

# Some Questions

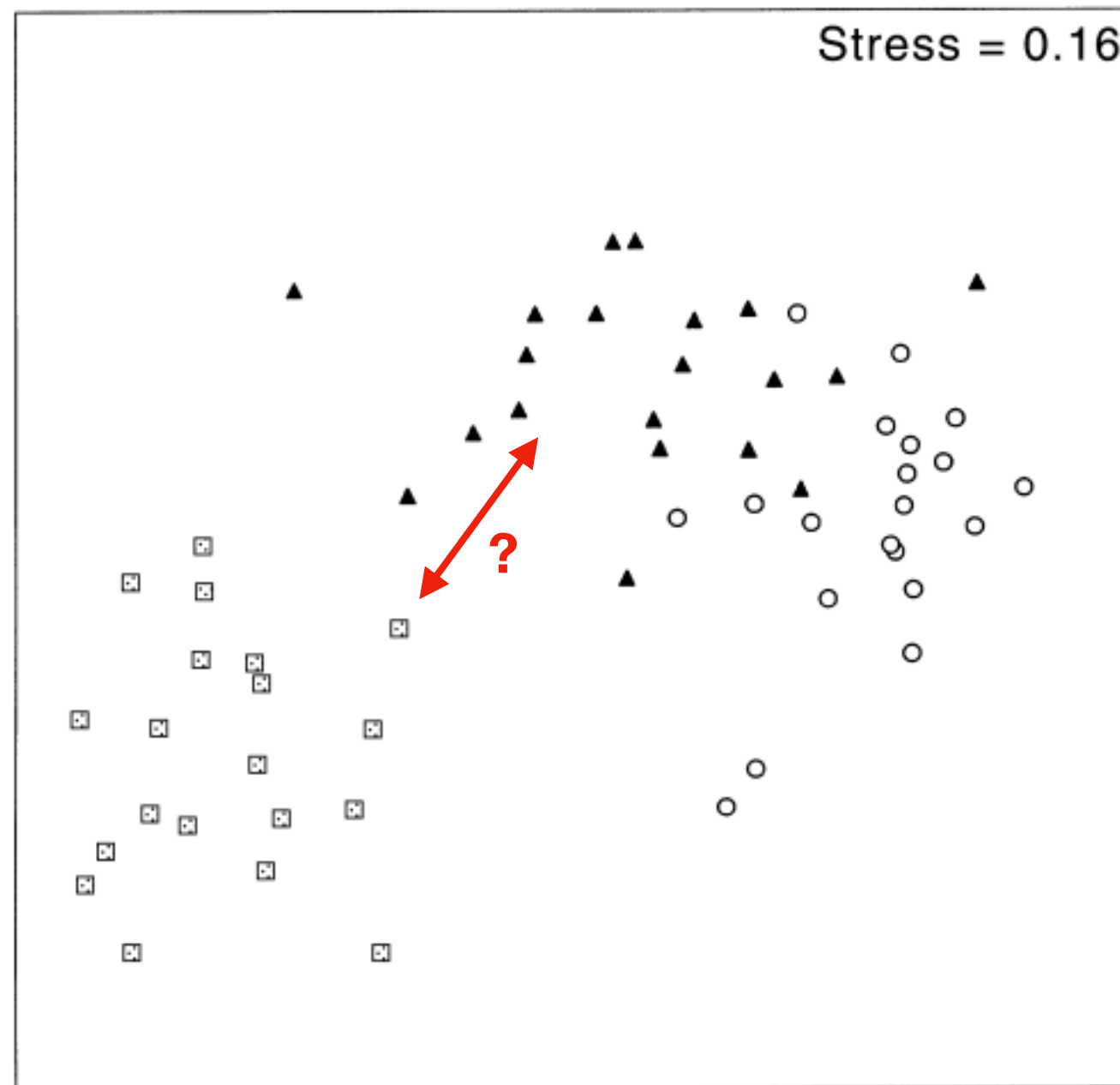
February 7, 2020

Josh Granek

# Questions

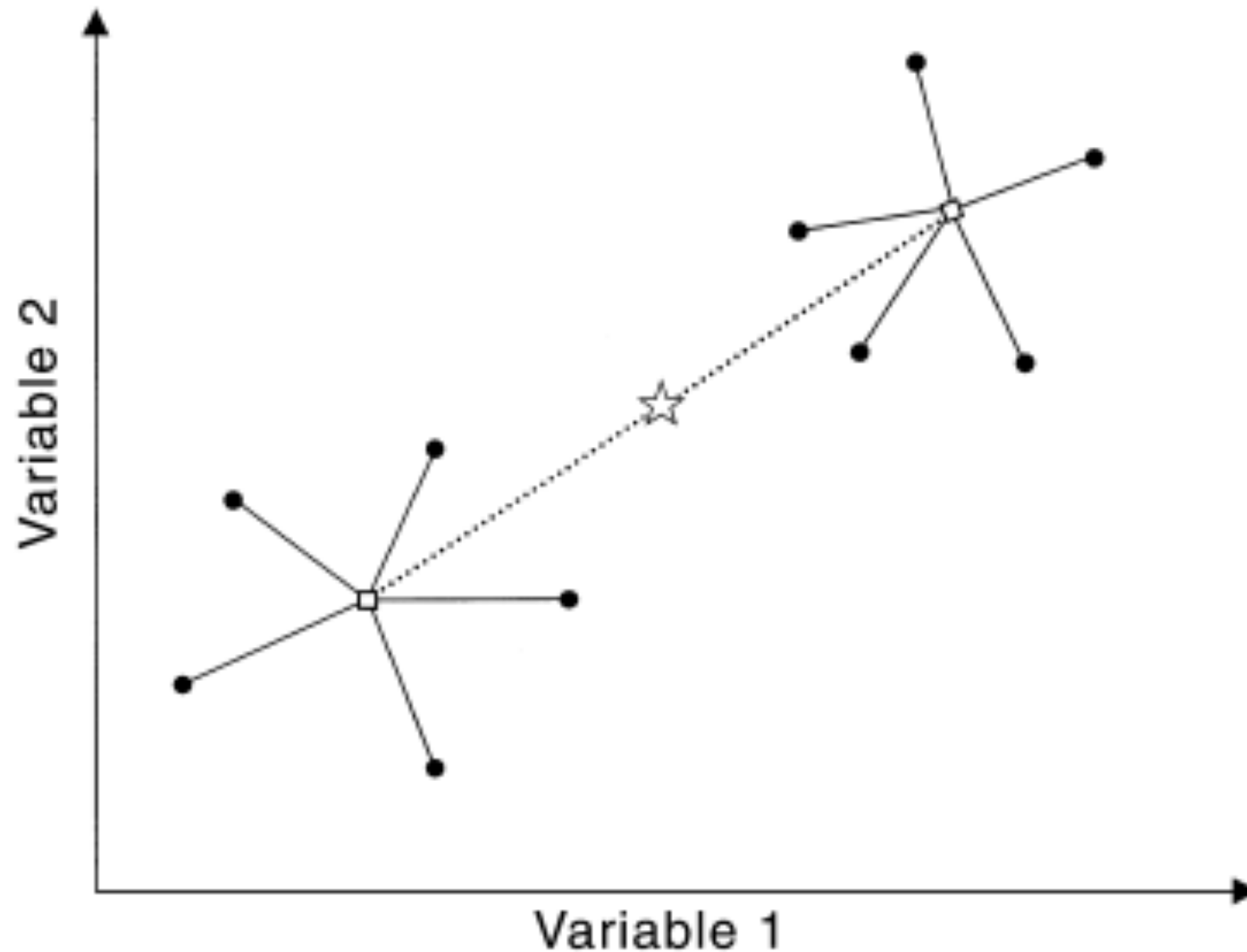
1. Is there a difference in the microbial communities present in condition A samples and condition B samples?
2. What **taxa** are different between condition A and condition B?
3. What **genes** are different between condition A and condition B?

