.

Pharmaceutical companies may be working on this but they haven't published data so this work of ours is prior art to any attempt to capture patents on popular strains or varietals.

This info could also be used to do things like lift the chemical pathway from kratom into yeast to make it easy to ferment tons of organic material with the active substances in kratom easily so it can never be stopped.

(Related: Plants CAN be patented…Despite what most people hear in forums)



^This graph is how we can inch the genome up the completeness curve. N50 is the mean length of the contigs (gap free stretches of ATC and Gs) in the genome. At 5X coverage we end up with N50s in 1-2 Kb range but this rapidly improves with more coverage. Doubling the coverage more than doubles the contiguity. We expect it will take 7 MiSeq runs to hit saturation and we will review each incremental run to assess strategy. We have access to Oxford nanopore sequencers. If funded, can layer in these reads to help mature a final assembly.

Open source Mitragyna speciosa

Despite the recent *AMP versus Myriad*[1] supreme court rulings regarding the lack of patent eligibility of naturally occurring gene sequences and the recent *Mayo versus Prometheus* ruling[2] regarding the patent eligibility of natural phenomena, cannabis plant and cannabinoid patents continue to issue. Cannabis remains federally illegal yet federal agencies are being awarded patents on their medicinal use (Hampson *et al*)[3]. Likewise, cannabis genotypes and their correlative chemotypes are now patent eligible despite these previous rulings. (US patents 6,630,507, 9,642,317, 9,095,554**,** 9370164, PP27,475)

Diamond versus Chakrabarty[4] clearly sets a patent eligibility precedence for human modified organisms and as a result the patent eligibility in the genomics and biotechnology sector has never been more confusing or more in flux. We believe the best policy in a rapidly changing

legal landscape is to be prolific with the publication of prior art to defend against current and future patent eligibility uncertainty. With the recent announcement of the FDAs desire to schedule more naturally occurring medicinal plants (*Mitragyna speciosa*), and the tendency of government agencies and corporations to patent such compounds or genotypes of said plants, we believe it is imperative to sequence and place public the *Mitragyna speciosa* genome to prevent future patent disputes over the medicinal activity of this natural plant. Mitragyna or "Kratom" has been reported to transition millions of people off of dangerous ethanol and opiates [5-10] and while cannabinoids are a safer alternative, they remain federally illegal.

A close examination of the opiate epidemic underscores the large regulatory hurdles put in place by the FDA. These restrictions create exhaustive barriers to entry to the drug market such that no corporation will approach the process without patent protection. Consider this incentive model combined with the PDUFA act of 1992[11] where the unelected FDA "defrays" their regulatory costs by taking payment directly from the companies they are supposed to regulate. This creates an enormous moral hazard in an agency that regulates 25 cents of every dollar. This poorly aligned incentive system leads to modification of safe compounds into patentable yet unsafe compounds that are killing over 100,000 US citizens annually. Peer to Peer rating technologies that regulate online economies for cars, hotels, and international crypto-currencies are demonstrating a safer decentralized model for communicating drug safety information.

In 2017, the FDA published a small sample of DNA sequence from *Mitragyna speciosa* in NCBI. Since patent applications usually require 18 months to publish, it is unclear if any patents have been applied for with this work. There do exist Mitragyna patent applications currently under review titled "Improved Methods For Making and Using Polynucleotide Sequences in the Synthesis of Alkaloid Compounds" and "Methods For Treating Withdrawal From Addictive Compounds". We believe the opiate epidemic is a serious issue and patents related to promising solutions to the problem are destructive to society. For this reason we believe it is imperative to push the entire genome public as soon as possible.

1. Myriad Av. https://en.wikipedia.org/wiki/Association_for_Molecular_Pathology_v._Myriad_Genetics,_Inc.
2. Wikipedia. https://en.wikipedia.org/wiki/Mayo_Collaborative_Services_v._Prometheus_Laboratories,_Inc.
3. Hampson AJ, Grimaldi M, Axelrod J, Wink D. Cannabidiol and (-)Delta9-tetrahydrocannabinol are neuroprotective antioxidants. Proceedings of the National Academy of Sciences of the United States of America. 1998 Jul 07;95(14):8268-73. PubMed PMID: 9653176. Pubmed Central PMCID: 20965.
4. Chakrabarty Dv. https://en.wikipedia.org/wiki/Diamond_v._Chakrabarty.
5. Leon F, Habib E, Adkins JE, Furr EB, McCurdy CR, Cutler SJ. Phytochemical characterization of the leaves of Mitragyna speciosa grown in U.S.A. Natural product communications. 2009 Jul;4(7):907-10. PubMed PMID: 19731590.
6. Halpenny GM. Mitragyna speciosa: Balancing Potential Medical Benefits and Abuse. ACS medicinal chemistry letters. 2017 Sep 14;8(9):897-9. PubMed PMID: 28947930. Pubmed Central PMCID: 5601368.

