“Using Genetic Sequencing to Unravel the Dynamics of Your Superorganism Body”

Weekly Bioinformatics Seminar Series

UC San Diego

La Jolla, CA

October 17, 2013

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

Abstract

The human body is host to 100 trillion microorganisms, ten times the number of cells in the human body, and these microbes contain 100 times the number of DNA genes that our human DNA does. The microbial component of this "superorganism" is comprised of hundreds of species spread over many taxonomic phyla. The human immune system is tightly coupled with this microbial ecology and in cases of autoimmune disease, both the host immune system and the microbial ecology can have excursions far from normal. I will review some of the known 163 SNPs in the human genome which pre-dispose the host to develop autoimmune inflammatory bowel disease (IBD). Motivated by a diagnosis that I have Crohn's disease, a form of IBD, I have been collecting massive amounts of data on my own body over the last five years. Analysis and graphing of this data demonstrates the episodic evolution of this coupled immune-microbial system. To decode the details of the microbial ecology requires high resolution genome sequencing feeding Big Data parallel supercomputers coupled to scalable visualization systems. The complexities of my time-varying microbial ecology will be compared to the NIH Human Microbiome Program data on people in states of health and disease.

I Arrived in La Jolla in 2000 After 20 Years in the Midwest and Decided to Move Against the Obesity Trend

Age 41



Age 51



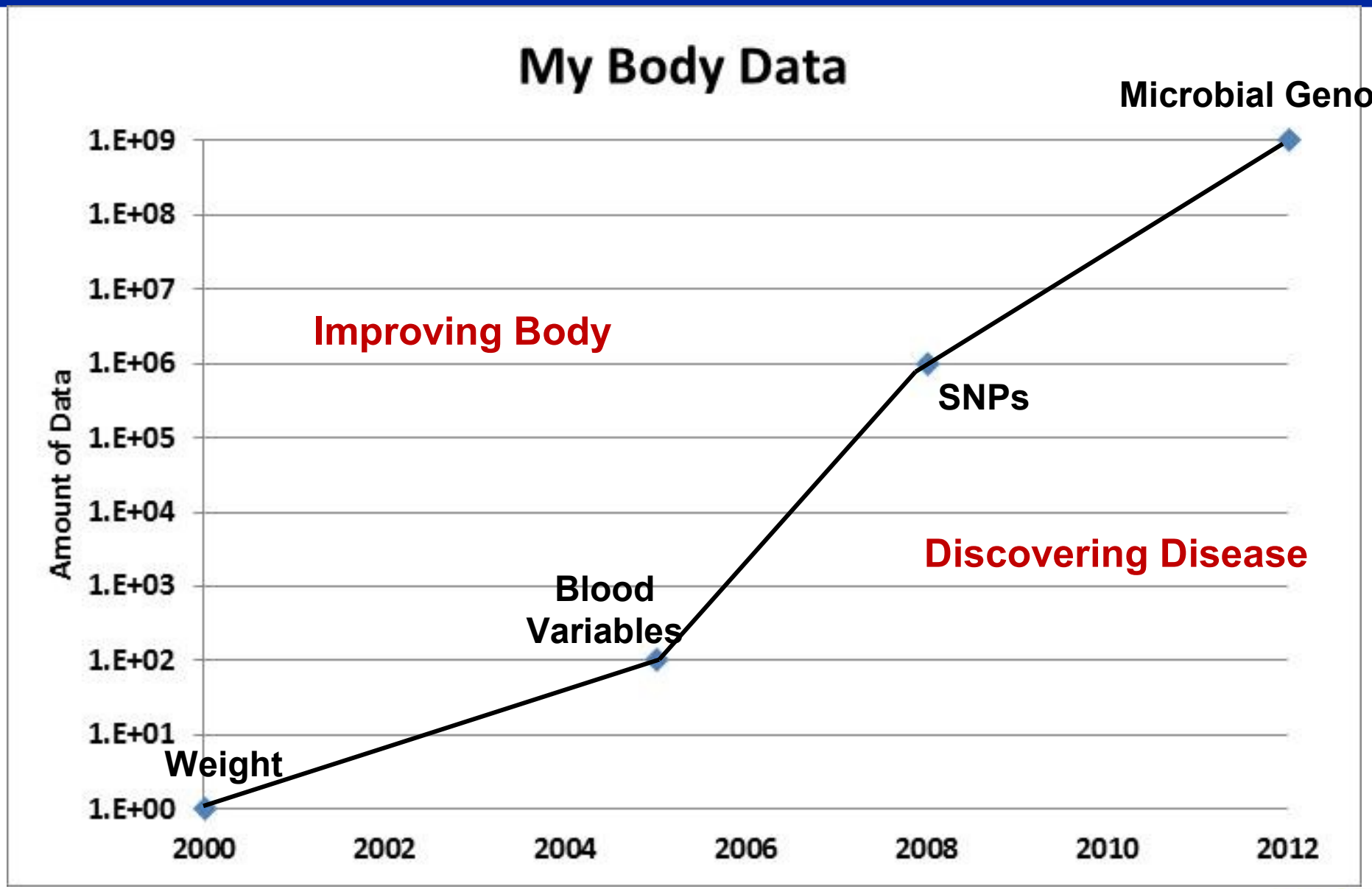
Age 61



I Reversed My Body's Decline By Quantifying and Altering Nutrition and Exercise

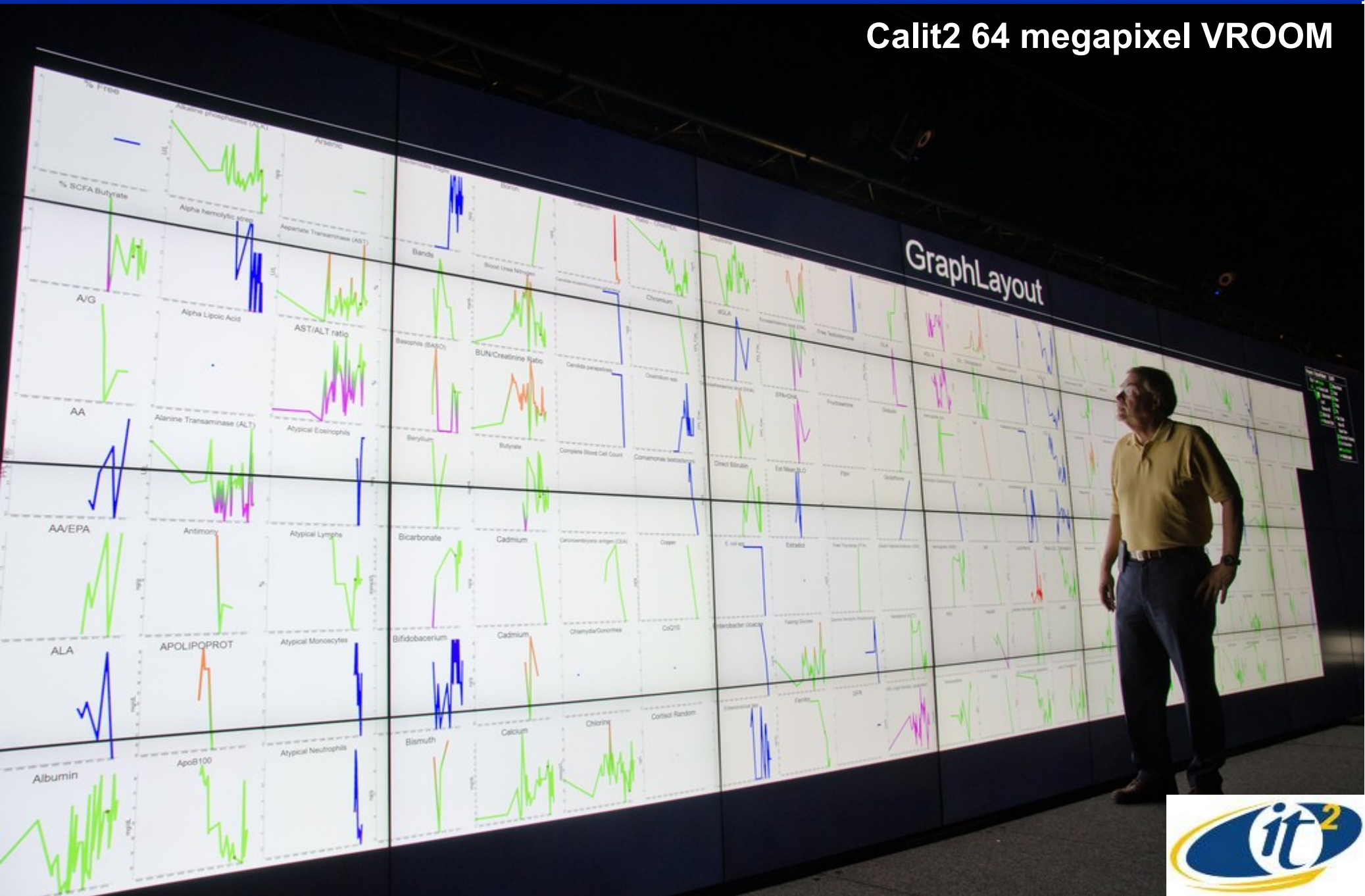
http://lsmarr.calit2.net/repository/LS_reading_recommendations_FiRe_2011.pdf

From One to a Billion Data Points Defining Me: The Exponential Rise in Body Data in Just One Decade!

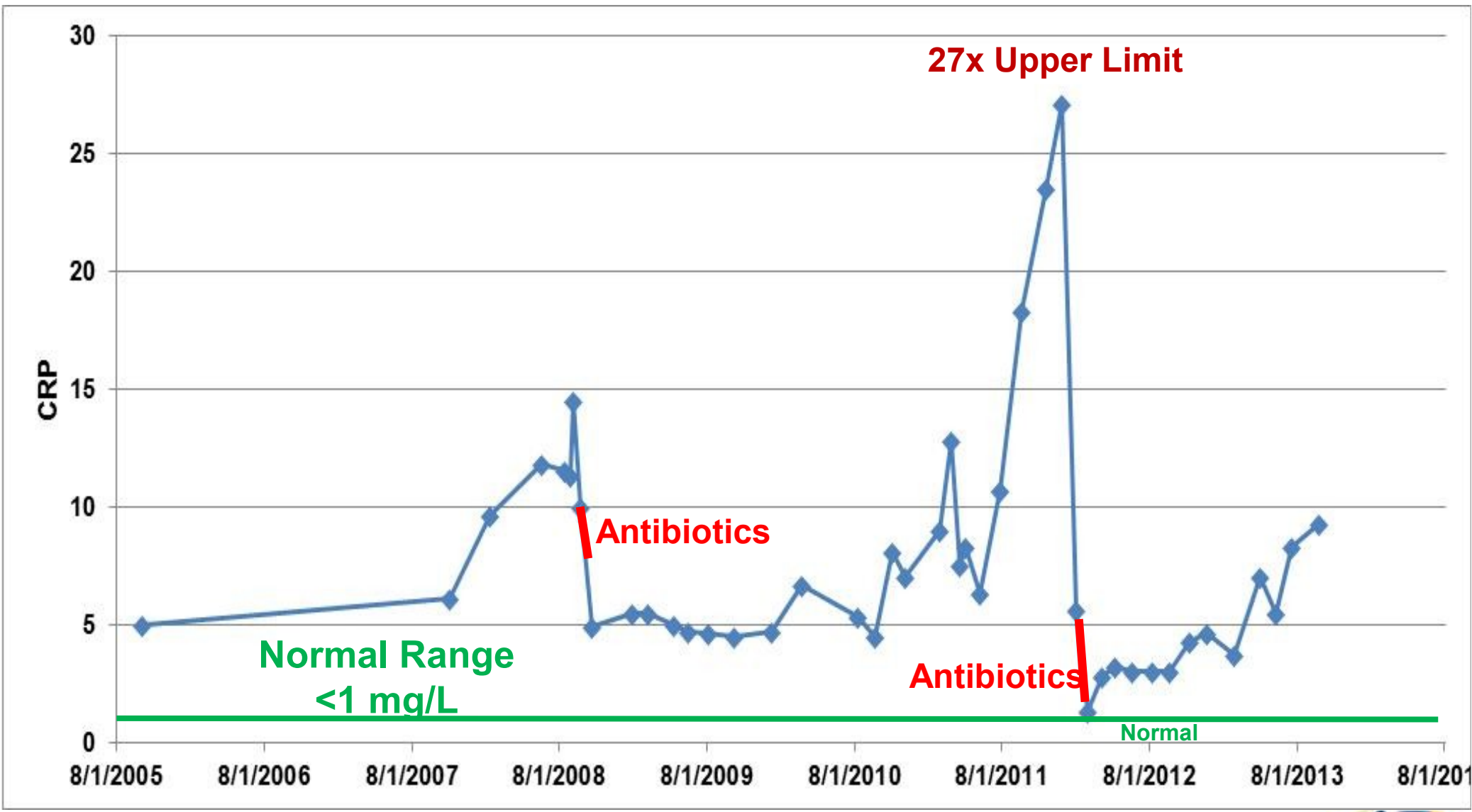


Each is a Personal Time Series
And Compared Across Population

Visualizing Time Series of 150 LS Blood and Stool Variables, Each Over 5-10 Years

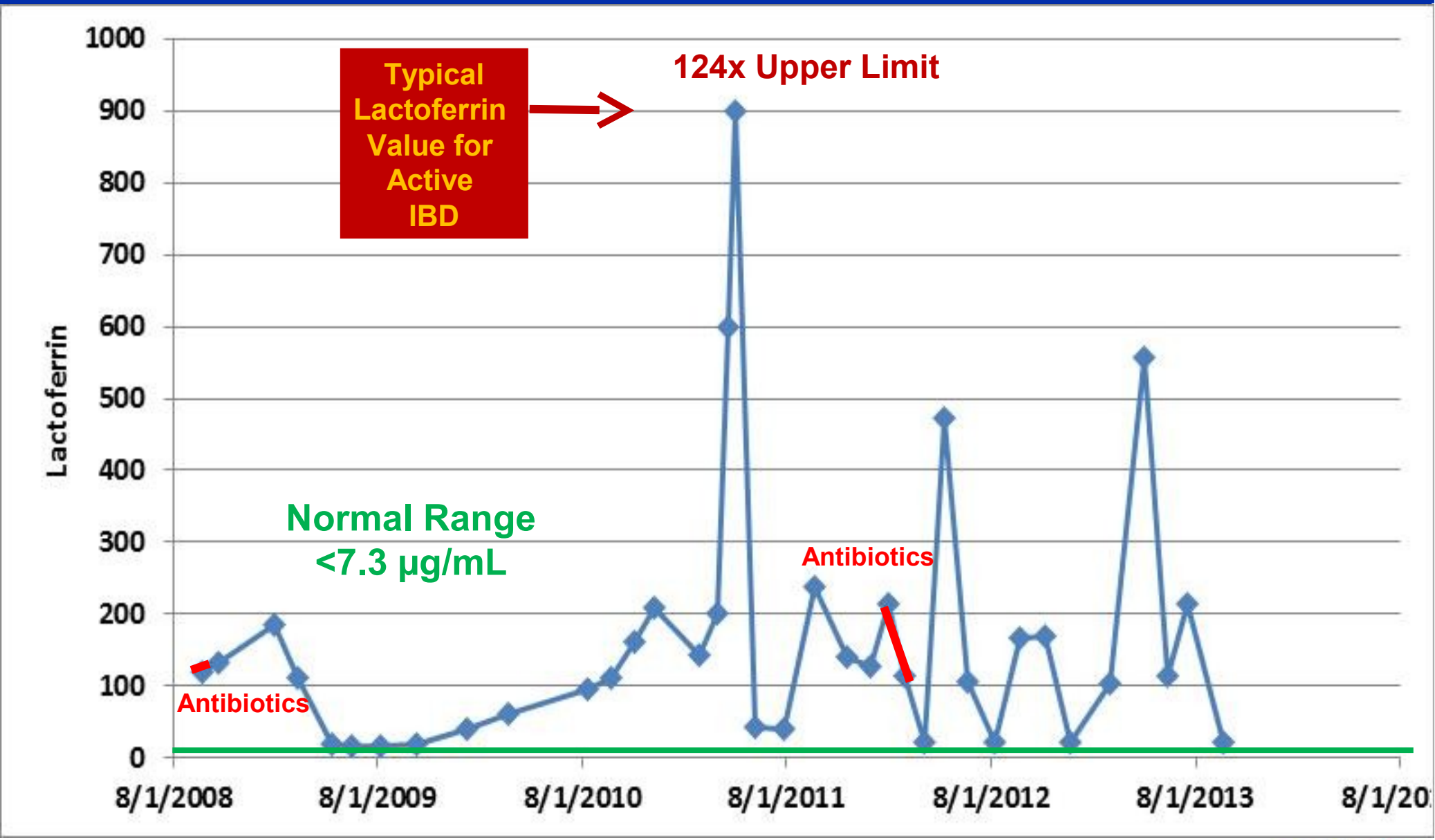


I Discovered I Had Episodic Chronic Inflammation by Tracking Complex Reactive Protein In My Blood Samples