# Are Communities Different?

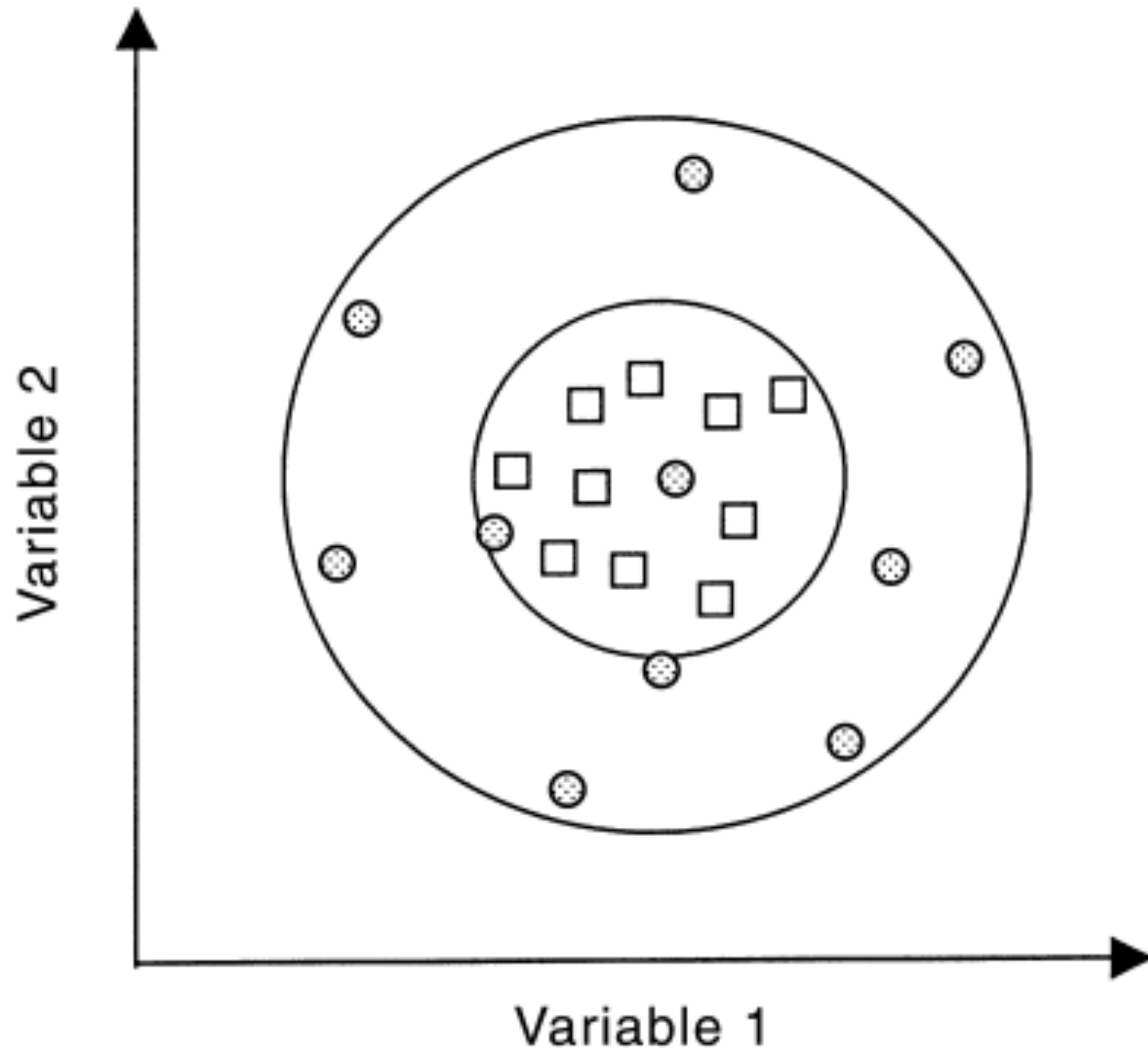


▲ Cats  
○ Dogs  
■ Rabbits

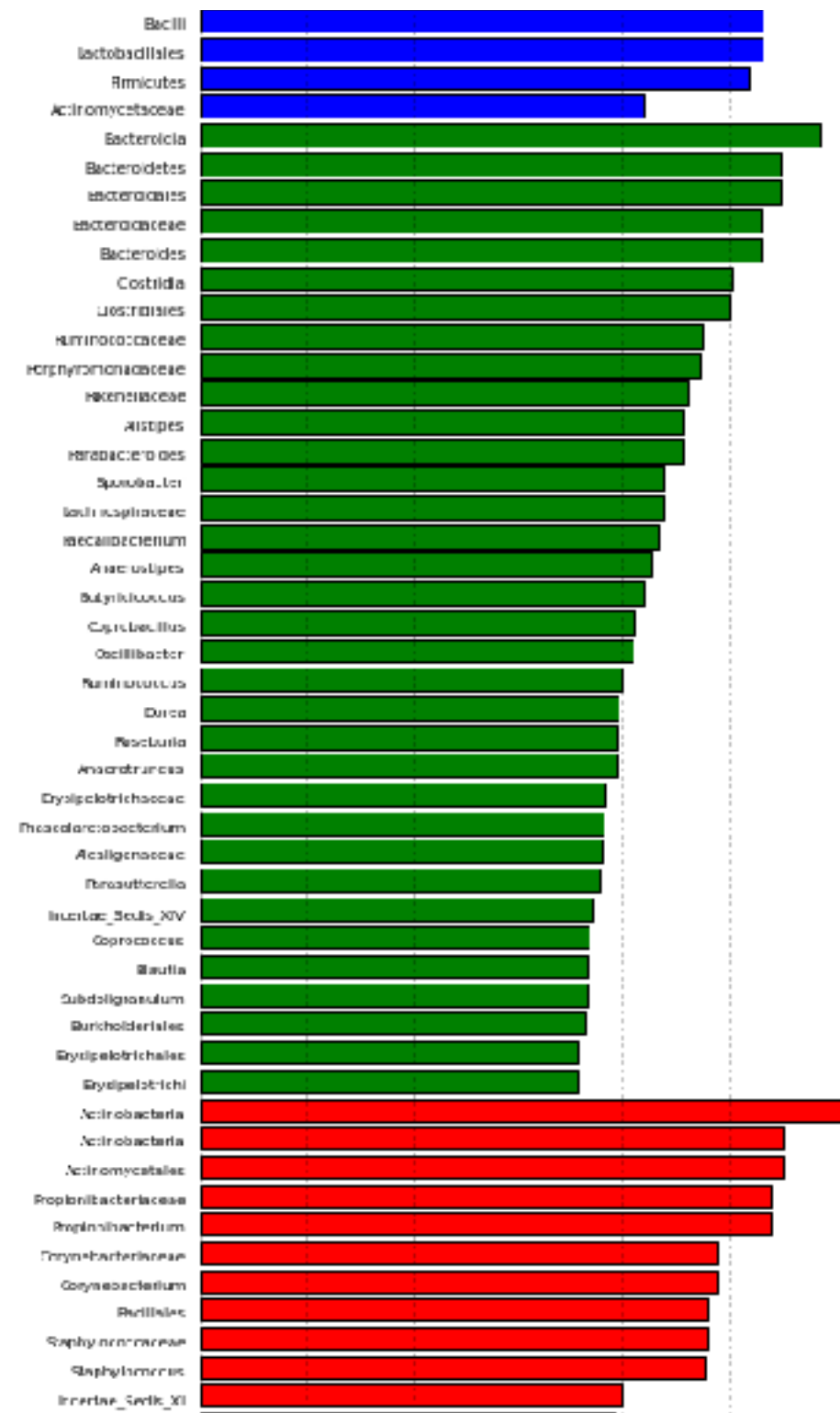
# PERMANOVA



# PERMANOVA Caveat



# What Taxa are Different?



# Biomarker Discovery

- LEfSe
- MetaBoot
- Metastats
- LIBSVM
- mRMR
- Regularized Low Rank-Sparse Decomposition (RegLRSD)

# What Genes are Different?



# Metagenomics

	What	Information	Analogy	Target Size	Cost
<b>Amplicon</b>	Marker Gene	Who is Present	Name	100bp - 1kb	Low
<b>Shotgun Metagenome</b>	Genomes	What Genes are Present	CV	100kb - 100Mb	High
<b>Shotgun Metatranscriptome</b>	All RNA	What Genes are Expressed	Twitter Feed	100kb - 100Mb	High

# Amplicon Sequencing

PCR amplify and sequence a marker gene

	Marker Gene
<b>Bacteria</b>	16s rRNA
<b>Fungi</b>	18s or ITS rRNA
<b>Archaea</b>	16s rRNA
<b>Protozoa</b>	18s rRNA
<b>Viruses</b>	?????