7. Kumarnsit E, Keawpradub N, Nuankaew W. Effect of Mitragyna speciosa aqueous extract on ethanol withdrawal symptoms in mice. Fitoterapia. 2007 Apr;78(3):182-5. PubMed PMID: 17335995.
8. Boyer EW, Babu KM, Adkins JE, McCurdy CR, Halpern JH. Self-treatment of opioid withdrawal using kratom (Mitragynia speciosa korth). Addiction. 2008 Jun;103(6):1048-50. PubMed PMID: 18482427. Pubmed Central PMCID: 3670991.
9. Babu KM, McCurdy CR, Boyer EW. Opioid receptors and legal highs: Salvia divinorum and Kratom. Clinical toxicology. 2008 Feb;46(2):146-52. PubMed PMID: 18259963.
10. Boyer EW, Babu KM, Macalino GE. Self-treatment of opioid withdrawal with a dietary supplement, Kratom. The American journal on addictions. 2007 Sep-Oct;16(5):352-6. PubMed PMID: 17882605.
11. FDA. https://en.wikipedia.org/wiki/Prescription_Drug_User_Fee_Act.

Mitragyna speciosa

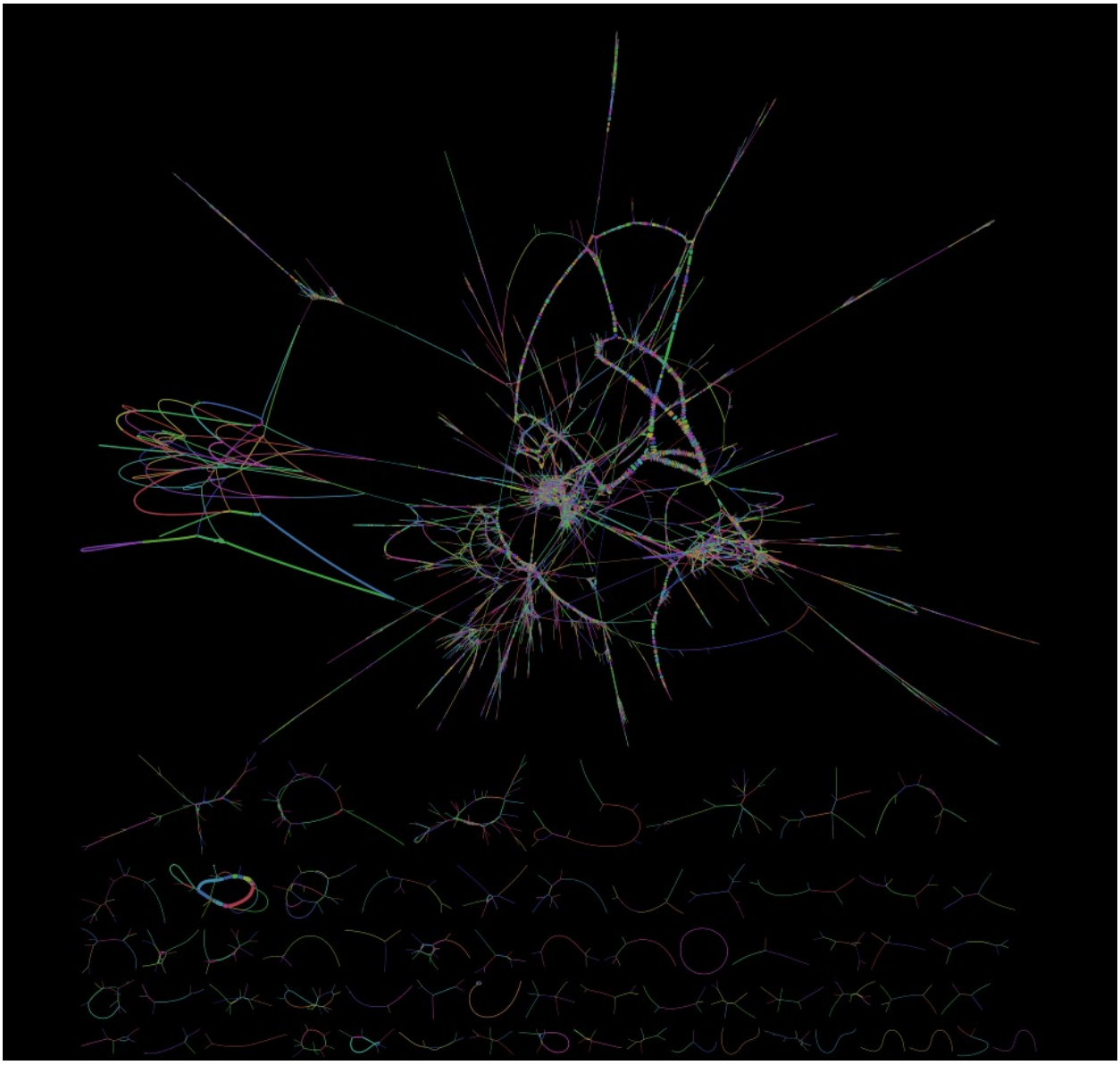
Presented within is a draft genome sequence of the Kratom genome (*Mitragyna speciosa*). This plant synthesizes compounds (mitragynine and 7-OH mitragynine) that have assisted millions of people abstain from opiate and alcohol abuse. Public sequencing of the *Mitragyna speciosa* genome can create prior art and thwart patent trolls that seek to exclusively own the benefits of this plant.