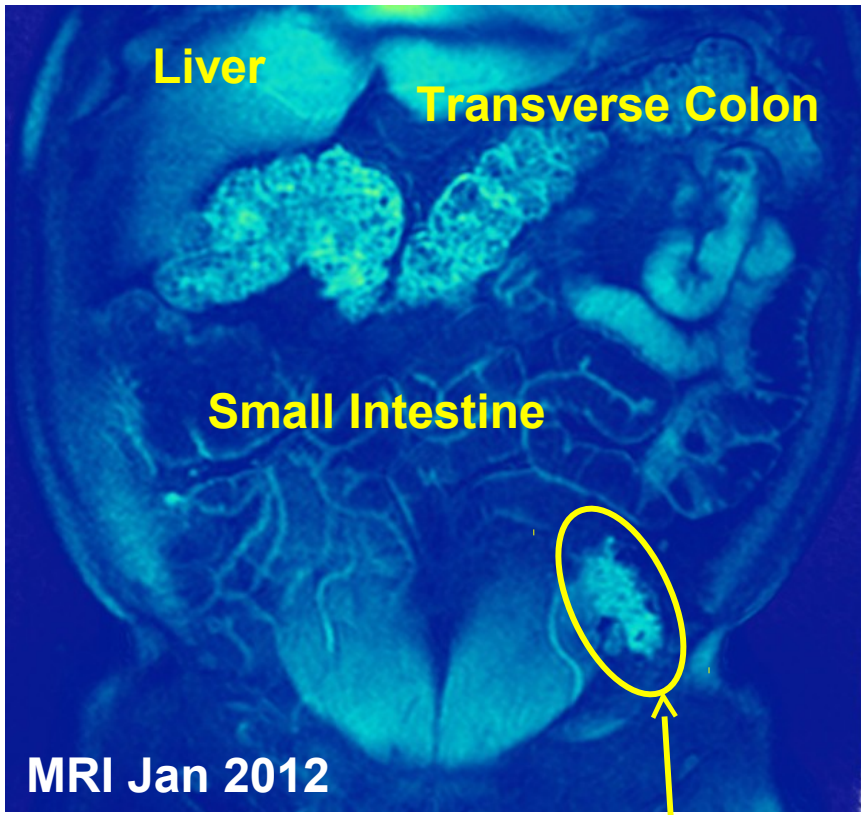
CRP is a Generic Measure of Inflammation in the Blood

By Adding Stool Samples, I Discovered I Had High Levels of the Protein Lactoferrin Shed from Neutrophils



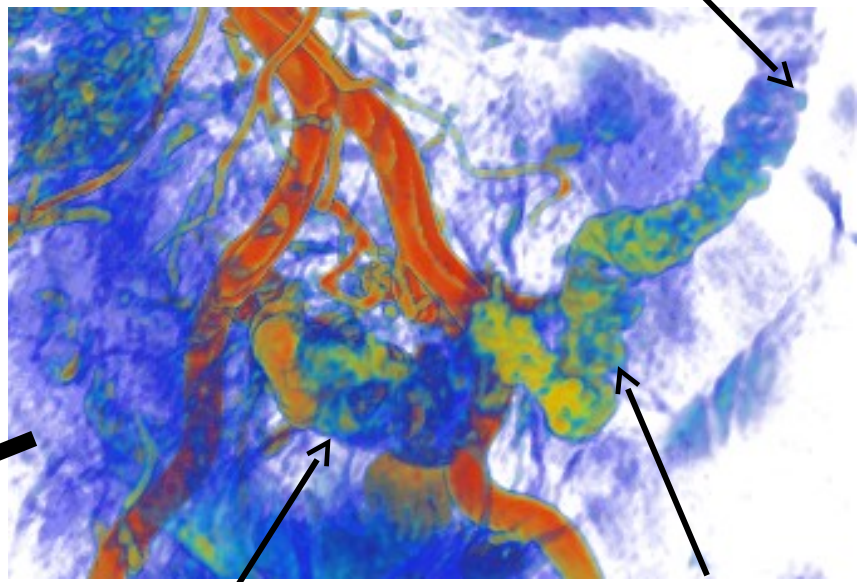
Lactoferrin is a Protein Shed from Neutrophils - An Antibacterial that Sequesters Iron

Confirming the IBD (Crohn's) Hypothesis: Finding the "Smoking Gun" with MRI Imaging



I Obtained the MRI Slices From UCSD Medical Services and Converted to Interactive 3D Working With Calit2 Staff & DeskVOX Software

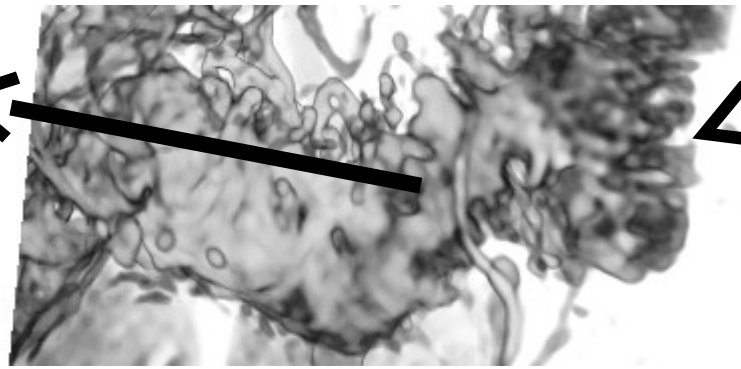
Descending Colon



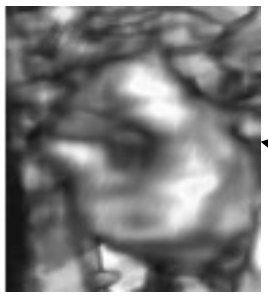
Major Kink

Sigmoid Colon Threading Iliac Arteries

Diseased Sigmoid Colon



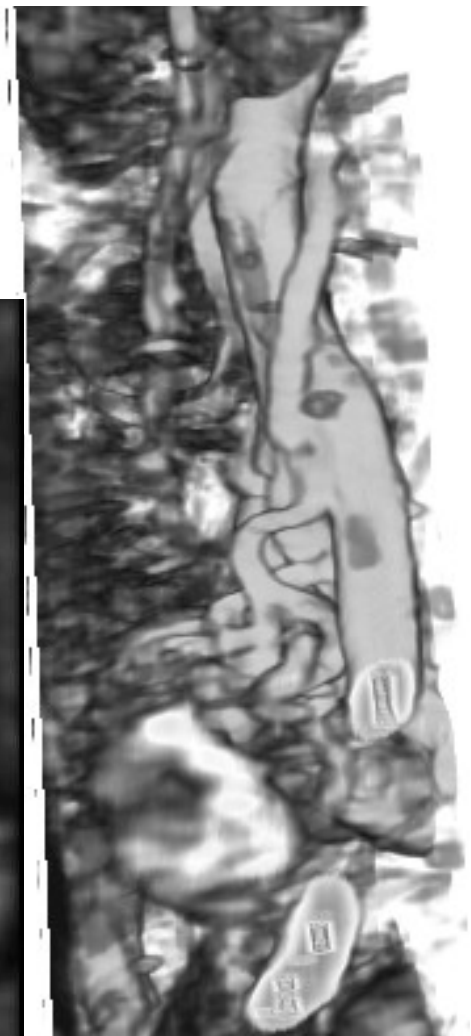
Cross Section



MRE Reveals Inflammation in 6 Inches of Sigmoid Colon
Thickness 15cm – 5x Normal Thickness

**"Long segment wall thickening
in the proximal and mid portions of the sigmoid colon,
extending over a segment of approximately 16 cm,
with suggestion of intramural sinus tracts.
Edema in the sigmoid mesentery
and engorgement of the regional vasa recta."
– MRI report**

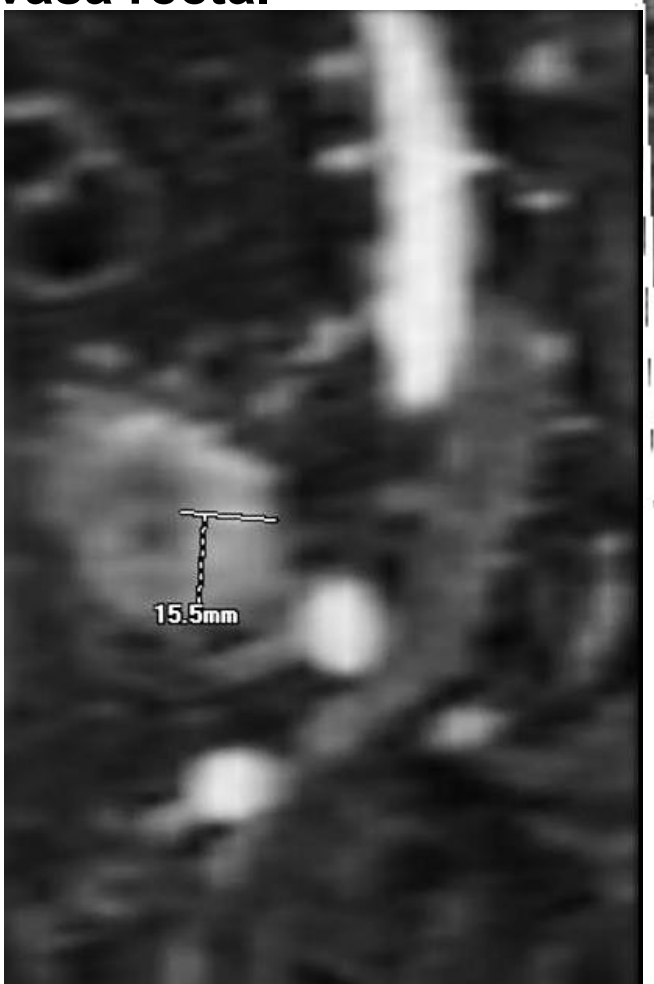
DeskVOX 3D Image



MAYO CLINIC

**Crohn's disease
affects the thickness
of the intestinal wall.**

**Having Crohn's disease
that affects your colon
increases your risk
of colon cancer.**

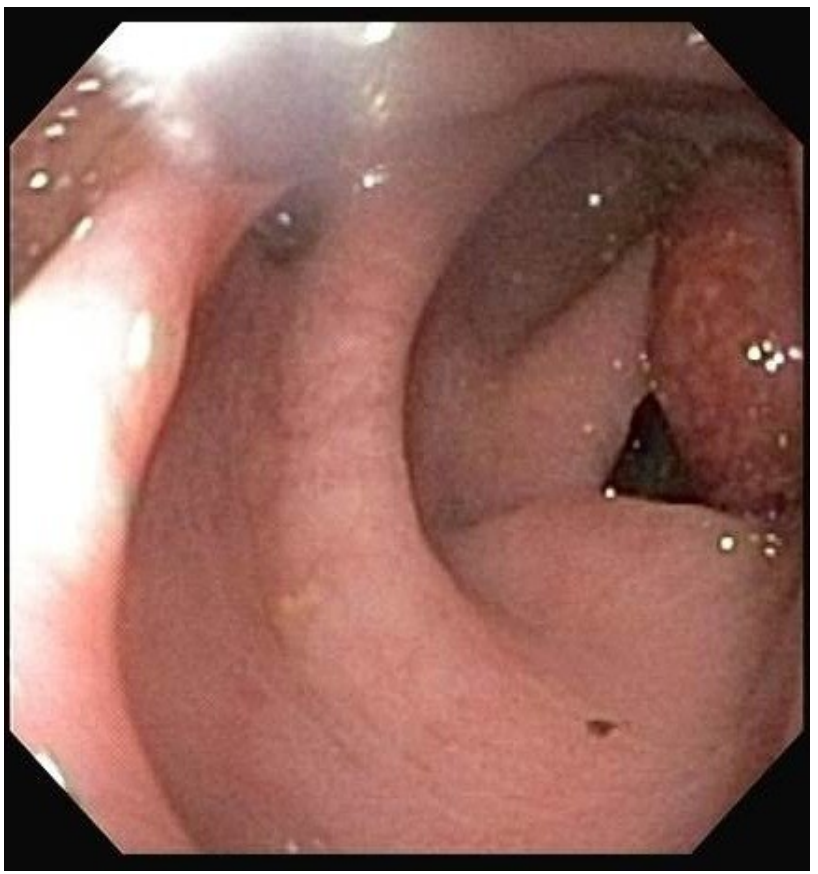


**Clinical MRI
Slice Program**

Colonoscopy Images Show Inflamed Pseudopolyps in 6 inches of Sigmoid Colon



Dec 2010



May 2011

Why Did I Have an Autoimmune Disease like IBD?

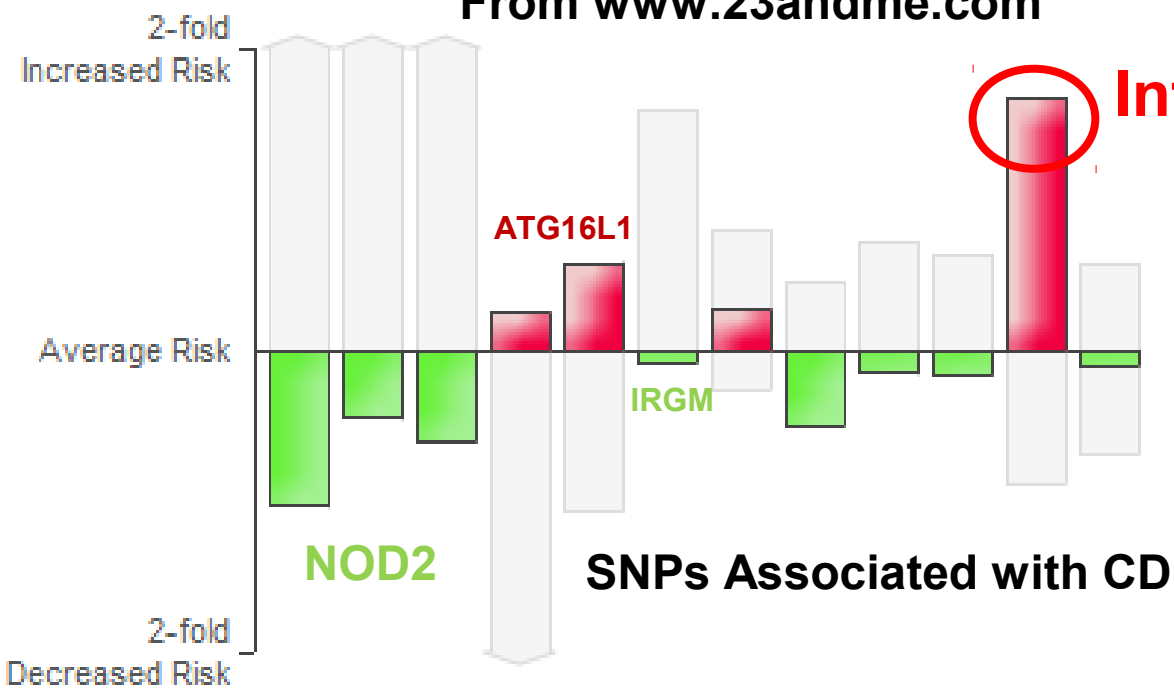
Despite decades of research,
the etiology of Crohn's disease
remains unknown.
Its pathogenesis may involve
a complex interplay between
host genetics,
immune dysfunction,
and microbial or environmental factors.
--*The Role of Microbes in Crohn's Disease*

So I Set Out to Quantify All Three!

Paul B. Eckburg & David A. Relman
Clin Infect Dis. 44:256-262 (2007)

I Wondered if Crohn's is an Autoimmune Disease, Did I Have a Personal Genomic Polymorphism?

