# “Mapping the Human Gut Microbiome in Health and Disease Using Sequencing, Supercomputing, and Data Analysis”

**Invited Talk Delivered by Mehrdad Yazdani, Calit2**

**Ayasdi Sponsored Lunch & Learn**

**American Society of Human Genetics (ASHG)**

**San Diego Convention Center**

**October 19, 2014**

**Dr. Larry Smarr**

**Director, California Institute for Telecommunications and Information Technology**

**Harry E. Gruber Professor,**

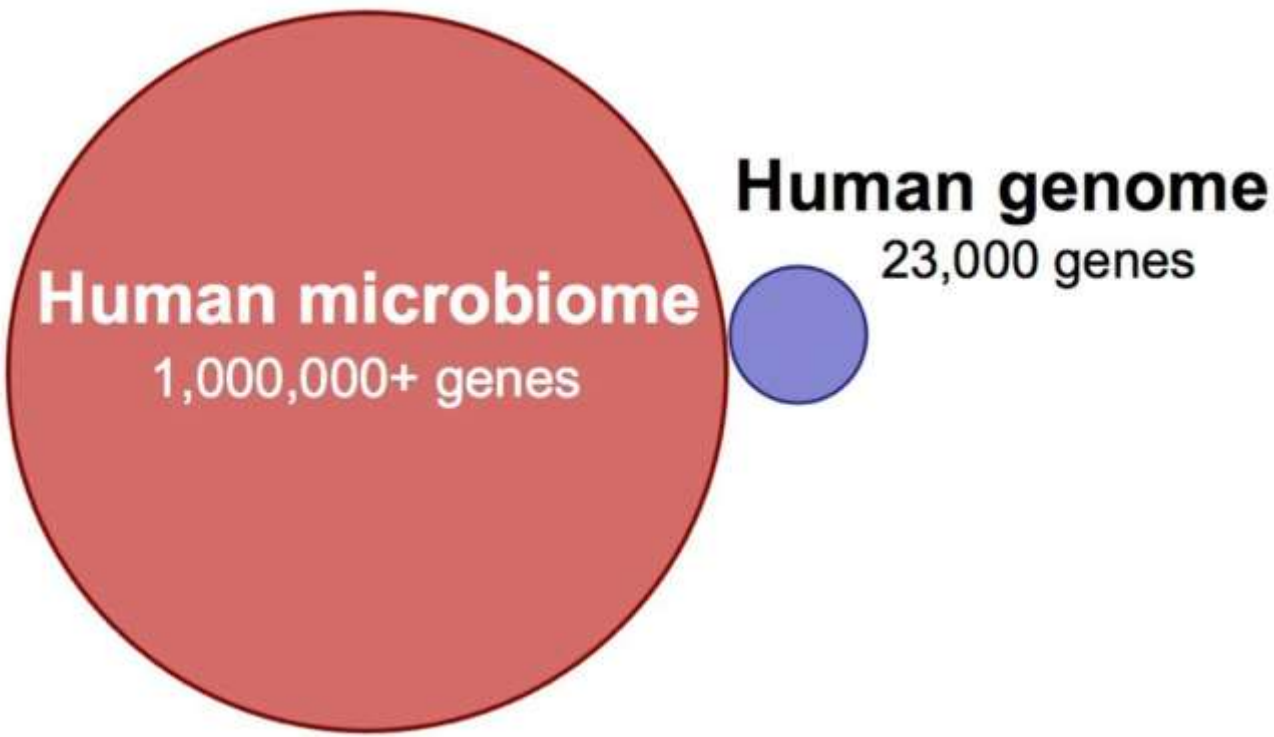
**Dept. of Computer Science and Engineering**