# Metagenomics

	What	Information	Analogy	Target Size	Cost
<b>Amplicon</b>	Marker Gene	Who is Present	Name	100bp - 1kb	Low
<b>Shotgun Metagenome</b>	Genomes	What Genes are Present	CV	100kb - 100Mb	High
<b>Shotgun Metatranscriptome</b>	All RNA	What Genes are Expressed	Twitter Feed	100kb - 100Mb	High

# Metagenomics

	What	Information	Analogy	Target Size	Cost	Discovery?
<b>Amplicon</b>	Marker Gene	Who is Present	Name	100bp - 1kb	Low	+/-
<b>Shotgun Metagenome</b>	Genomes	What Genes are Present	CV	100kb - 100Mb	High	++
<b>Shotgun Metatranscriptome</b>	All RNA	What Genes are Expressed	Twitter Feed	100kb - 100Mb	High	++

# PICRUS<sub>t</sub>

- **What I Have:**

250bp sequence from v4 region of 16s  
rRNA gene

- **What I Want:**

1. All the genes in the sample

2. The relative abundance of all the genes in  
the sample

# Inferring Gene Content

## 16s rRNA v4

```
GCGAGCGTTGTCCGGAATTATTGGGCGTAAAGAGCGTGTAGGCGGTTTCGGT  
AAGTCTGCCGTGAAAACCTGGGGCTCAACCCCGGGCGTGCGGTGGATACTG  
CCGGGCTAGAGGATGGTAGAGGCGAGTGGAATTCCCGGTGTAGCGGTGAAA  
TGCGCAGATATCGGGAGGAACACCAGTAGCGAAGGCGGCTCGCTGGGCCAT  
TCCTGACGCTGAGACGCGAAAGCTAGGGG
```

Rubrobacter

The screenshot shows a web browser window with the URL [www.ncbi.nlm.nih.gov/genome/?term=Rubrobacter](http://www.ncbi.nlm.nih.gov/genome/?term=Rubrobacter). The page is the NCBI Genome database search results for "Rubrobacter". The search bar shows "Genome" and "Rubrobacter". Below the search bar, there are links for "Create alert", "Limits", and "Advanced". The "Display Settings" dropdown is set to "Summary". The "Search results" section shows "Items: 5". The first result is "1. [Rubrobacter xylanophilus](#)", which is a "Cellulose-degrading bacterium". The details for this result are: Kingdom: Bacteria; Subgroup: Actinobacteria; Sequence data: genome assemblies:1; Chromosome: 1; Date: 2006/08/09; ID: 1131.

NCBI Resources How To

Genome Genome Rubrobacter

Create alert Limits Advanced

Display Settings: Summary Send to:

**Search results**

Items: 5

☐ [Rubrobacter xylanophilus](#)

1. Cellulose-degrading bacterium

Kingdom: Bacteria; Subgroup: Actinobacteria

Sequence data: genome assemblies:1

Chromosome: 1

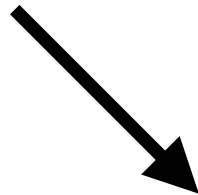
Date: 2006/08/09

ID: 1131

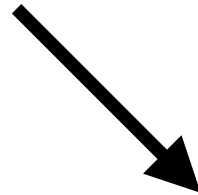
# Inferring Gene Content

## 16s rRNA v4

GCGAGCGTTAATCGGAATTACTGGGCGTAAAGGGCGCGTAGGCGGTGAAGT  
AAGTCGGGTGTGAAAGCCCCGGGCTCAACCTGGGAACTGCATCCGATACTG  
CTTCGCTAGAGTATGGTAGAGGGAAGCGGAATTCCGGGTGTAGCGGTGAAA  
TGCGTAGATATCCGGAGGAACACCAGTGGCGAAGGCGGCTTCCTGGACCAA  
TACTGACGCTGAGGCGCGAAAGCGTGGGG

