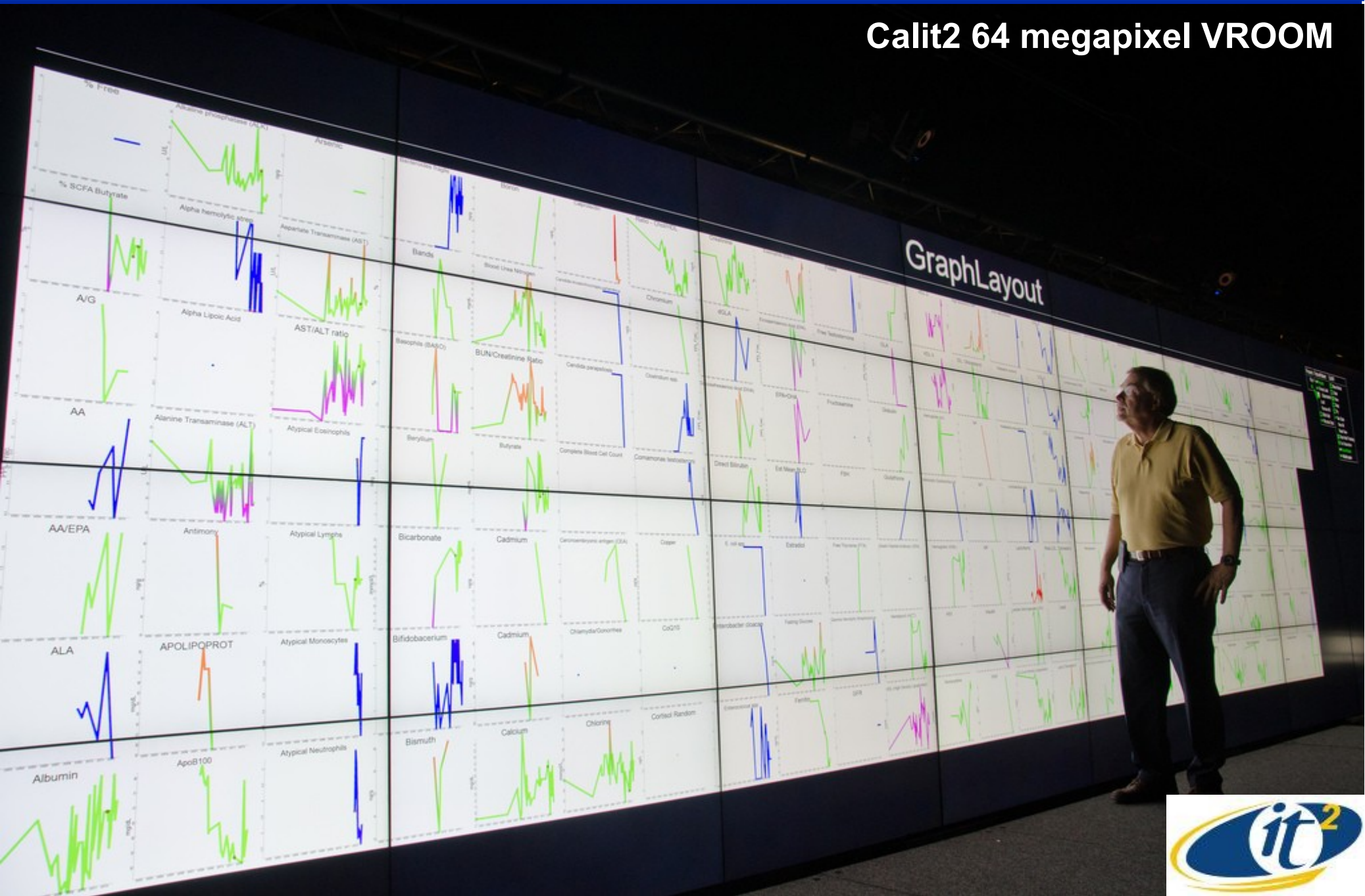
"Quantifying the Time Progression of the Interaction of the Human Immune System with the Gut Microbiome"

