“Quantifying the Time Progression of a Human Autoimmune Disease using Genome Sequencing and Supercomputers”

University of California, San Francisco
San Francisco, CA
December 3, 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
http://lsmarr.calit2.net

Abstract

The human body contains ten times the number of microbe cells as human cells 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 IBD. Motivated by a diagnosis that I have Crohn's disease, 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. I have also evaluated the relative abundances of Fusobacteria species and E. coli strains that have been hypothesized to be related to colon cancer. To decode the details of the microbial ecology required high resolution metagenomics sequencing at the Venter Institute, several CPU-decades of supercomputer time, 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 IBD.

UCSF Has a Vision of a Future Precision Medicine

In May 2013, the world's foremost thinkers, creators, and innovators convened on the campus of UCSF to make precision medicine a reality.

"It is only by patients demanding that health improve, that we think the precision medicine vision will actually take place."

-- UCSF Chancellor Susan Desmond-Hellmann, MD, MPH



Where I Believe We are Headed: Predictive, Personalized, Preventive, & Participatory Medicine

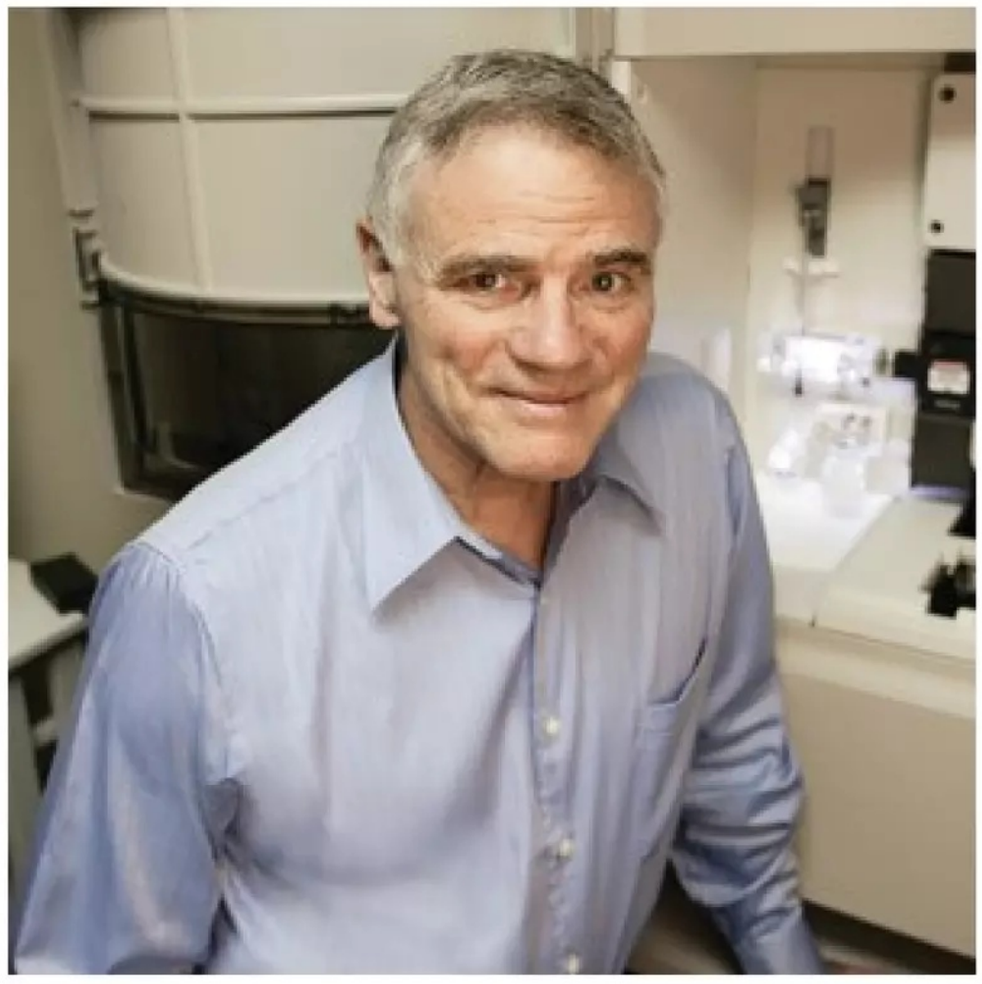
Newsweek

I am Lee Hood's Lab Rat!

A Doctor's Vision of the Future of Medicine

Leroy Hood
NEWSWEEK

From the magazine issue dated Jul 13, 2009



www.newsweek.com/2009/06/26/a-doctor-s-vision-of-the-future-of-medicine.html

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



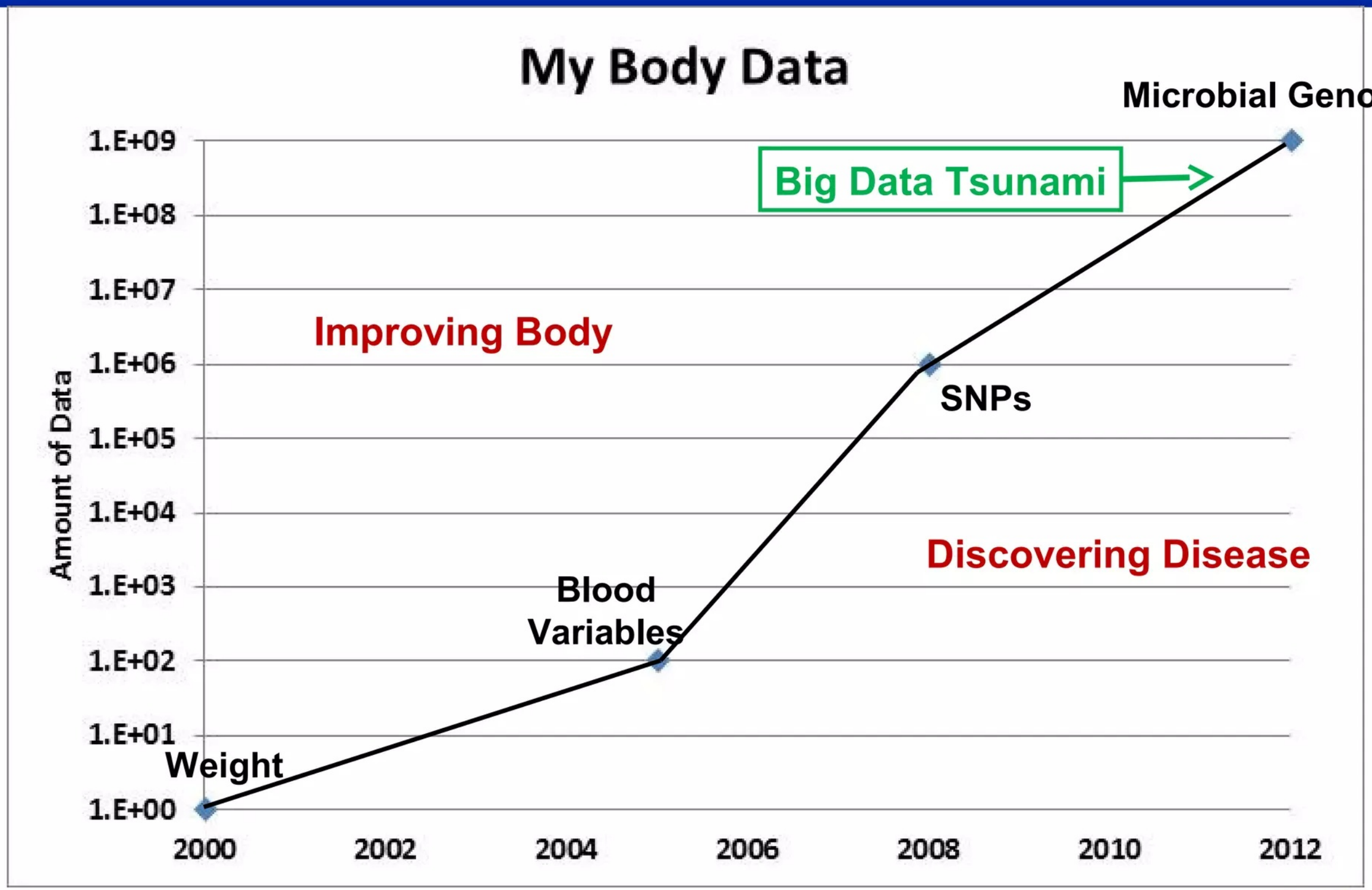
Age 61



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

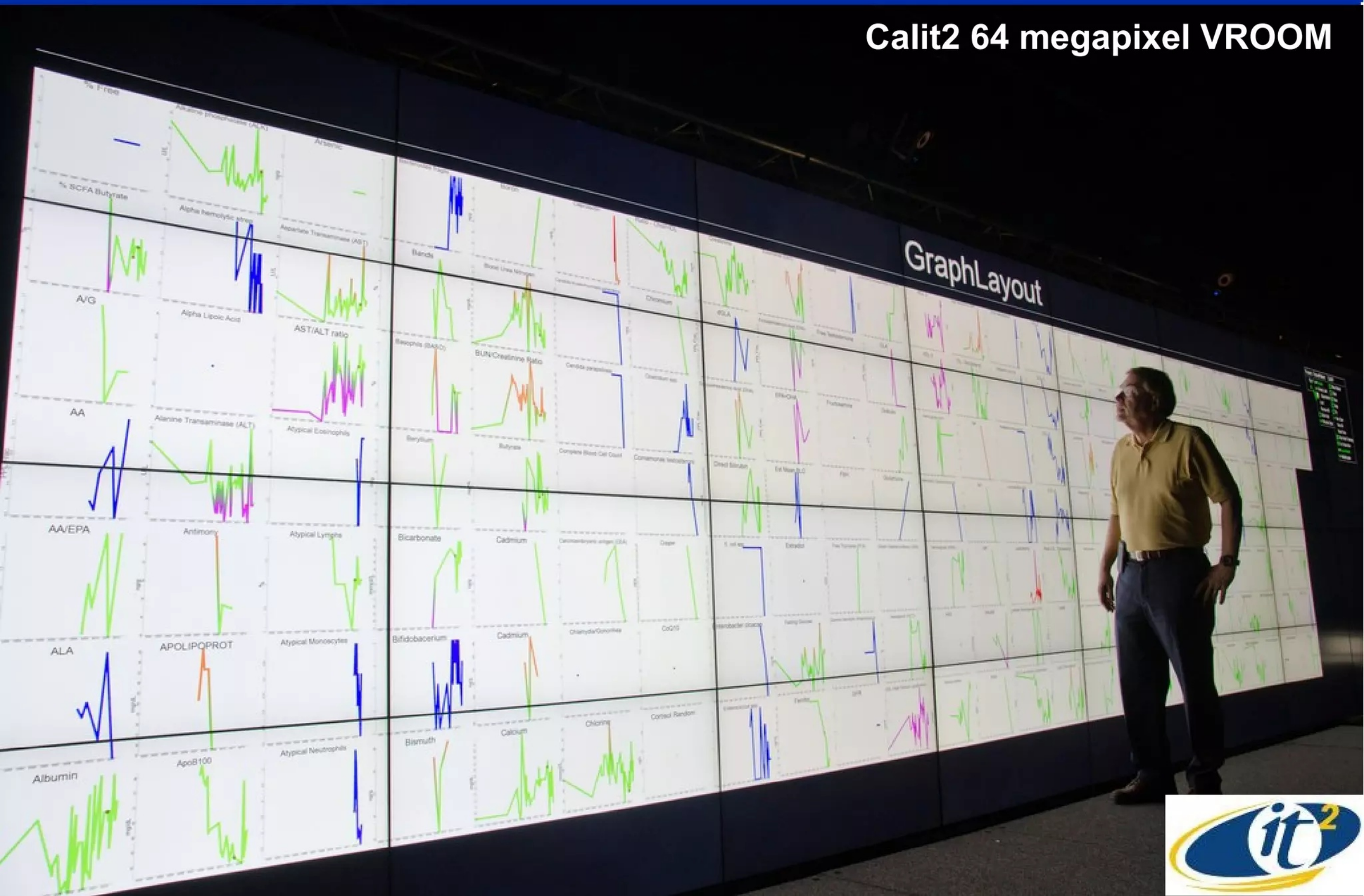
http://lsmarr.calit2.net/repository/LS_reading_recommendations_FiRe_2011.pdf

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

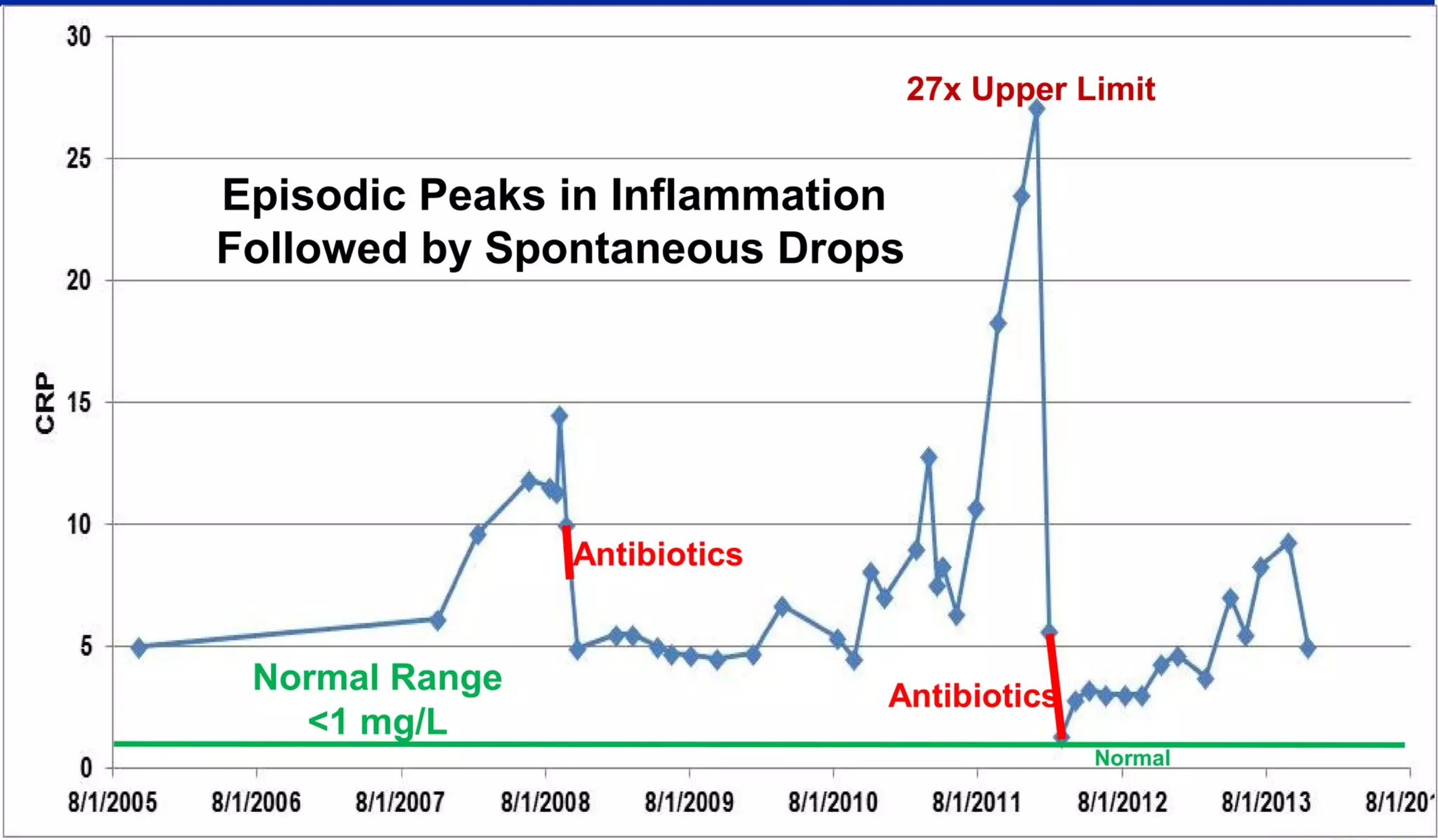


Visualizing Time Series of 150 LS Blood and Stool Variables, Each Over 5-10 Years

Calit2 64 megapixel VROOM

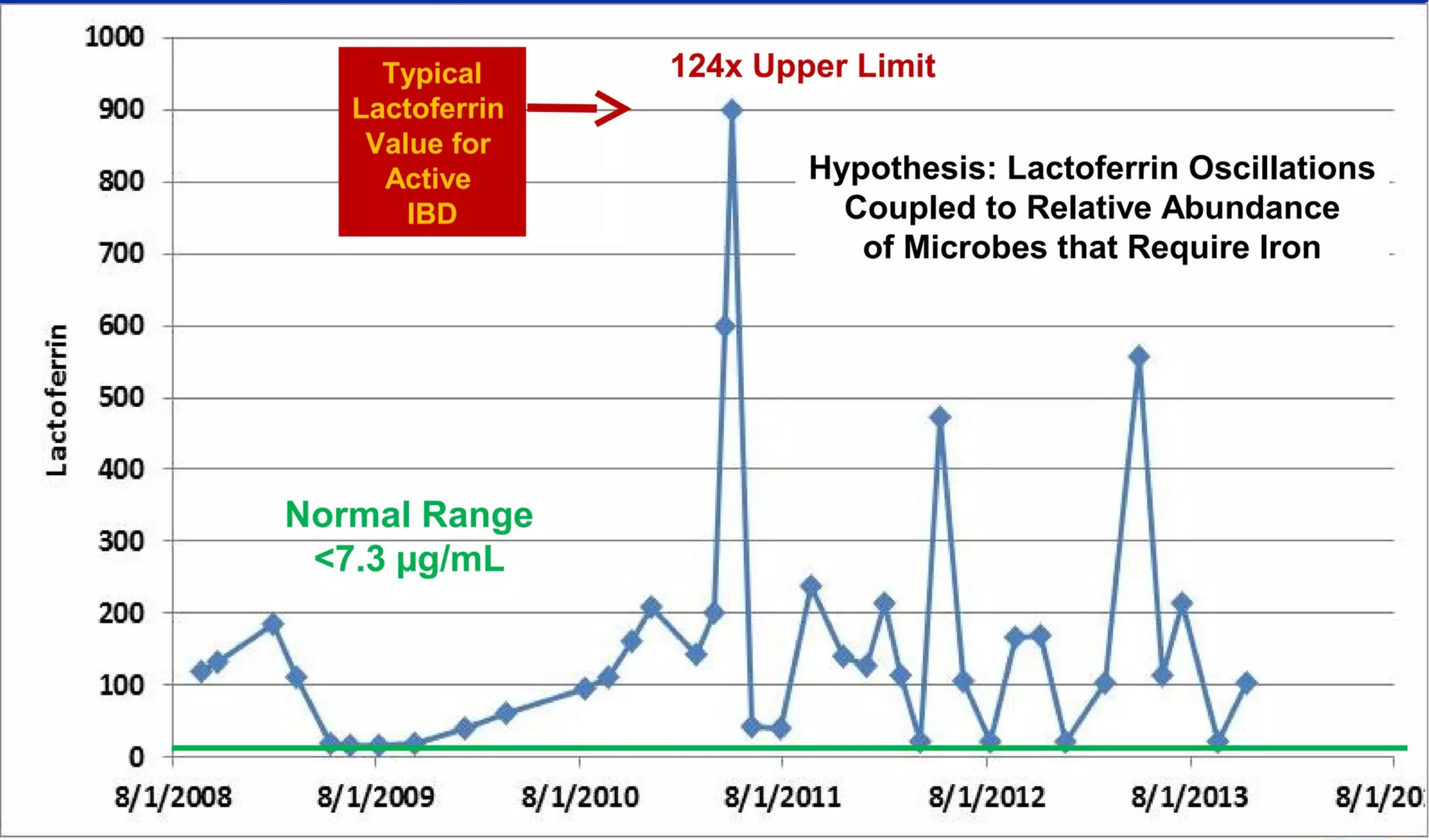


Only One of My Blood Measurements Was Far Out of Range--Indicating Chronic Inflammation