**Jacobs School of Engineering, UCSD**

**http://lsmarr.calit2.net**

# Inclusion of the Microbiome Will Radically Change Medicine and Wellness

## Your Body Has 10 Times As Many Microbe Cells As Human Cells



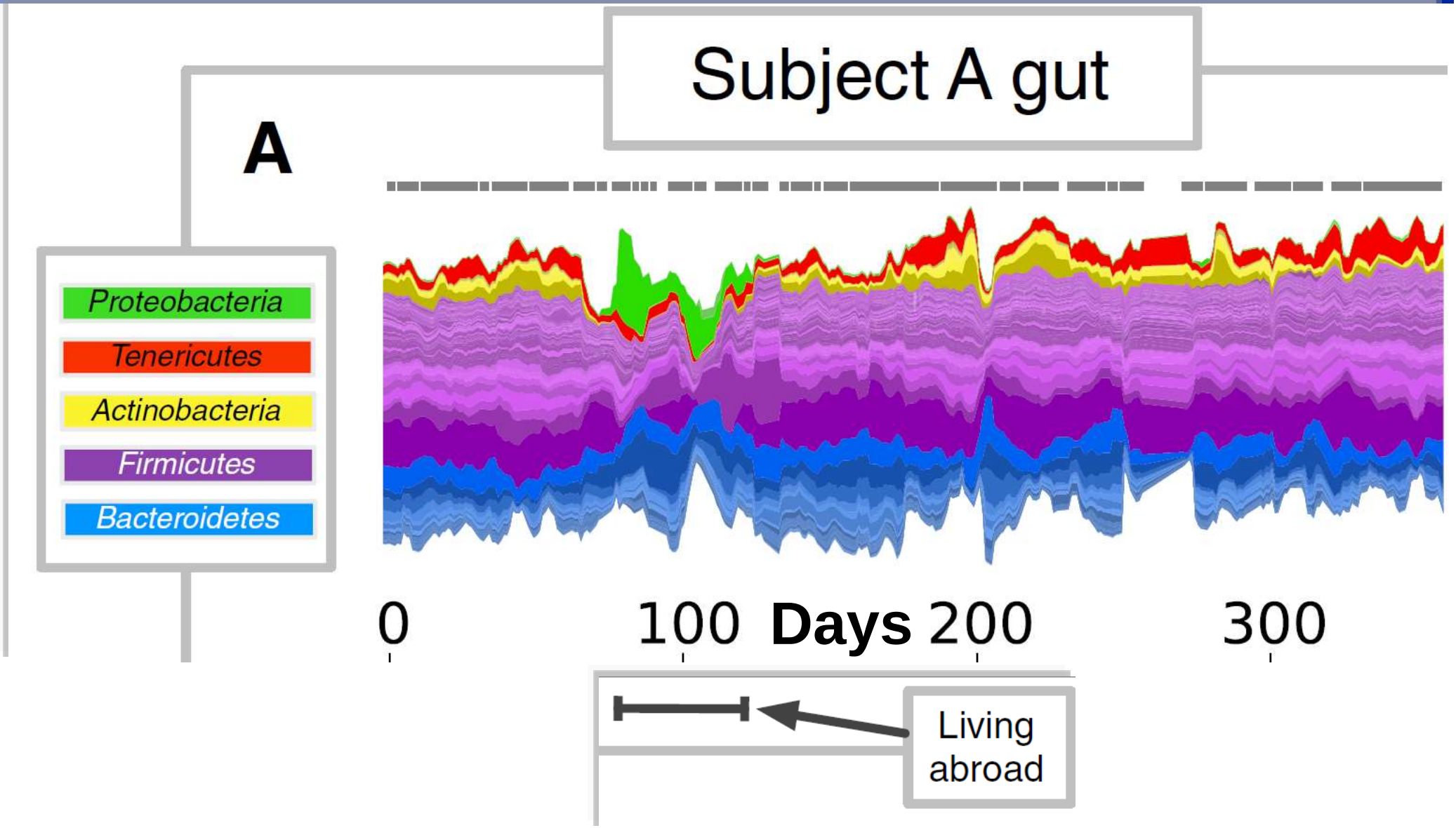

**99% of Your DNA Genes Are in Microbe Cells Not Human Cells**

**Challenge:**
**Map Out Microbial Ecology and Function**

J. Craig Venter™ INSTITUTE

it²

# A Year of Sequencing a Healthy Gut Microbiome Daily - Remarkable Stability with Abrupt Changes




**Genome Biology (2014)**
**David, et al.**

# To Map Out the Dynamics of Autoimmune Microbiome Ecology
## Couples Next Generation Genome Sequencers to Big Data Supercomputers