From www.23andme.com



Polymorphism in Interleukin-23 Receptor Gene — 80% Higher Risk of Pro-inflammatory Immune Response

Now Comparing 163 Known IBD SNPs with 23andme SNP Chip

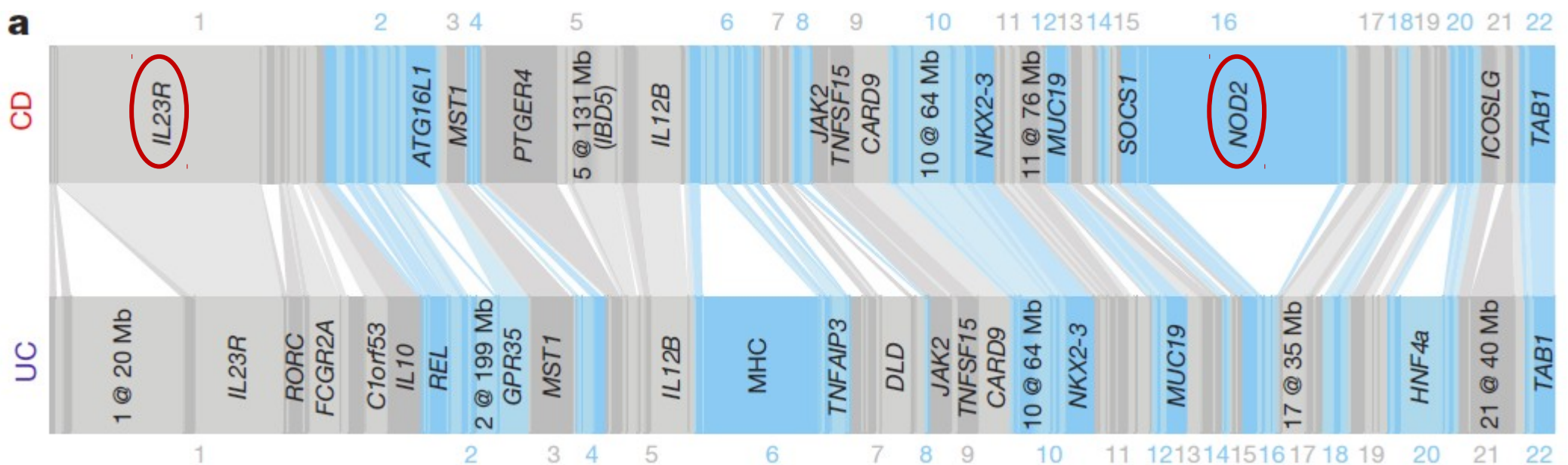
News and Views

Nature Medicine **13**, 26 - 28 (2007)
doi:10.1038/nm0107-26

IL-23: a master regulator in Crohn disease

Markus F Neurath[1]

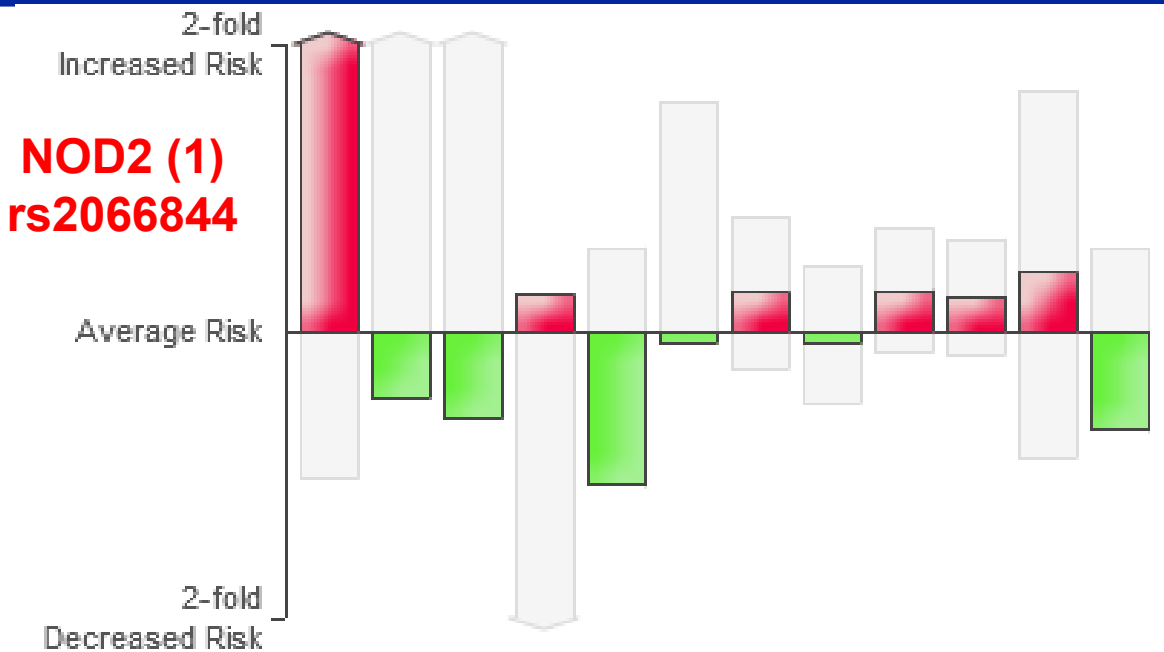
Variance Explained by Each of the 163 SNPs Associated with IBD



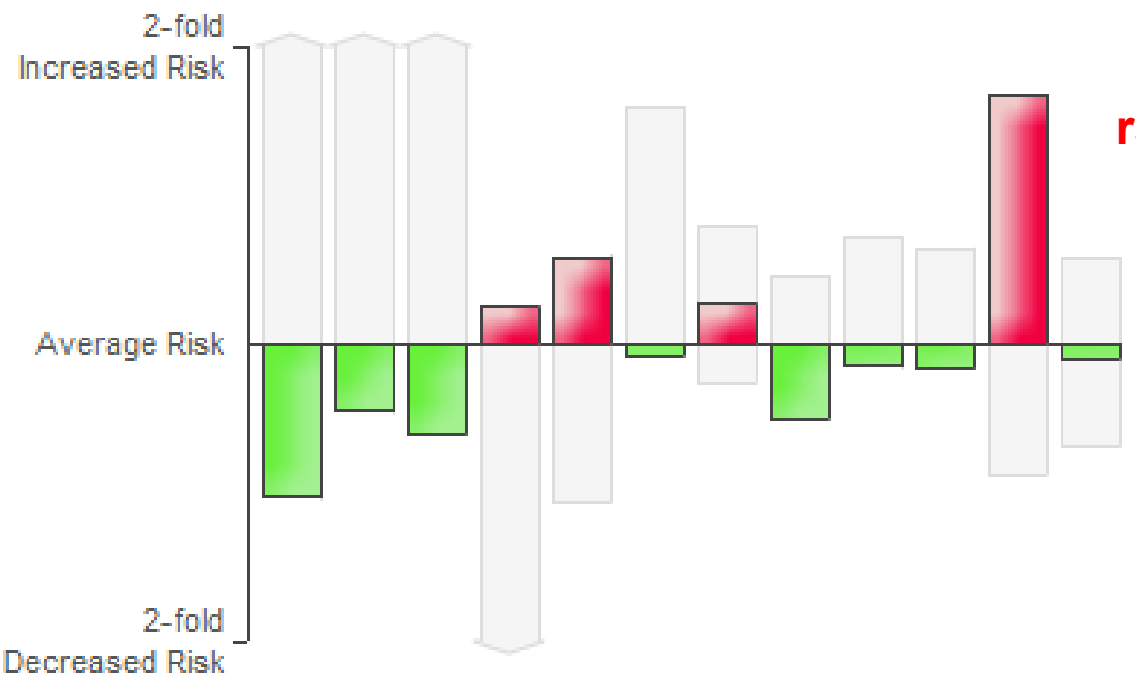
- The width of the bar is proportional to the variance explained by that locus
- Bars are connected together if they are identified as being associated with both phenotypes
- Loci are labelled if they explain more than 1% of the total variance explained by all loci

"Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease," Jostins, et al. Nature 491, 119-124 (2012)

Crohn's May be a Related Set of Diseases Driven by Different SNPs



Female
CD Onset
At 20-Years Old



Il-23R
rs1004819

Me-Male
CD Onset
At 60-Years Old

I Had My Full Human Genome Sequenced in 2012 - 1 Million/Year by 2015

Next Step: Compare Full Genome With IBD SNPs



PGP Used Complete Genomics, Inc. to Sequence my Human DNA

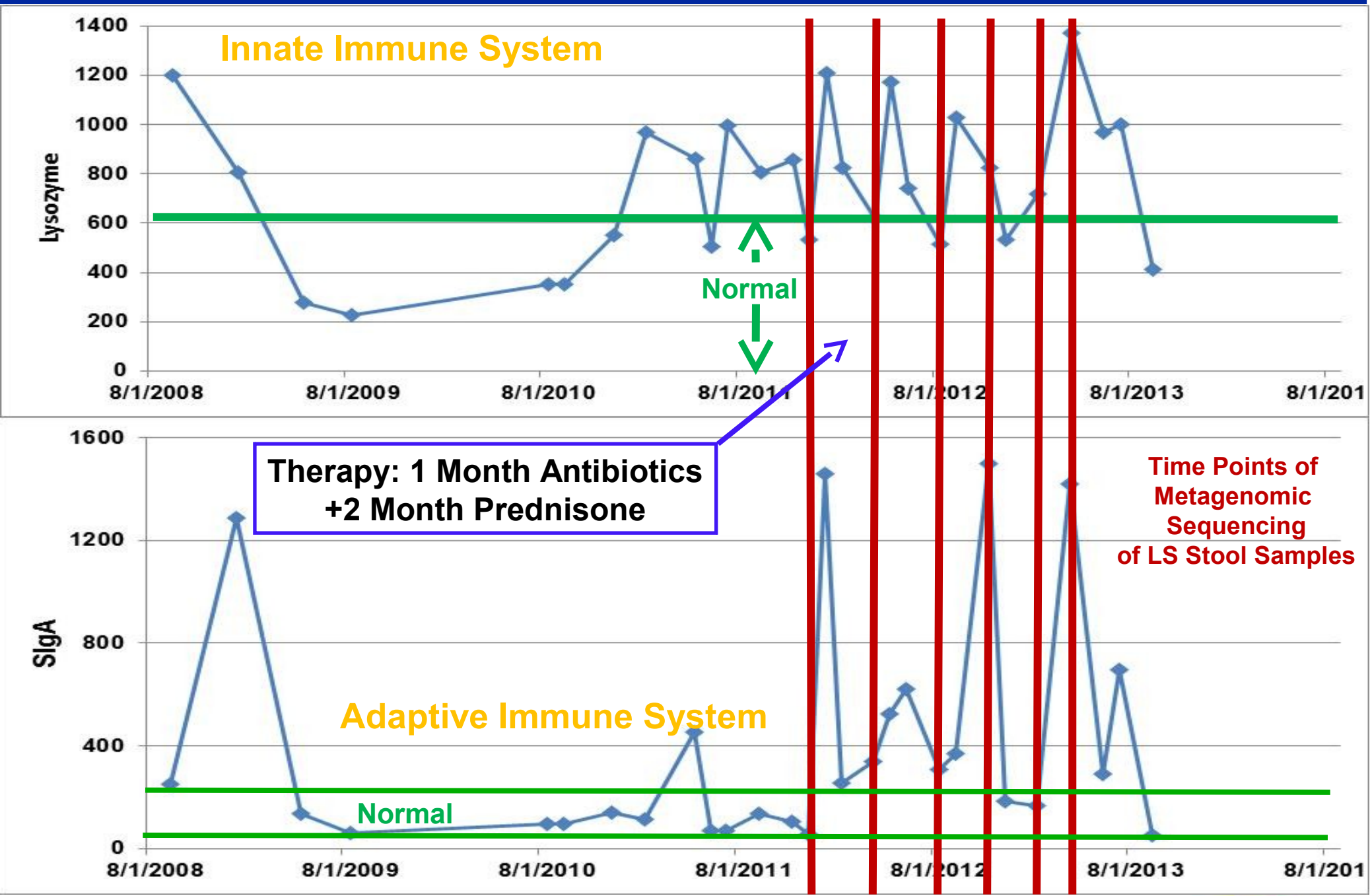
My Anonymized Human Genome is Available for Download

Personal Genome Project

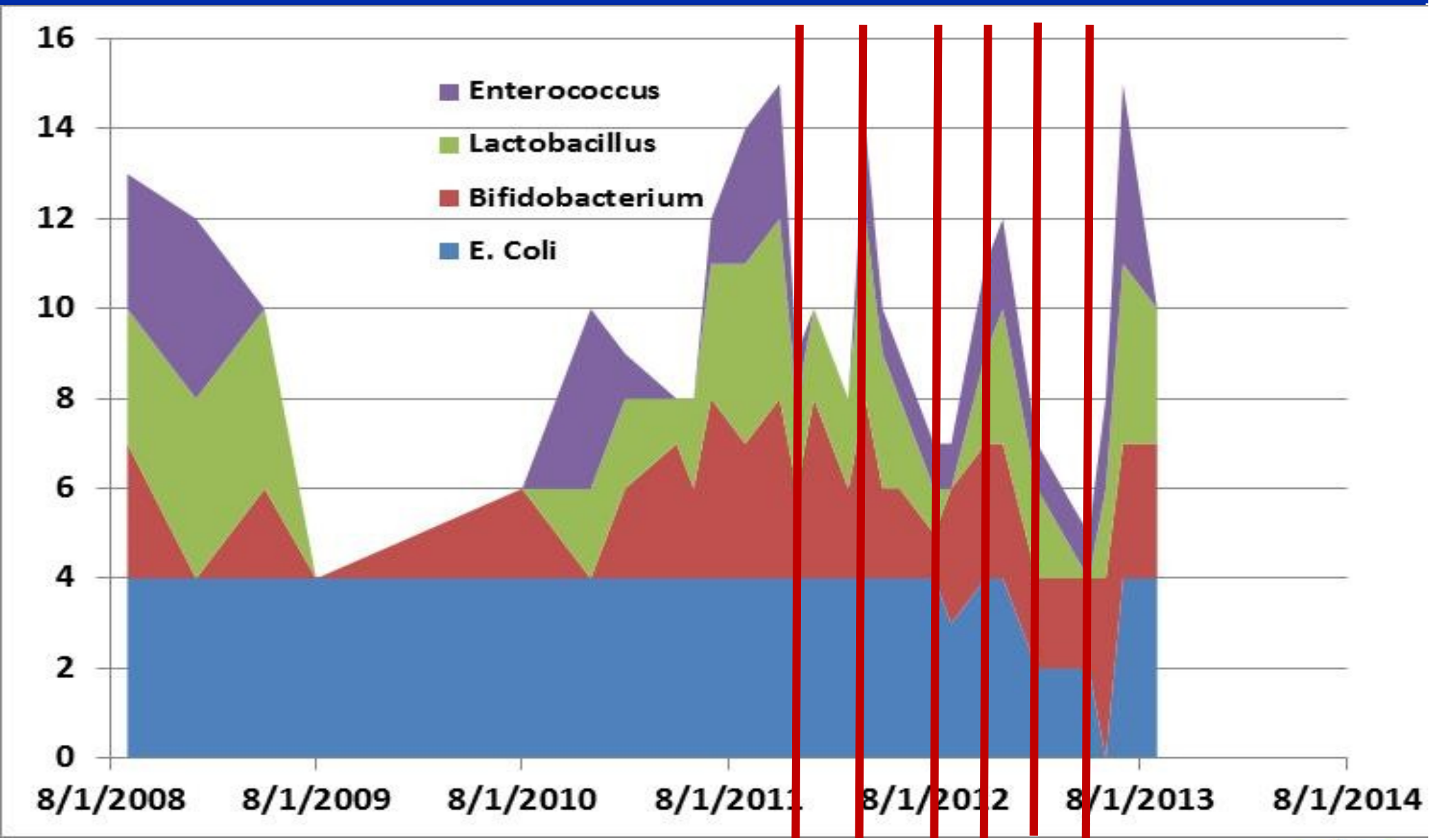


www.personalgenomes.org

Fine Time Resolution Sampling Reveals Unexpected Dynamics of Innate and Adaptive Immune System



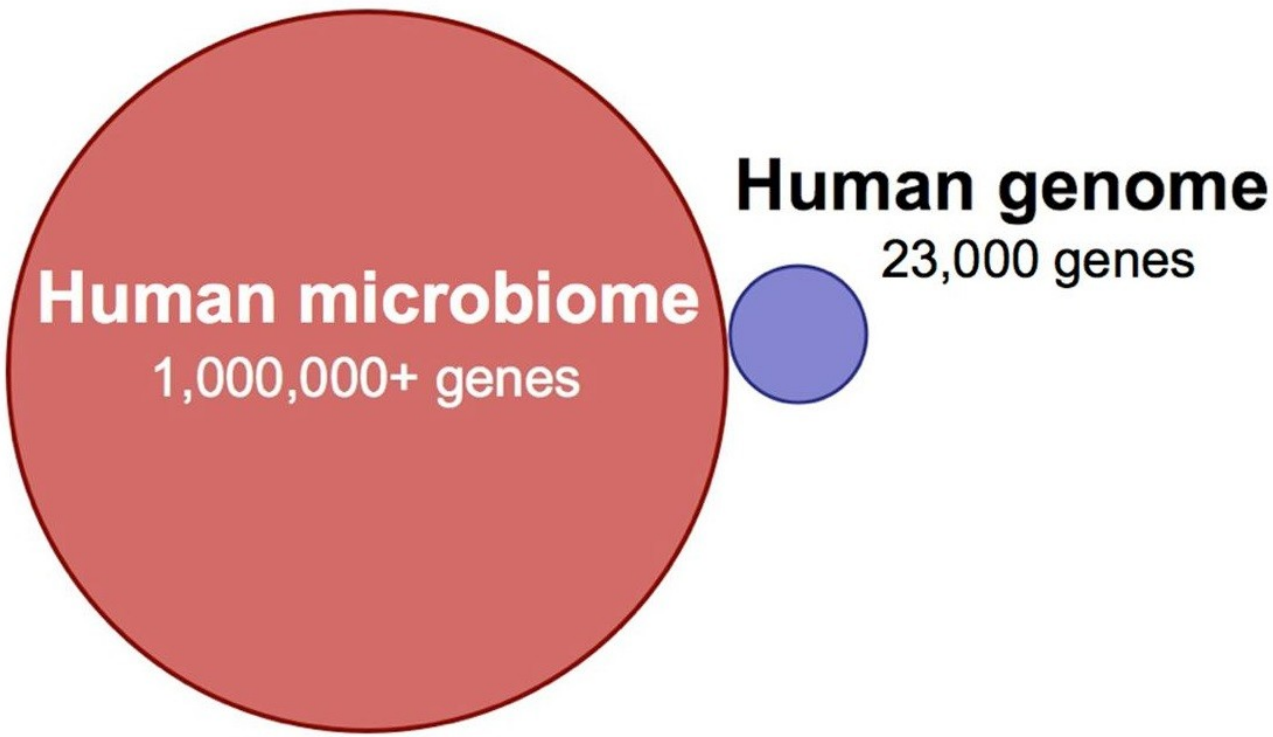
LS Cultured Bacterial Abundance Reveals Dynamic Microbiome Dysfunction