Complex Reactive Protein (CRP) is a Blood Biomarker for Detecting Presence of Inflammation

Adding Stool Tests Revealed Oscillatory Behavior in an Immune Variable

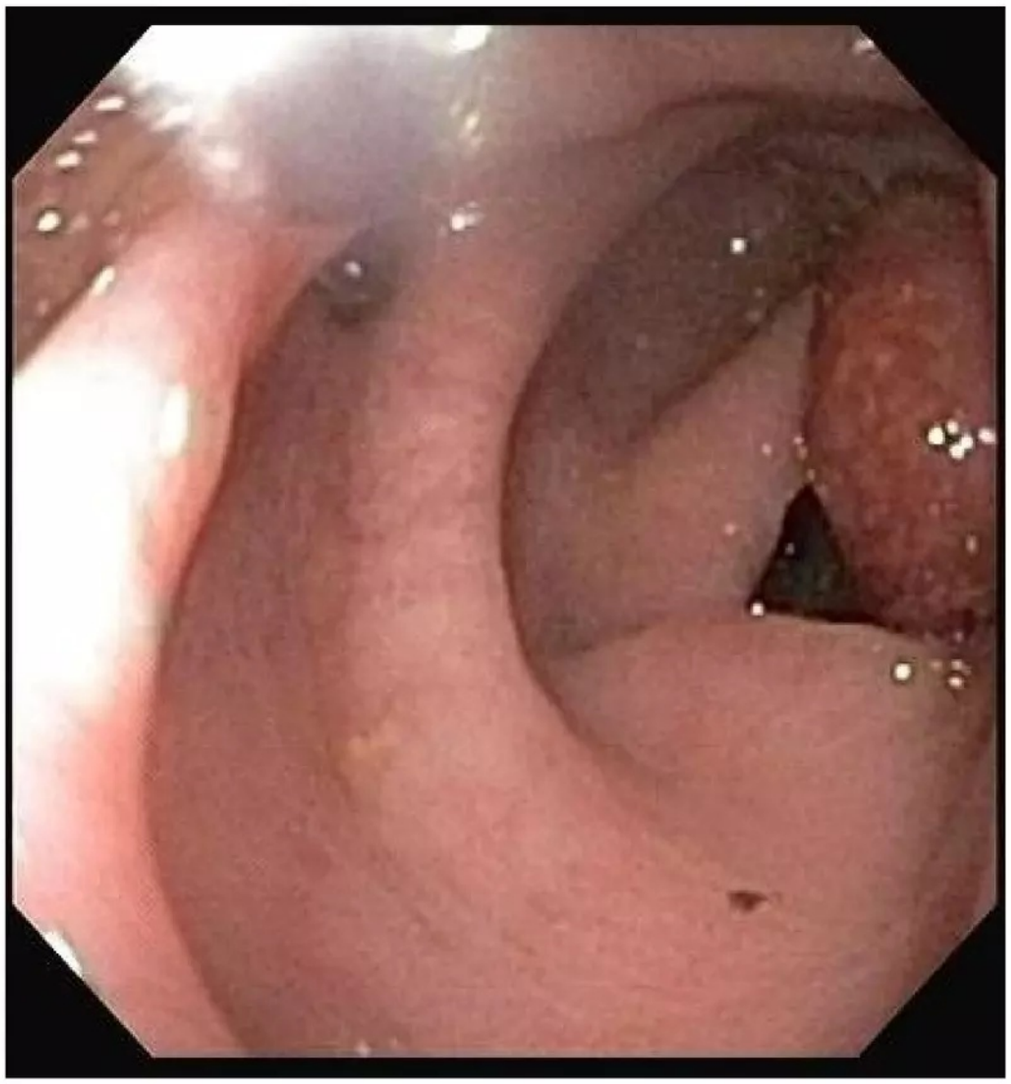


Lactoferrin is a Protein Shed from Neutrophils -
An Antibacterial that Sequesters Iron

Colonoscopy Images Show Inflamed Pseudopolyps in 6 inches of Sigmoid Colon



Dec 2010



May 2011

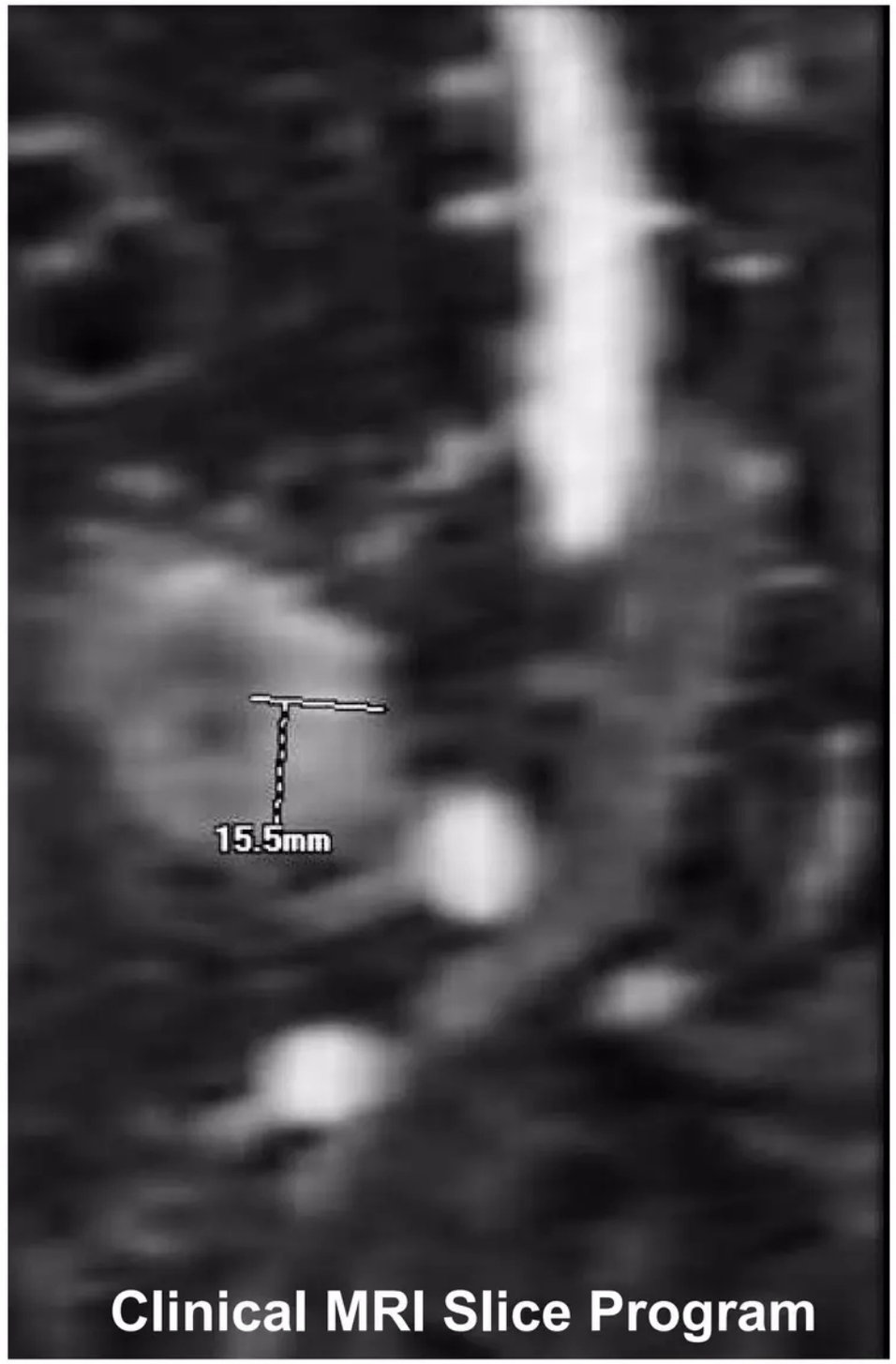
Confirming the Colonic Crohn's Hypothesis: Finding the "Smoking Gun" with MRI Imaging

"Long segment wall thickening in the proximal and mid portions of the sigmoid colon, extending over a segment of ~16 cm, with suggestion of intramural sinus tracts. Edema in the sigmoid mesentery and engorgement of the regional vasa recta."
– MRI report, Cynthia Santillan, M.D. UCSD Jan 2012

MAYO CLINIC

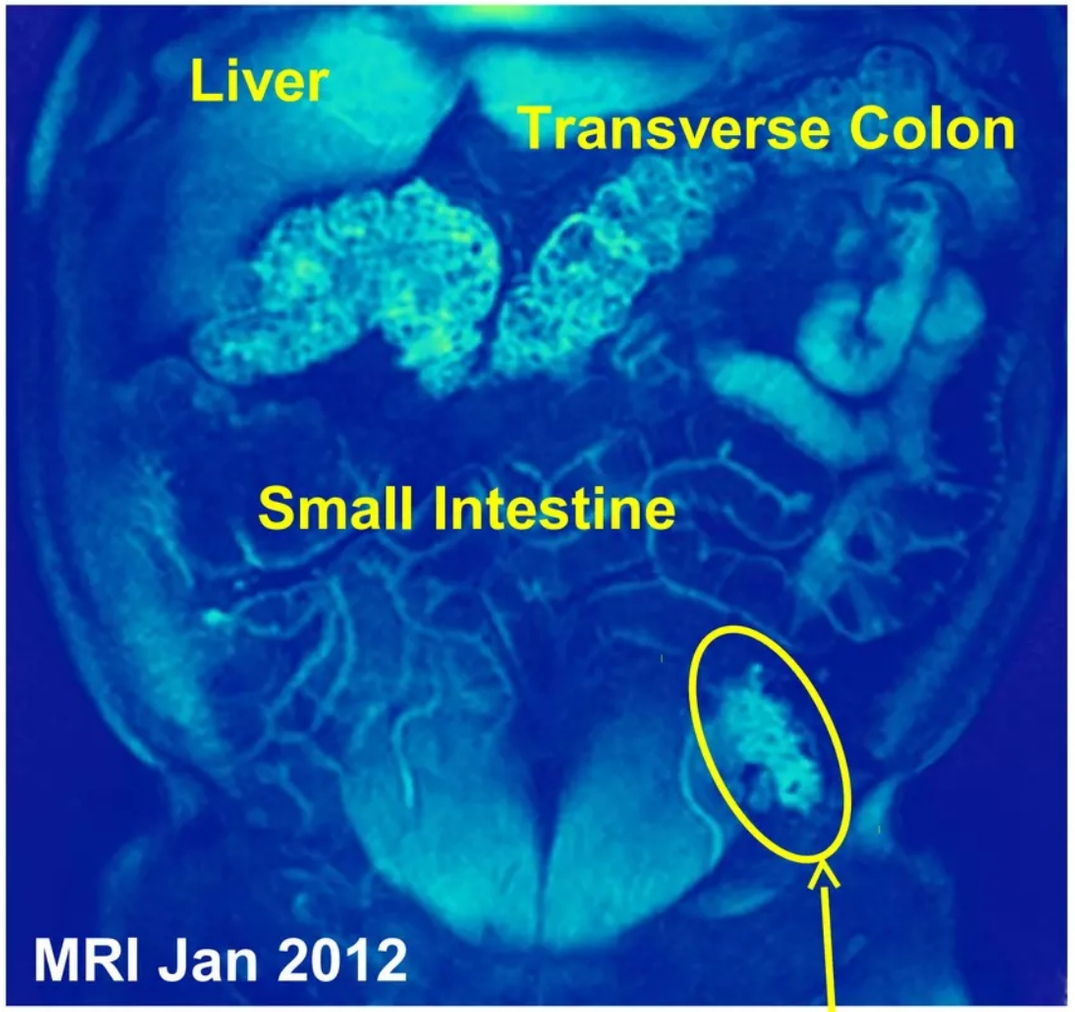
Crohn's disease affects the thickness of the intestinal wall.

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



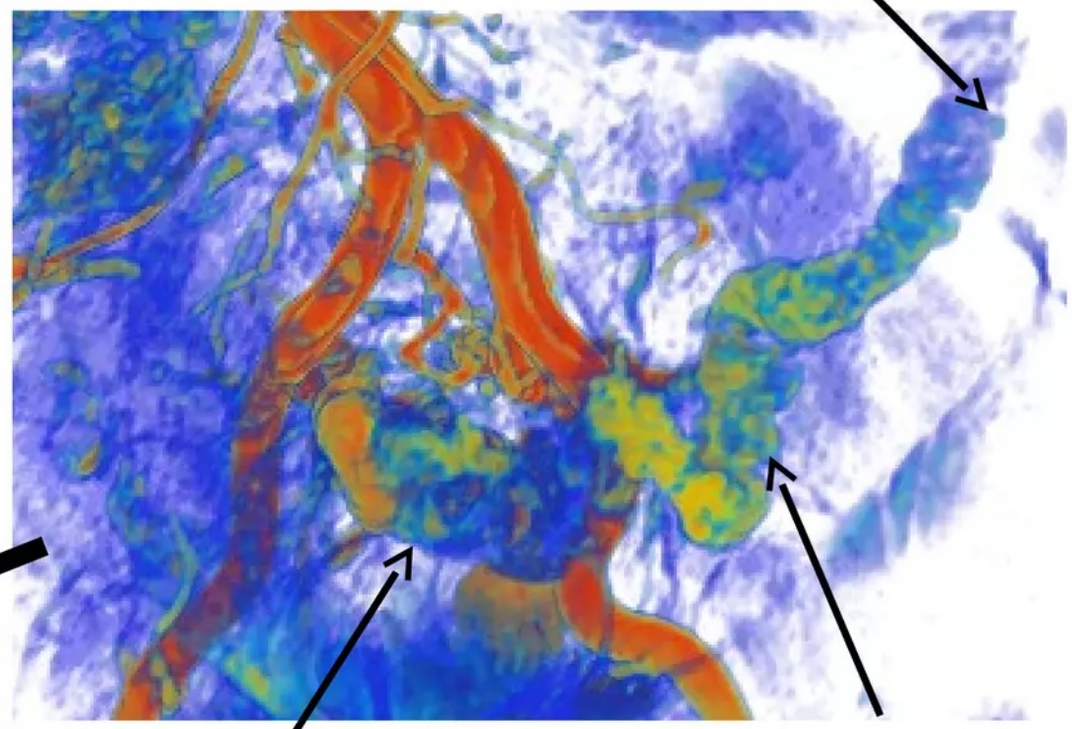
Reveals Inflammation in 6 Inches of Sigmoid Colon
Thickness 15cm – 5x Normal Thickness