This is an incremental draft assembly of the first 42 million reads to gauge the size of the genome and its complexity. It is 65% AT and kmer analysis implies there are over 327 million bases in its genome.

Already large mitochondrial and chloroplast contigs are assembling and 100% of the 40 Mitragyna sequences in NBCI have strong and perfect ITS BLAST hits this assembly. While this expands the DNA sequence of this organism more than **10 fold from what was known previously**, ***we still need your help***. This is not a complete genome and it likely requires **10x more** sequence to establish a complete prior art picture. Crypro-currency addresses are included to bring more support for this important plant.

=-=-=-

Assembly graph of first batch of kratom DNA. Click for full-size image:



=--==--=-

These reference the old torrents, but are proof that we did this sequencing and shared it on the torrents on Nov 24, 2017 at 00:19:17 UTC:

DOWNLOAD, SHARE AND SEED THE FIRST BATCH OF KRATOM DNA WE SEQUENCED (It's a magnitude more bases than as NCBI / NIH released):

TORRENT MAGNET LINK

Please seed and tell two friends, this helps keep Kratom free.

This is the folder with sequence data plus associated files, blast, assembly, NIH's "pretty good for government work" attempt at sequencing, etc.

SHA256s_of_All_Files_in_that folder

Kratom DNA bonus files (Individual raw read files, Forward reads separated from Reverse Reads):
TORRENT MAGNET LINK.

SHA 256 OF file Mitragyna_speciosa_R1.fastq:
4A9D0B4716134C0EA869EC6CBB0A6A653F9ED3EF0243B399F985E122EA3427CE

SHA 256 OF file Mitragyna_speciosa_R2.fastq:
A4384C44B5597090C53D20AE480159B6569FDFC15F3235EDA1B439A45A899E86

898E8CB78BBC50B151698CB7993D0091814843FF

Get our small zip of all peer review text files via Torrent here (please seed!)

File name: FINAL-peer_review-by-Bryan-J_Jones_PhD.zip
Size: 381kB (389854 bytes)
SHA-256 of the whole zip is:
141CC6BD9199A38EE759620975523E745AAC0017B387C8981C9329976357B4D8

NEW TO TORRENTS? Download and install BitTorrent, then click those links above and our genome data files will start downloading.

It may take many hours, depending on your connection. These are big files.

Once downloaded, just leave the computer with BitTorrent running when you're not using it to "seed" the files to others.

Direct download removed due to lack of donations for expenses like hosting. Please use Torrents.

NameCoin blockchain AS A NOTARY FOR PROOF OF FIRST EXISTENCE OF OUR DATA:

We put proof of the two most important files in this first sequence on the Namecoin blockchain the day we released this. November 23, 2017 (Thanksgiving):

FILE 1:
File name: Mitragyna_speciosa_400bp_trimpaired_rp.fastq.gz

Value in Namecoin blockchain: **Kratom_DNA_Project-sequence1_Nov-23-2017_size_2point63GB_2827182770-bytes _Name_Mitragyna_speciosa_400bp_trimpaired_rp_dot_fastq__SHA256_ 7304756227F795B64F5E5208401C794D090234824B6E0A5EB1A176D04D8C54F8-WE_ARE_ALL_SATOSHI_NOW**

Registered on Namecoin blockchain at Nov 23, 2017 at 21:27:27 UTC in block 371804

Link on block explorer: https://namecha.in/name/i/Kratom_DNA_Project-sequence1_Nov-23-2017_size_2point63GB_2827182770-bytes%20_Name_Mitragyna_speciosa_400bp_trimpaired_rp_dot_fastq__SHA256_%207304756227F795B64F5E5208401C794D090234824B6E0A5EB1A176D04D8C54F8-WE_ARE_ALL_SATOSHI_NOW

FILE 2:
File name: mitragyna_400bp_paired_assembly.fa

Value in Namecoin blockchain: **Kratom_DNA_Project-sequence1_Nov-23-2017_size_304point99MiB_**319801332**-bytes _Name_** mitragyna_400bp_paired_assembly.fa **__SHA256_ 8E160BFD4CD36DFE96F2C3CA551C10AB53AF4B295D732EC6C3C1AD3E8E390D66-WE_ARE_ALL_SATOSHI_NOW**

Registered on Namecoin blockchain at Nov 24, 2017 at 00:19:17 UTC in block 371819

Link on block explorer: https://namecha.in/name/i/Kratom_DNA_Project-sequence1_Nov-23-2017_size_304point99MiB_319801332-bytes%20_Name_%20mitragyna_400bp_paired_assembly.fa%20__SHA256_%208E160BFD4CD36DFE96F2C3CA551C10AB53AF4B295D732EC6C3C1AD3E8E390D66-WE_ARE_ALL_SATOSHI_NOW

=-=-=

=-=-=

First Ph.D. Peer Review of Kratom Genome Project DNA – please torrent

Facebook

Twitter Share / Save

Peer review of our data was performed by Bryan J. Jones, Ph.D.