Time Points of Metagenomic Sequencing of LS Stool Samples

Next: Analyze the Dynamics of My Microbiome Ecology- 85% of the Species Can Not Be Cultured

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



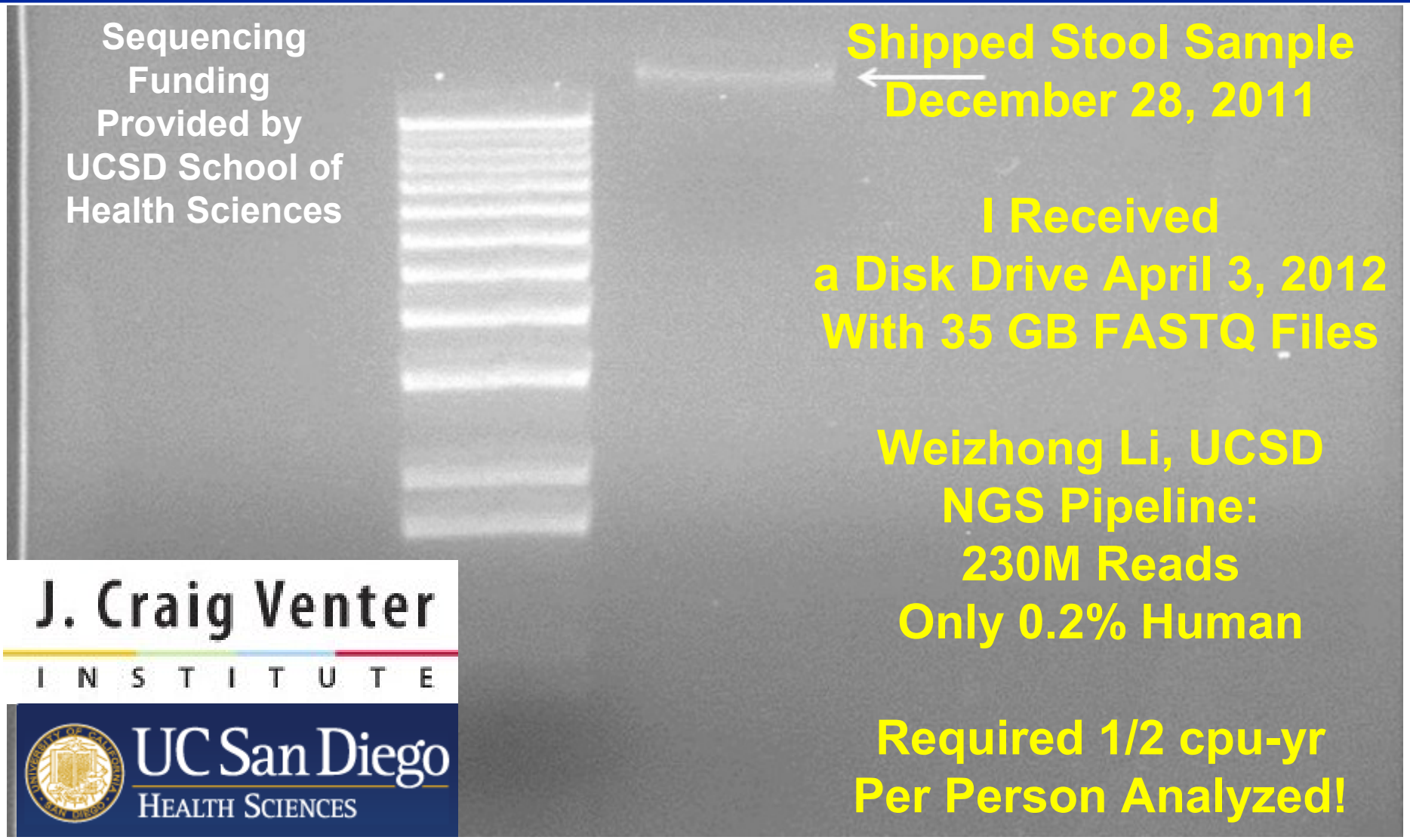
99% of Your DNA Genes Are in Microbe Cells Not Human Cells

Inclusion of the Microbiome Will Radically Change Medicine

J. Craig Venter™ INSTITUTE

it²

To Map My Gut Microbes, I Sent a Stool Sample to the Venter Institute for Metagenomic Sequencing



Gel Image of Extract from Smarr Sample-Next is Library Construction

Manny Torralba, Project Lead - Human Genomic Medicine
J Craig Venter Institute

January 25, 2012

We Created a Reference Database Of Known Gut Genomes

- **NCBI April 2013**
  - 2471 Complete + 5543 Draft Bacteria & Archaea Genomes
  - 2399 Complete Virus Genomes
  - 26 Complete Fungi Genomes
  - 309 HMP Eukaryote Reference Genomes
- **Total 10,741 genomes, ~30 GB of sequences**

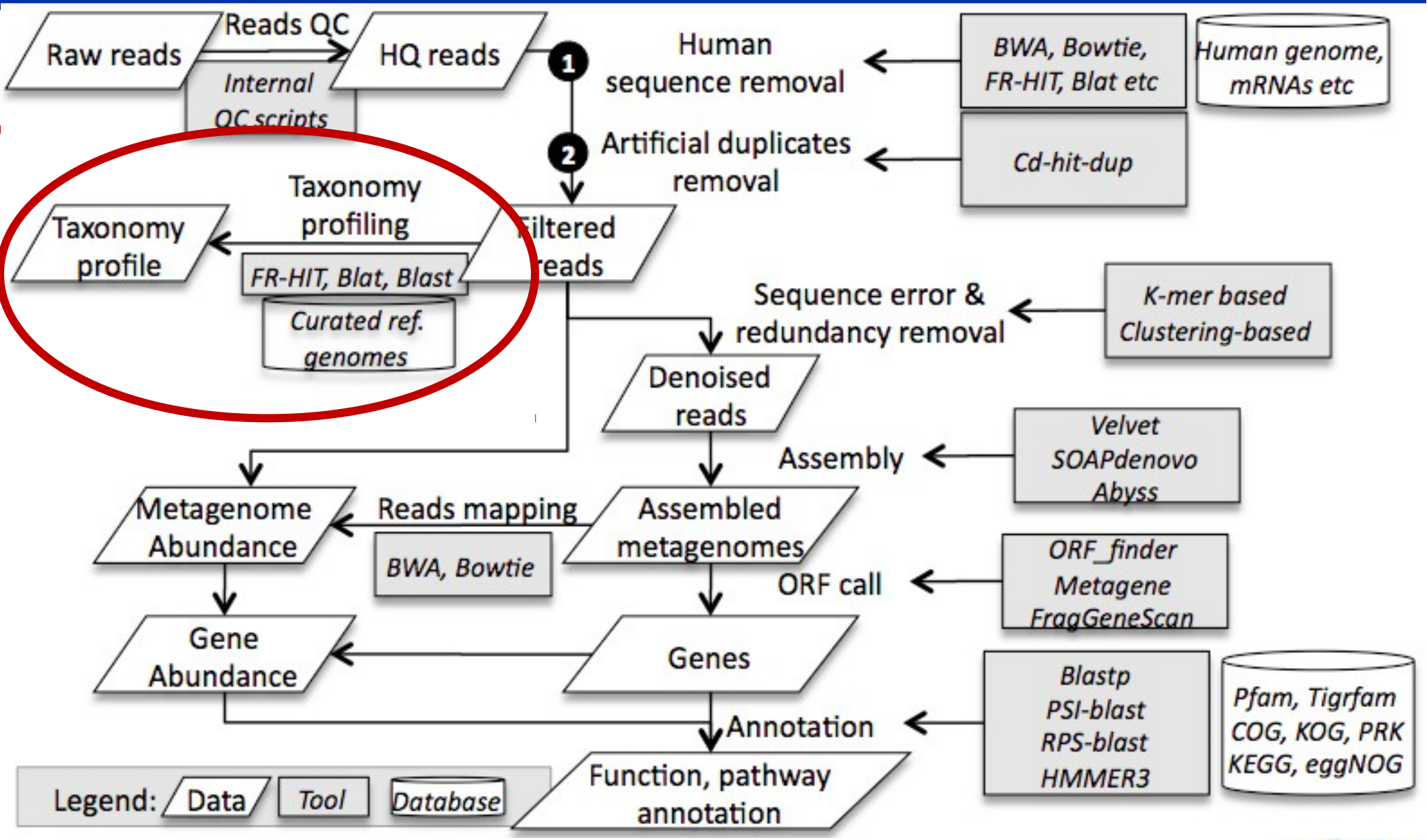
Now to Align Our 12.5 Billion Reads Against the Reference Database

J. Craig Venter INSTITUTE

Source: Weizhong Li, Sitao Wu, CRBS, UCSD

it²

Computational NextGen Sequencing Pipeline: From "Big Equations" to "Big Data" Computing



National Human Genome Research Institute

HMP

PI: (Weizhong Li, CRBS, UCSD):
NIH R01HG005978 (2010-2013, $1.1M)

CRBS

it²

We Used SDSC's Gordon Data-Intensive Supercomputer to Analyze a Wide Range of Gut Microbiomes

- ~180,000 Core-Hrs on Gordon
  - KEGG function annotation: 90,000 hrs
  - Mapping: 36,000 hrs
    - Used 16 Cores/Node and up to 50 nodes
  - Duplicates removal: 18,000 hrs
  - Assembly: 18,000 hrs
  - Other: 18,000 hrs
- Gordon RAM Required
  - 64GB RAM for Reference DB
  - 192GB RAM for Assembly
- Gordon Disk Required
  - Ultra-Fast Disk Holds Ref DB for All Nodes
  - 8TB for All Subjects



SDSC

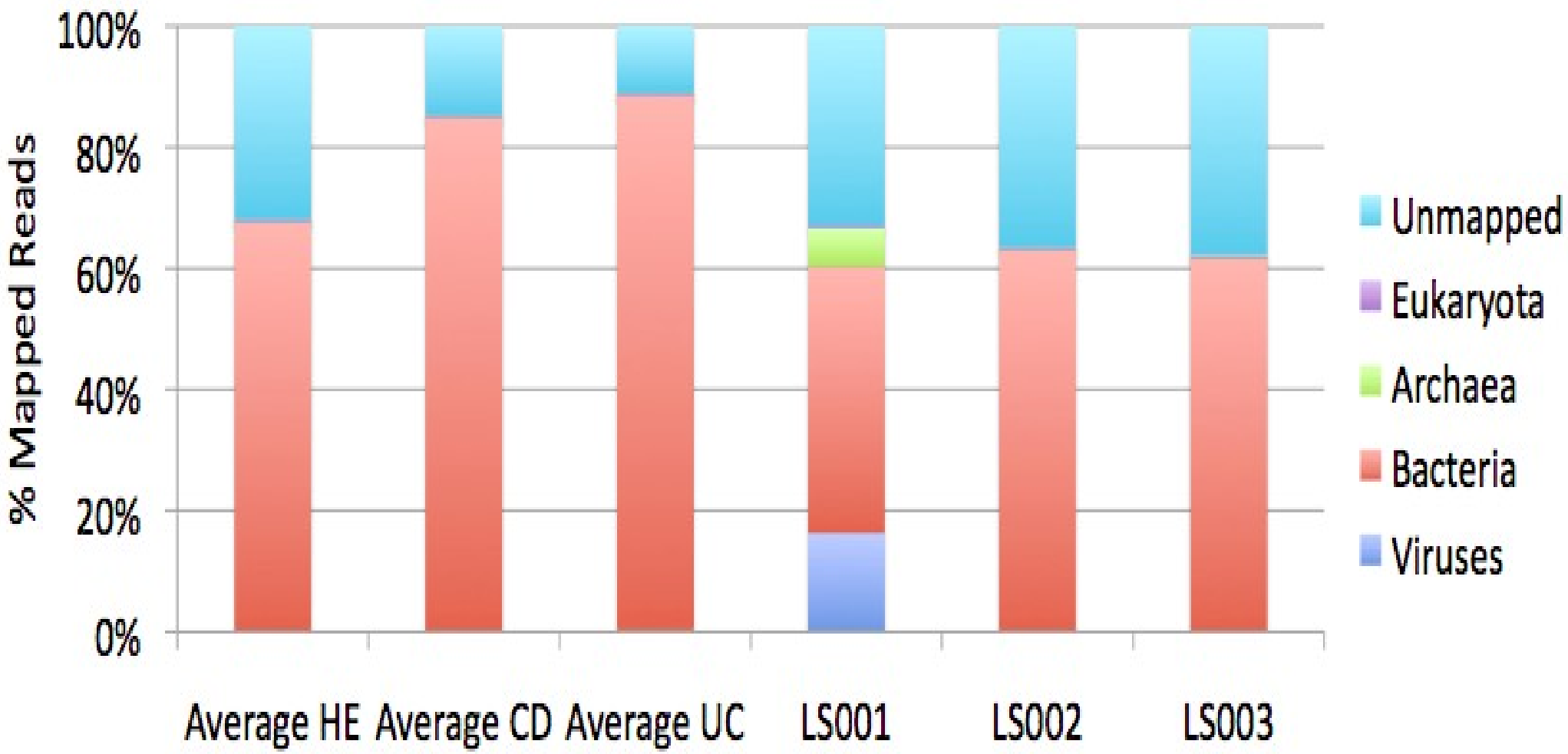
Enabled by
a Grant of Time
on Gordon from SDSC
Director Mike Norman

NSF NIH HMP

J. Craig Venter INSTITUTE

CRBS it²

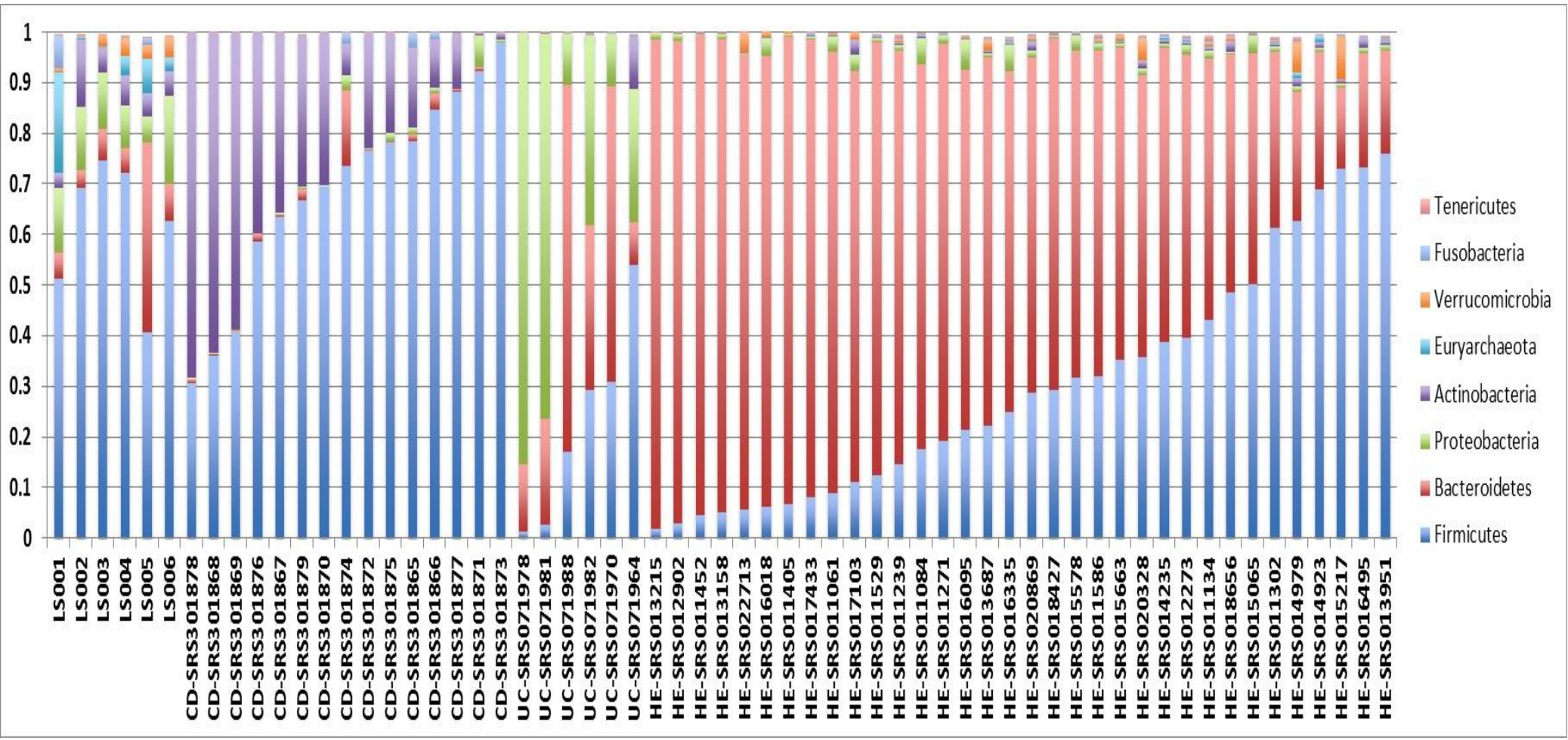
A Significant Fraction of the Reads Do Not Map Onto The Reference Genome Set