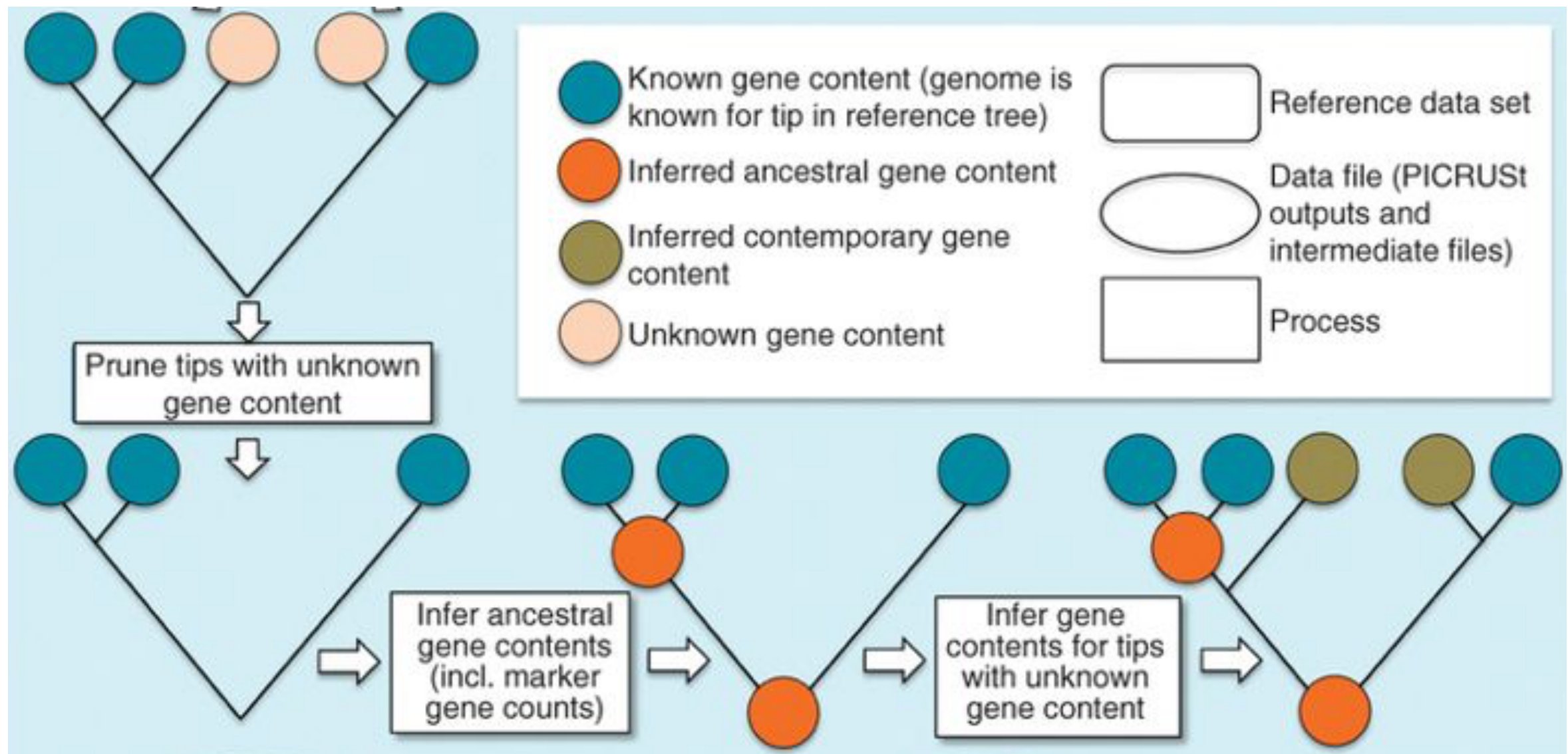


?



?

# Inferring Gene Content





graneek JA - PubMed - NCBI

www.ncbi.nlm.nih.gov/pubmed/?term=graneek+JA

NCBI Resources How To Sign In to NCBI

PubMed  
US National Library of Medicine  
National Institutes of Health

graneek JA Search

Create RSS Create alert Advanced Help

Article types  
Clinical Trial  
Review  
Customize ...