Dr. Jones confirmed in detail that:

1) Since the authors chose to remain anonymous, confirm you are not reviewing your own work or affiliated with the Kratom Genome Project.
2) Confirm the genome is Kratom with whatever methods are deemed appropriate. FastQC, BLAST and ITS regions were suggested.
3) Confirm The Kratom Genome Project sequenced X amount more than what is publicly available.
4) Four-day deadline

He published his results here on ResearchGate.net and also here on his LinkdIn.

Or read PDF here.

Or read on the page without leaving this website.

Get our small zip of all review text files via Torrent here (please seed!)

Please seed and tell two friends, this helps keep Kratom free.

That 381kB zip has peer review drafts, the final, the entire email thread between the peer reviewer and the Kratom Genome Project, and a list of the SHA-256 values for each file.

FINAL – M speciosa Genome verification – by Bryan J. Jones Ph.D.

File name: FINAL-peer_review-by-Bryan-J_Jones_PhD.zip
Size: 381kB (389854 bytes)
SHA-256 of the whole zip is:
141CC6BD9199A38EE759620975523E745AAC0017B387C8981C9329976357B4D8

Value in Namecoin Blockchain:
Kratom_DNA_Peer_Review1_Nov-29-2017_size_381kB_389854-bytes_Name_FINAL-peer_review-by-Bryan-J_Jones_PhD.zip_SHA256_141CC6BD9199A38EE759620975523E745AAC0017B387C8981C9329976357B4D8-WE_ARE_ALL_SATOSHI_NOW

Registered in block 372762, Nov 29, 2017, at 16:12:37 UTC.

Link on block explorer: https://namecha.in/name/p/Kratom_DNA_Peer_Review1_Nov-29-2017_size_381kB_389854-bytes_Name_FINAL-peer_review-by-Bryan-J_Jones_PhD.zip_SHA256_141CC6BD9199A38EE759620975523E745AAC0017B387C8981C9329976357B4D8-WE_ARE_ALL_SATOSHI_NOW

Here is our genome data that he reviewed.

Here is the previously published NCBI / NIH data he compared it to.

Dr. Jones' short form CV:
Postdoctoral researcher at the University of Minnesota.
Publication and citation record:
https://scholar.google.com/citations?user=eOFC7bEAAAAJ&hl=en
Latest publication http://pubs.acs.org/doi/abs/10.1021/acs.biochem.7b00571

His background is in biochemistry with genetics, experience working with NCBI databases and BLAST. He also wrote a program for designing stabilizing mutations that interfaces with BLAST and Entrez.

We held an open call for a subtitle candidate to review the data. Many applied. Rather than wait months or years for normal peer review channels, we paid Dr. Jones for his time (but not to influence his results. He's a respected science and isn't going to throw a game).

He got it done in 3 days. And it cost us 1.5 cents US worth of Namecoin coin to irrefutably prove on the Namecoin blockchain that he published the document the day we say here that he did.

Bryan was paid 1/10th of one Bitcoin for his time. Half was paid to him up front, and half upon completion. The total .1 BTC was worth about 800 USD when offered (400 x 2 payments). It was worth nearly 1000 USD a few days later when he turned in his report and was paid the rest. It was enough to get it done quickly.

Some will protest "But it's not pure to pay money!" Well, there is already money in peer review. The 1000 USD paid to a RAID Reviewer for prompt service is comparable to the 1000 USD document processing fee here: https://f1000research.com/for-authors/article-processing-charges

And that's just to publish in the most decentralized way near the mainstream. Open-access journal charge around $1,000–$2,000 to publish. Then you have to attract peer reviewers.

We call our new peer review technique *RAID Review* (Reverse Anon Incentivized Direct Review). We believe RAID Review will probably not go over well with the established community, but is required to get things done when there are forces coming at it from every side.

Even beyond that, RAID Review has the power to streamline discovery and implementation.

For more on RAID Review, please read *Fixing what's wrong in Peer Review w/ RAID Review (Reverse Anon Incentivized Direct Review).*

Oh, by the way, See http://www.raidreviews.org where the Kratom Genome Project just nailed our RAID Review plans to the front door of science.

Chemotype certificate for Red Vein Thai kratom sample we sequenced.

=-=-=-=--===-

BLAST Analysis of our Initial Sequence

Facebook
Twitter

There are two sources of *Mitragyna speciosa* data that we are aware of.

A search of NCBI will provide 40 sequences. These are well annotated sequences that sum to 344,569bp including two chloroplast sequences, ITS sequences and many genes of interest. They are derived from multiple different researchers from around the world using multiple different sequencing technologies. Anyone can download these as a FASTA file posted here.

Below is a BLASTN output table of these 40 sequences aligned to our assembly. All of them have hits. Some SNPs are expected as the Red Vein Thai sample sequenced is not the same as previous strains sequenced.