Source: Weizhong Li, CRBS, UCSD

Phyla Gut Microbial Abundance Without Viruses: LS, Crohn's, UC, and Healthy Subjects

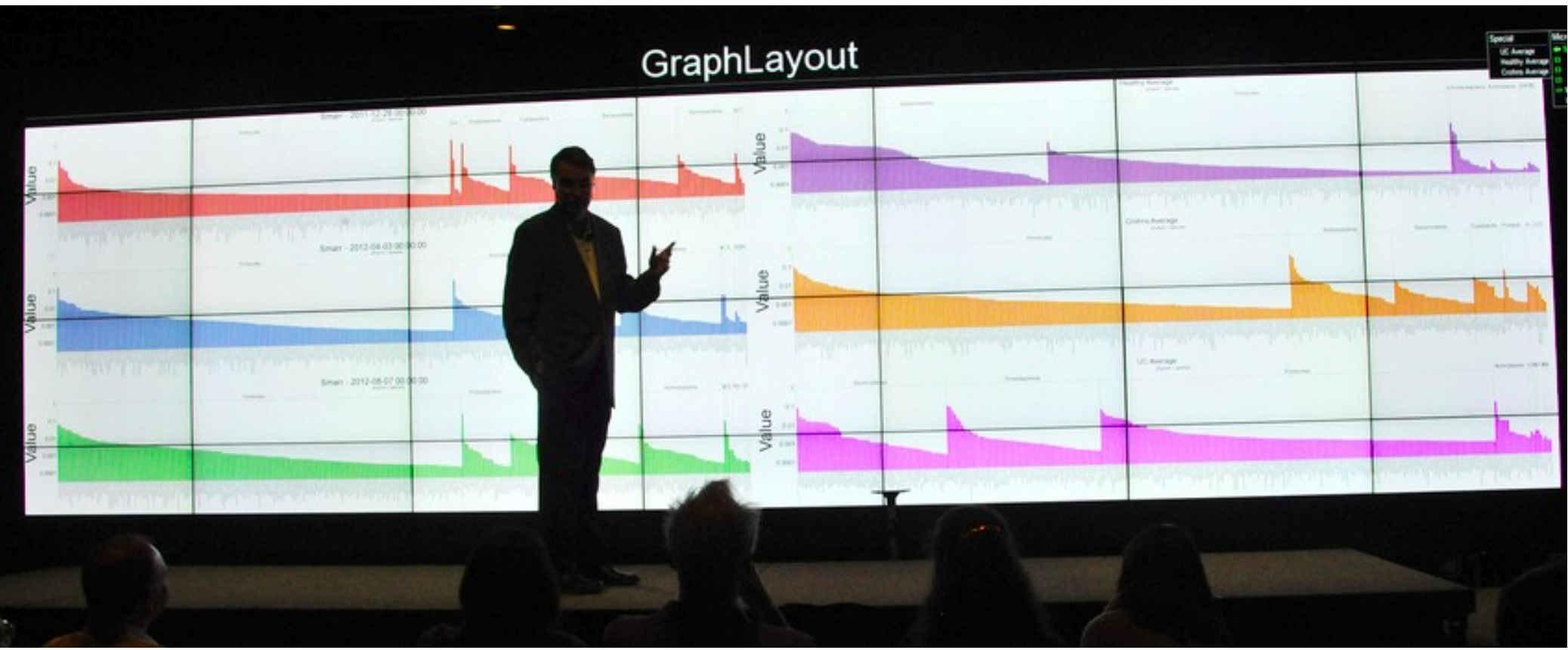
Source: Weizhong Li, Sitao Wu, CRBS, UCSD



LS **Crohn's** **Ulcerative Colitis** **Healthy**

Toward Noninvasive Microbial Ecology Diagnostics

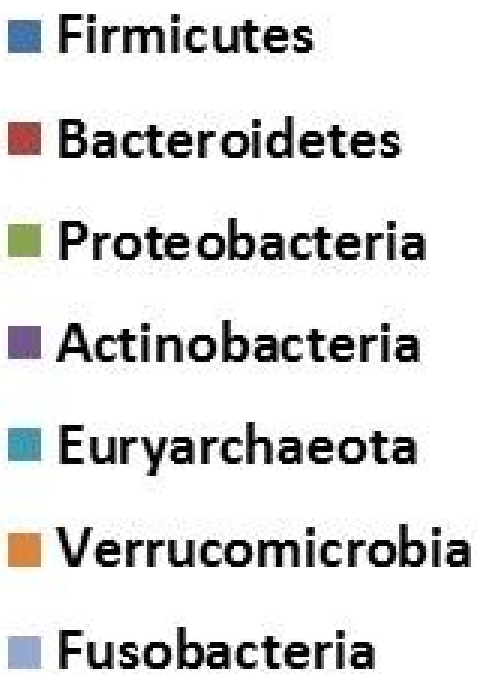
Using Scalable Visualization Allows Comparison of the Relative Abundance of 200 Microbe Species

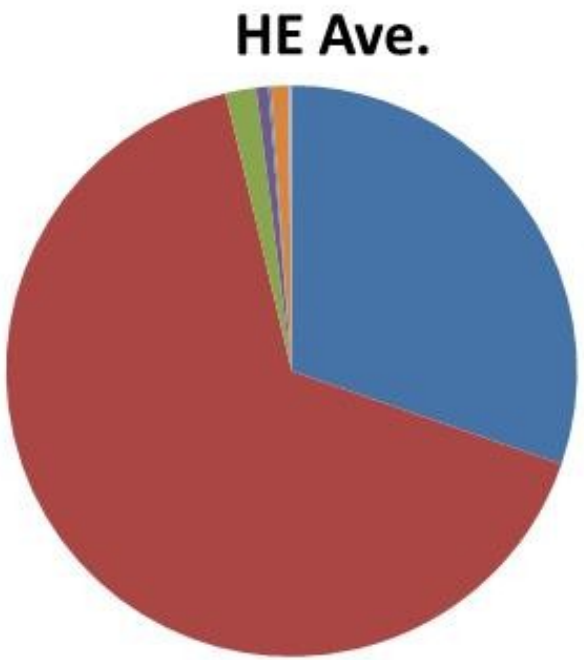


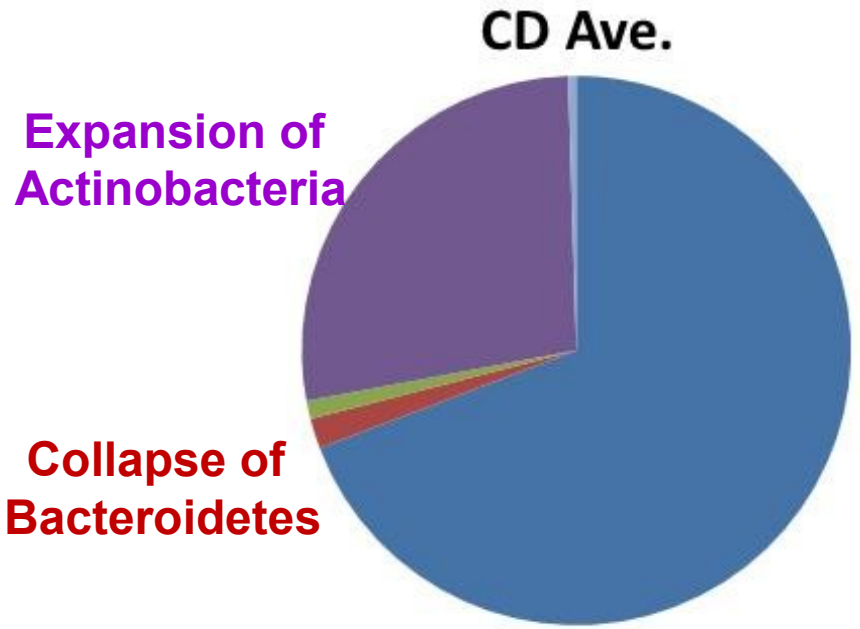
Comparing 3 LS Time Snapshots (Left)
with Healthy, Crohn's, UC (Right Top to Bottom)

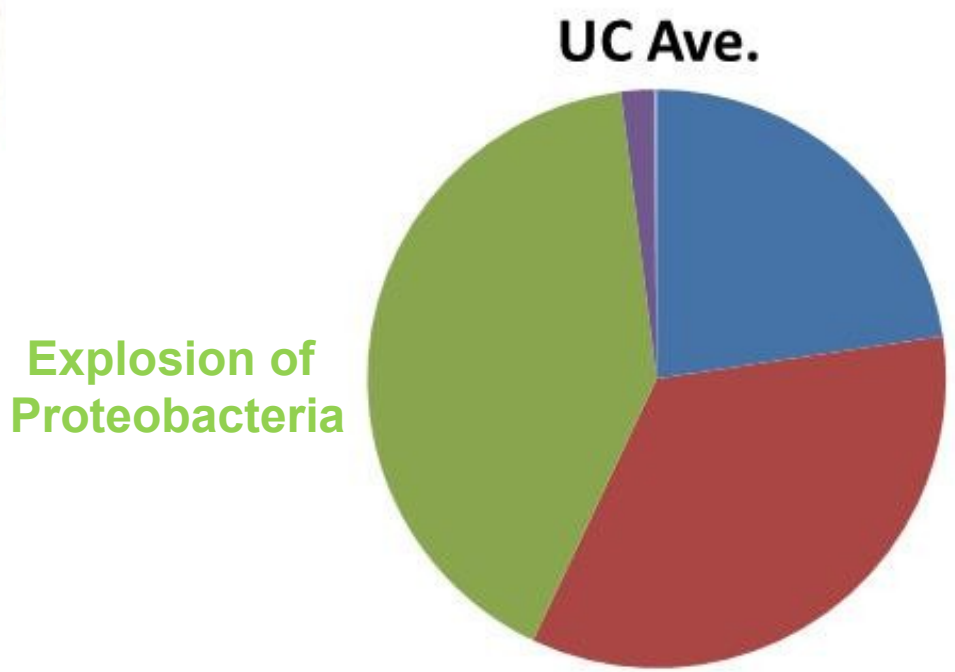
Calit2 VROOM-FuturePatient Expedition

Comparison of 35 Healthy to 15 CD and 6 UC Gut Microbiomes at the Phyla Level

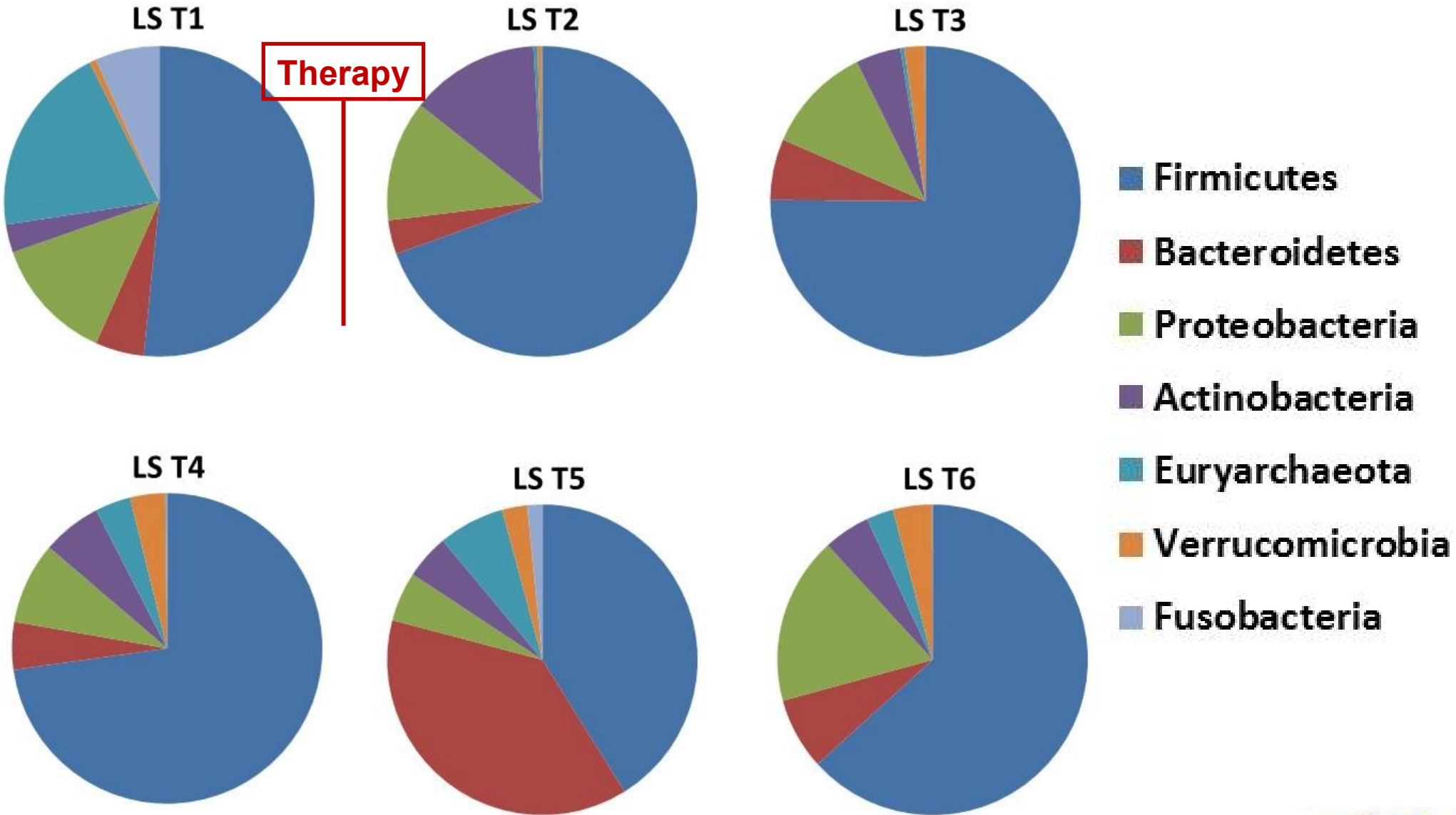






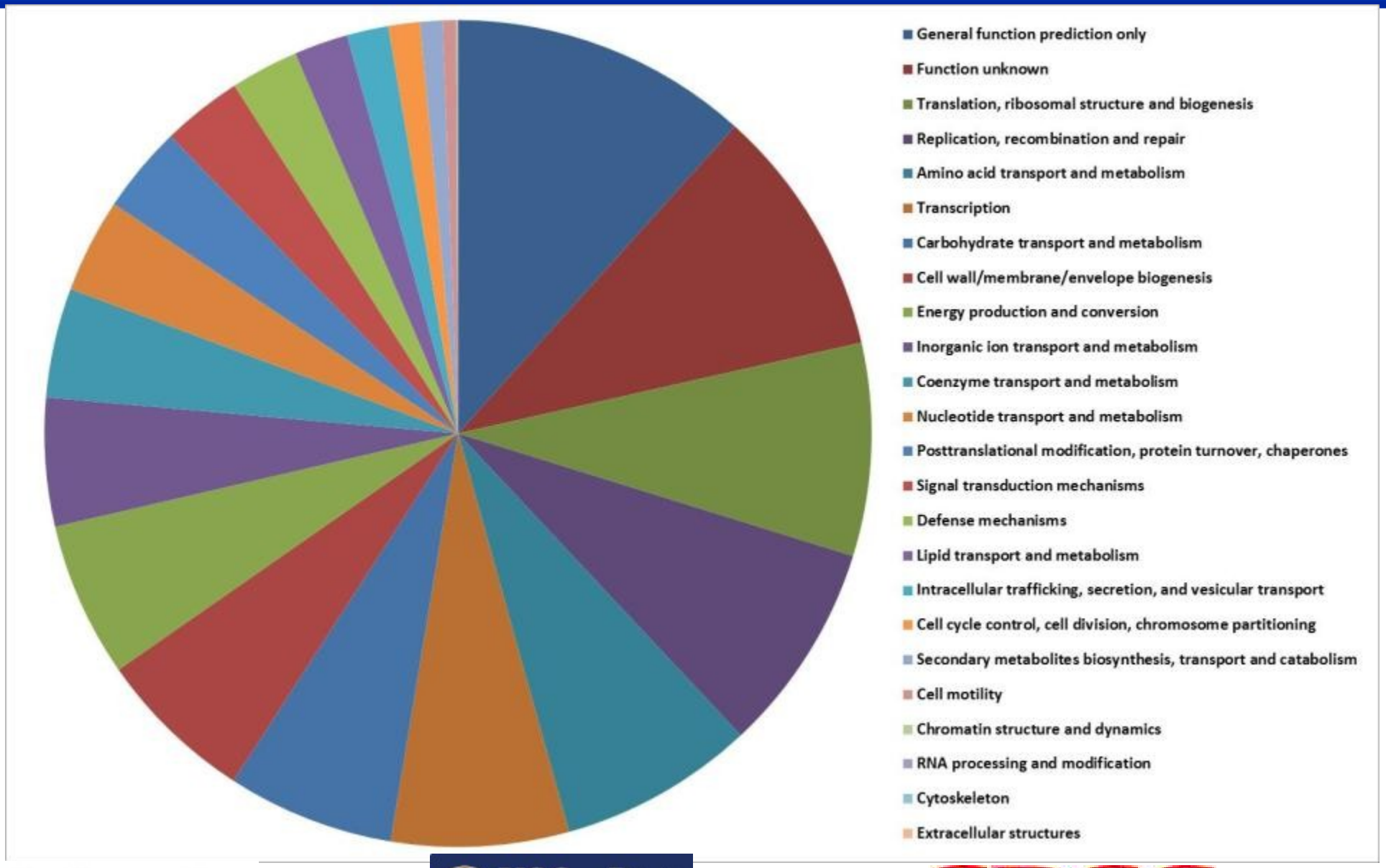


Time Series Reveals Autoimmune Dynamics of Gut Microbiome by Phyla



Six Metagenomic Time Samples Over 16 Months

From Taxonomy to Function: Analysis of LS Clusters of Orthologous Groups (COGs)