- **Metagenomic Sequencing**
  - JCVI Produced ~150 Billion DNA Bases From Seven of LS Stool Samples Over 1.5 Years
  - We Downloaded ~3 Trillion DNA Bases From NIH Human Microbiome Program Data Base
    - 255 Healthy People, 21 with IBD
- **Supercomputing (Weizhong Li, JCVI/HLI/UCSD):**
  - ~180,000 Core-Hours SDSC's Gordon
  - ~35,000 Core-Hours Dell HPC Cloud
- **Produced Relative Abundance of ~10,000 Bacteria, Archaea, Viruses in ~300 People**
  - ~3Million Filled Spreadsheet Cells

Illumina HiSeq 2000 at JCVI



SDSC Gordon Data Supercomputer

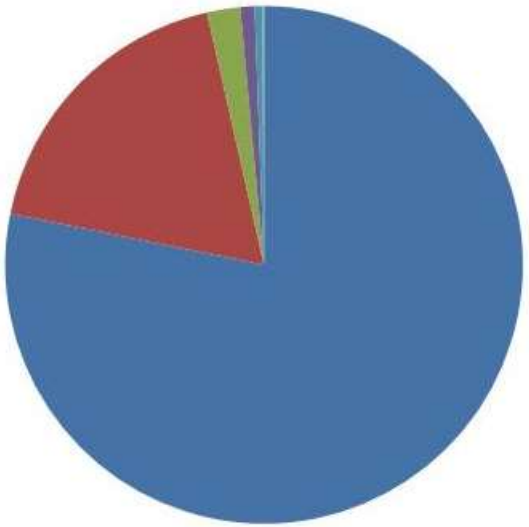



Dell Solutions Center
Industry Solutions Lab
SANGER DSU

J. Craig Venter INSTITUTE SDSC DELL NIH HMP CRBS it²

# We Found Major State Shifts in Microbial Ecology Phyla Between Healthy and Two Forms of IBD

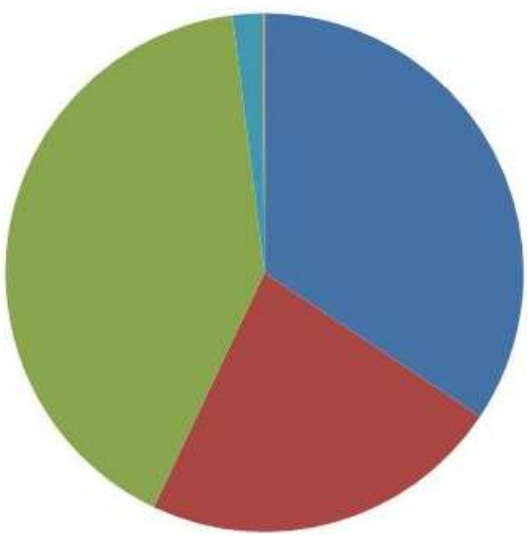
**Average HE**



- Bacteroidetes
- Firmicutes
- Proteobacteria
- Verrucomicrobia
- Actinobacteria
- Fusobacteria
- Euryarchaeota