Rows: 40

Query	Number of HSPs	Lowest E-value
JF412821.1	20	
JF412820.1	17	
JF412819.1	16	
JF412818.1	13	
JF412817.1	6	
JF412816.1	13	
JF412815.1	6	
JF412814.1	6	
JF412813.1	12	
JF412812.1	8	
JF412811.1	20	
JQ038374.1	14	
JQ038373.1	25	
JQ038372.1	28	
HM543187.1	12	
AB249645.1	6	
NC_034698.1	48	
KY085908.1	48	
KC737823.1	11	
KC737766.1	10	
KC737719.1	10	
KC737658.1	10	
KC737618.1	6	
KC737576.1	4	
JF412827.1	11	
JF412826.1	11	

The ironclad evidence that this sequence is in fact *Mitragyna speciosa* is the 100% perfect BLAST hit to the published ITS sequence AB249645.1

```
Database: /Users/kevinmc/CLCdatabases/mitragyna_400
Query= AB249645.1
Length=608


Sequences producing significant alignments:

mitra_S1_R1_001_(paired)_contig_56  No definition l
mitra_S1_R1_001_(paired)_contig_166426  No definiti
mitra_S1_R1_001_(paired)_contig_164498  No definiti
mitra_S1_R1_001_(paired)_contig_4336  No definition
mitra_S1_R1_001_(paired)_contig_111030  No definiti
mitra_S1_R1_001_(paired)_contig_267547  No definiti

ALIGNMENTS
>mitra_S1_R1_001_(paired)_contig_56 No definition l
          Length=9283

 Score = 1097.7 bits (1216),  Expect = 0E00
 Identities = 608/608 (100%), Gaps = 0/608 (0%)
 Strand = Plus/Minus

Query  1     TCGAATCCTGCAAAACGCACGACCGTGAACCCGTGTTA
             ||||||||||||||||||||||||||||||||||||||
Sbjct  8051  TCGAATCCTGCAAAACGCACGACCGTGAACCCGTGTTA

Query  61    TGGAGACTAAGCCCTCCCTTCCTTCCCGGCGCTCCCCG
             ||||||||||||||||||||||||||||||||||||||
Sbjct  7991  TGGAGACTAAGCCCTCCCTTCCTTCCCGGCGCTCCCCG

Query  121   TAACTCAAACCCGGCGCGGAACGCGCCAAGGAAAACTC
             ||||||||||||||||||||||||||||||||||||||
```

There is also a SRA archive of 1.3M reads from the FDA. You will need to download specialized software and be knowledgable with command line interfaces to download this data.

Run	Spots	Bases	Size	GC content	Published	Access Type
SRR5602600	1.3M	658.7Mbp	420.7M	35.4%	2017-05-25	public

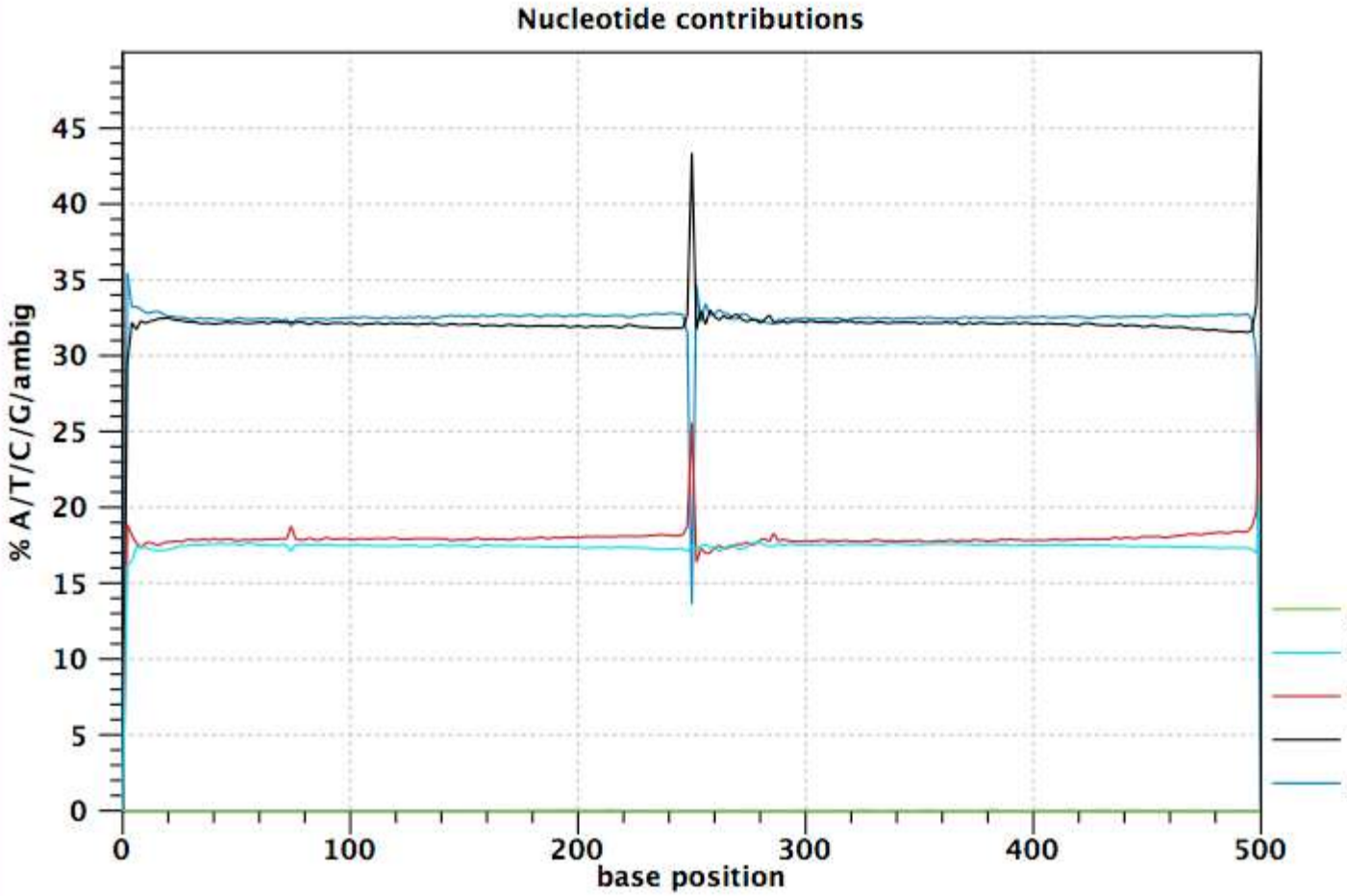
This run has 2 reads per spot:
L=248, 100% L=248, 100%