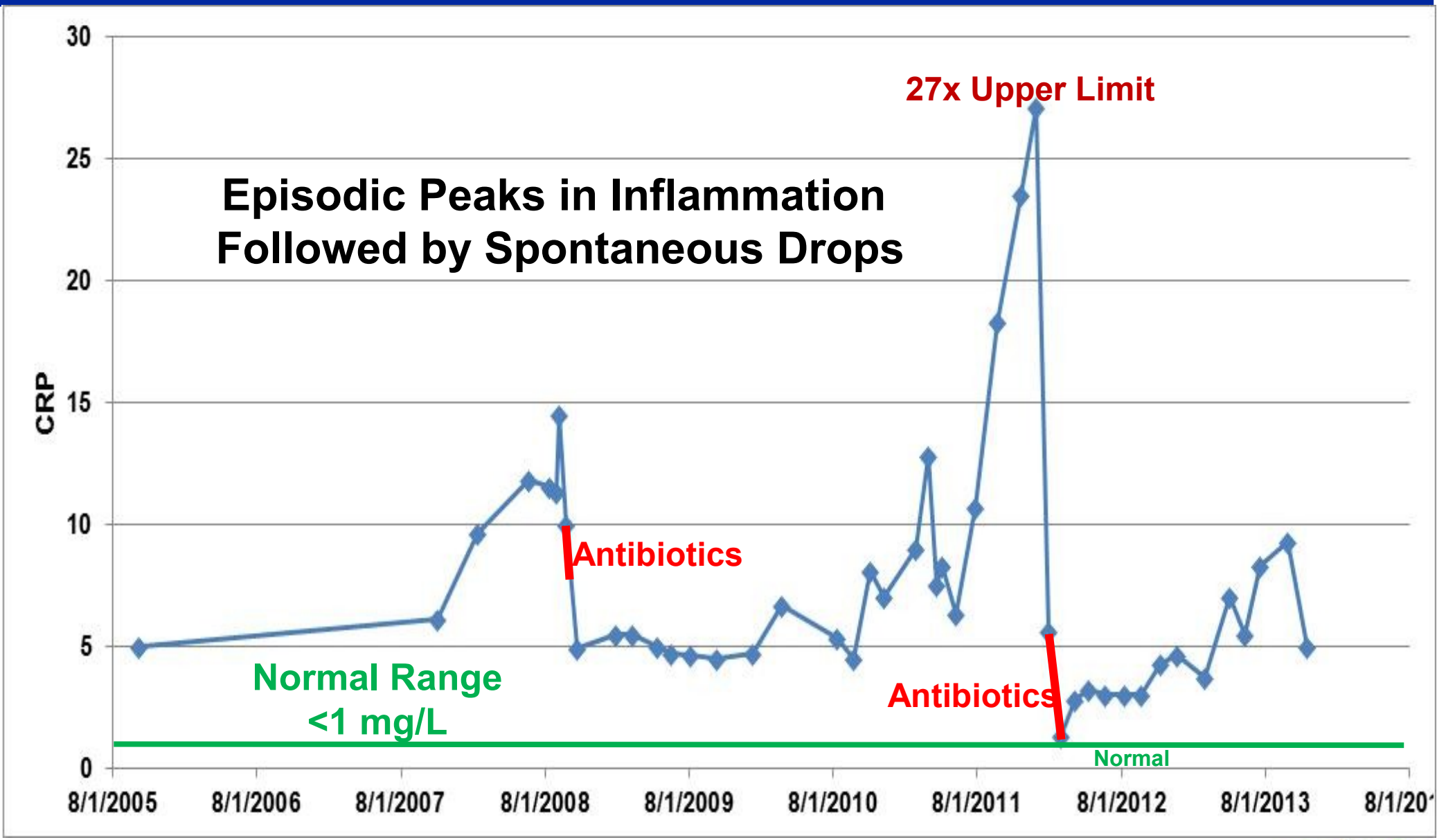
Research Council Presentation
UC San Diego Health Sciences
March 5, 2014