Converting MRI Slice Views To Interactive 3D Virtual Reality



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

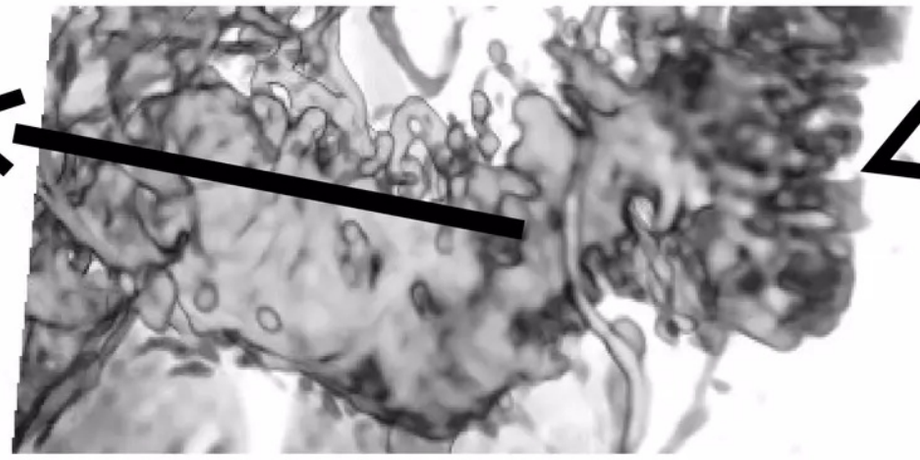
Descending Colon



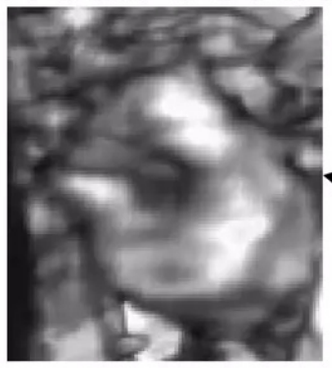
Major Kink

Sigmoid Colon Threading Iliac Arteries

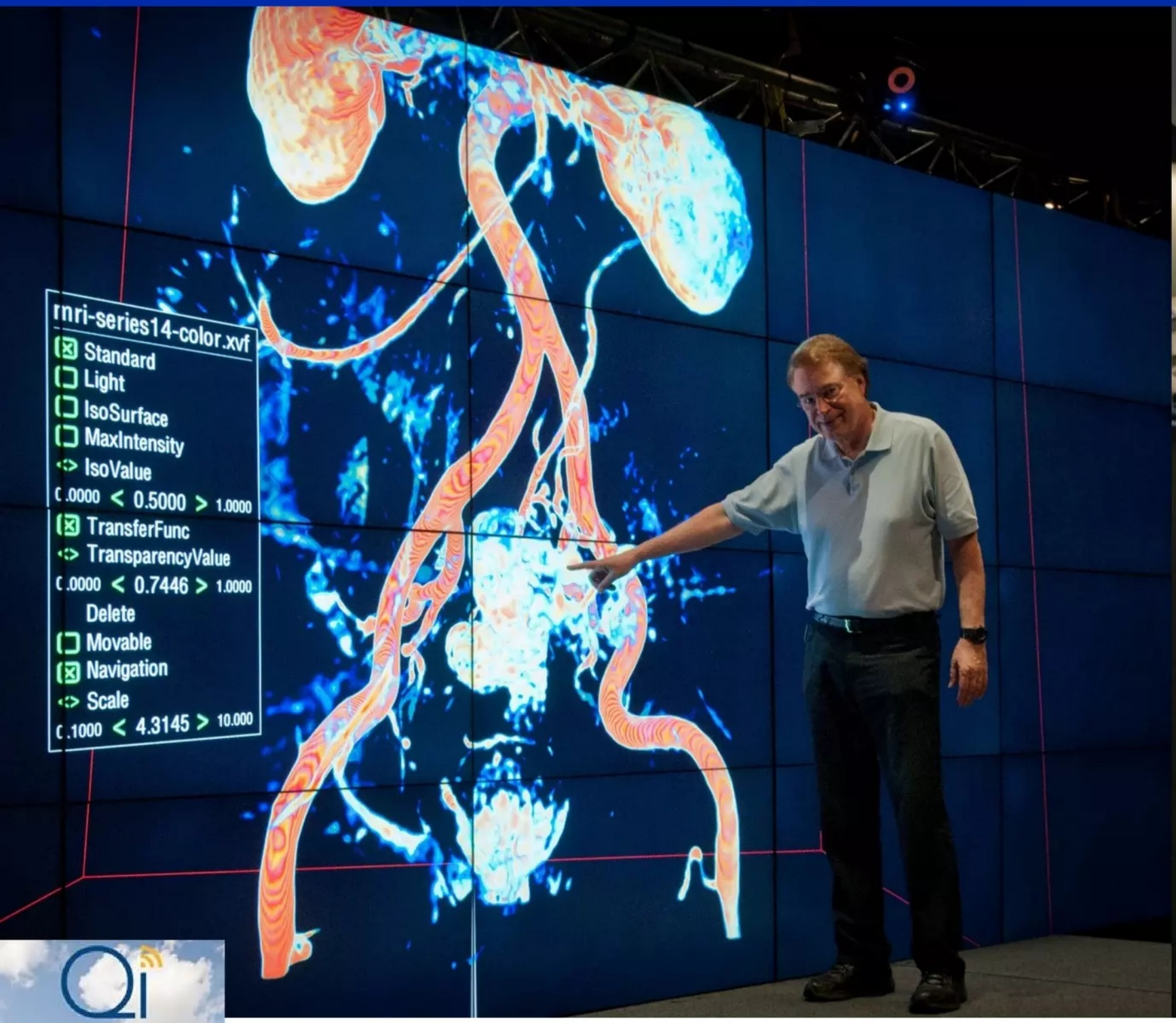
Diseased Sigmoid Colon



Cross Section



Exploring My Anatomy Digitally Enables 3D Printing of the Diseased Organ







Research: Calit2 FutureHealth Team

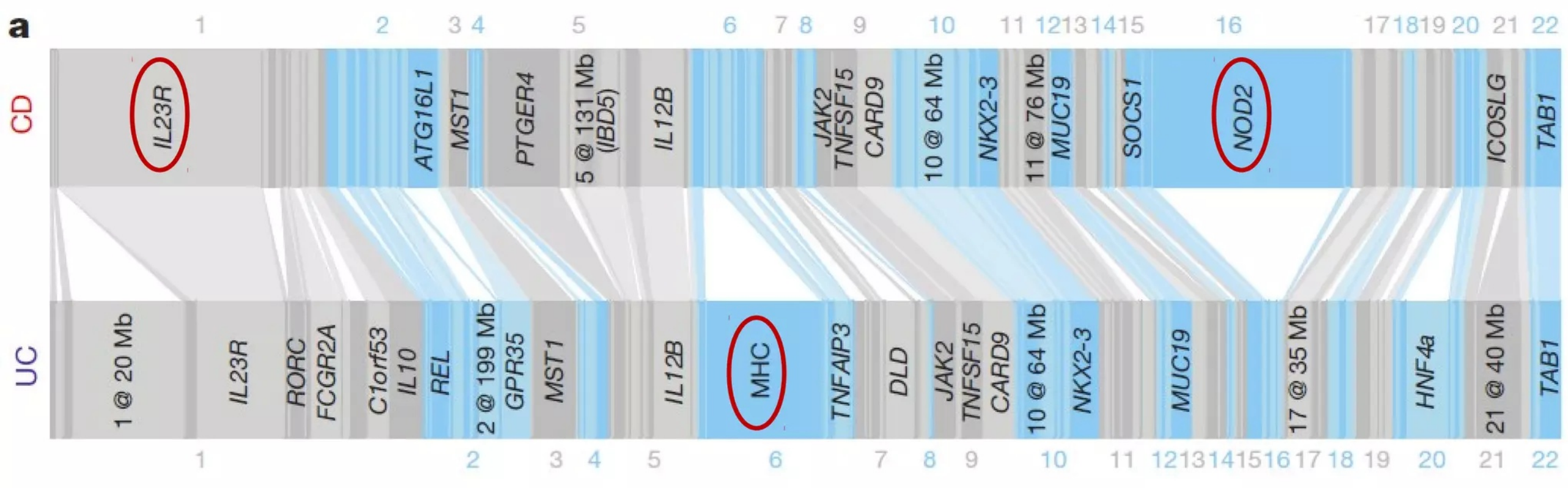
Why Did I Have an Autoimmune Disease like IBD?

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)

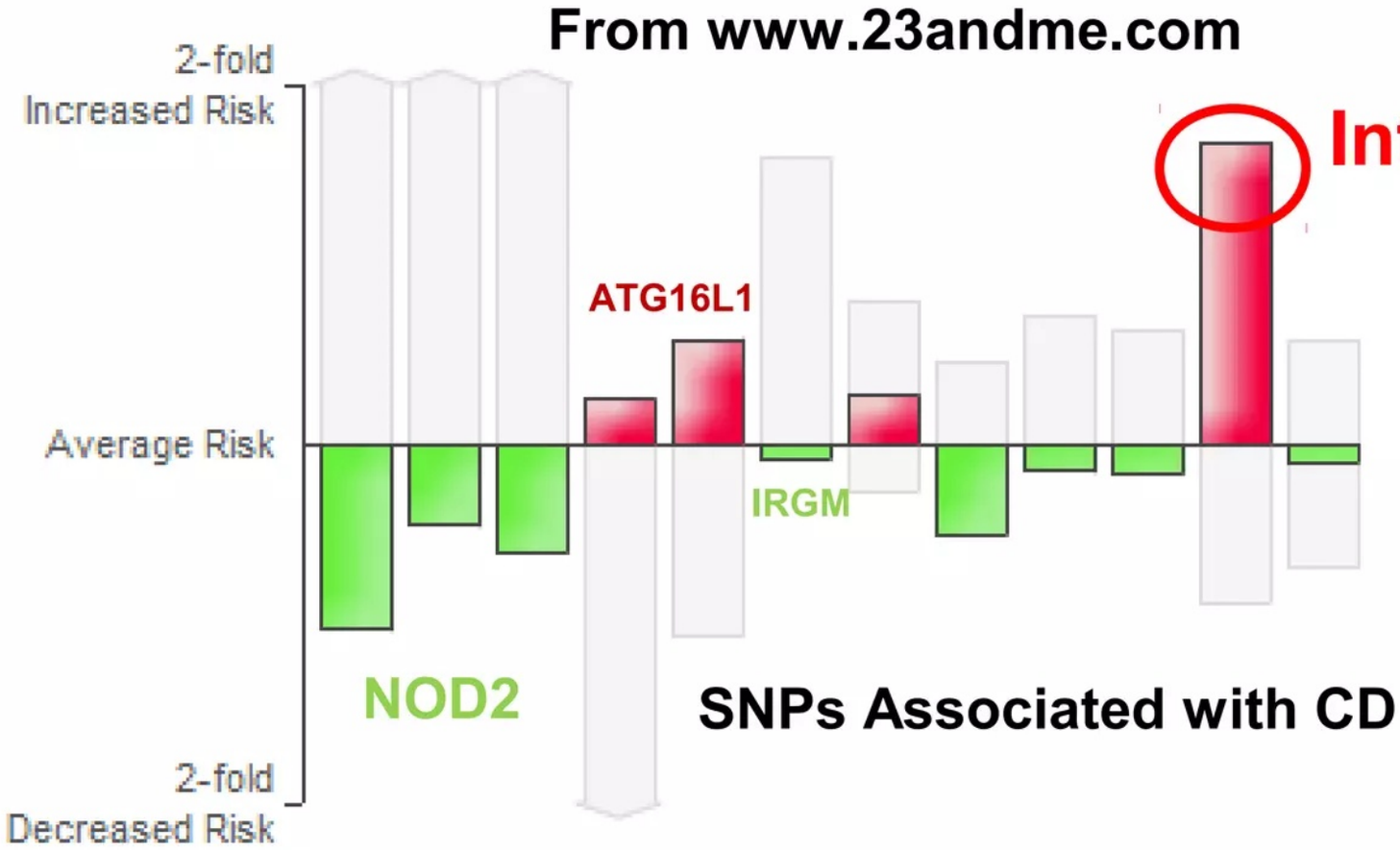
I Compared my 23andme SNPs With the 163 Known SNPs Associated with IBD



- 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)

I Found I Had One of the Earliest Known SNPs Associated with Crohn's Disease



Polymorphism in Interleukin-23 Receptor Gene — 80% Higher Risk of Pro-inflammatory Immune Response rs1004819

News and Views

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

IL-23: a master regulator in Crohn disease

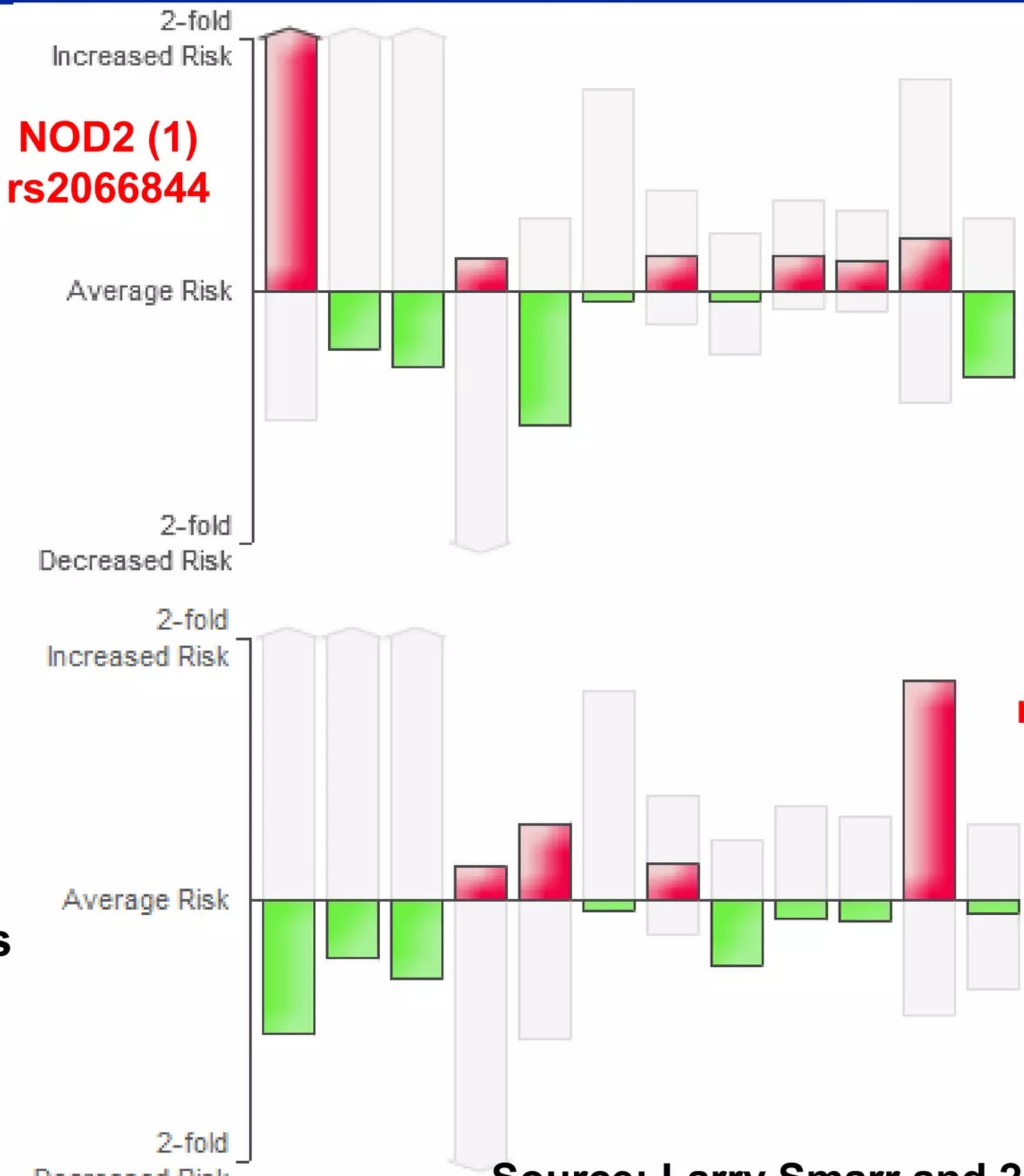
Markus F Neurath[1]

A Genome-Wide Association Study Identifies *IL23R* as an Inflammatory Bowel Disease Gene

Richard H. Düerr, *et al.*
Science **314**, 1461 (2006)

There Is Likely a Correlation Between CD SNPs and Where and When the Disease Manifests

Subject with Ileal Crohn's



NOD2 (1)
rs2066844

Female
CD Onset
At 20-Years Old

Subject with Colon Crohn's

Il-23R
rs1004819

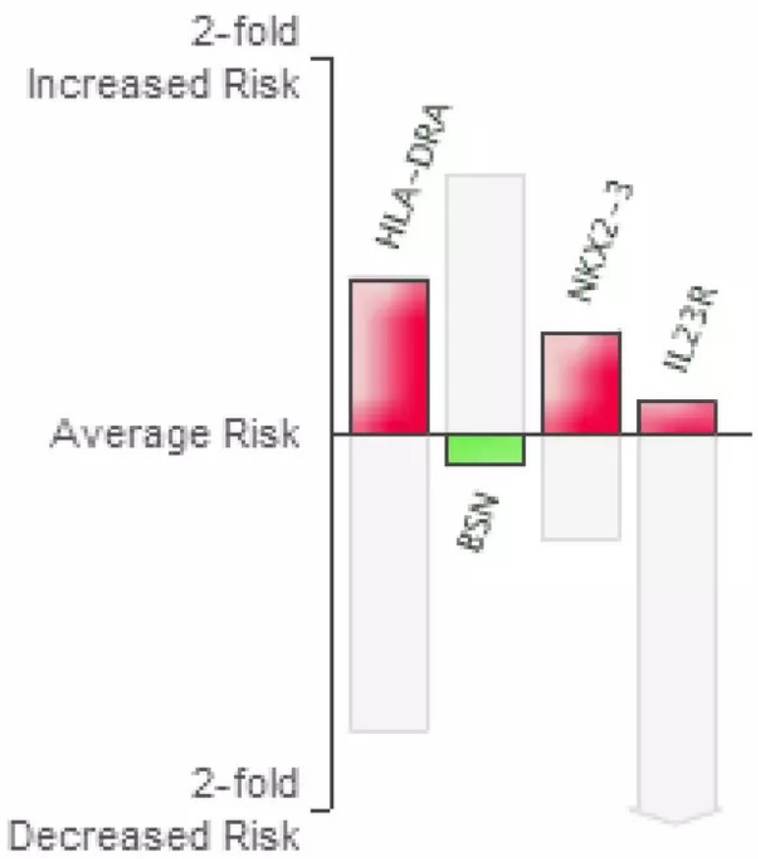
Me-Male
CD Onset
At 60-Years Old

Source: Larry Smarr and 23andme

I Also Had an Increased Risk for Ulcerative Colitis, But a SNP that is Also Associated with Colonic CD



I Have a 33% Increased Risk for Ulcerative Colitis HLA-DRA (rs2395185)

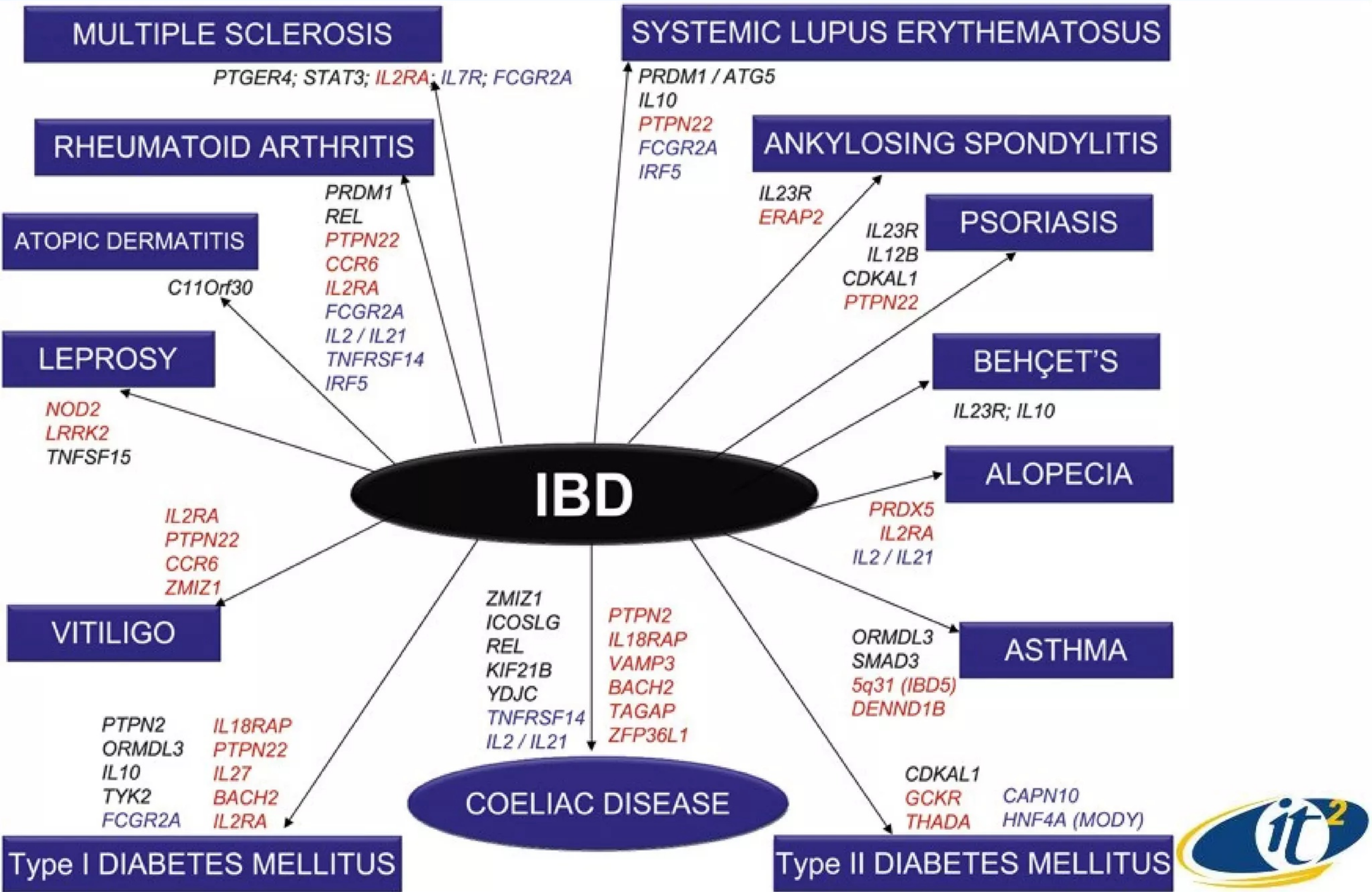


I Have the Same Level of HLA-DRA Increased Risk as Another Male Who Has Had Ulcerative Colitis for 20 Years