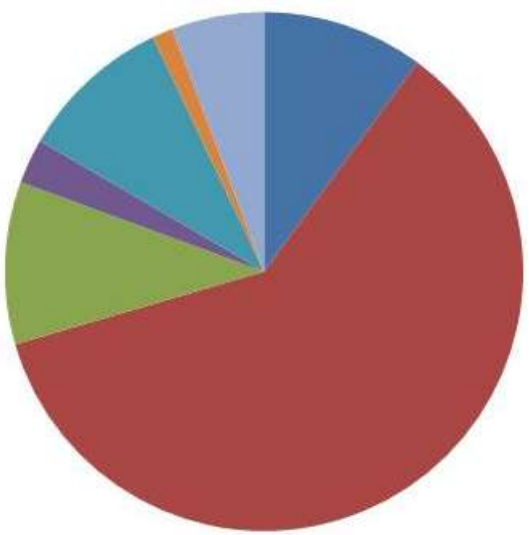
**Most Common Microbial Phyla**

**Average Ulcerative Colitis**



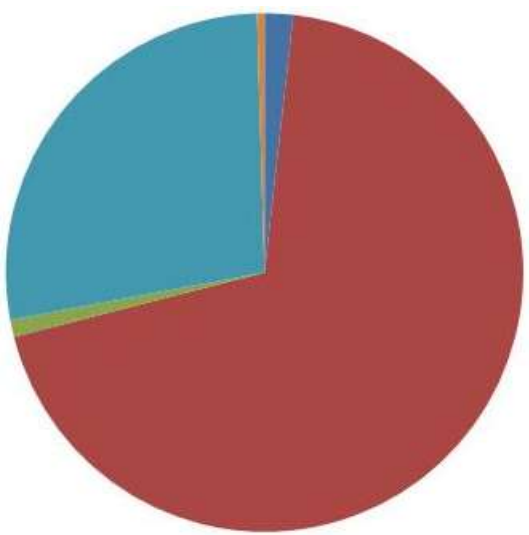
**Explosion of Proteobacteria**

**Average LS**



**Hybrid of UC and CD**
**High Level of Archaea**

**Average Crohn's Disease**



**Collapse of Bacteroidetes**
**Explosion of Actinobacteria**

# Using Ayasdi to Discover Hidden Patterns in Our Data

## Ayasdi Advanced Analytics

topological data analysis



# Categorical Data Lens to Separate Healthy from Disease States

All Ileal Crohn's



All Healthy



All Healthy





All Healthy



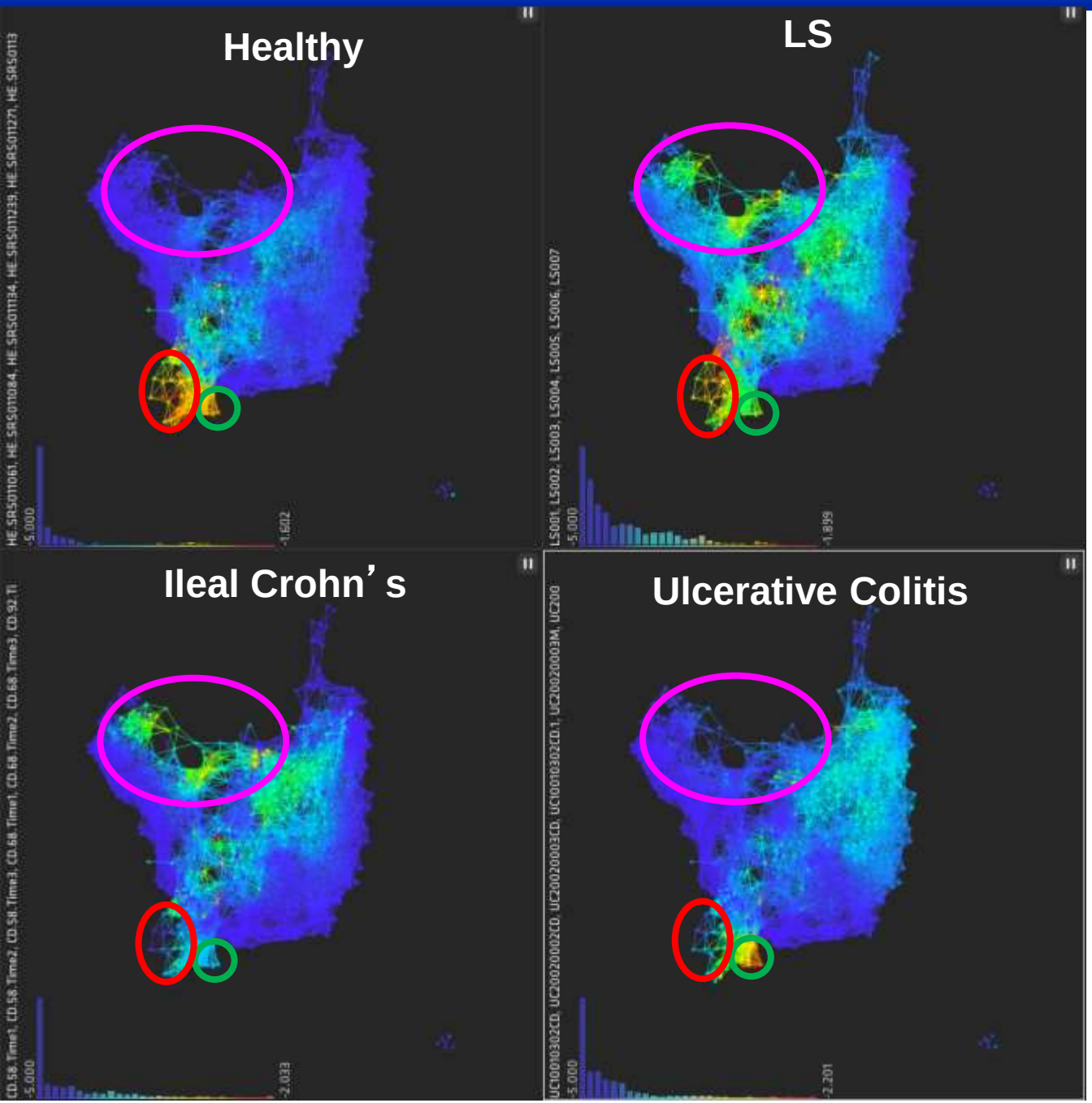
Healthy, Ulcerative Colitis, and LS

Ayasdi Cure™
Turn Data into Therapies

# Group Comparisons using Ayasdi's Statistical Tools

| Column Name | KS Statistic ▼1 | KS p-value | T-test p-value |
|---|---|---|---|
| Actinomyces.sp..oral.taxon.448 | 1.000 | <.001 | <.001 |
| Bacteroides.fragilis | 1.000 | <.001 | <.001 |