We have attempted to assemble this data with little success. After further inspection it appears the FastQ file has the Forward and Reverse reads concatenated into a single 499 base pair read.

```
@SRR5602600.1 1 length=499
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTTTTTGGTTTATAAAGGTATGAGTACTGGAGGATACATGCAAAAGTATCTGGTAGAAAACATGACAT
ATAATAAATAAAGCAATAACTCAGAATGCTACACAATAAATAAATTGAACTATCAAGTAAAAACGCACACAATAAAGCTATGATGAAATGATAACGAACAAAT
TTTATTGGAGTAACAAGTTCTTTATTCTTTTCCATTCTTTATTATTGAAGAGTTAATATTGTTATTTTCCTTATCCCTTTTGCTTTTGTTTCTCGTTATTTTT
TCTTTCTTTTTTTGTTTAGTTTATTTTGTTCTTTTTTTTT
+SRR5602600.1 1 length=499
>>>>111>>>>>//AAA/>/>/////</<<</////<0<0.--..<1000000000/0000000//.../000000000//0000000000000//.0/0/0(
//9///////9/////////////;///////-/-/////////////////:////////////---------/////////////;///////9///-A1>AAF1
12BDGDF1FFBGDFFHHD22D1DDB2EGFG1DB2BDG2DD2DGFGGF@DG1>12@BF22B2BF22BBB@11221111BB1111111110B0222/////B221
11111111>11---///000;0000000//:00<009-;-
@SRR5602600.2 2 length=499
ATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAGGTTTAAAGGGAGACGTAGGAGCAAGTCGTTGACTAGCGTGTAGTAAAAGGGGACTAGTAGACCTTCC
ATAAAAATATTAGTTTGTATTTGTAAAACGAGTTTTGGGGTAACTCGGAATACTTTAACGTAATCTCAACATGTGTTTAGTAAATGTTGATAATAAGGTGTCC
ACAATCCTTTCTGGTGACCATTTCCAAAAAAATAGAGTTTTGAAAAAACTGTACGTGTGTGAAAATTAATCTGATTTTTATTGTAATATTTTATTTCGTAAA
GGCCATTTTCCTTTTTATACTGGTTTTATTTTTAATCTAT
+SRR5602600.2 2 length=499
>11>>3110>>>////AA////>/>//<<<//</!0??1111111//...>..0.10.</!00<../.000//-;../00000:0..-.-.//;090900;/.
/////9/-/9//9/9//;//;/9///9//---9-9;------////-----/////////-9---9/99/-/;/;/9/-////////;/////////33>3>CA
EFGFFHHHG4BBGBF1B3EHH3GGHEFFGA@?GGD33B?FG?FEFFACG1GGF2/0<</@01G>111111?1?1@F1<F/1<11111111111>F1=10000.
--/0//00000;90//.////////:AE/;/;/;.//9/9/
@SRR5602600.3 3 length=497
CTCTATTATATGAAAGCATCAATAAAATACTTTCTACAGTCCTCAAACATTCGTAAAAAGATGTTTGAGCAATTGTAGCCTCCTTTTTCAACTACAAAGAAGC
AGCTAACAAGTTGCCGGGTTTTCCTTGCGTCCCATTGCATCATATGAAACAACCAACCTATTTCATTTTTACTCTCGAAGCCAAAAATATTCACTCTACGGTT
TAATATTTTTTATGCATTTACATTTGGATGTTAATTTACCGAATGTAATTAGGTTAAAAATGTGCGTCTTTTGTAATTTTTGTATGTTTGTATGTTTATTATT