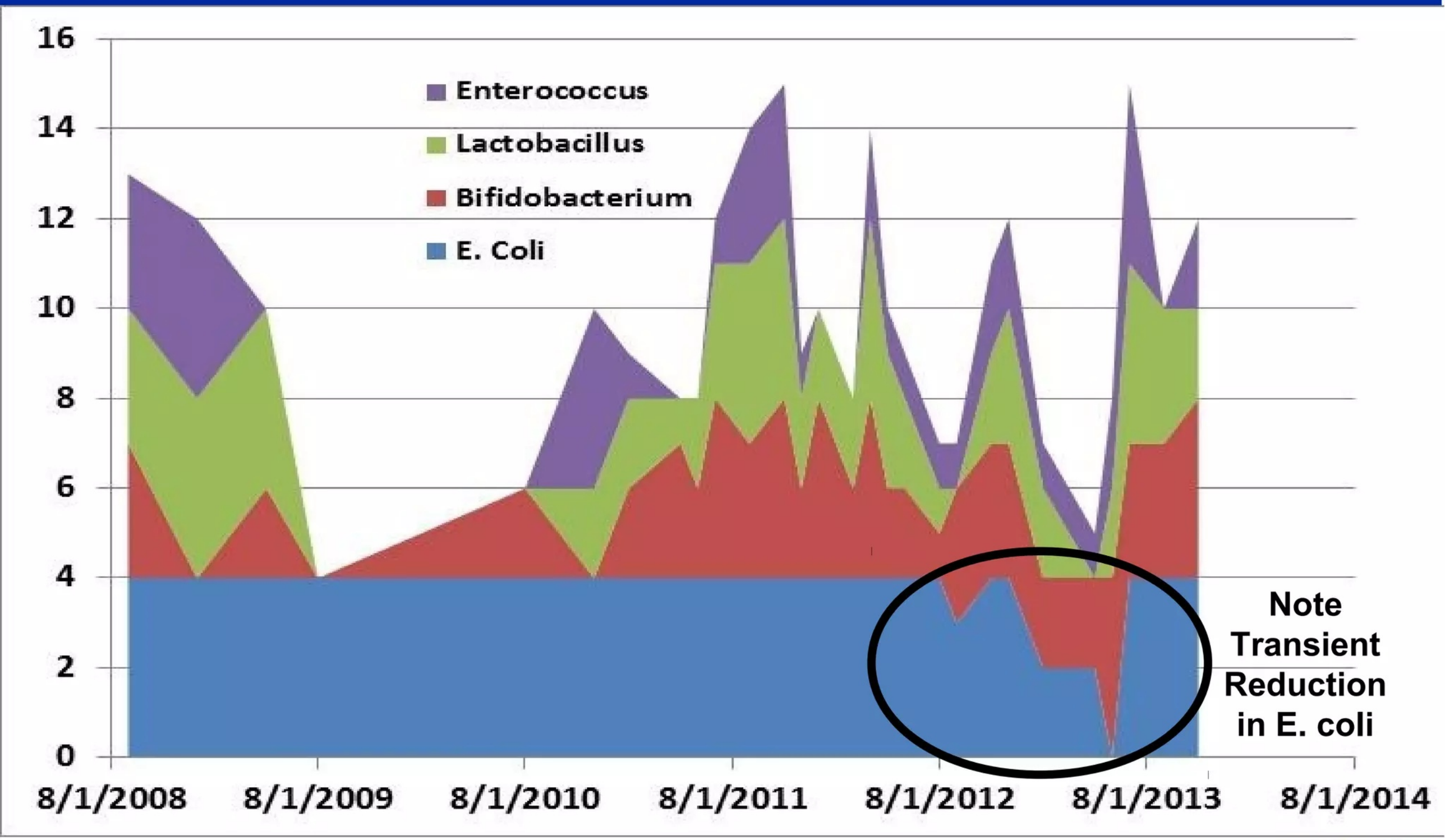
"Our results suggest that at least for the SNPs investigated [including HLA-DRA], colonic CD and UC have common genetic basis."
-Waterman, et al., IBD 17, 1936-42 (2011)

Autoimmune Disease Overlap from SNP GWAS

Gut Lees, et al. 60:1739-1753 (2011)

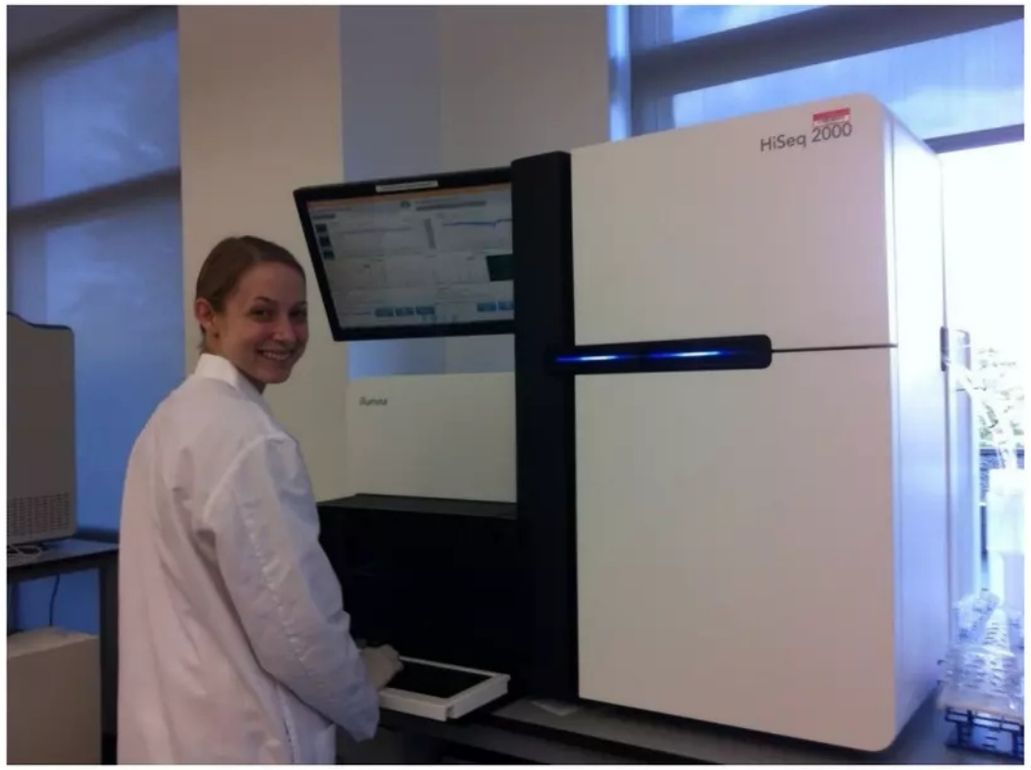


LS Cultured Bacterial Abundance Reveals Oscillations As Well



To Map Out the Dynamics of My Microbiome Ecology I Partnered with the J. Craig Venter Institute

- **JCVI Did Metagenomic Sequencing on Six of My Stool Samples Over 1.5 Years**
- **Sequencing on Illumina HiSeq 2000**
  - **Generates 100bp Reads**
  - **Run Takes ~14 Days**
  - **My 6 Samples Produced**
    - **190.2 Gbp of Data**
- **JCVI Lab Manager, Genomic Medicine**
  - **Manolito Torralba**
- **IRB PI Karen Nelson**
  - **President JCVI**



Illumina HiSeq 2000 at JCVI



Manolito Torralba, JCVI



Karen Nelson, JCVI

J. Craig Venter INSTITUTE

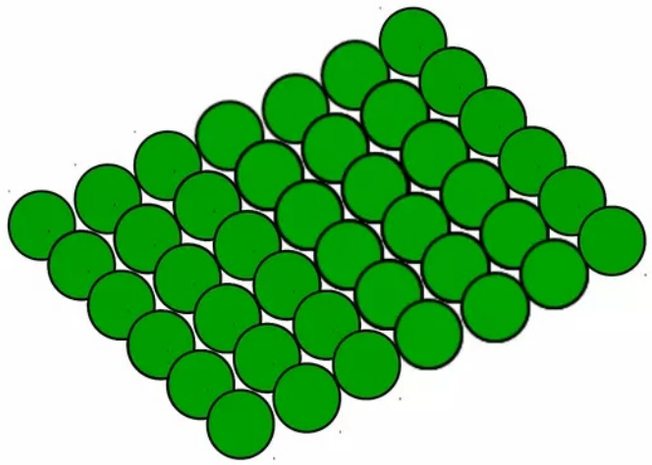
it²

We Downloaded Additional Phenotypes from NIH HMP For Comparative Analysis

Download Raw Reads
~100M Per Person

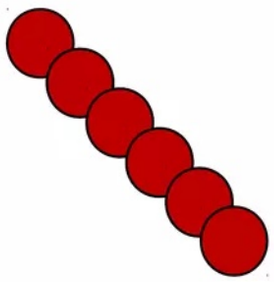
"Healthy" Individuals

35 Subjects
1 Point in Time



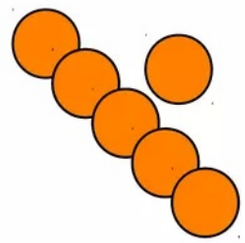
Larry Smarr

6 Points in Time

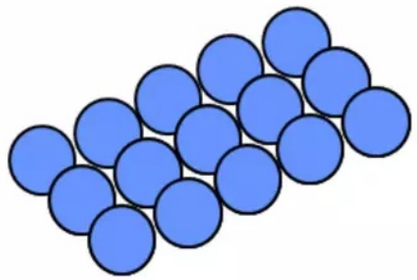


IBD Patients

2 Ulcerative Colitis Patients,
6 Points in Time



5 Ileal Crohn's Patients,
3 Points in Time



NIH

HMP

Total of 5 Billion Reads

Source: Jerry Sheehan, Calit2
Weizhong Li, Sitao Wu, CRBS, UCSD

it²

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**

NCBI
National Center for Biotechnology Information

HMP

Now to Align Our 5 Billion Reads Against the Reference Database

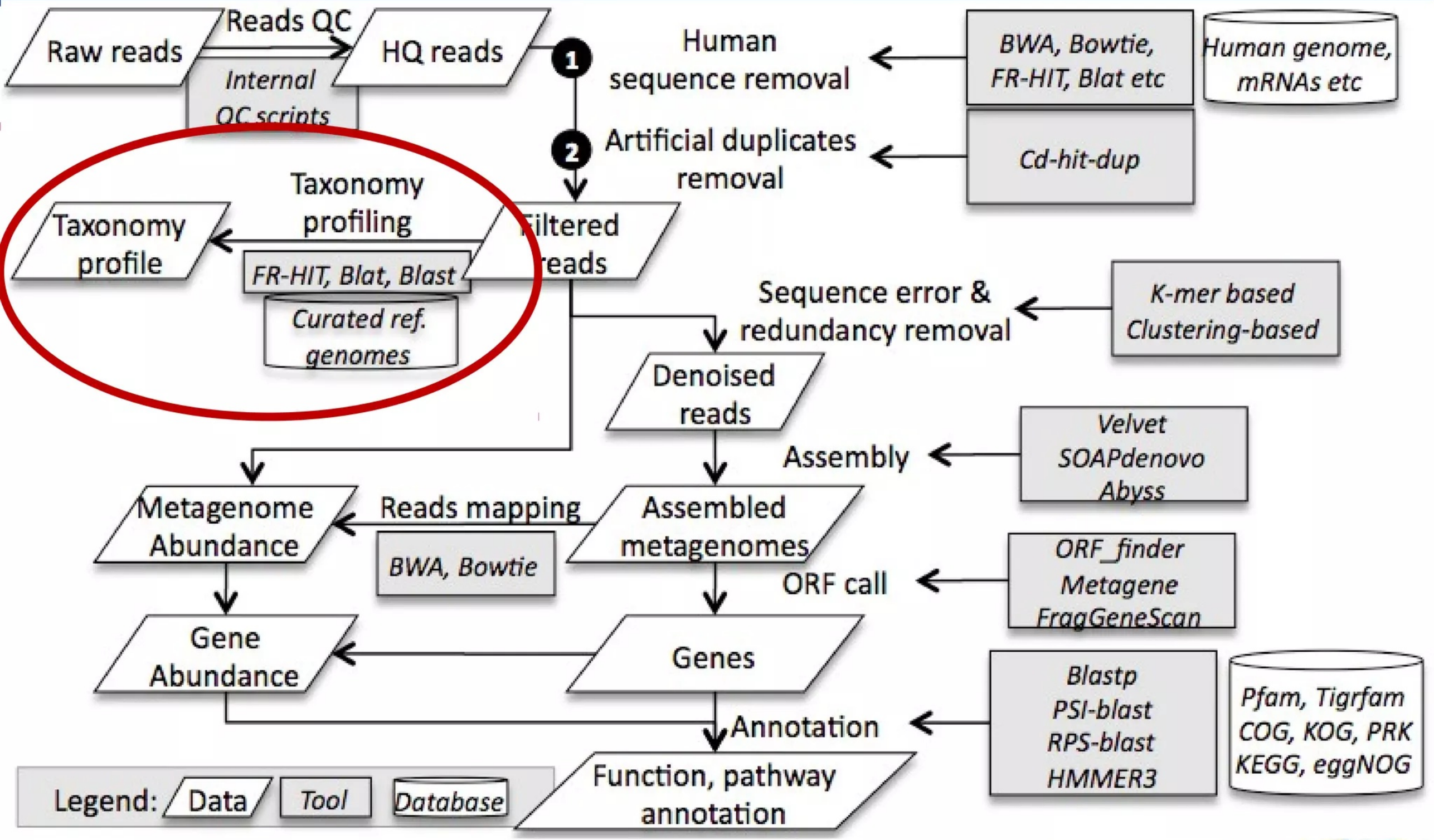
NIH

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

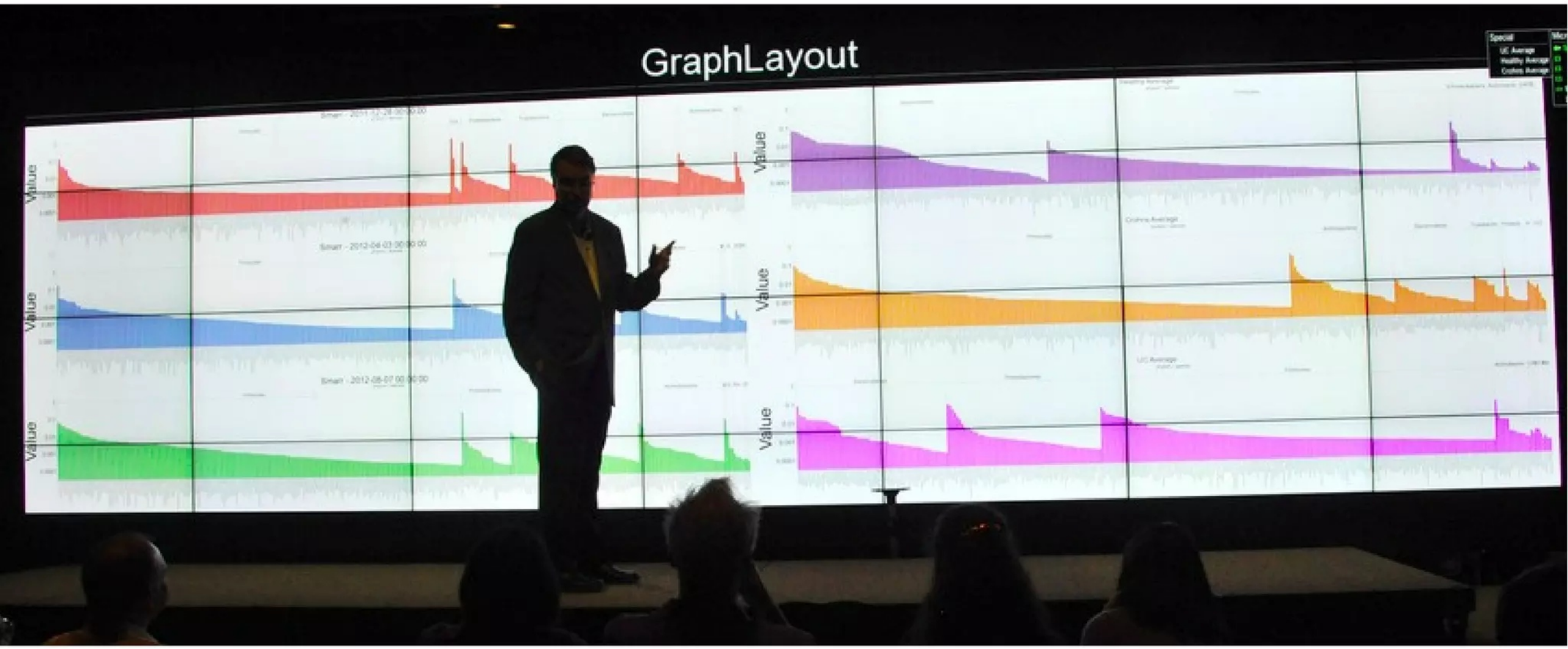
Enabled by a Grant of Time on Gordon from SDSC Director Mike Norman

NSF NIH HMP

J. Craig Venter INSTITUTE

CRBS it²

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

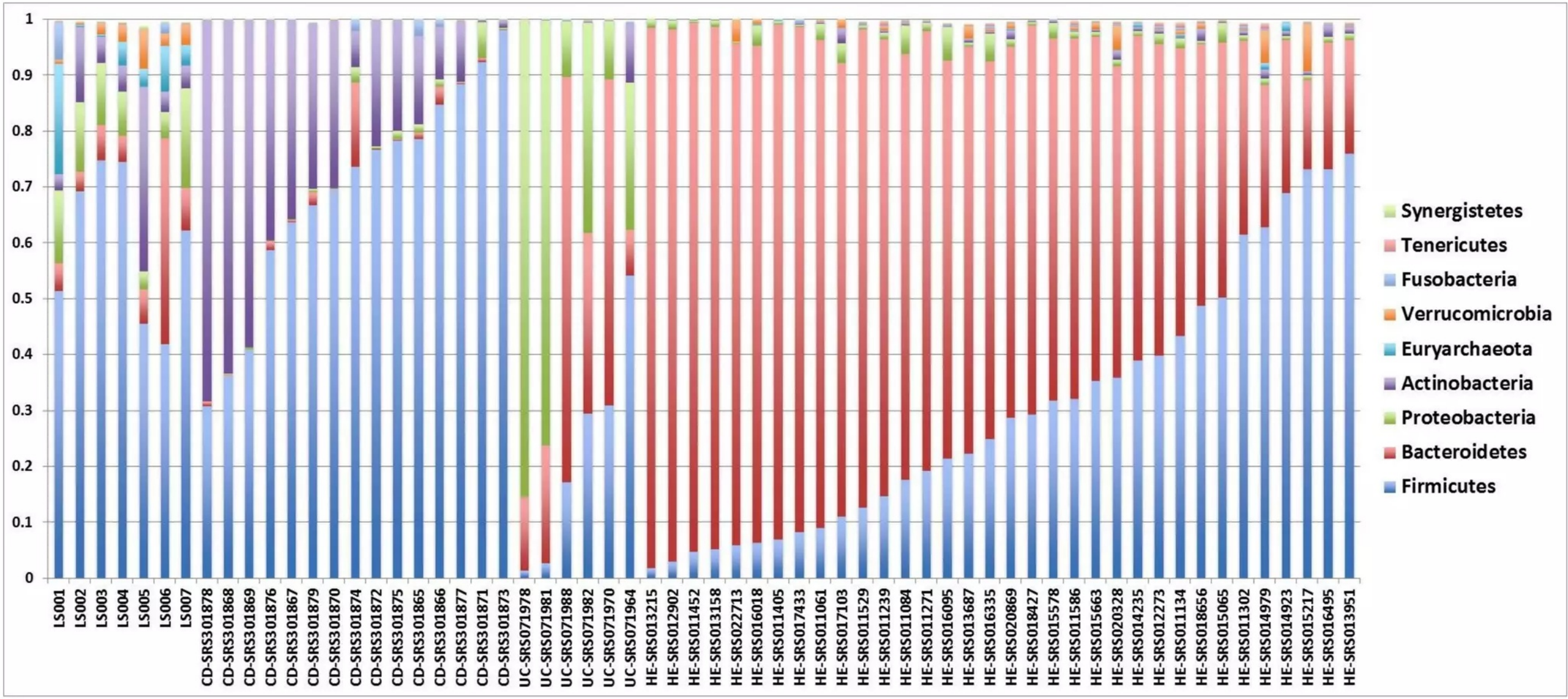


Comparing 3 LS Time Snapshots (Left)
with Healthy, Crohn's, UC (Right Top to Bottom)