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

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

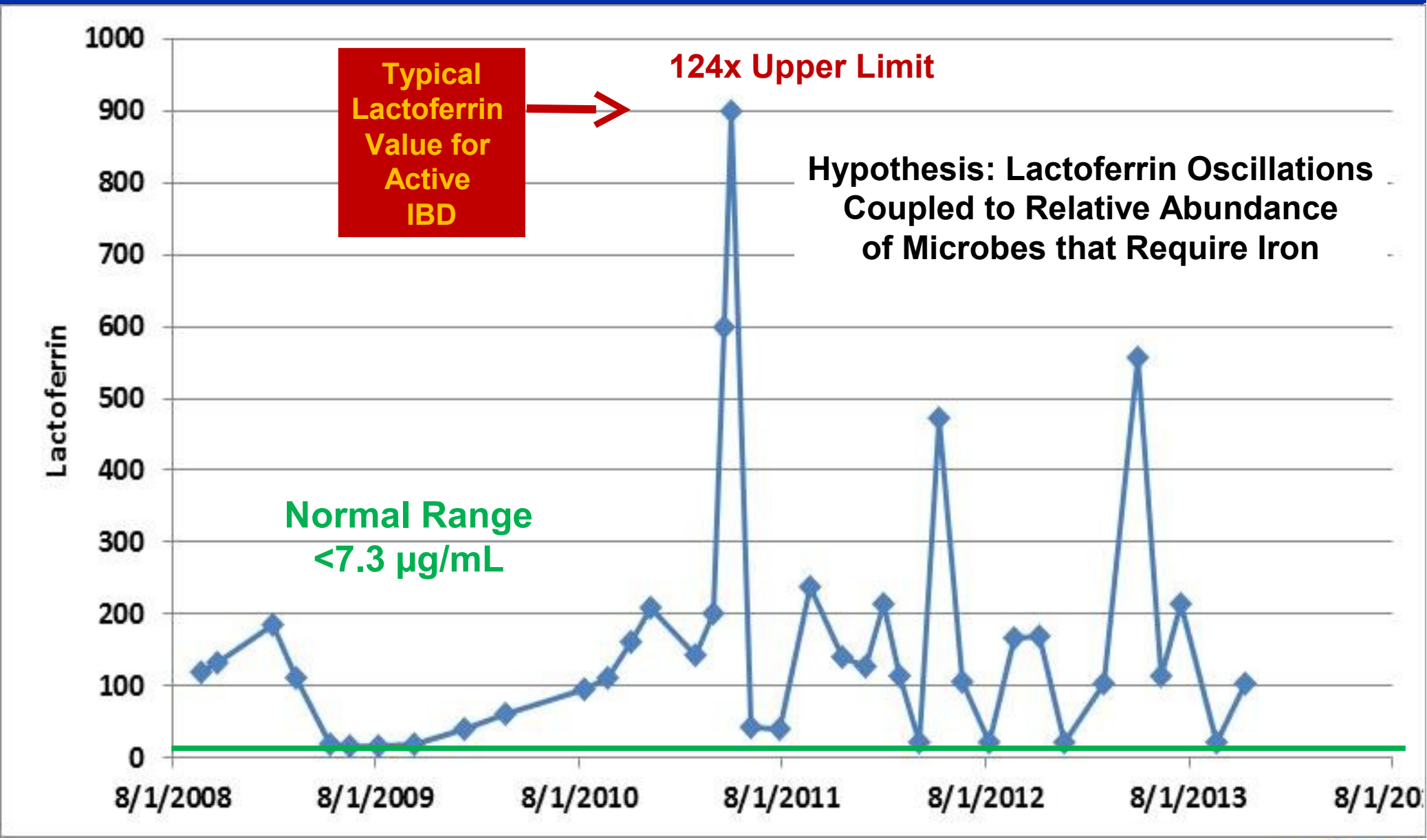


Only One of My Blood Measurements Was Far Out of Range--Indicating Chronic Inflammation