J. Craig Venter INSTITUTE

UC San Diego Health Sciences

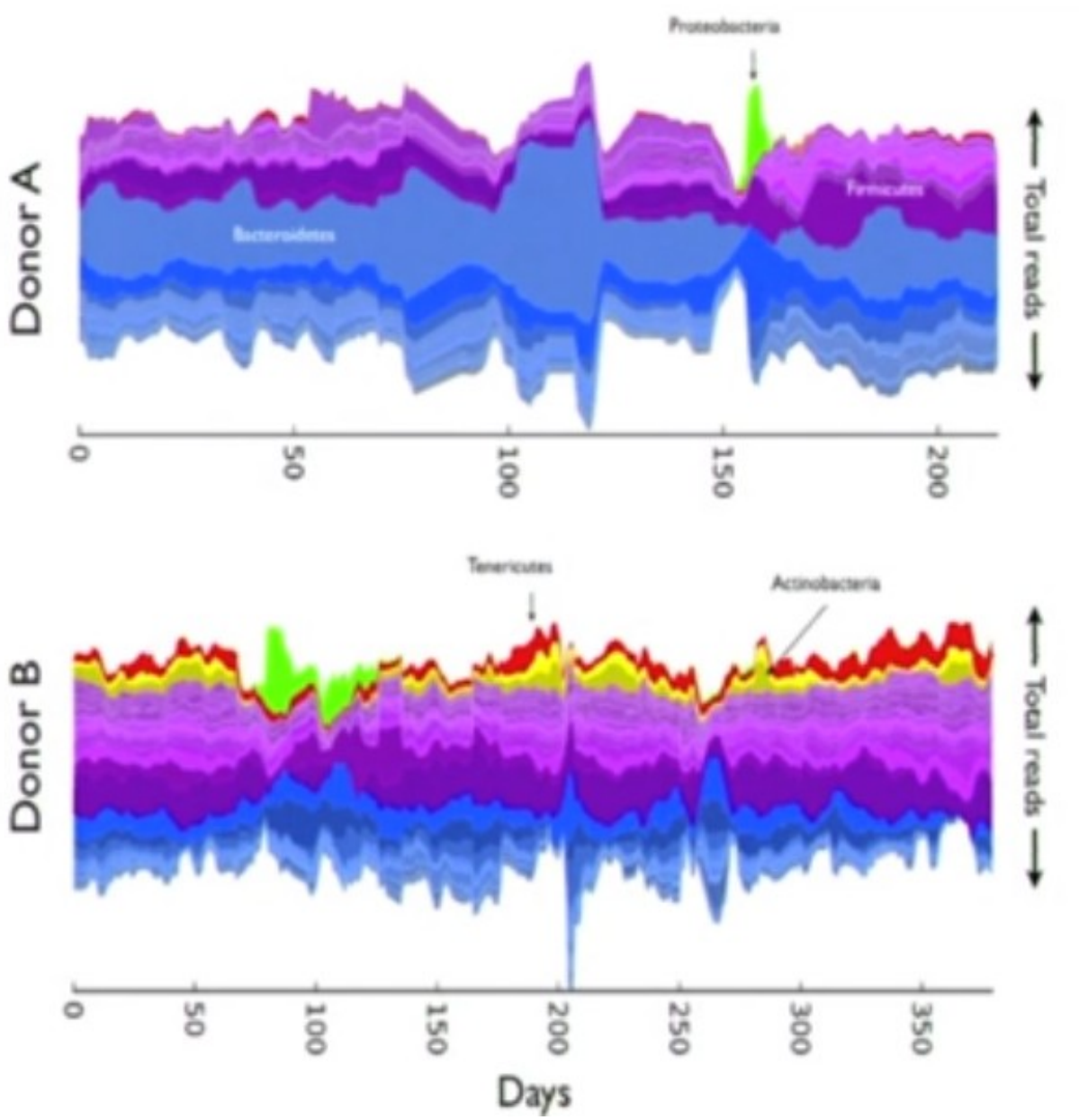
SDSC

it²

Analysis: Weizhong Li & Sitao Wu, UCSD

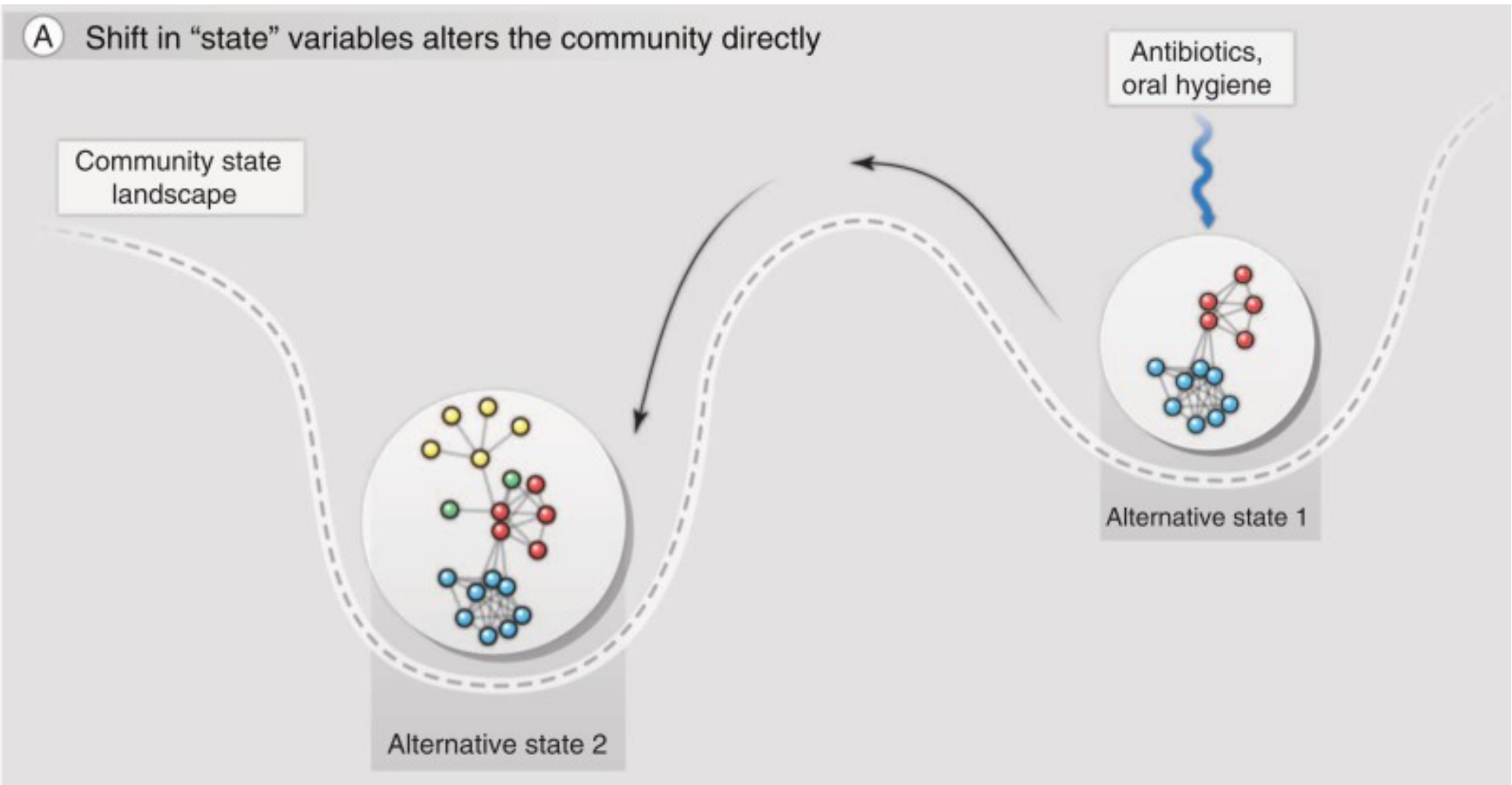
The Adult Healthy Gut Microbiome Is Remarkably Stable Over Time

Gut communities over time



Source: Eric Alm, MIT

Lessons from Ecological Dynamics I: Gut Microbiome Has Multiple Relatively Stable Equilibria



"The Application of Ecological Theory Toward an Understanding of the Human Microbiome,"
Elizabeth Costello, Keaton Stagaman, Les Dethlefsen, Brendan Bohannan, David Relman
Science 336, 1255-62 (2012)

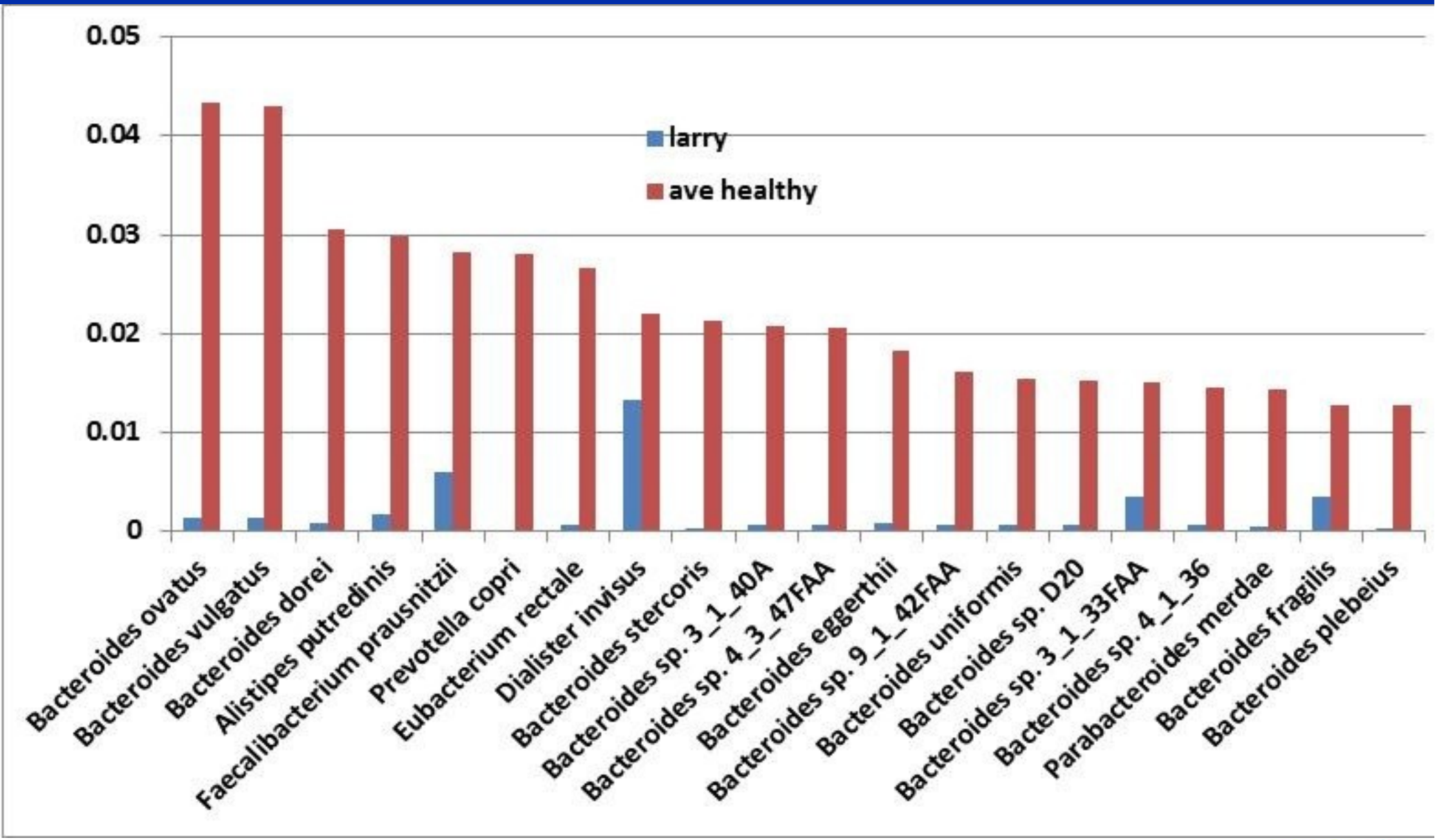
Lessons From Ecological Dynamics II:
Invasive Species Dominate After Major Species Destroyed



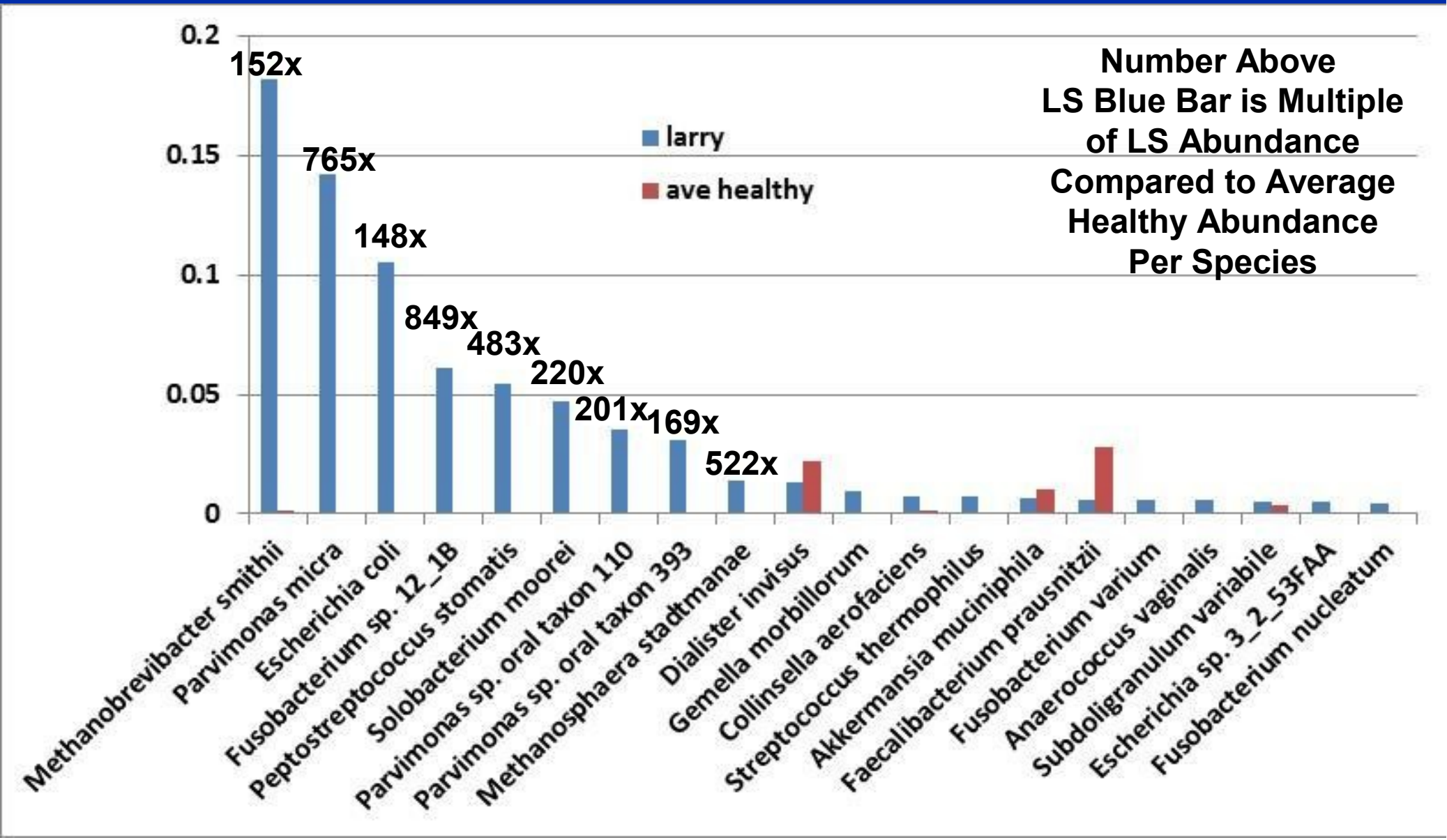
”In many areas following these burns invasive species are able to establish themselves, crowding out native species.”