Calit2 VROOM-FuturePatient Expedition

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

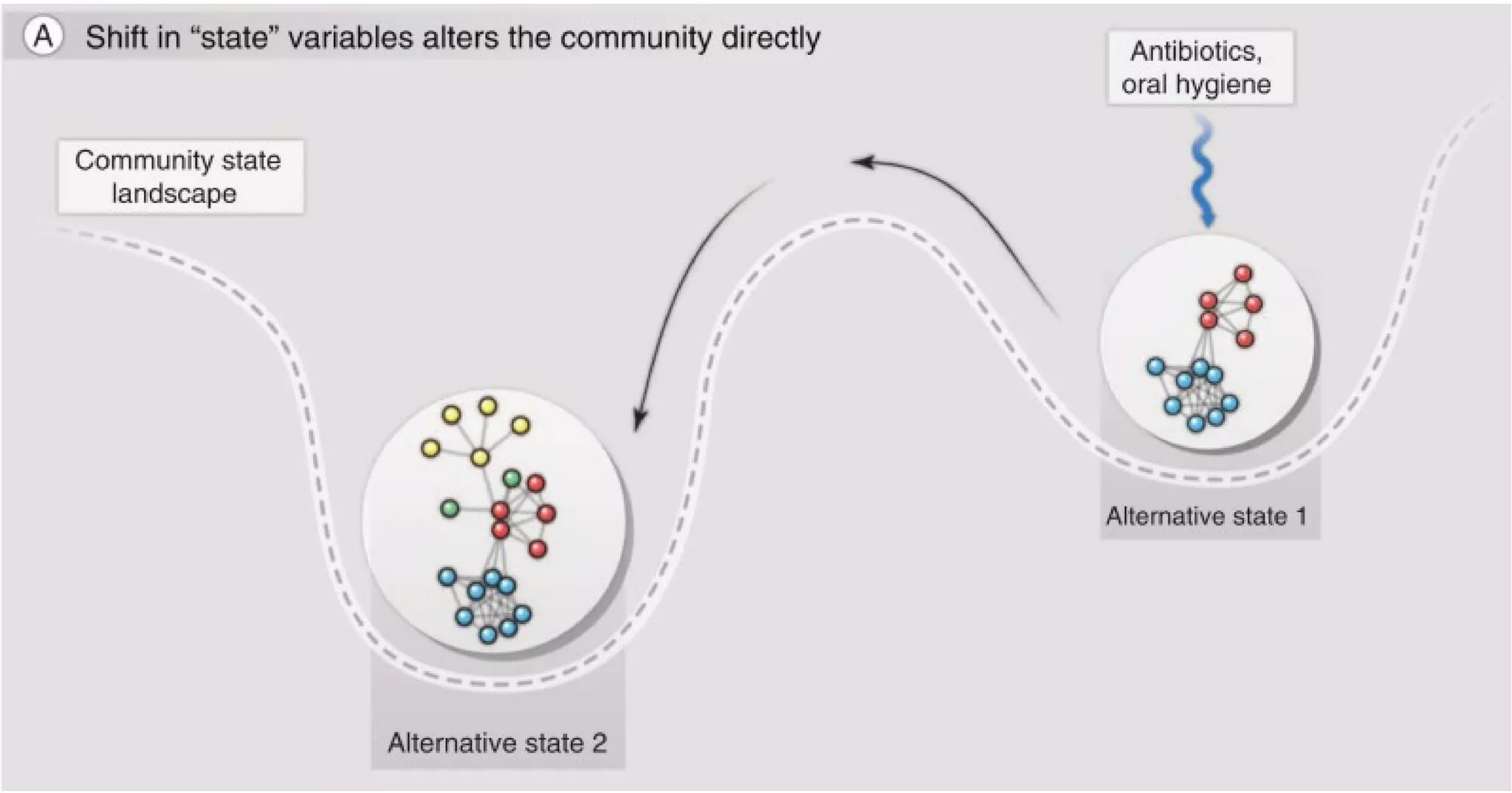
Source: Weizhong Li, Sitao Wu, CRBS, UCSD



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

Toward Noninvasive Microbial Ecology Diagnostics

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



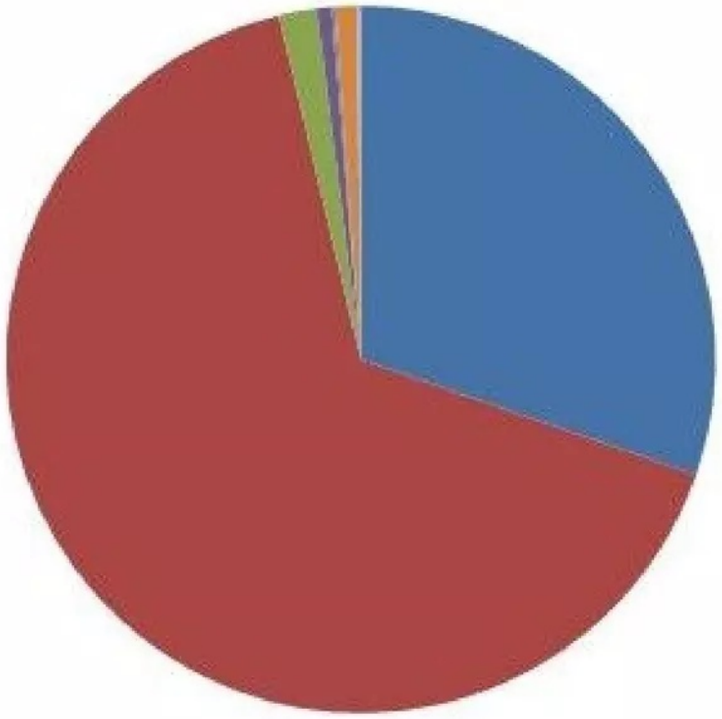
"The Application of Ecological Theory Toward an Understanding of the Human Microbiome,"
Elizabeth Costello, Keaton Stagaman, Les Dethlefsen, Brendan Bohannan, David Relman
Science 336, 1255-62 (2012)

Multiple Microbial Equilibrium Revealed by Comparing 35 Healthy to 15 CD and 6 UC Gut Microbiomes

Microbial Phyla

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

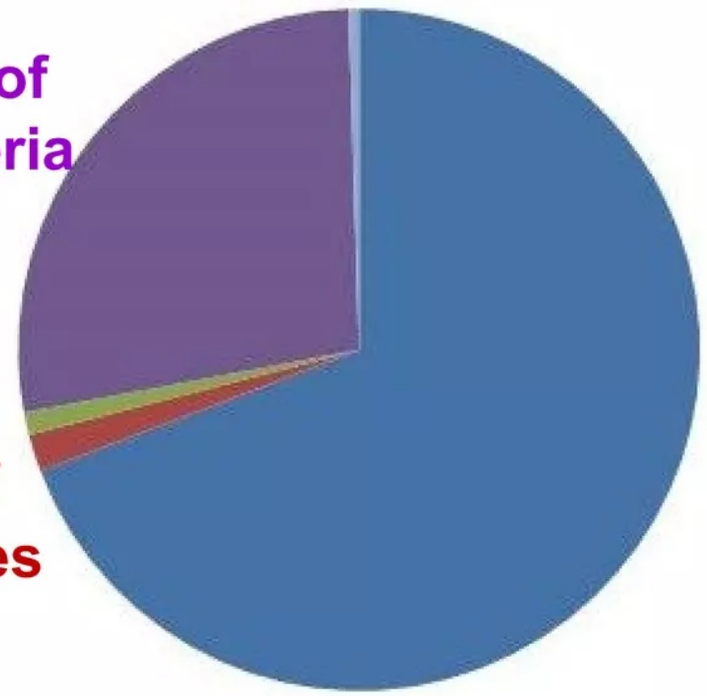
HE Ave.



CD Ave.

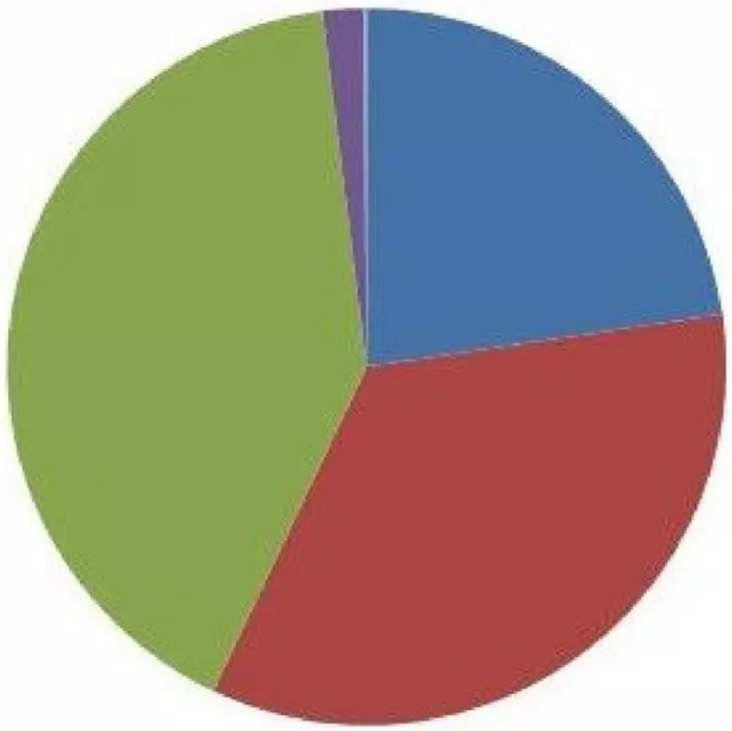
Expansion of Actinobacteria

Collapse of Bacteroidetes



UC Ave.

Explosion of Proteobacteria



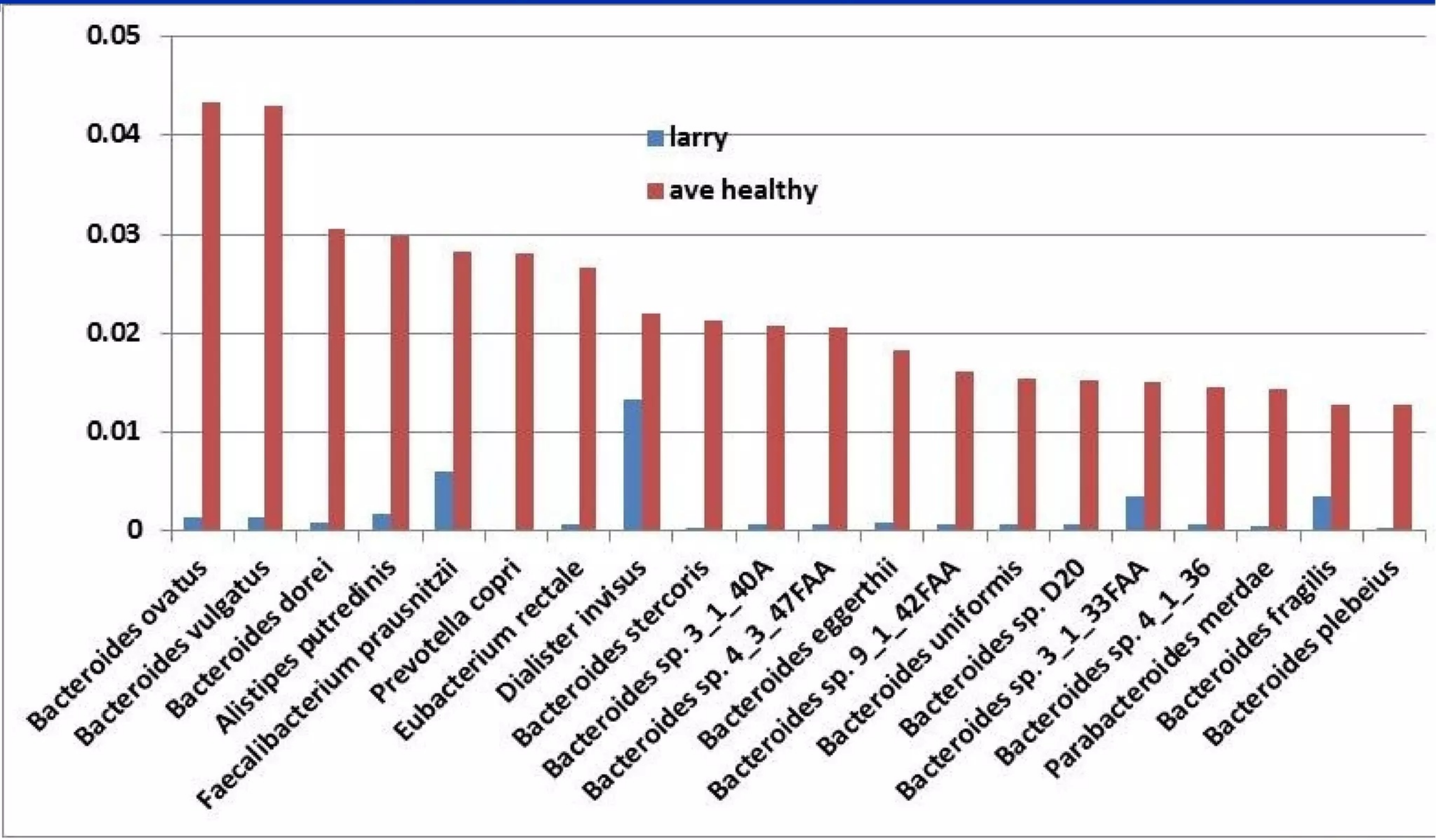
Lessons From Ecological Dynamics II:
Invasive Species Dominate After Major Species Destroyed



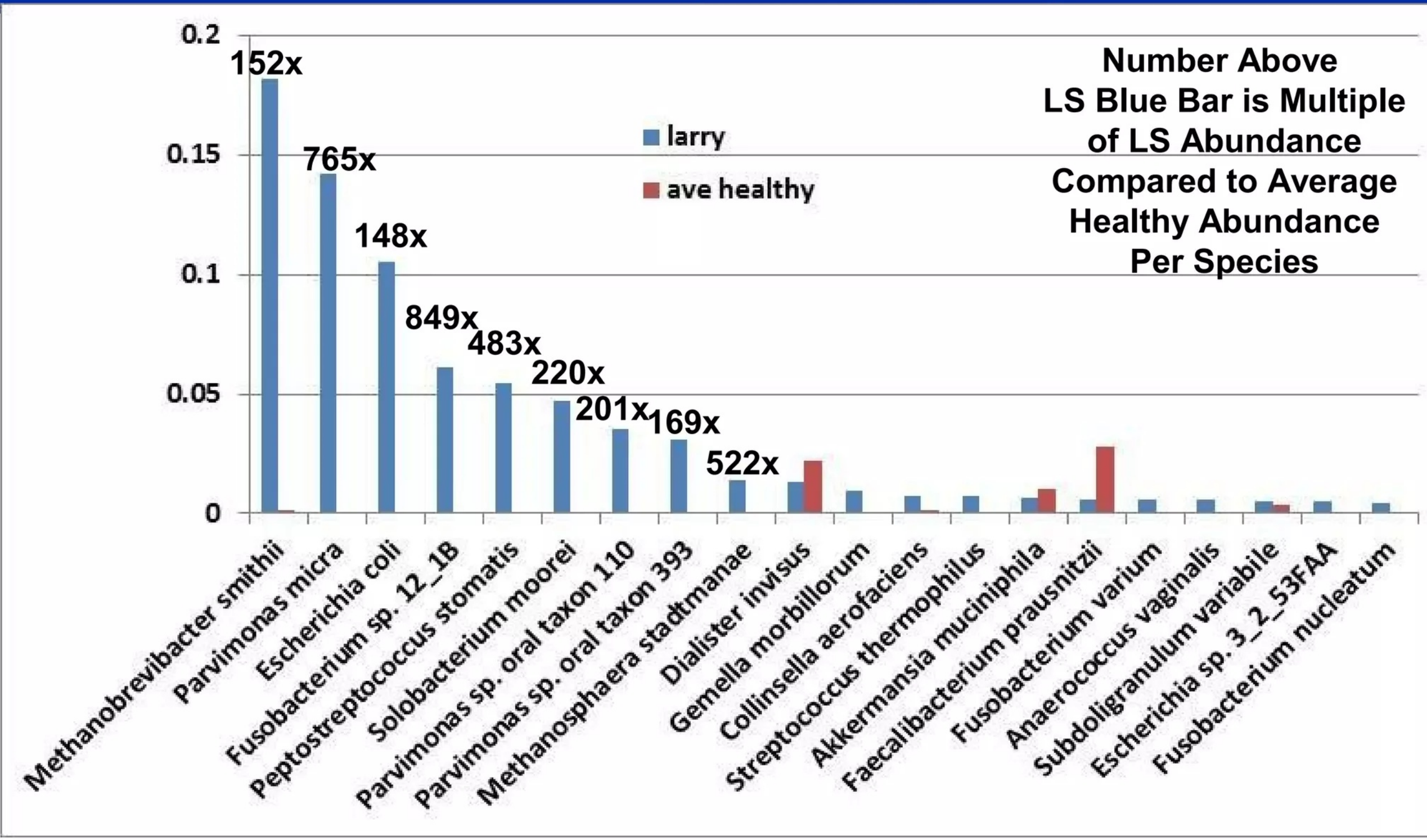
”In many areas following these burns
invasive species are able to establish themselves,
crowding out native species.”