| Bacteroides.dorei | 1.000 | <.001 | <.001 |
| Tannerella.sp..6_1_58FAA_CT1 | 1.000 | <.001 | <.001 |
| Bacteroides.coprophilus | 1.000 | <.001 | <.001 |
| Bacteroides.faecis | 1.000 | <.001 | <.001 |
| Parabacteroides.johnsonii | 1.000 | <.001 | <.001 |
| Paraprevotella.xylaniphila | 1.000 | <.001 | <.001 |
| Prevotella.disiens | 1.000 | <.001 | <.001 |
| Alloprevotella.tannerae | 1.000 | <.001 | <.001 |
| Prevotella.buccalis | 1.000 | <.001 | <.001 |
| Bacteroides.salyersiae | 1.000 | <.001 | <.001 |
| Prevotella.oris | 1.000 | <.001 | <.001 |
| Bacteroides.oleiciplenus | 1.000 | <.001 | <.001 |
| Prevotella.timonensis | 1.000 | <.001 | <.001 |
| Actinomyces.odontolyticus | 0.9938 | <.001 | <.001 |
| Actinomyces.sp..oral.taxon.180 | 0.9938 | <.001 | <.001 |
| Alistipes.finegoldii | 0.9938 | <.001 | <.001 |
| Bacteroides.sp..2_1_7 | 0.9938 | <.001 | <.001 |
| Bacteroides.salanitronis | 0.9938 | <.001 | <.001 |
| Prevotella.bivia | 0.9938 | <.001 | <.001 |