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

Adding Stool Tests Revealed Oscillatory Behavior in an Immune Variable

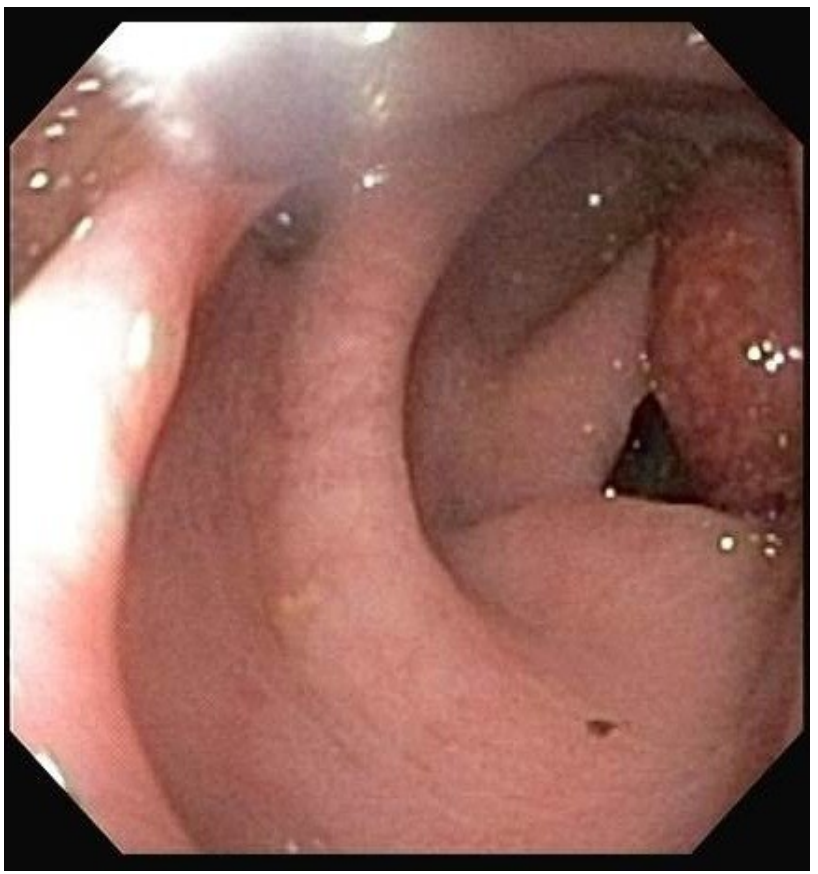


Lactoferrin is a Protein Shed from Neutrophils - An Antibacterial that Sequesters Iron

Colonoscopy Images Show Inflamed Pseudopolyps in 6 inches of Sigmoid Colon



Dec 2010



May 2011

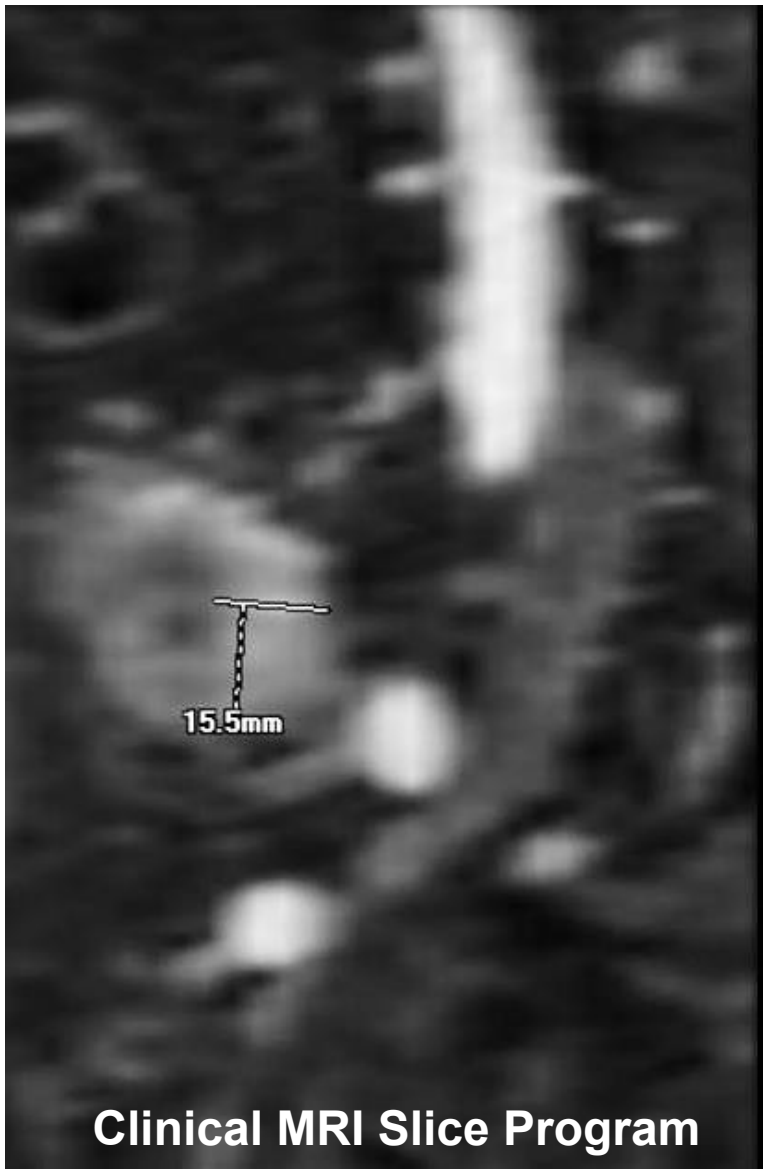
Confirming the Colonic Crohn's Hypothesis: Finding the "Smoking Gun" with MRI Imaging