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

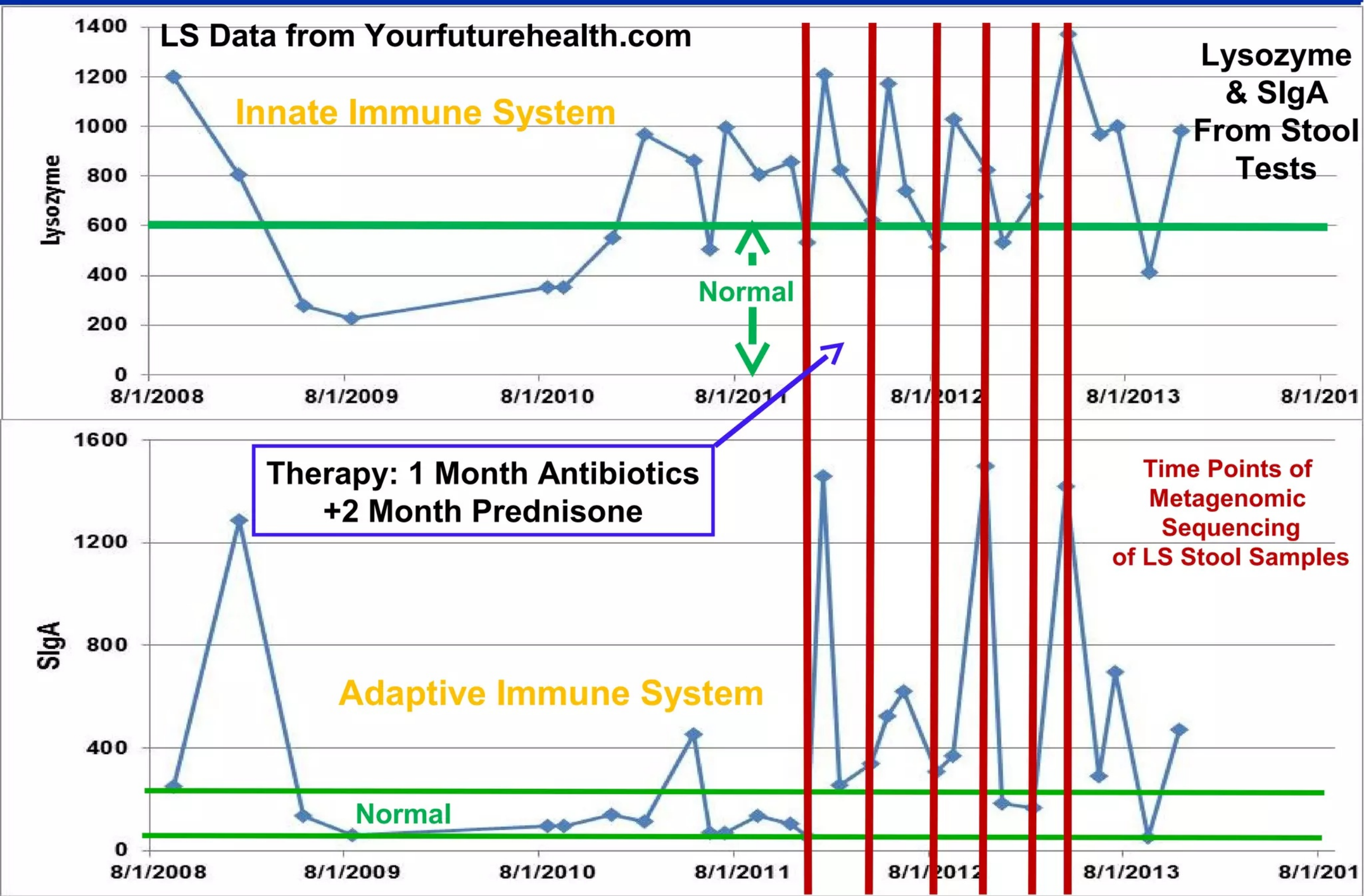


Top 20 Most Abundant Microbial Species In LS vs. Average Healthy Subject

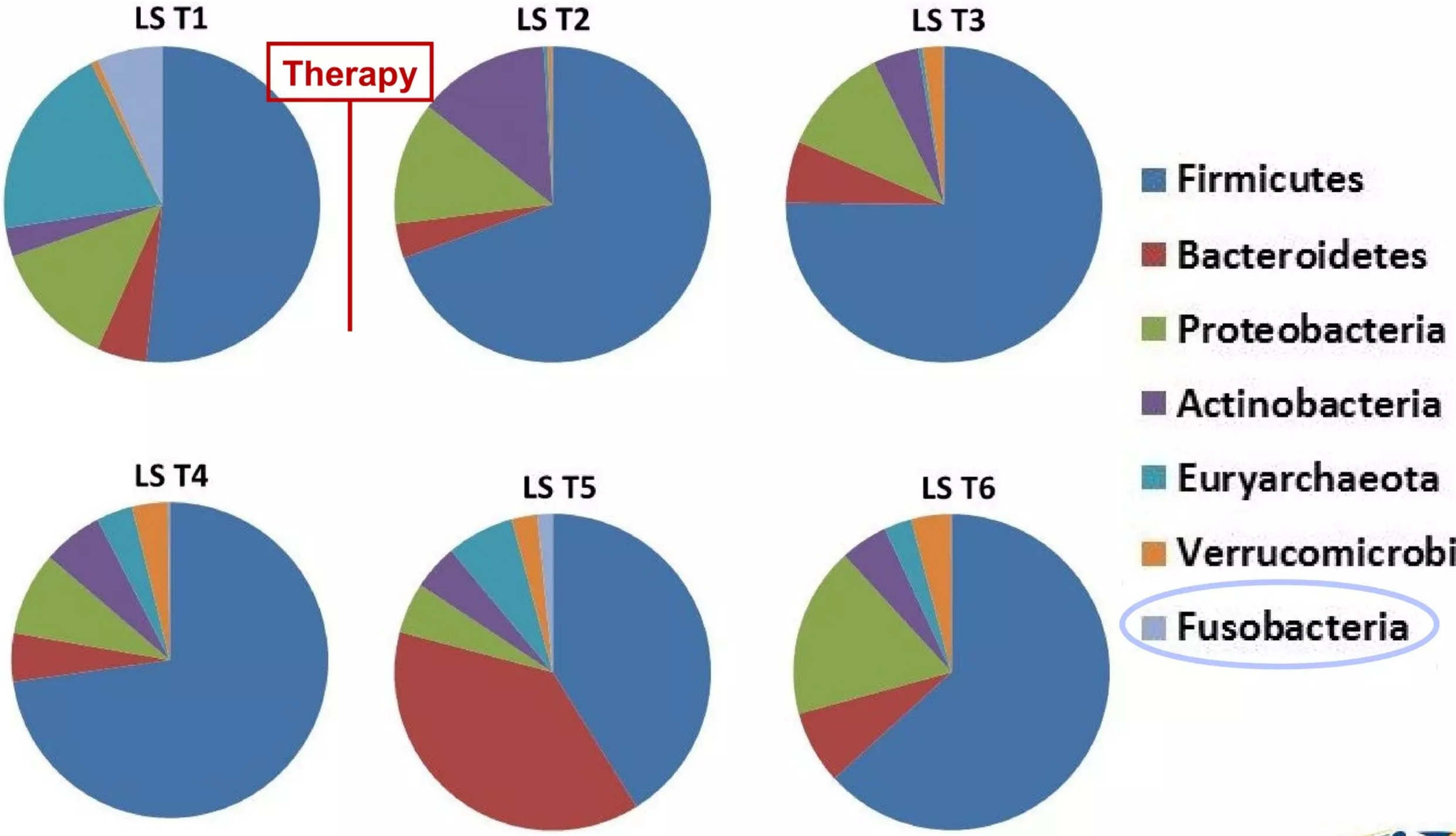


Source: Sequencing JCVI; Analysis Weizhong Li, UCSD
LS December 28, 2011 Stool Sample

Comparing Changes in Gut Microbiome Ecology with Oscillations of the Innate and Adaptive Immune System



Time Series Reveals Autoimmune Dynamics of Gut Microbiome by Phyla



Six Metagenomic Time Samples Over 16 Months

Fusobacteria Are Found To Be More Abundant In Colonrectal Carcinoma (CRC) Tissue

Two Cancer Studies Find Bacterial Clue in Colon

The New York Times

By GINA KOLATA
Published: October 17, 2011

Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma

Aleksandar D. Kostic, **et al.**

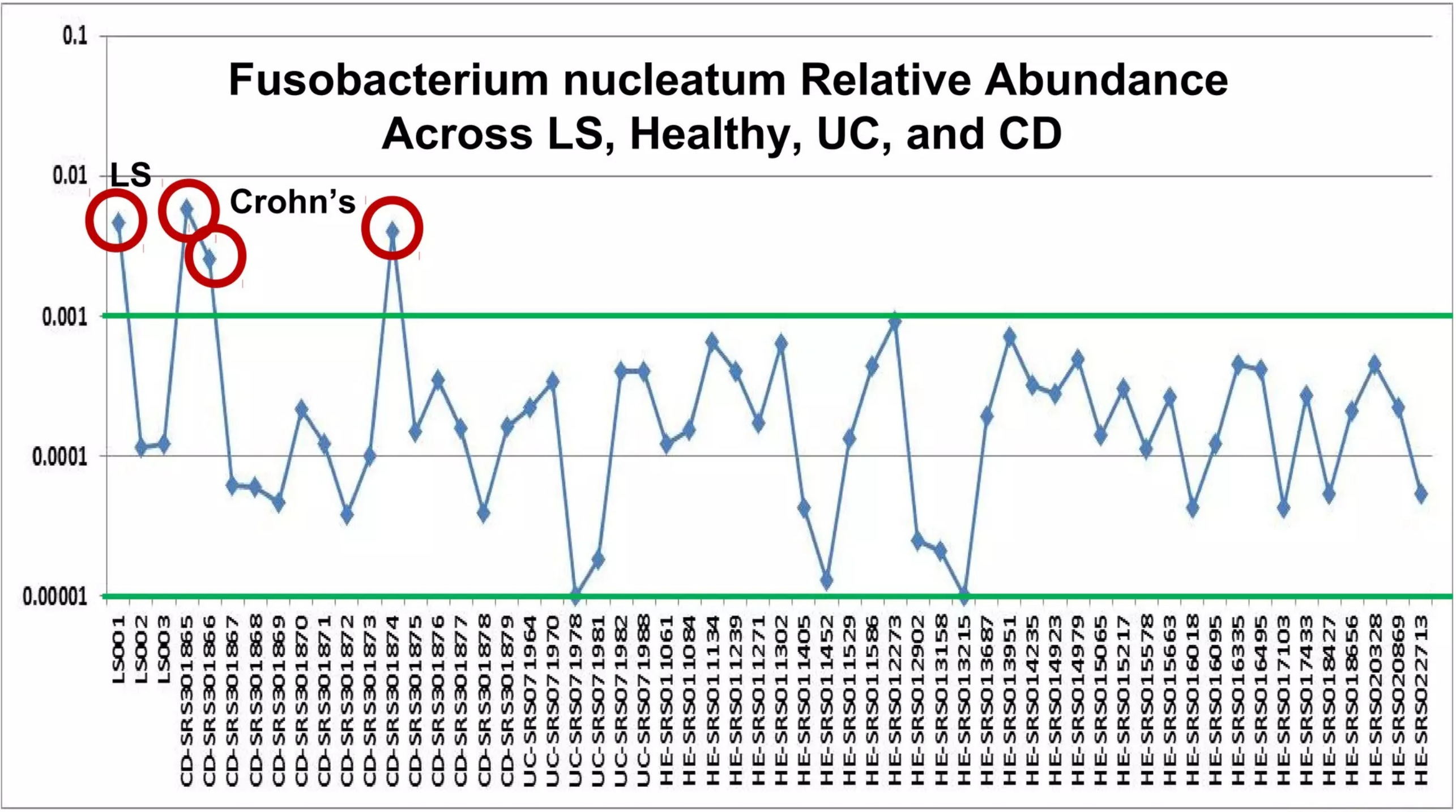
Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma

Mauro Castellarin, **et al.**

CSH PRESS

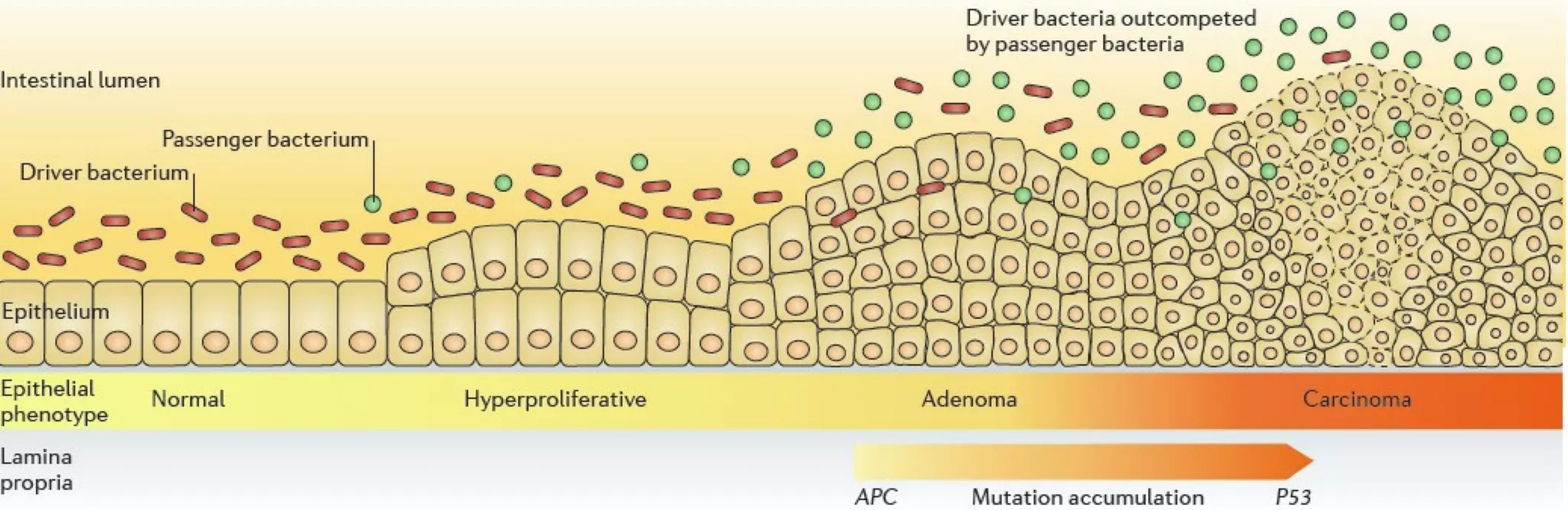
GENOME RESEARCH

Could the Presence of Fusobacterium Nucleatum Be an Early Indicator of a Transition to CRC?



The Bacterial Driver-Passenger Model for Colorectal Cancer Initiation

Is Fusobacterium nucleatum a “Driver” or a “Passenger”

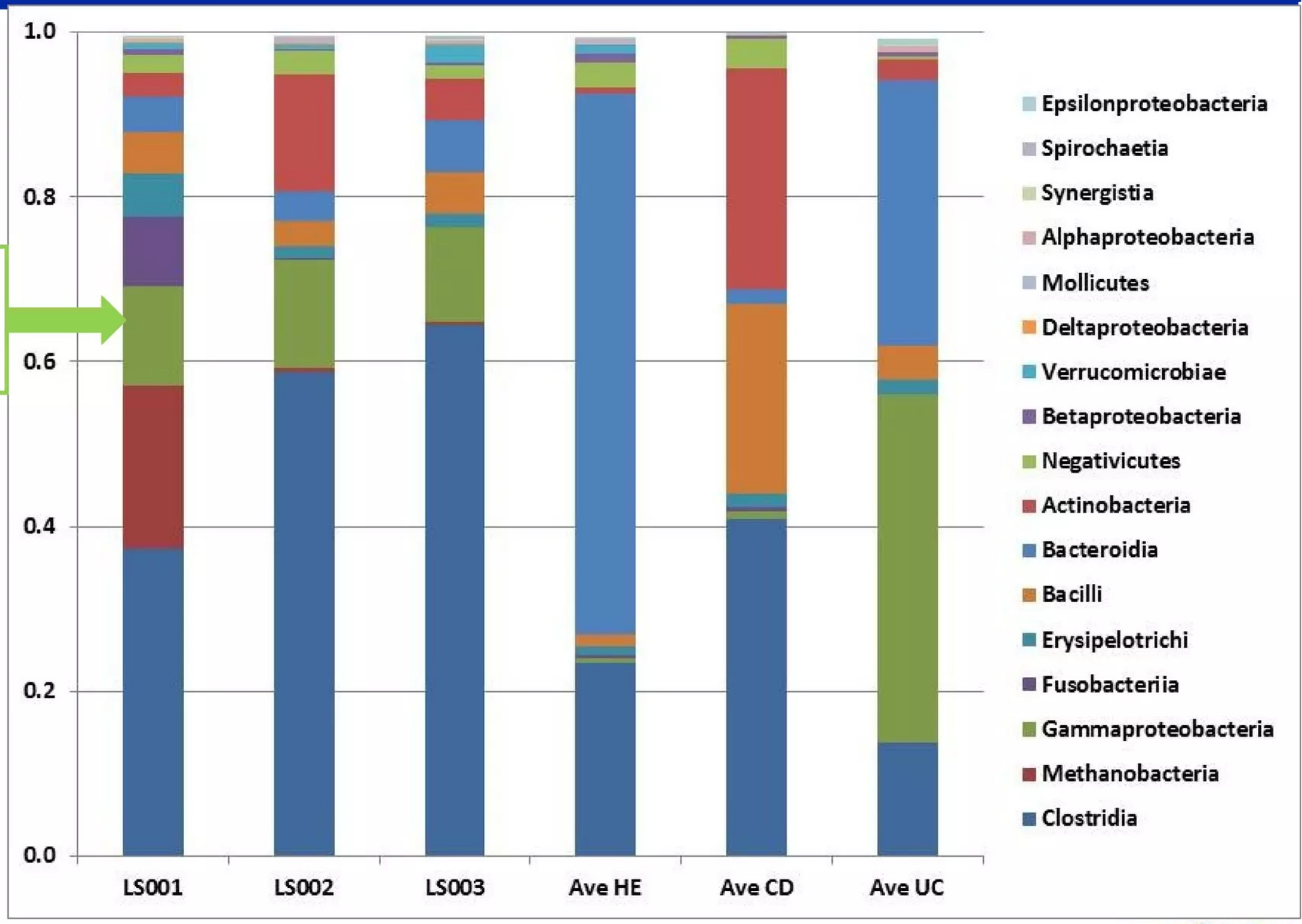


“Early detection of Colorectal Cancer (CRC) is one of the greatest challenges in the battle against this disease & the establishment of a CRC-associated microbiome risk profile could aid in the early identification of individuals who are at high risk and require strict surveillance.”