# Ayasdi Enables Discovery of Differences Between Healthy and Disease States Using Microbiome Species

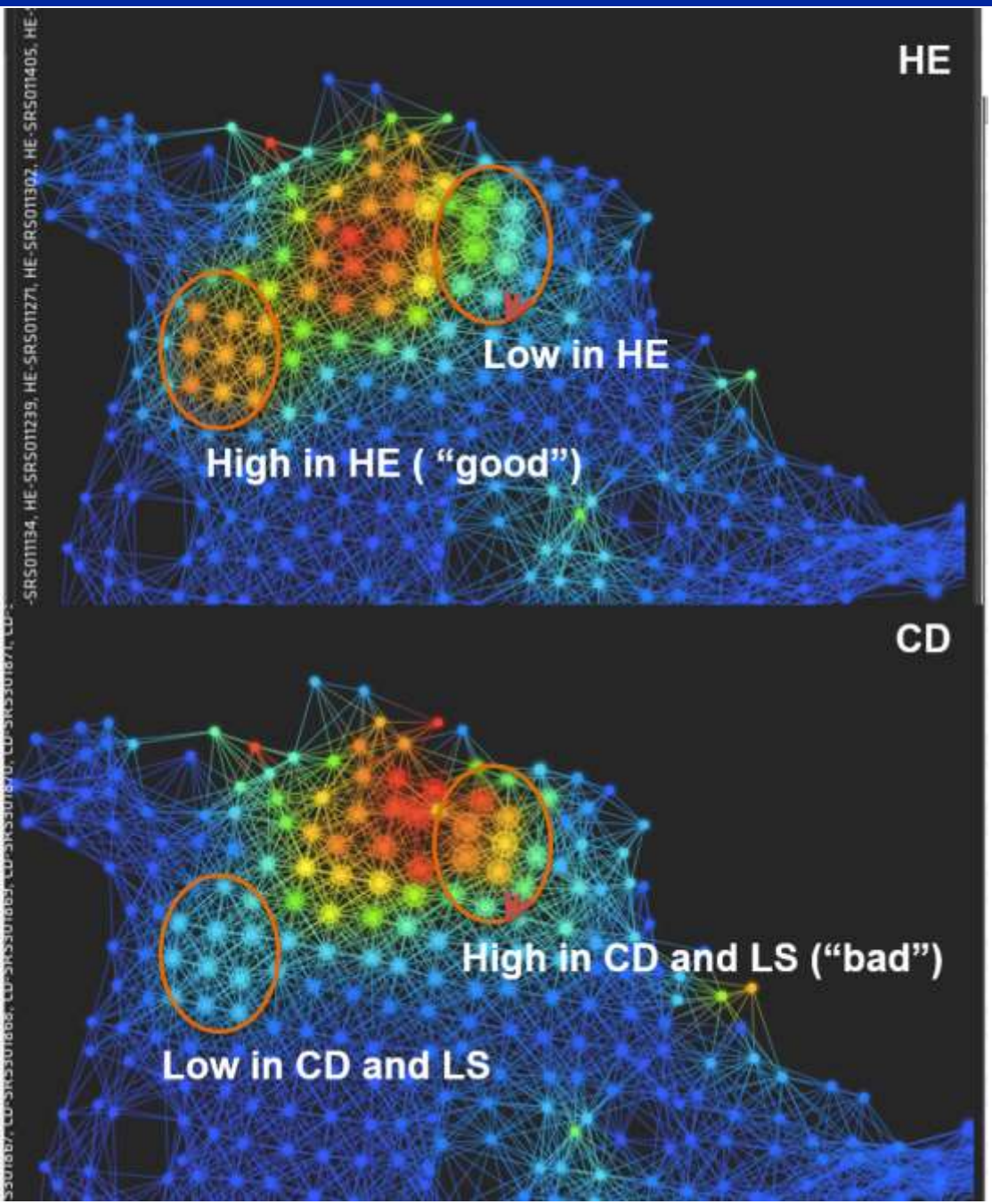



**Ayasdi Cure™**
Turn Data into Therapies

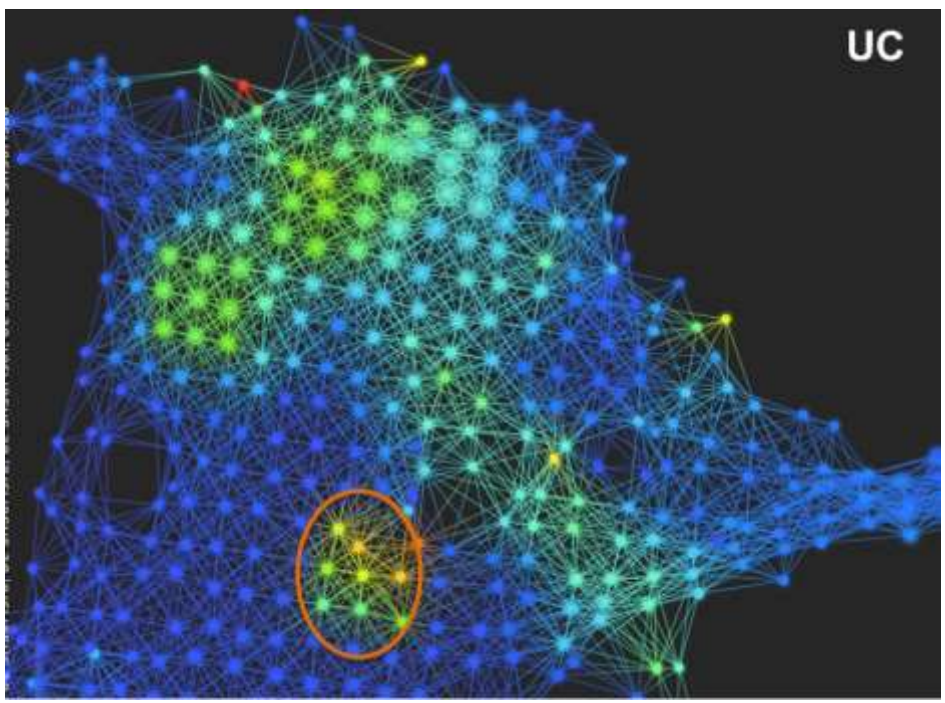
- **High in Healthy and LS**
- **High in Healthy and Ulcerative Colitis**
- **High in Both LS and Ileal Crohn's Disease**