TTTAGACTTTTGTTCTTGTTTTGGGTGCTTTCTGATAT
+SRR5602600.3 3 length=497
AA1AA3@3BD@333BA1BFGGGB3A11FG1DDGHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHGGHHGHHHHGHHHHHHHHDGHHHHHHHHHHHHHHHHHHHHHHC
GHHHHHHHFBBGGFHGGCAFGGGH1FGGGCGGCHFHGHHHHHHHHHHHH0GHHHGGHGCF/:CHFFHFHG0GHHHHECGGEAFGFEGGEFGGGGGFGAABBBFA
HEHHHHHHGHGHHHGHHHHBFFGHH5D3GHFFHHHHHH5FAA>FGHHFHHFHFHHHHB43BFFG411/?FGBG2?444BB?/B4B4FHGB24BFGFF4FF4GH
;:B0/00<;G00/;0;00;09BF-..;9F9F00000;0
@SRR5602600.4 4 length=498
CAATTAATTTTGGACCTGAGAAAAAGCTATTAATTCTTCCATAAAGAAGATCCTAACTGTTTCAATGACATTTGAACTCTCATAACAGCTTGTTAATTGTCCT
AAACACTATAATCGTCTGAAATGCTAAGACTGCTCAGTTTTACTTGTCTTATAACCTTCAATTTCCACACCTAATTTCATTGATGTGAAAAGTTAAAAGTTGC
AAATTGAAGGTTATAAGACAAGTAAAACTGAGCAGTCTTAGCATTTCAGACGATTATAGTGTTTACTTGTACAATTGAAGCAAGATTGTTTTAGTGCTTTAGA
TGAAACAGTTAGGTTCTTCTTTATGGAAGAATTAATAGC
+SRR5602600.4 4 length=498
AAAAB5BFFFFFFBGGGGGGGGFHHGHBGFFGHHHHHHHHHHHHHGHGGHHFHHGGFHHHFFGHHHHFHHGHHHHHDFFHHHHHHHHHHHHHHHGHHHHHDFGG
FGHHFHHHHHHGHGHHHFHHHHFHHHHHHHHHHHHHHHGFHGHHHHHGHFHGFGHGFHFFGFH>GHFHFHGGHFFHGHHHHHHHHFBGGHHFH<DG>AAAAFFF
HHHHEEFGHHFHHHHHHGFHHHGHHHHFHHHHHHHGFFBGHHHHDFHHHHFEEEFFHGHF44FCDGHHHF4E?FDFH333?FFHHHGBFHFHHGFGDHHHEHE
B000=D//DG0<G0000CGH0;CFEH0;:C:0CGFH0C0
@SRR5602600.5 5 length=498
CATTCTCTTTATAATCTAAAATTTGATGGTTATTTTTTTCTTTGTAGCATTTTAATAGGTGGGATATATTTCATTGATGCTGGTGCAGATGCCCTGTGGAACA
TACTGATGTGATGCTGAAAATGTGAATTAAGGTAAACCCAGTTATGATAAAGCTGCAACACTGAAAAGAGAATTAAGGAAAACACAATTATAATAAGAATTTA
TTTTTATCTGAACAACTGCATAACTATTAGGTTGATTGACCAATTTTTTTAATAGCTTAACTACTTGATTTTTTTTAACAGTAACACAAGCACTTGATTAATT
CAGCTAGCTTTCTGTATGTTGTCTCTGATTTGTCCTCTA
+SRR5602600.5 5 length=498
>?AAA5@DDFFFGFGGGGGFGGAGH6FFH4D5GHGHHGGHHHHHGHHHHHHHHHGHHHHGBFFGHGHHHHHHHHHHHGHHHFHHGHHHHHFHHGHHGHHHGBFGF
FFHFHHHHHHHHHHHHBGHHHHFGG2GHHEHHDFFHGHGGHGHHHHHGHHHHHGHFF1GHGGHFFGHFEFGHHHHGHFHFBFFHHGEHFFFHH0;F3>3>AFF
HHHHGHHEFHBGHHHHGHGHHHHGHGHFGGFDGFGHHFHHHHHHGHGHGCHHHHHHFH3BFGHCFF33BFHHGEGG/?3B2@F2F><</@/2@CFD1FGD1FH
GCB0:CC:;C009;90;B999:000FBFF009;00;CBC
@SRR5602600.6 6 length=499
```