Tjalsma, et al. Nature Reviews Microbiology v. 10, 575-582 (2012)

LS Time Series Gut Microbiome Classes vs. Healthy, Crohn's, Ulcerative Colitis

Class Gamma-proteobacteria



E. coli claims the driving seat for cancer

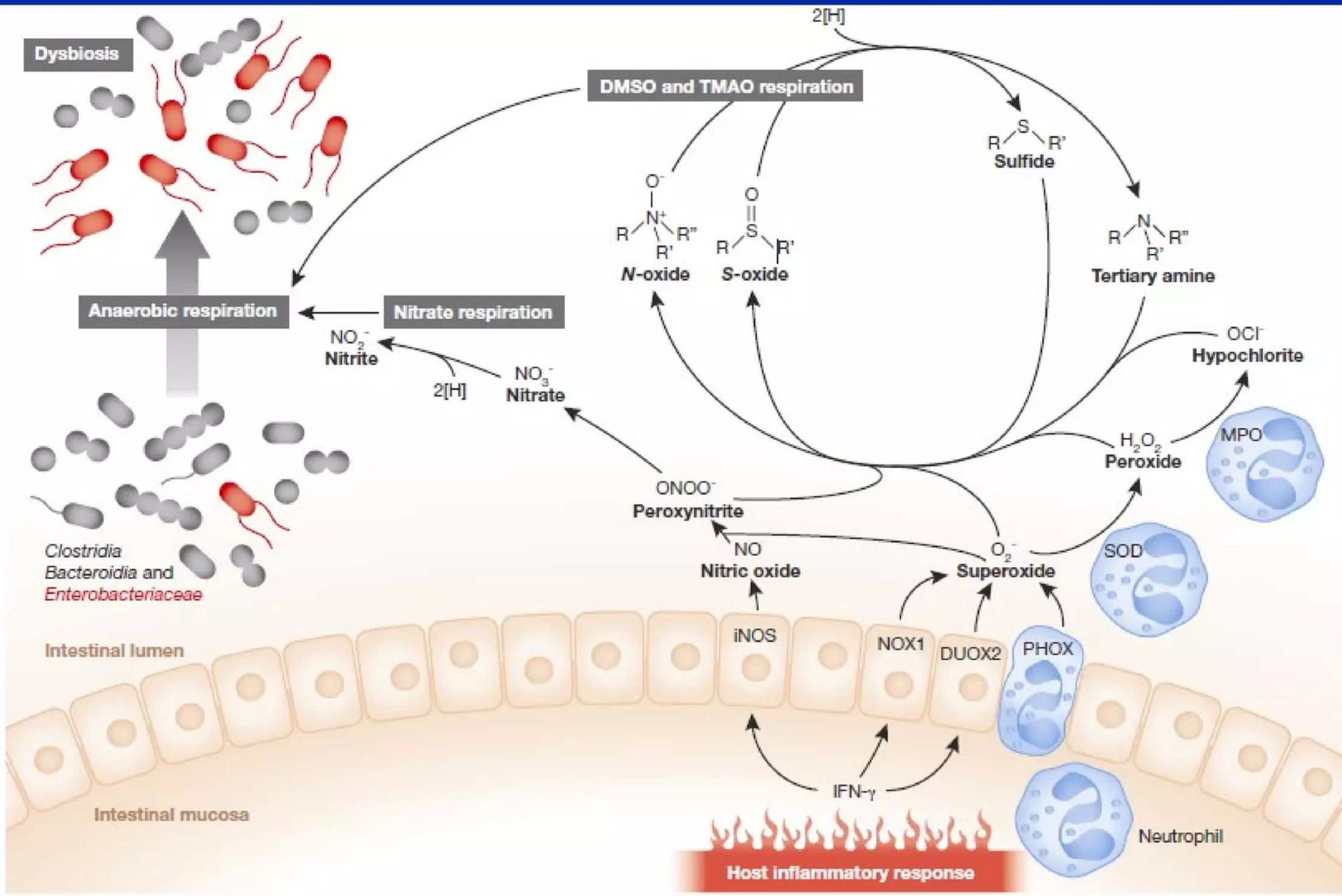
"Arthur *et al.* provide evidence that inflammation
alters the intestinal microbiota
by favouring the proliferation of genotoxic commensals,
and that the *Escherichia coli*
genotoxin colibactin promotes colorectal cancer (CRC)."



Christina Tobin Kåhrström
Associate Editor,
Nature Reviews Microbiology

VOLUME 12 | OCTOBER 2012
NATURE REVIEWS | CANCER

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



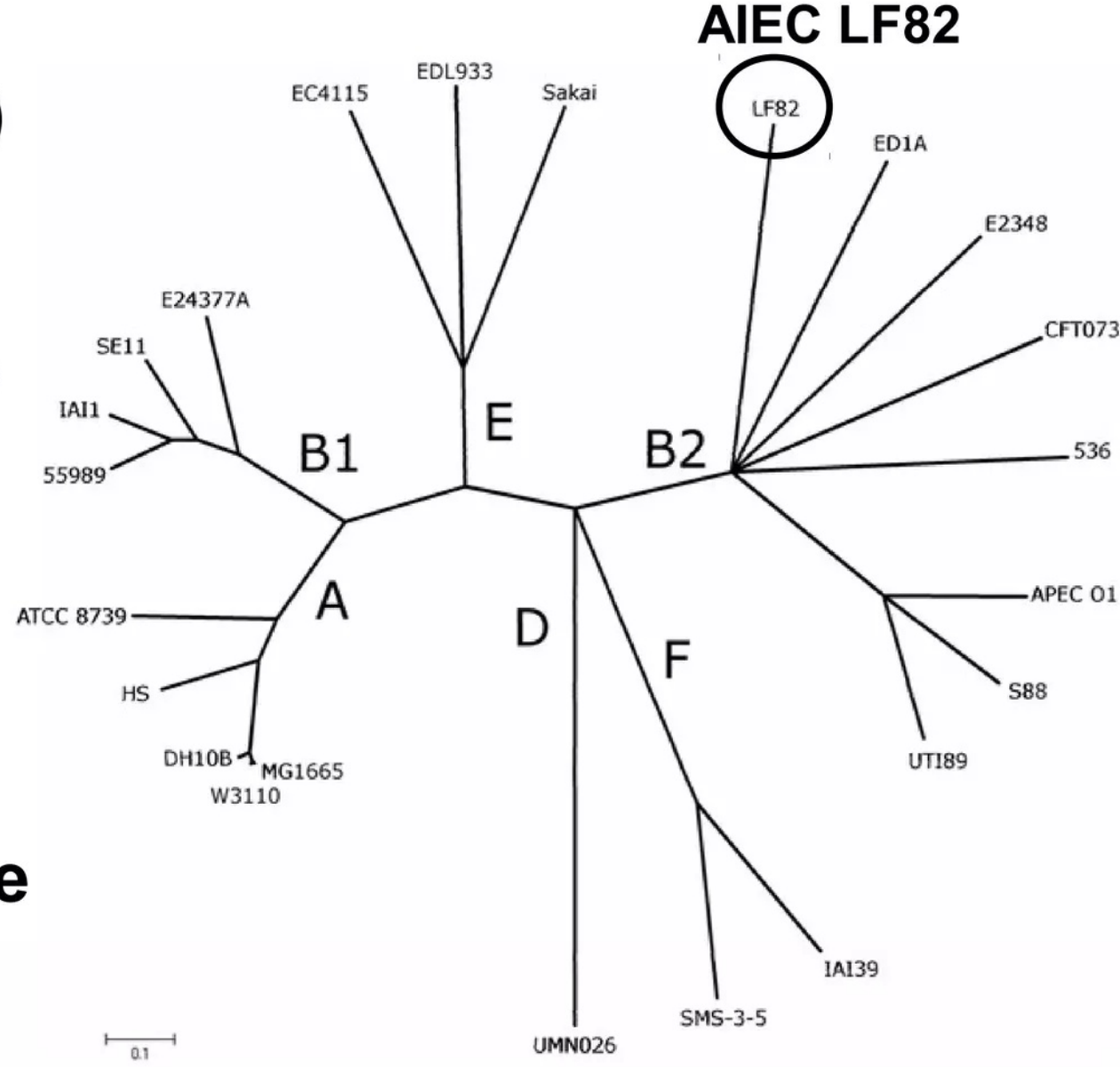
Sebastian E. Winter, Christopher A. Lopez & Andreas J. Bäumler,
EMBO reports VOL 14, p. 319-327 (2013)

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)

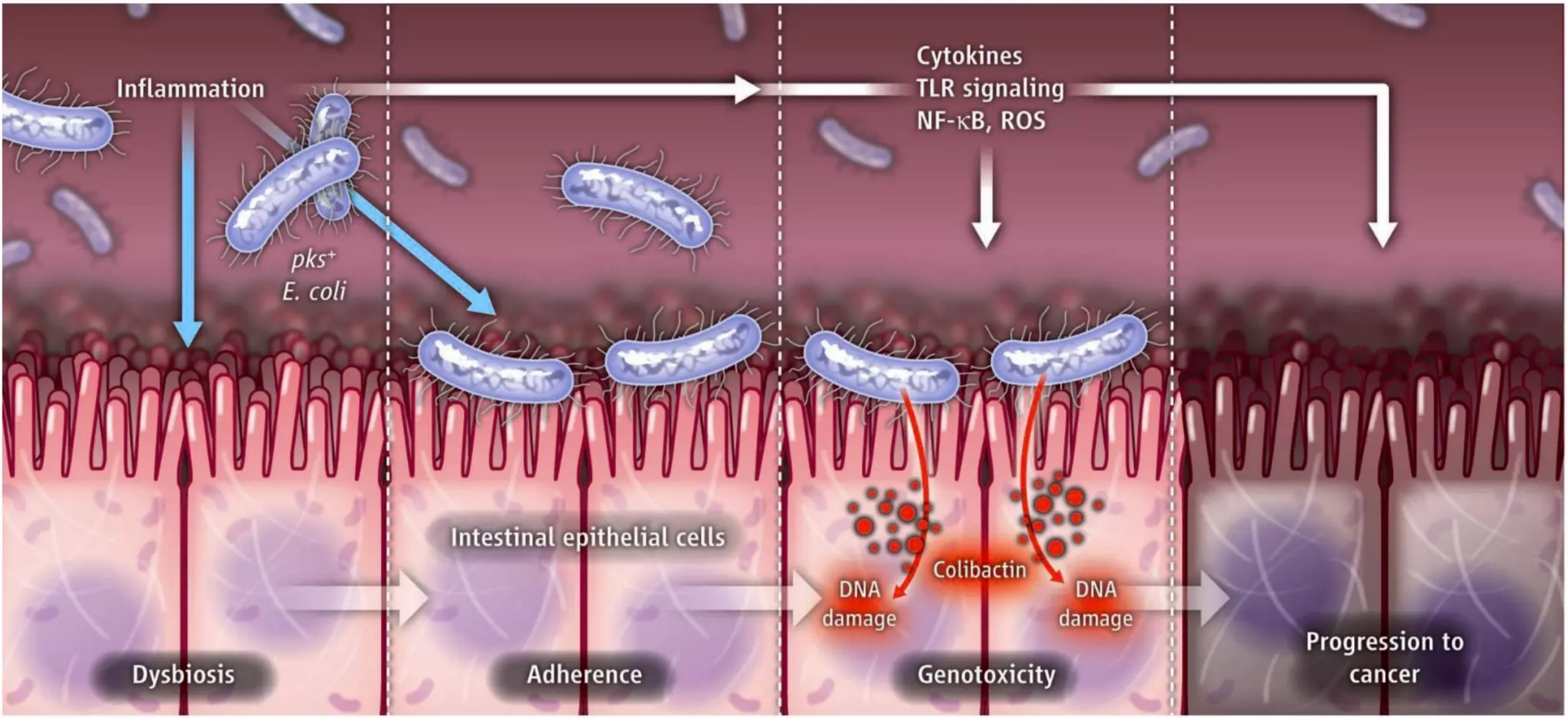


E. coli/Shigella Phylogenetic Tree
Miquel, et al.
PLOS ONE, v. 5, p. 1-16 (2010)

***Family Containing E. coli**

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

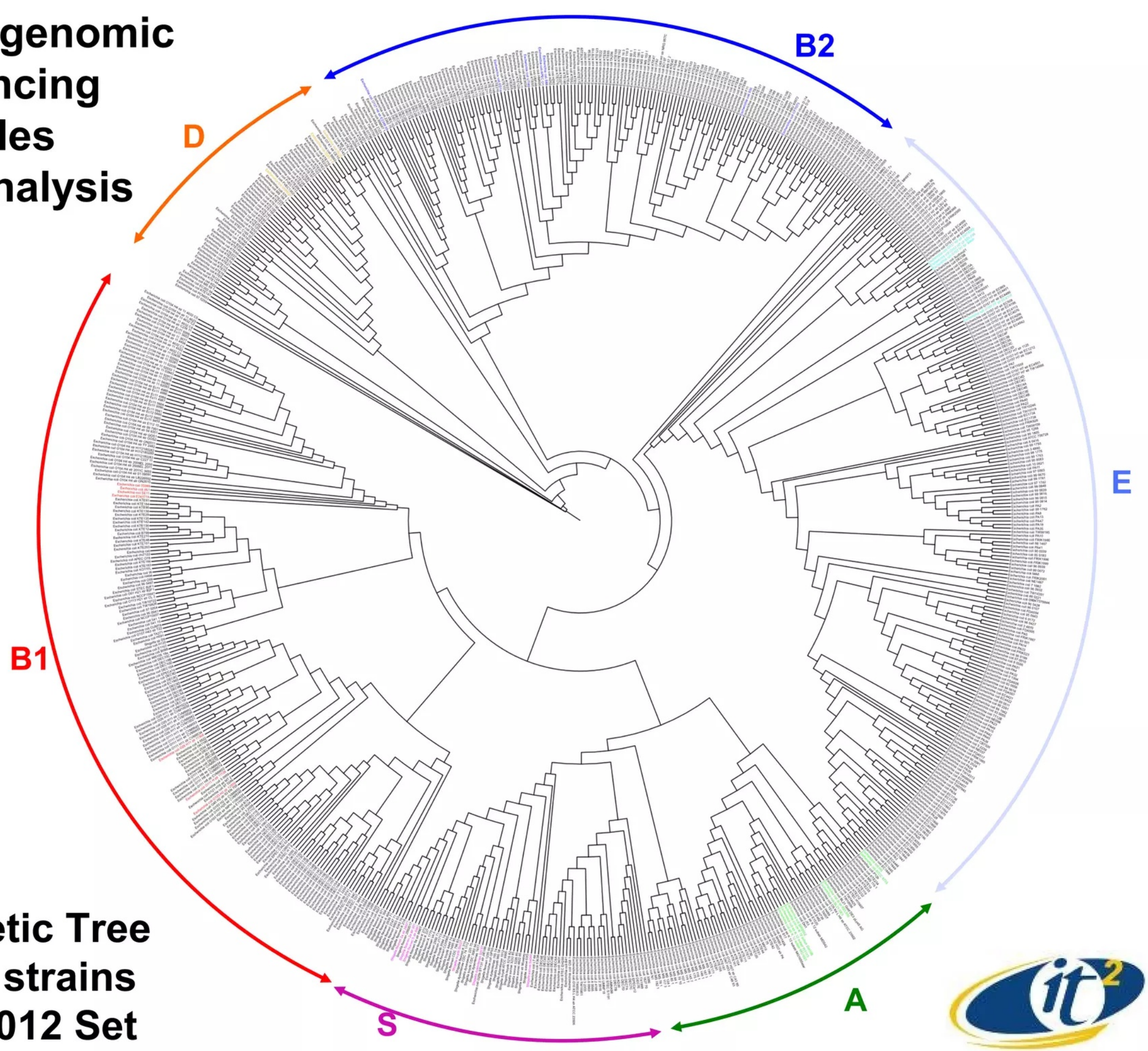
Escherichia coli Strain NC101



Science
AAAS

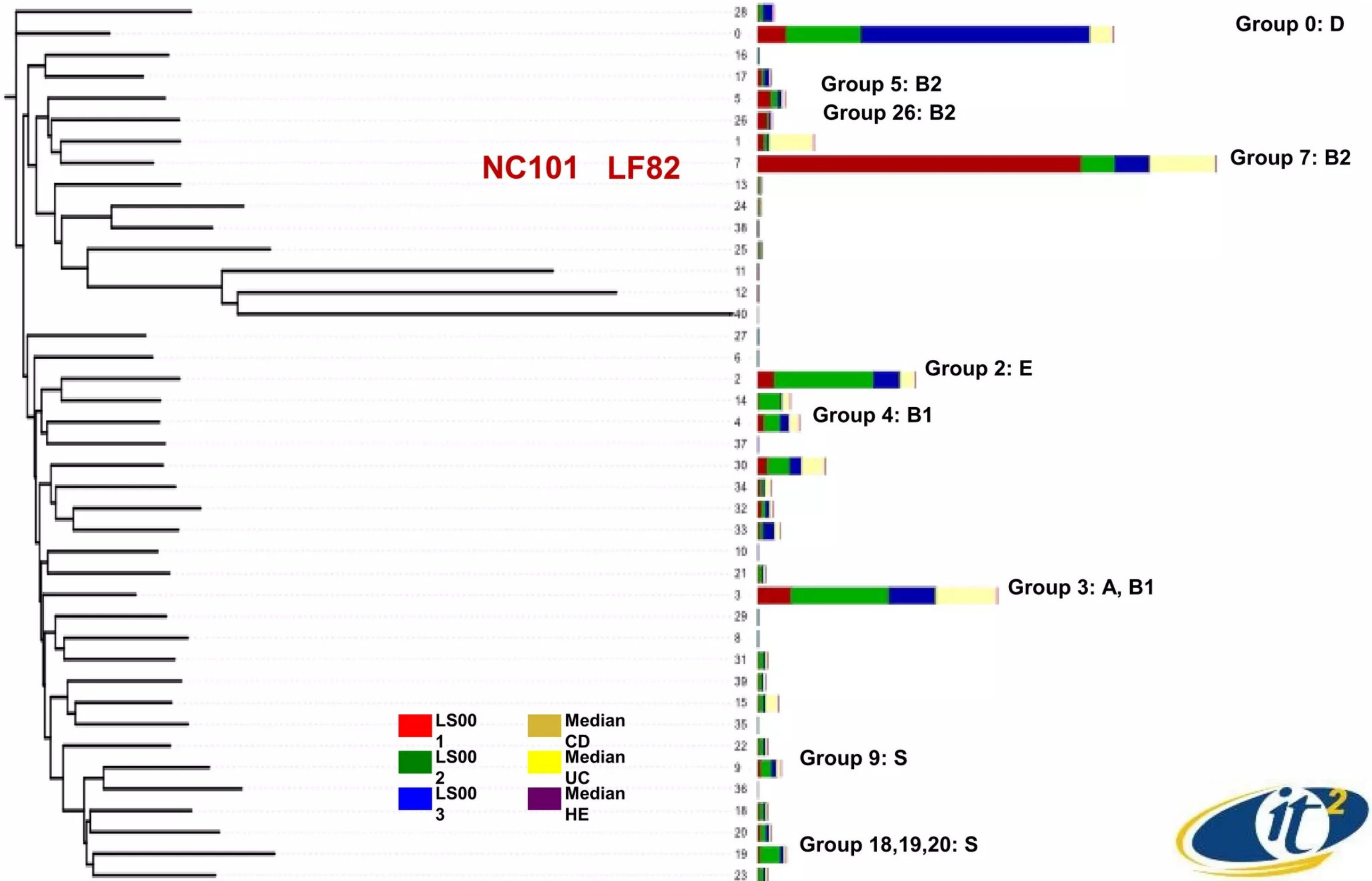
Bacteria Deliver a Genotoxic Hit
Robert F. Schwabe and Timothy C. Wang
Science **338**, 52 (2012);
DOI: 10.1126/science.1229905