Source: Ponderosa Pine Fire Ecology
http://cpluhna.nau.edu/Biota/ponderosafire.htm

Almost All Abundant Species (≥1%) in Healthy Subjects Are *Severely Depleted* in Larry's Gut Microbiome



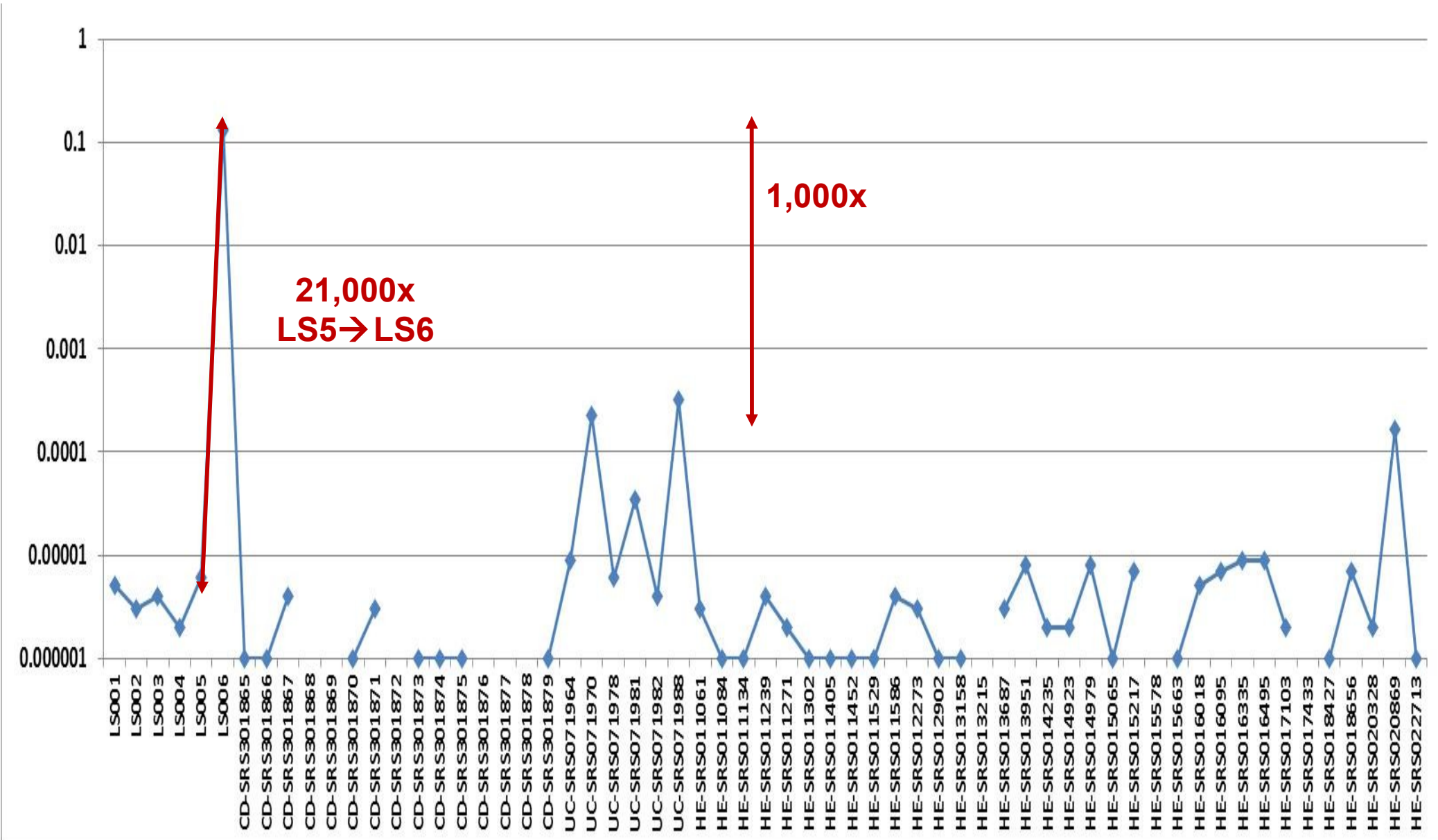
Top 20 Most Abundant Microbial Species In LS vs. Average Healthy Subject



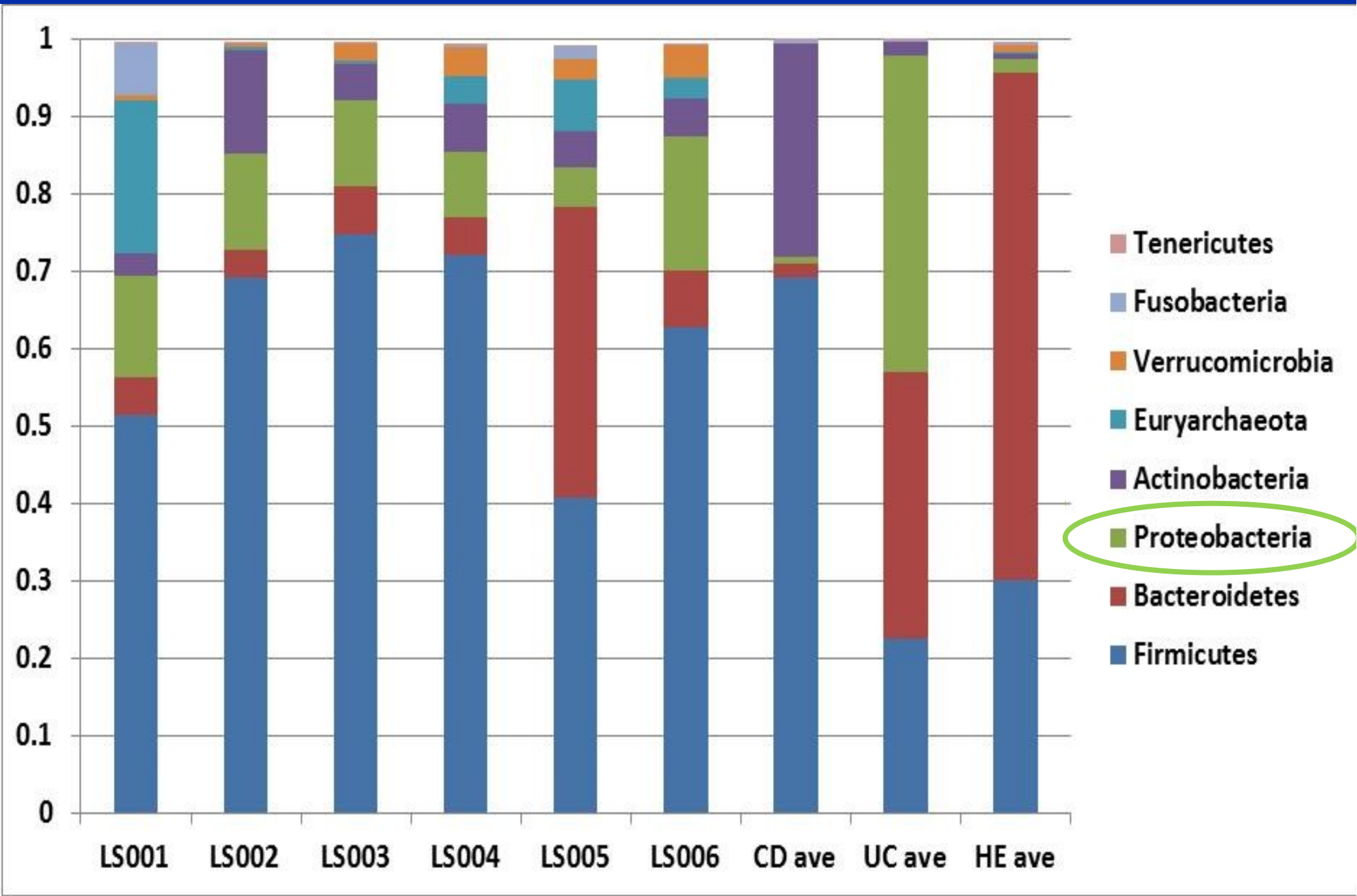
Source: Sequencing JCVI; Analysis Weizhong Li, UCSD
LS December 28, 2011 Stool Sample

The Dramatic Bloom of Enterobacteriaceae bacterium 9_2_54FAA

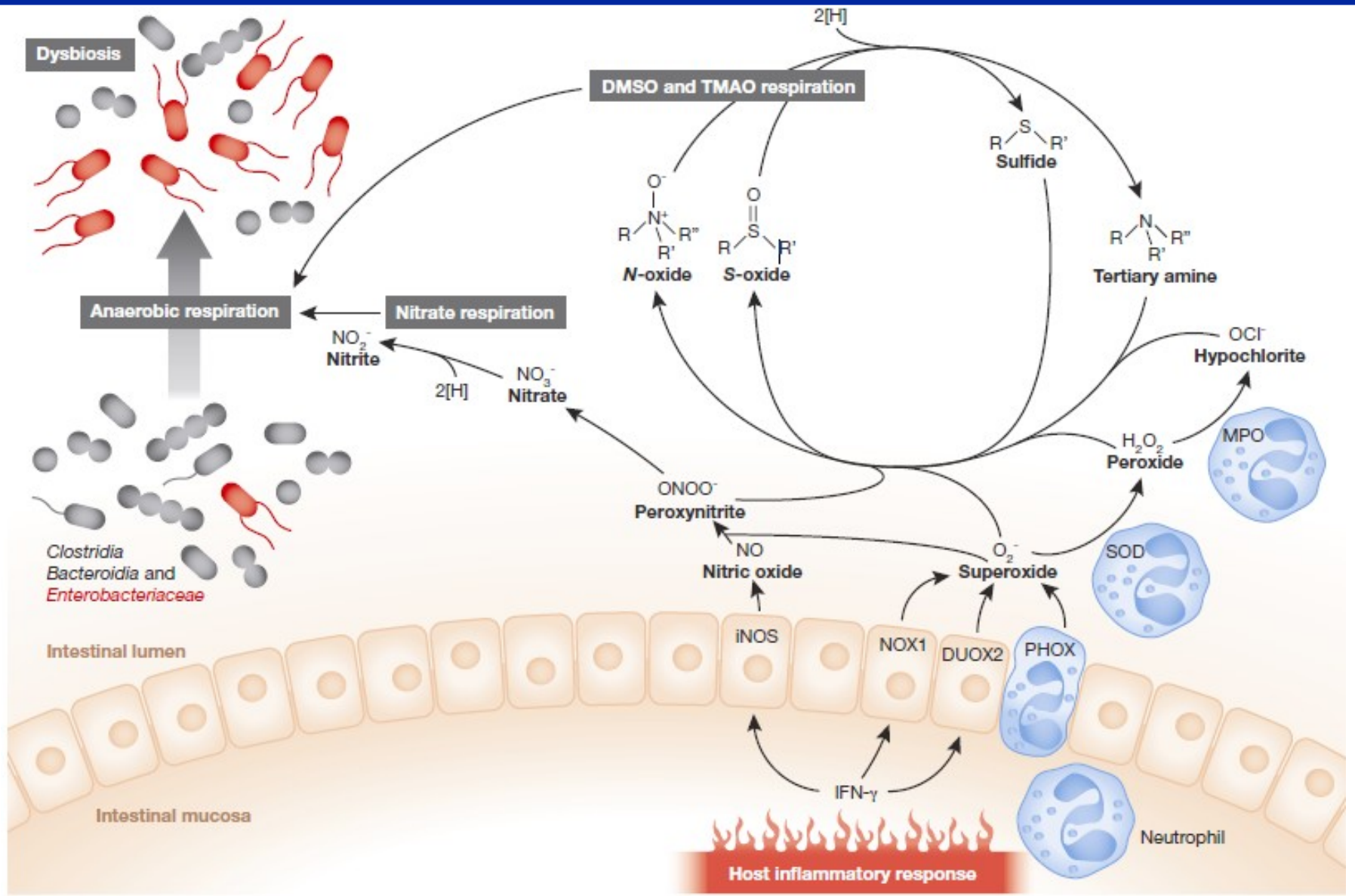
This Microbe is a Proteobacteria Targeted by the NIH HMP



Focusing in on the Dynamical Change Within Proteobacteria



Inflammation Enables Anaerobic Respiration Which Leads to Phylum-Level Shifts in the Gut Microbiome



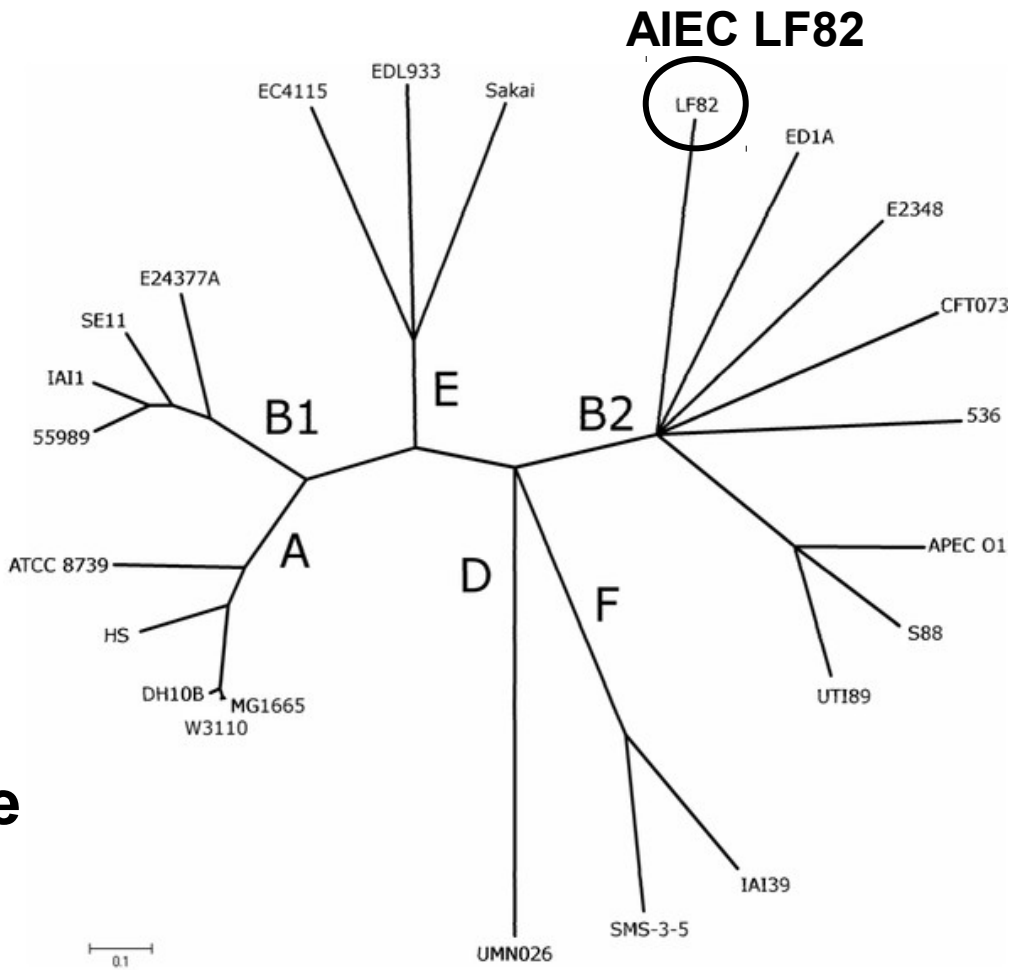
Sebastian E. Winter, Christopher A. Lopez & Andreas J. Bäumler, EMBO reports VOL 14, p. 319-327 (2013)

Does Intestinal Inflammation Select for Pathogenic Strains That Can Induce Further Damage?