Text availability  
Abstract  
Free full text  
Full text

PubMed  
Commons  
Reader comments  
Trending articles

Publication dates  
6 years  
10 years  
Custom range...

Species  
Humans  
Other Animals

Clear all  
Show additional filters

Summary 20 per page Sort by Most Recent

Search results  
Items: 17

☐ Evidence for distinct brain networks in the control of rule-based motor behavior.  
1. Graneek JA, Sergio LE.  
J Neurophysiol. 2015 Aug;114(2):1298-309. doi: 10.1152/jn.00233.2014. Epub 2015 Jul 1.  
PMID: 26133796  
[Similar articles](#)

☐ Rapid mapping of insertional mutations to probe cell wall regulation in *Cryptococcus neoformans*.  
2. Esther SK, Graneek JA, Alspaugh JA.  
Fungal Genet Biol. 2015 Sep;89:8-21. doi: 10.1016/j.fgb.2015.08.003. Epub 2015 Jun 23.  
PMID: 26112692  
[Similar articles](#)

☐ Integrating chemical mutagenesis and whole-genome sequencing as a platform for forward and reverse genetic analysis of *Chlamydia*.  
3. Kokes M, Dunn JD, Graneek JA, Nguyen BD, Barker JR, Valdivia RH, Bastidas RJ.  
Cell Host Microbe. 2015 May 13;17(5):716-25. doi: 10.1016/j.chom.2015.03.014. Epub 2015 Apr 23.  
PMID: 25920970 Free PMC Article  
[Similar articles](#)

☐ Antifungal drug resistance evoked via RNAi-dependent epimutations.  
4. Calc S, Shertz-Wall C, Lee SC, Bastidas RJ, Nicolás FE, Graneek JA, Mieczkowski P, Torres-Martínez S, Ruiz-Vázquez RM, Cardenas ME, Heltman J.  
Nature. 2014 Sep 25;513(7516):555-9. doi: 10.1038/nature13575. Epub 2014 Jul 27.  
PMID: 25079329 Free PMC Article  
[Similar articles](#) 1 comment

☐ Decoupled visually-guided reaching in optic ataxia: differences in motor control between canonical and non-canonical orientations in space.  
5. Graneek JA, Pisella L, Stemmerger J, Vighetto A, Rossetti Y, Sergio LE.  
PLoS One. 2013 Dec 31;8(12):e80136. doi: 10.1371/journal.pone.0080136. eCollection 2013.  
PMID: 24302036 Free PMC Article  
[Similar articles](#)

☐ The genetic architecture of biofilm formation in a clinical isolate of *Saccharomyces cerevisiae*.  
6. Graneek JA, Murray D, Kayıkcı Ö, Magwene PM.  
Genetics. 2012 Feb;190(2):597-603. doi: 10.1534/genetics.112.142067. Epub 2012 Nov 19.  
PMID: 23172859 Free PMC Article  
[Similar articles](#)

☐ The role of the caudal superior parietal lobule in updating hand location in peripheral vision: further evidence from optic ataxia.  
7. Graneek JA, Pisella L, Blangero A, Rossetti Y, Sergio LE.  
PLoS One. 2012;7(10):e48619. doi: 10.1371/journal.pone.0048619. Epub 2012 Oct 5.  
PMID: 23071689 Free PMC Article  
[Similar articles](#)

☐ Pleiotropic signaling pathways orchestrate yeast development.  
8. Graneek JA, Kayıkcı Ö, Magwene PM.  
Curr Opin Microbiol. 2011 Dec;14(6):676-81. doi: 10.1016/j.mib.2011.06.004. Epub 2011 Sep 28. Review.  
PMID: 21962291 Free PMC Article  
[Similar articles](#)

Send to: Filters: Manage Filters

Find related data  
Database: Select  
Find items

Search details  
graneek JA [Author]  
Search See more...

Recent Activity  
Turn Off Clear

graneek JA (17) PubMan

graneek J (20) PubMed

Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. PubMed

Scott Harrison (71) PubMed

See more...