Deep Metagenomic Sequencing Enables Strain Analysis

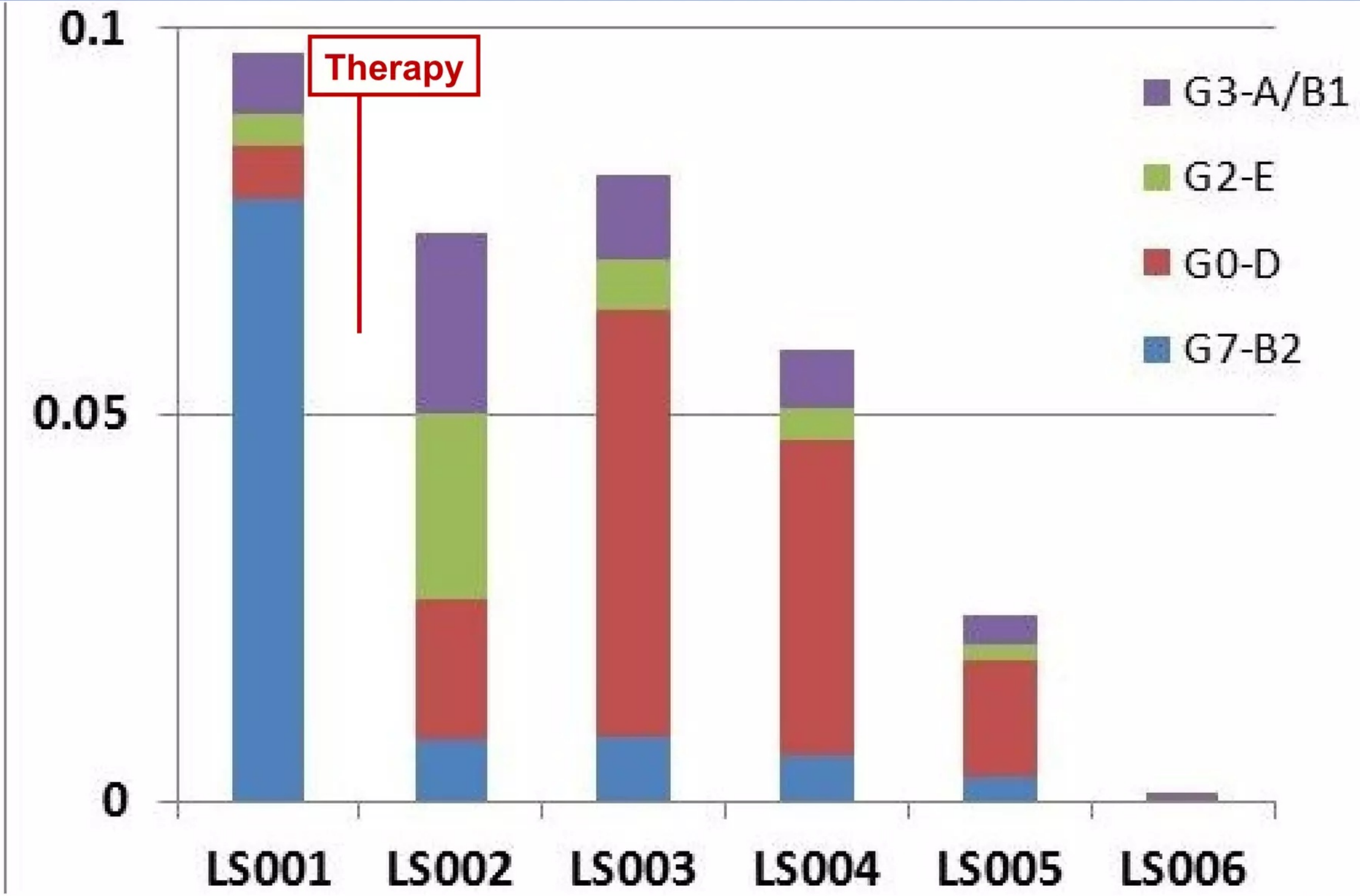


Phylogenetic Tree
778 Ecoli strains
=6x our 2012 Set

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

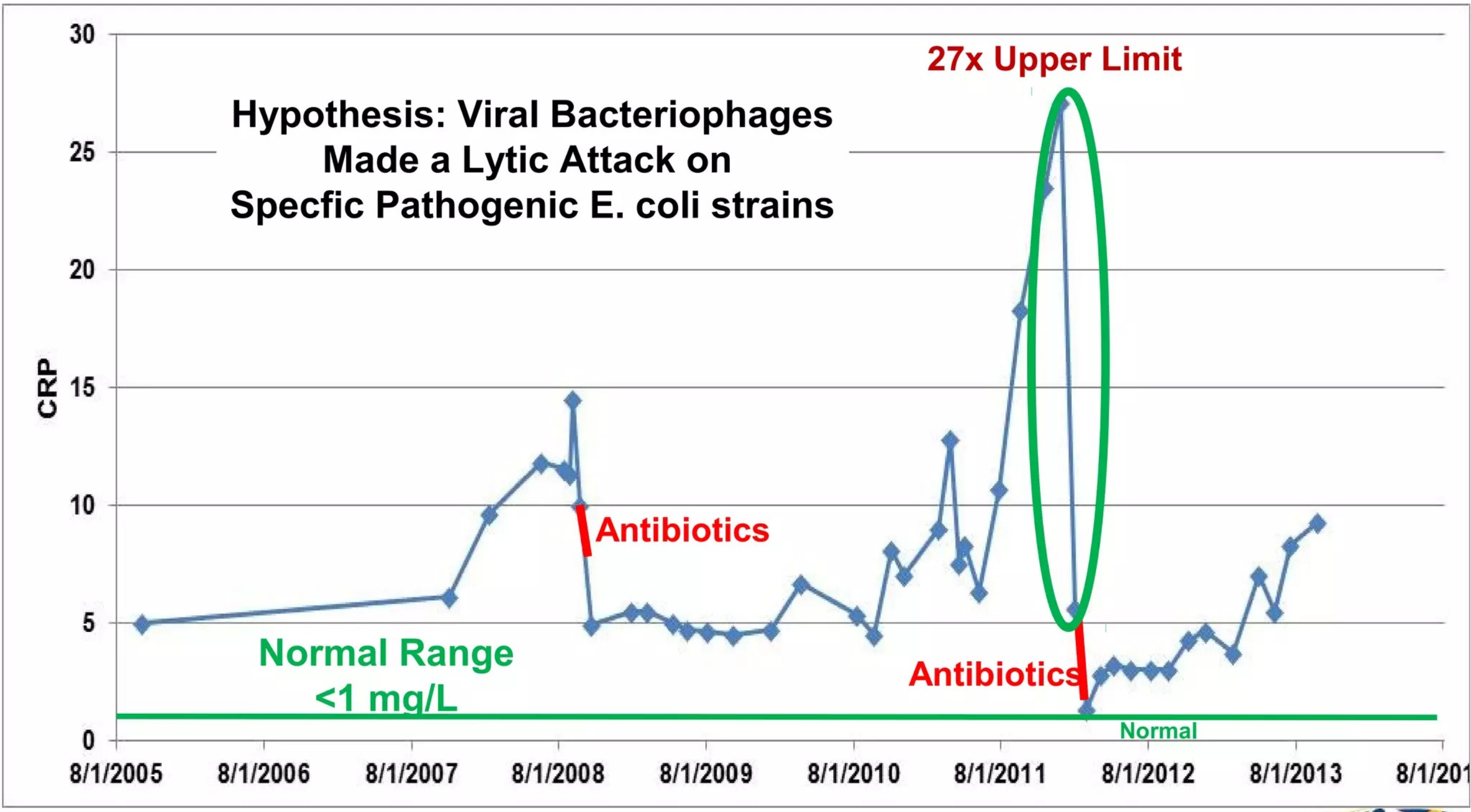


Reduction in E. coli Over Time
With Major Shifts in Strain Abundance



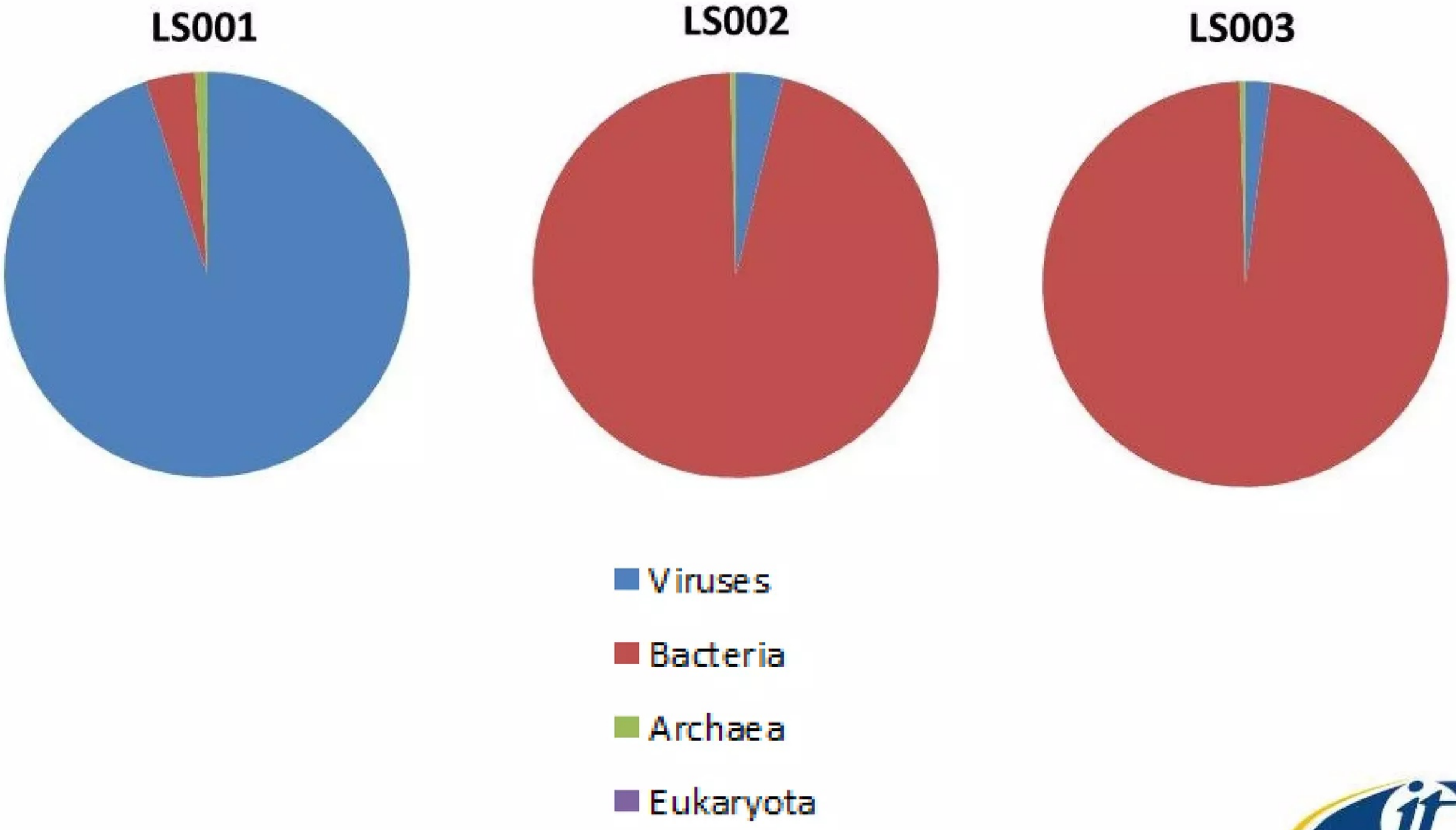
Strains >0.5% Included

What Caused the Dramatic Drop in My Inflammation Before Taking Antibiotics?

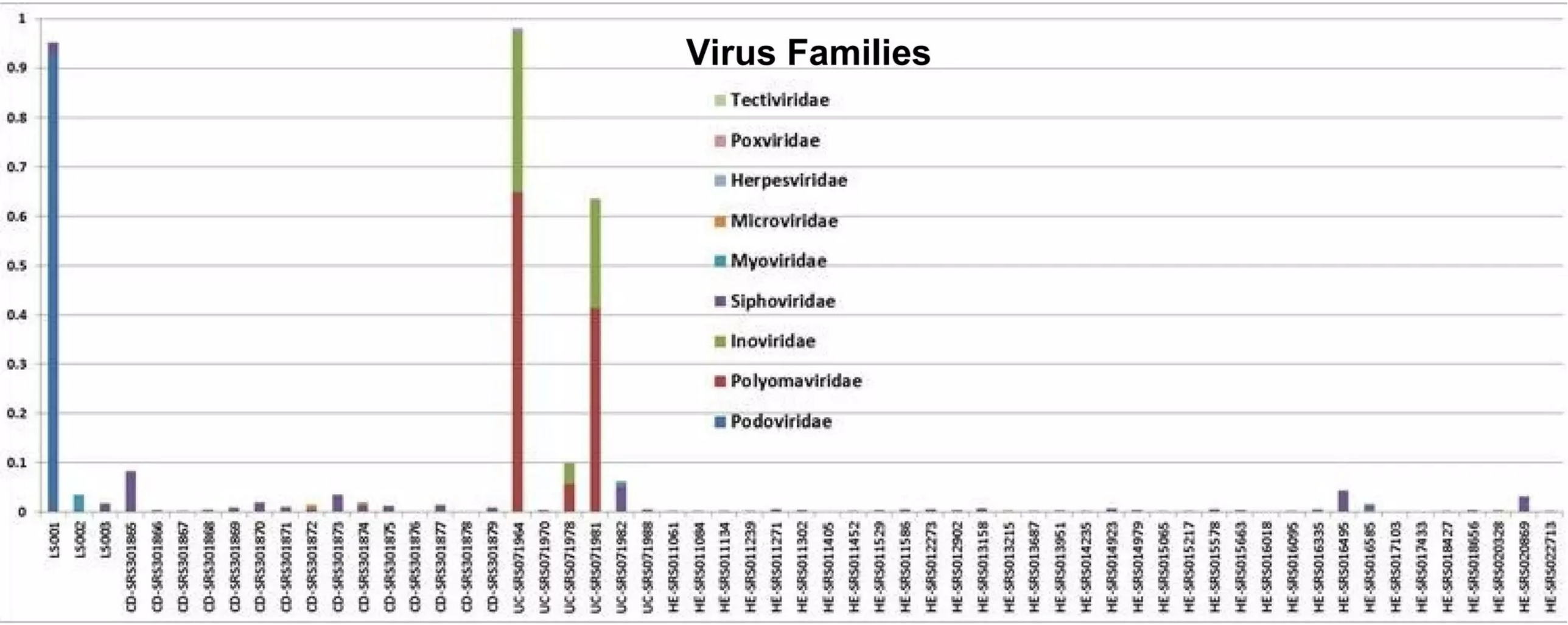


CRP is a Generic Measure of Inflammation in the Blood

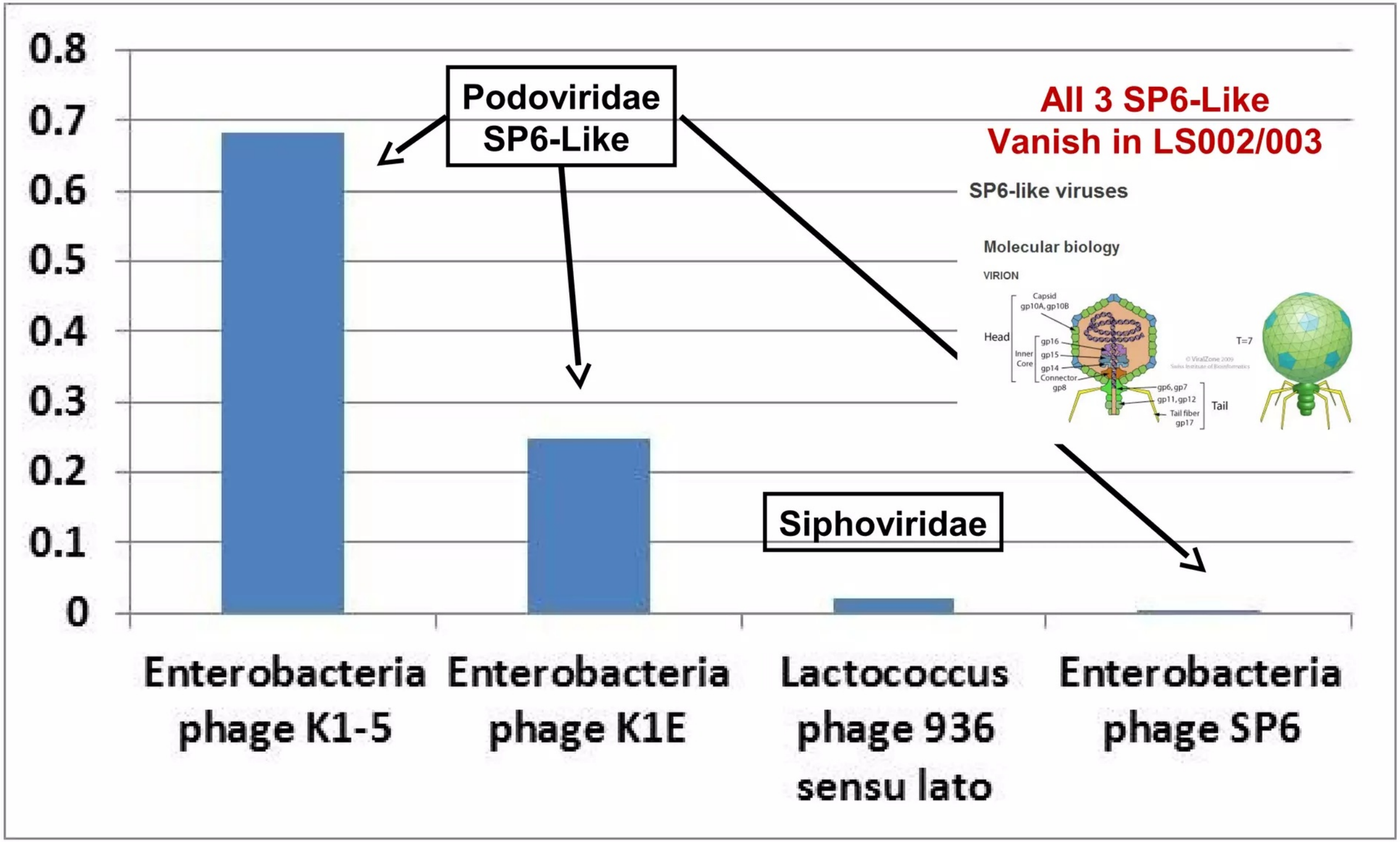
Radical Shift in Relative Abundance After Therapy



LS001 Viral Abundance is Similar to Some UC Patients, But Different Families



LS001 Relative Abundance of Viruses Among All Virus, Bacteria, Archaea, Eukaryota



Abundance >0.1%
Out of 493 Viral Reference Species

The Disruption of Consumer Health Data Gathering Is Growing Rapidly

Blood Variable Time Series

WellnessFX

InsideTracker

Stool Variable Time Series

YFH™

Human Genetic Variations

23andMe

MicrobiomeTime Series

μBiome

it²

From Quantified Self to National-Scale Biomedical Research Projects

DELSA

DATA-ENABLED LIFE SCIENCES ALLIANCE

Accelerate the impact of

Data-Enabled Life Sciences Research

on the pressing needs of the global society.

My Anonymized Human Genome is Available for Download

Personal Genome Project

The Quantified Human Initiative is an effort to combine our natural curiosity about self with new research paradigms. **Rich datasets of two individuals, Drs. Smarr and Snyder**, serve as 21st century personal data prototypes.



www.delsaglobal.org

www.personalgenomes.org

it²

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

UCSD Health Sciences Team

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