**Using Multidimensional Scaling Lens with Correlation Metric**

# Moving from Ecological Taxonomy to Cellular Pathways




**Dataset from Larry Smarr Team**
**With 60 Subjects (HE, CD, UC, LS)**
**Each with 10,000 KEGGs -**
**600,000 Cells**




**Source: Pek Lum, Chief Data Scientist, Ayasdi**

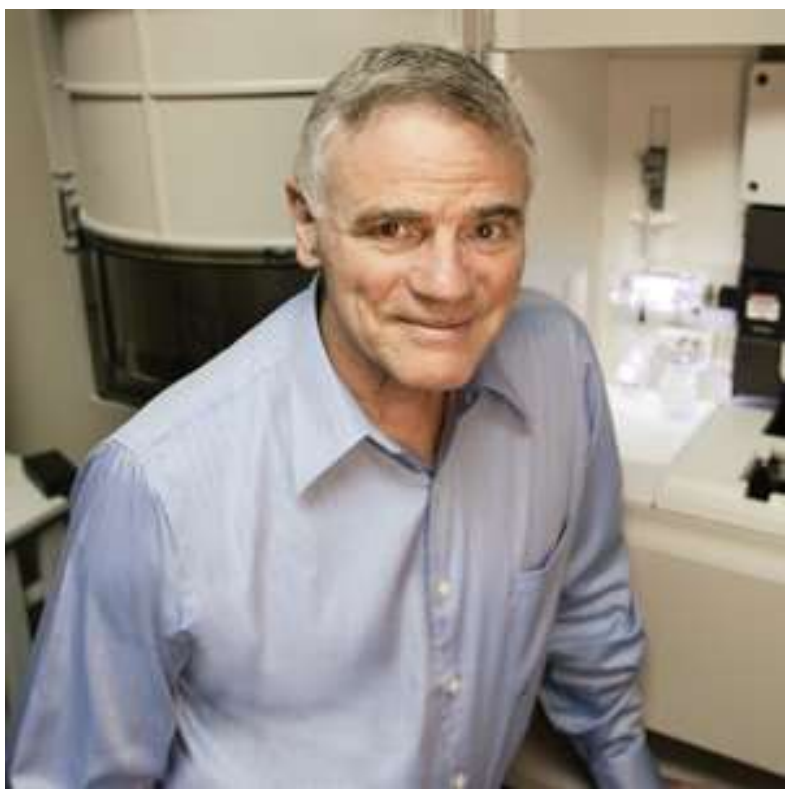
Ayasdi Cure™
Turn Data into Therapies

# Next Step: Apply What We Have Learned to New Microbiome Datasets

- **Larry Smarr is a Member of the Pioneer 100**
- **Our Team Now Has the Gut Microbiomes of the Pioneer 100**
- **We Plan to Analyze Them for Differences Using Ayasdi Tools**
- **Do Metagenomics on Those Who Are Outliers**

PIONEER 100
HUNDRED PERSON WELLNESS PROJECT
Institute for Systems Biology

**Will Grow to 1000, then 10,000**



**http://isbmolecularme.com/tag/100-pioneers/**

# Thanks to Our Great Team!

## UCSD Metagenomics Team

**Weizhong Li**
**Sitao Wu**

## Calit2@UCSD Future Patient Team

**Jerry Sheehan**
**Tom DeFanti**
**Kevin Patrick**
**Jurgen Schulze**
**Andrew Prudhomme**
**Philip Weber**
**Fred Raab**
**Joe Keefe**
**Ernesto Ramirez**




**Ayasdi**
Devi
Sanjnan
Pek

## JCVI Team

**Karen Nelson**
**Shibu Yooseph**
**Manolito Torralba**

## SDSC Team

**Michael Norman**
**Mahidhar Tatineni**
**Robert Sinkovits**

## UCSD Health Sciences Team

**William J. Sandborn**
**Elisabeth Evans**
**John Chang**
**Brigid Boland**
**David Brenner**