"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.

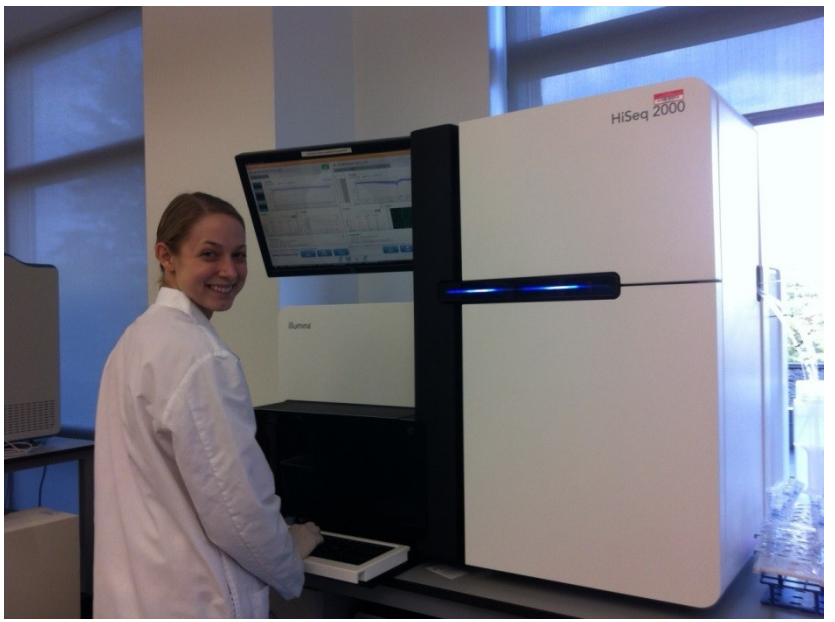
Having Crohn's disease that affects your colon increases your risk of colon cancer.



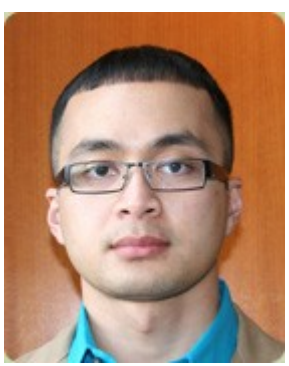
Reveals Inflammation in 6 Inches of Sigmoid Colon
Thickness 15cm – 5x Normal Thickness