"Adherent-invasive E. coli (AIEC) are isolated more commonly from the intestinal mucosa of individuals with Crohn's disease than from healthy controls."

"Thus, the mechanisms leading to dysbiosis might also select for intestinal colonization with more harmful members of the Enterobacteriaceae* —such as AIEC— thereby exacerbating inflammation and interfering with its resolution."

Sebastian E. Winter , et al.,
EMBO reports VOL 14, p. 319-327 (2013)

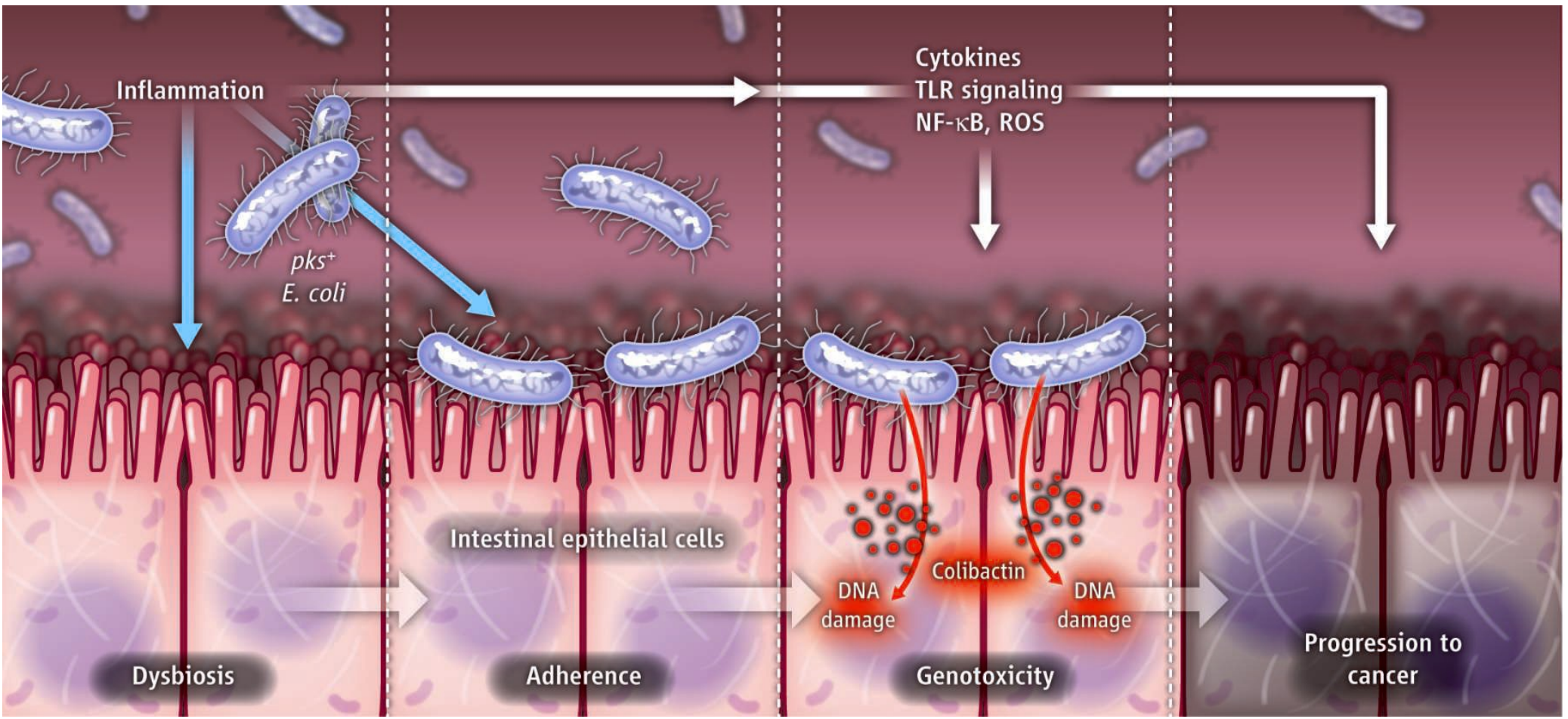


E. coli/Shigella Phylogenetic Tree
Miquel, et al.
PLOS ONE, v. 5, p. 1-16 (2010)

*Family Containing E. coli

Chronic Inflammation Can Accumulate Cancer-Causing Bacteria in the Human Gut

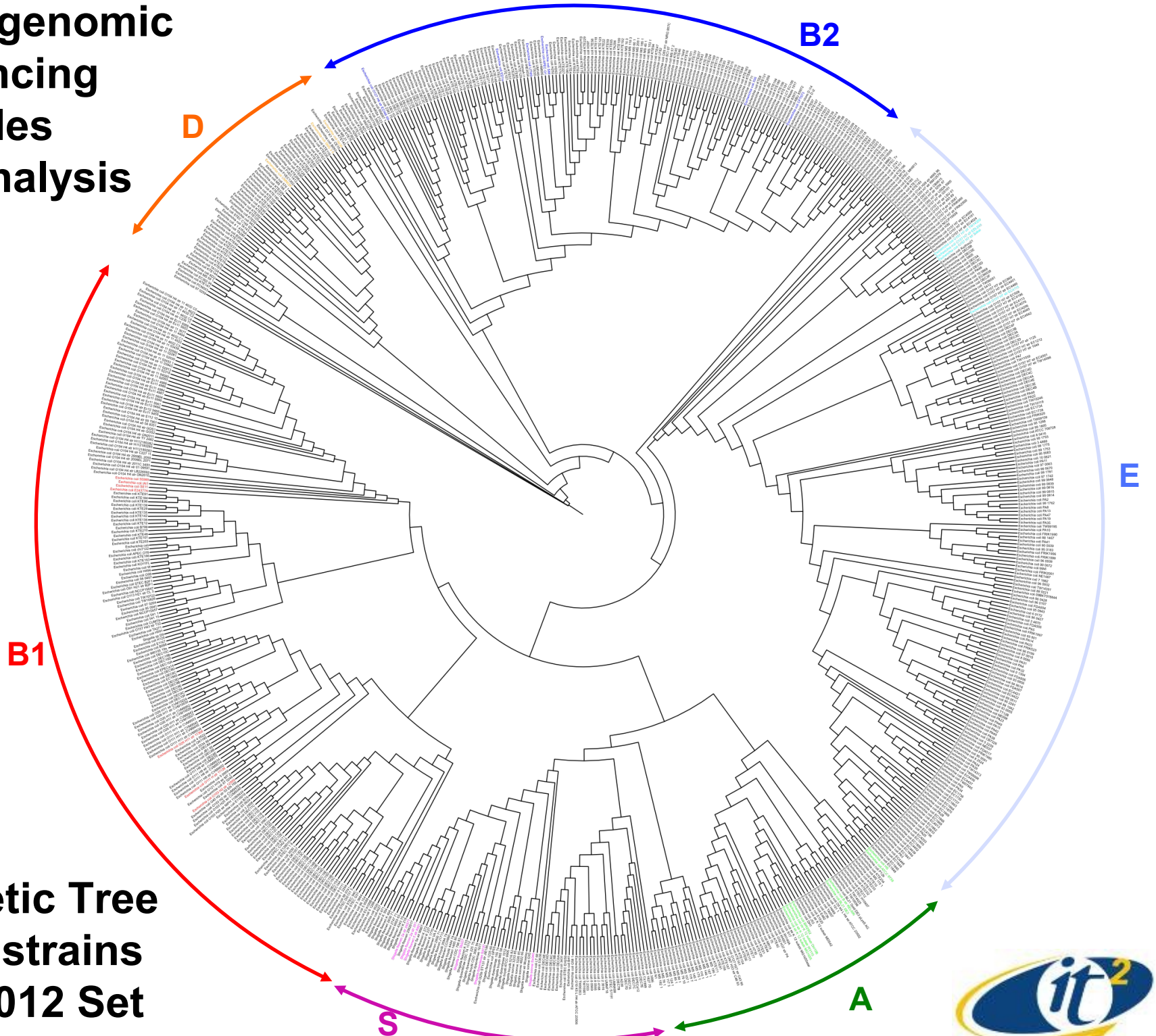
Escherichia coli Strain NC101



Science
AAAS

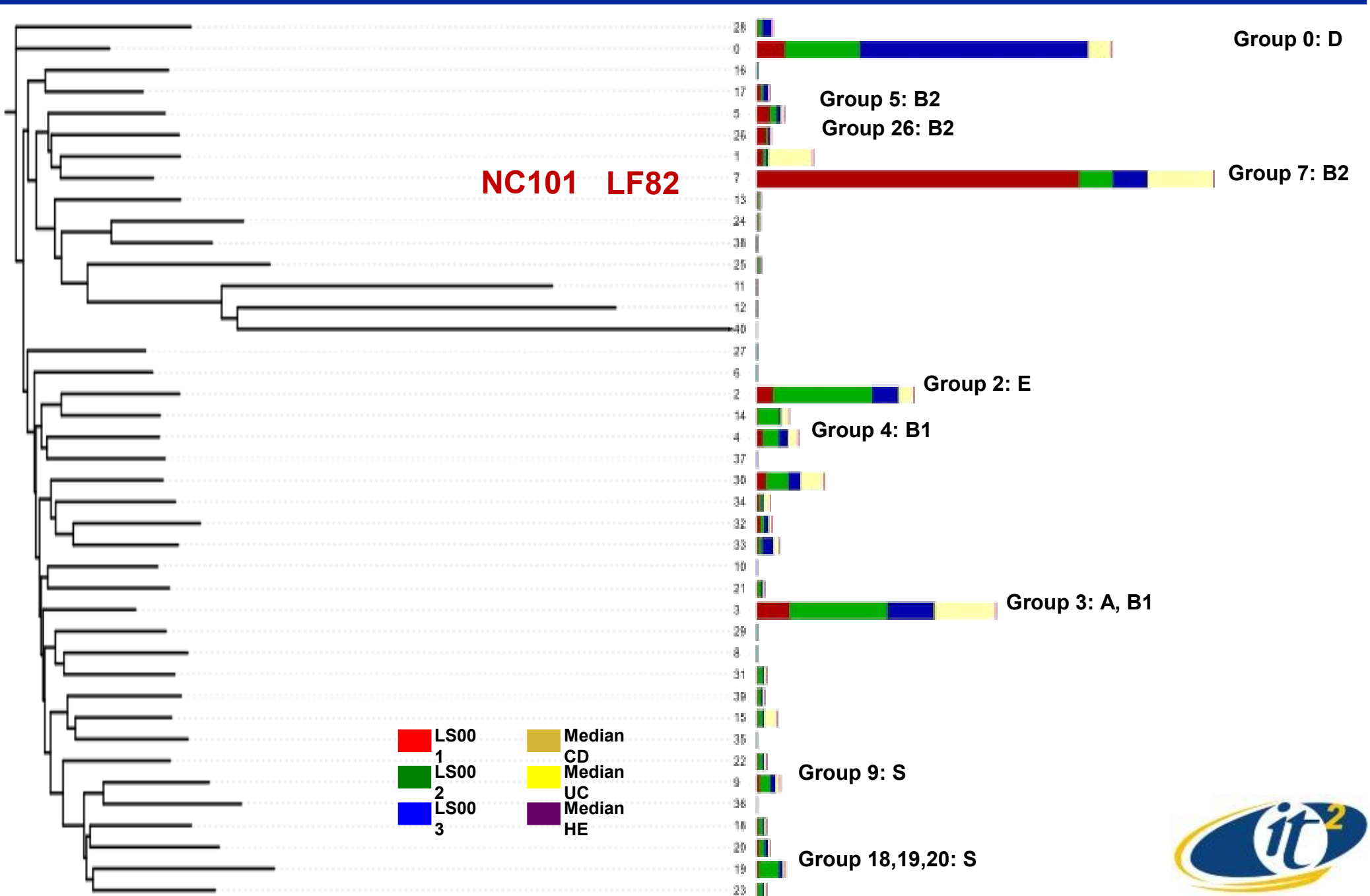
Bacteria Deliver a Genotoxic Hit
Robert F. Schwabe and Timothy C. Wang
Science **338**, 52 (2012);
DOI: 10.1126/science.1229905

Deep Metagenomic Sequencing Enables Strain Analysis

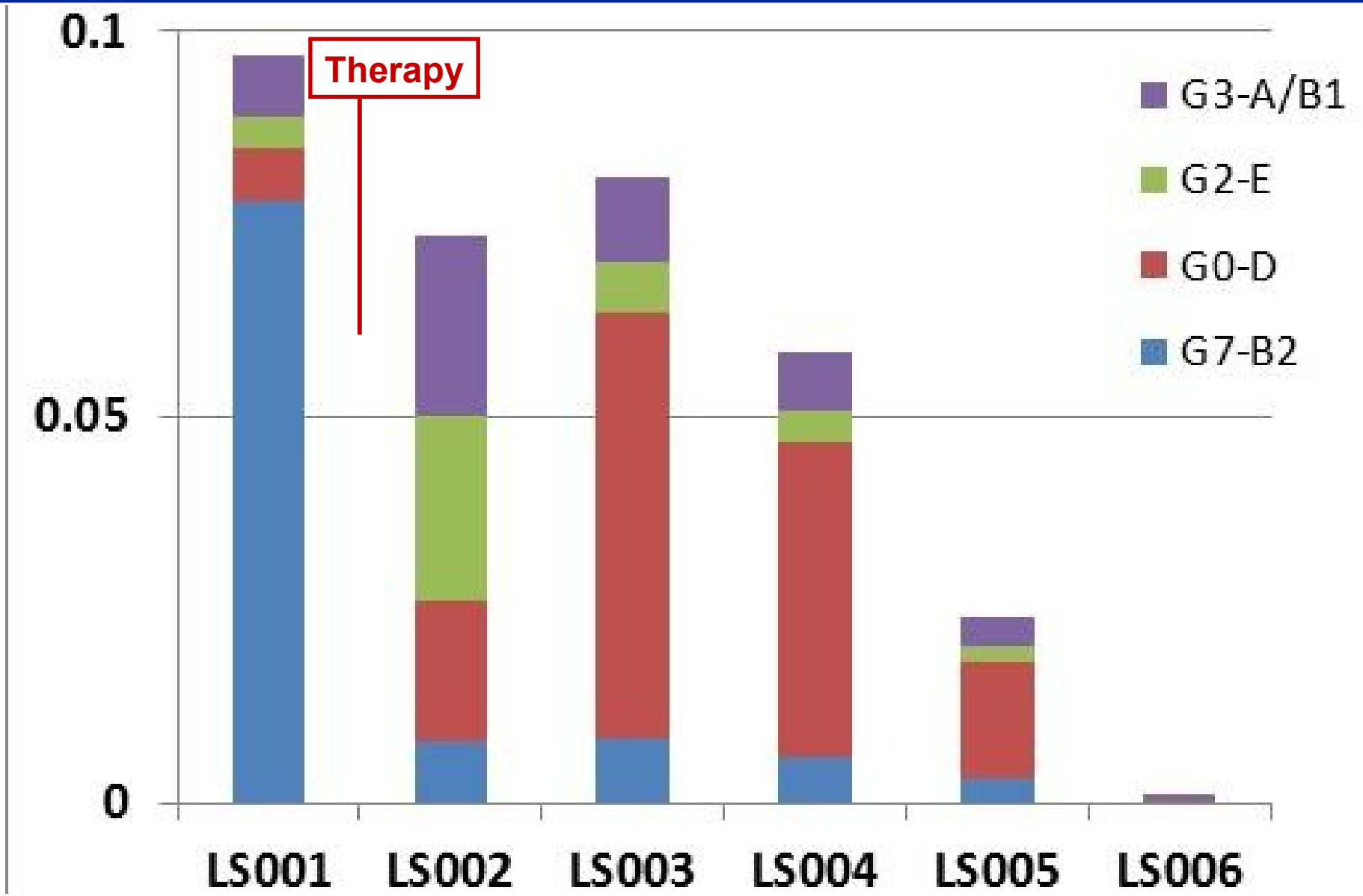


Phylogenetic Tree
778 Ecoli strains
=6x our 2012 Set

We Divided the 778 E. coli Strains into 40 Groups, Each of Which Had 80% Identical Genes



Reduction in E. coli Over Time
With Major Shifts in Strain Abundance



Strains >0.5% Included

Next Step: Time Series of Metagenomic Gut Microbiomes and Immune Variables in an N=100 Clinic Trial

Goal: Understand
The Coupled Human Immune-Microbiome Dynamics
In the Presence of Human Genetic Predispositions

Thanks to Our Great Team!

UCSD Metagenomics Team

Weizhong Li
Sitao Wu

JCVI Team

Karen Nelson
Shibu Yooseph
Manolito Torralba

Calit2@UCSD Future Patient Team

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



SDSC Team

Michael Norman
Mahidhar Tatineni
Robert Sinkovits