This can be seen with an AT analysis over Read Length which demonstrates a spike at the strand flipping base at 250bp. This is consistent with the SRA table above. In order to make use of this data one needs to decouple these Forward and Reverse reads so assemblers do not attempt to assemble each read as contiguous 499bp reads and adapter trim the strand flipping base and the 1st and last base which appear to be adapter derived.

3.2 Nucleotide contributions

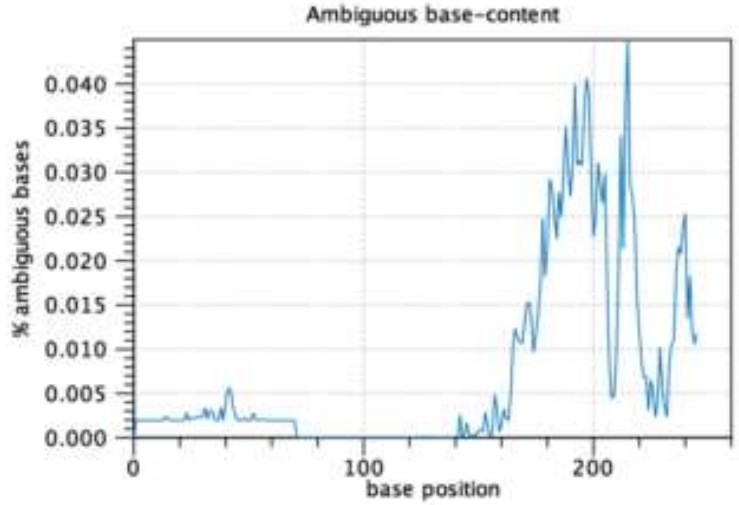


Coverages for the four DNA nucleotides and ambiguous bases.
x: base position
y: number of nucleotides observed per type normalized to the total number of nucleotides observed at that positi

The unix command cut can help trim the first 200 bases off the reads.

`cut -c 1-200 in.fastq >out.fastq` These reads will map to the assembly with elevated error rates on the 3 prime end of the reads.

3.4 Ambiguous base-content



Combined coverage of ambiguous bases.
x: base position
y: number of ambiguous bases observed at current position normalized to the total number of bases observed at that position

Mapping the highest quality first 100 bases of these reads to the assembly produces 62% of the reads mapping. This implies that we are still on the most productive part of the shotgun sequencing curve. One more run should substantially improve the quality of the assembly.

1 Mapping summary report

1.1 Summary statistics

	Count	Percentage of reads	Ave
References	335,915	-	
Mapped reads	832,410	62.70%	
Not mapped reads	495,124	37.30%	
Total reads	1,327,534	100.00%	

One can assemble the unmapped reads and BLAST them against NCBI and one will find small hits to various plant Mito and Chloroplast DNA. These are likely repetitive or NUMT DNA that is not assembling at lower coverages.

Use the NCBI sratoolkit Fastq-dump -I –split-files SRR5602600 to separate these reads more cleanly.

-=-=-=-=-=

PDF of analysis results from Wonderland Labs: 112775-Certificate-ID:

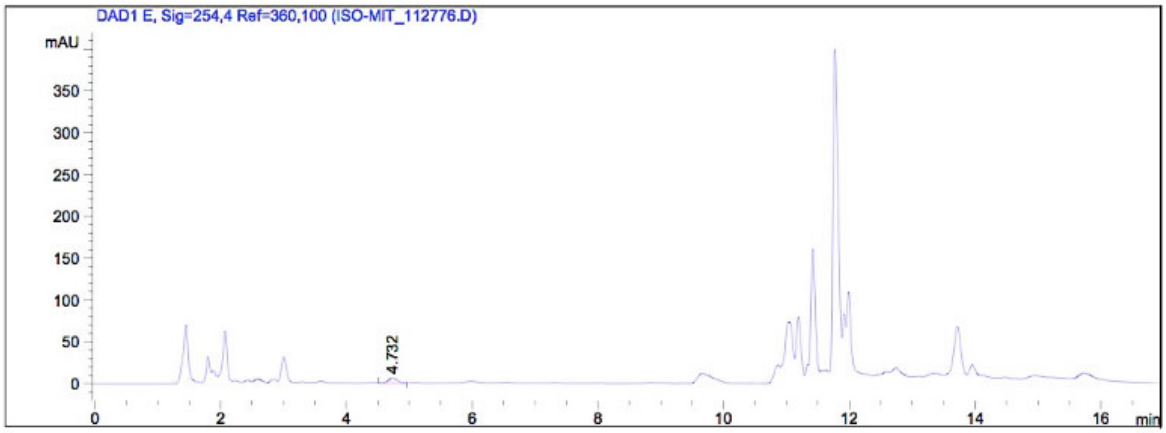
WONDERLAND LABS

Certificate of Analysis

1 DECEMBER 2017

Red Vein Thai / Kratom

Client:	Kratom Genome	Acquired By:	James C.
Sample Type:	Powdered	Date Acquired:	11/29/2017
Sample Weight:	2000mg	Acquisition Method:	MIT_KN1-R02-18M
Dilution:	1.0	Testing ID:	ISO-MIT_112776.D
Vial:	Vial 2	Date Processed:	12/01/2017
Injection Volume:	8.3 µl	Run Time:	18 min
Sample Set:	112775	Calibration ID:	999-11111



#	Analyte	RT (min)	Area	Height	Alkaloid	Total %
1	Mitragynine	4.732	66.56819	6.39543	$C_{23}H_{30}N_2O_4$	0.14
2	7-Hydroxymitragynine	N/A	N/A	N/A	$C_{23}H_{30}N_2O_5$	N/A

Download PDF of analysis results from Wonderland Labs: 112775-Certificate-ID

Wonderland Labs web site.

Verify certiricate with Wonderland Labs here.

The USPTO has issued cannabis patents that claim both genotype and chemotype. Prior art is strongest when presented with both.
A good read on Mitragynine and 7-hydroxyMitragynine.
http://www.emcdda.europa.eu/publications/drug-profiles/kratom

Category: Kratom DNA

=-=---=-