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

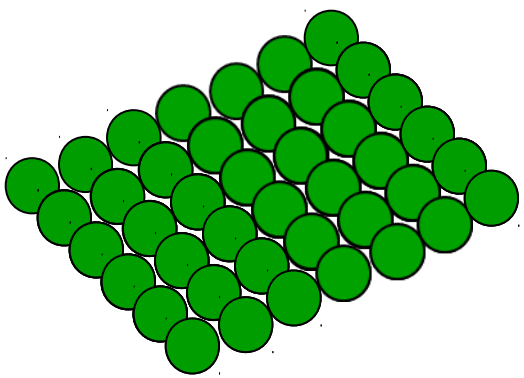
it2

We Downloaded Additional Phenotypes from NIH HMP For Comparative Analysis

Download Raw Reads
~100M Per Person

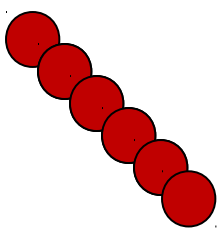
"Healthy" Individuals

35 Subjects
1 Point in Time



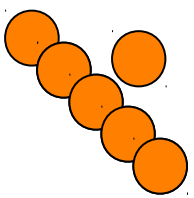
Larry Smarr

6 Points in Time

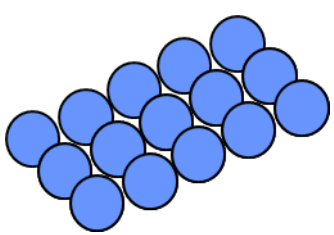


IBD Patients

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

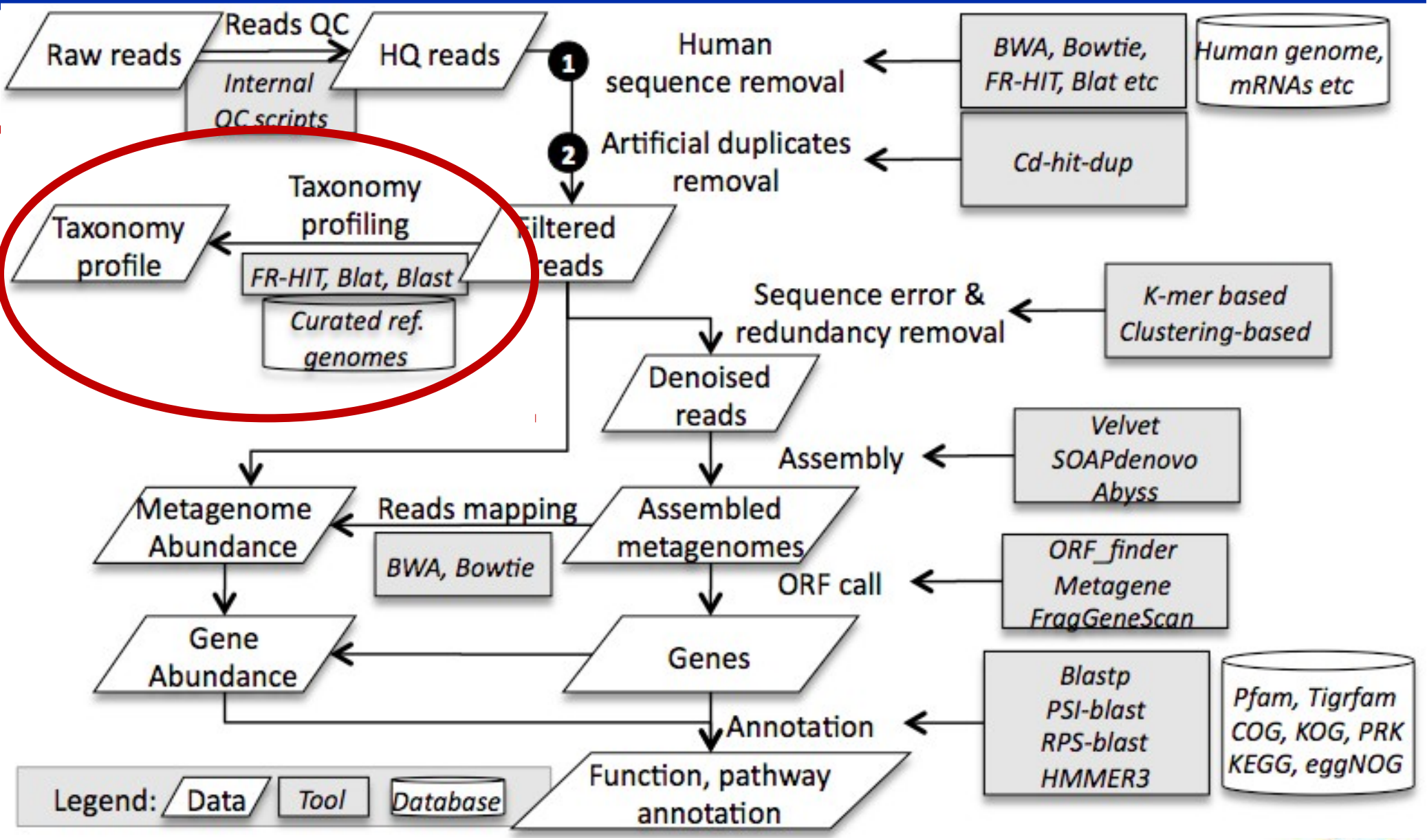
NIH

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

We Recently Used Dell's Supercomputer to Analyze An Additional 219 HMP and 110 MetaHIT Samples