# 16s rRNA v4

GCGAGCGTTGTCCGGAATTATTGGGCGTAAAGAGCGTGTAGGCGGTTTCGGT  
AAGTCTGCCGTGAAAACCTGGGGCTCAACCCCGGGCGTGCGGTGGATACTG  
CCGGGCTAGAGGATGGTAGAGGCGAGTGAATTCCCGGTGTAGCGGTGAAA  
TGCGCAGATATCGGGAGGAACACCAGTAGCGAAGGCGGCTCGCTGGGCCAT  
TCCTGACGCTGAGACGCGAAAGCTAGGGG

Rubrobacter  
(genus)

www.ncbi.nlm.nih.gov/genome/?term=Rubrobacter

NCBI Resources How To

Genome

Create alert Limits Advanced

Display Settings: Summary

Send to:

## Search results

Items: 5

- ☐ [Rubrobacter xylanophilus](#)  
1. Cellulose-degrading bacterium  
Kingdom: Bacteria; Subgroup: Actinobacteria  
Sequence data: genome assemblies:1  
Chromosome: 1  
Date: 2006/08/09  
ID: 1131
- ☐ [Rubrobacter radiotolerans](#)  
2. **Rubrobacter radiotolerans** overview  
Kingdom: Bacteria; Subgroup: Actinobacteria  
Sequence data: genome assemblies:2  
Chromosome: 1; Plasmids: 3  
Date: 2014/05/01  
ID: 11149
- ☐ [Rubrobacter indicooceani](#)  
3. Kingdom: Bacteria; Subgroup: Actinobacteria  
Sequence data: genome assemblies:1  
Chromosome: 1; Plasmids: 2  
Date: 2018/09/13  
ID: 72442
- ☐ [Rubrobacter aplysinae](#)  
4. Kingdom: Bacteria; Subgroup: Actinobacteria  
Sequence data: genome assemblies:1  
Date: 2015/08/18  
ID: 38493

Genome

Genome

Limits Advanced

[Organism Overview](#); **Genome Assembly and Annotation report [15657]**; [Genome Tree report \[8363\]](#); [Plasmid Annotations](#)

## Escherichia coli

Partial: [All](#) Anomalous: [All](#) Levels: ☒ All ☒ Complete [776] ☒ Chromosome [89] ☒ Scaffold [4852] ☒ Contig [9940]

Organism/Name	Strain	CladeID	BioSample	BioProject	Assembly	Level	Size (Mb)
<a href="#">Escherichia coli IAI39</a>	<a href="#">IAI39</a>	19668	<a href="#">SAMEA3133234</a>	<a href="#">PRJNA33411</a>	<a href="#">GCA_000026345.1</a>	●	5.13207
<a href="#">Escherichia coli str. K-12 substr. MG1655</a>	<a href="#">K-12 substr. MG1655</a>	19668	<a href="#">SAMN02604091</a>	<a href="#">PRJNA225</a>	<a href="#">GCA_000005845.2</a>	●	4.64165
<a href="#">Escherichia coli O83:H1 str. NRG 857C</a>	<a href="#">NRG 857C</a>	19668	<a href="#">SAMN02603727</a>	<a href="#">PRJNA41221</a>	<a href="#">GCA_000183345.1</a>	●	4.89488
<a href="#">Escherichia coli C104:H4 str. 2011C-3493</a>	<a href="#">2011C-3493</a>	19668	<a href="#">SAMN01831188</a>	<a href="#">PRJNA81095</a>	<a href="#">GCA_000299455.1</a>	●	5.43741
<a href="#">Escherichia coli UMN026</a>	<a href="#">UMN026</a>	19668	<a href="#">SAMEA3133233</a>	<a href="#">PRJNA33415</a>	<a href="#">GCA_000026325.2</a>	●	5.3582
<a href="#">Escherichia coli C157:H7 str. Sakai</a>	<a href="#">Sakai substr. RIMD 0509952</a>	19668	<a href="#">SAMN01911278</a>	<a href="#">PRJNA226</a>	<a href="#">GCA_000008865.2</a>	●	5.5946

# PICRUSt is ...

- PICRUSt is to Bacteria
- Googling is to People