- **Dell's Sanger cluster**
  - 32 nodes, 512 cores,
  - 48GB RAM per node
  - 50GB SSD local drive, 390TB Lustre file system

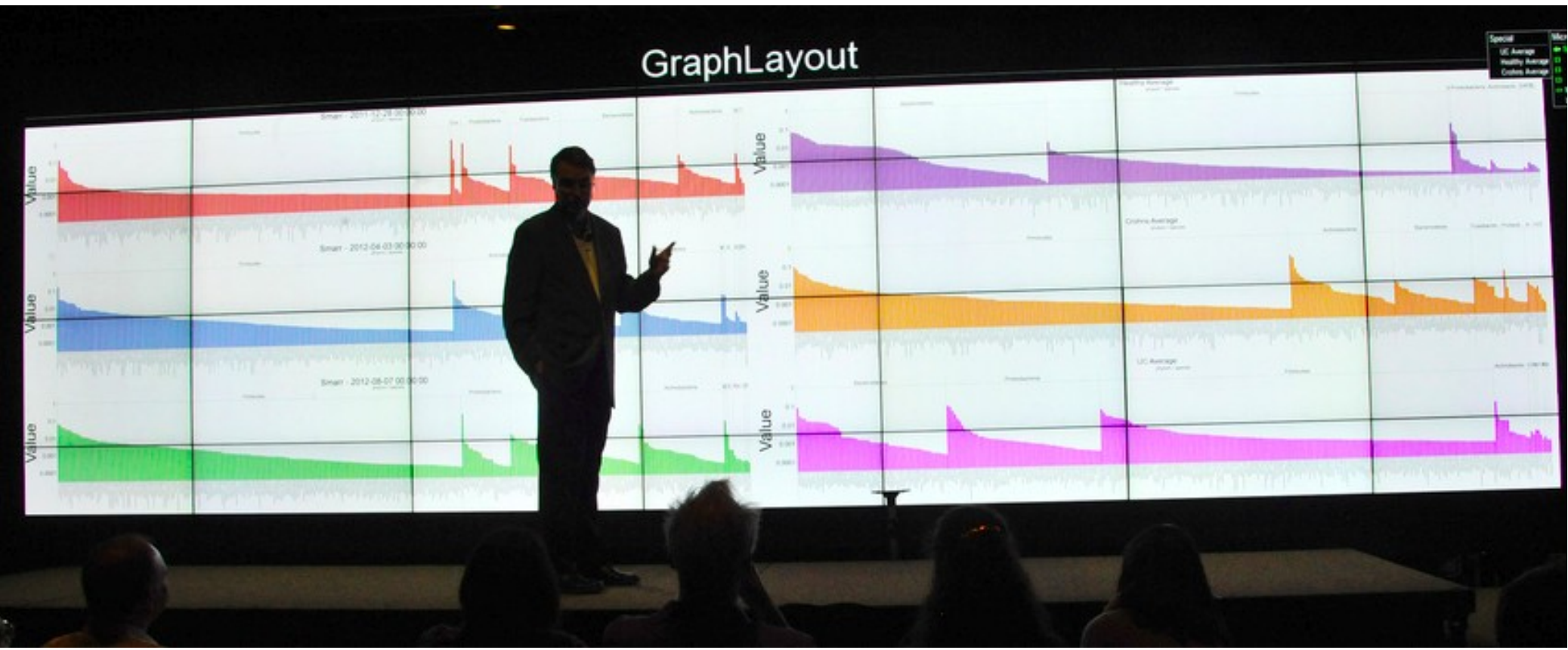


Dell Solutions Center
Industry Solutions Lab
SANGER DSU

- **We used faster but less sensitive method with a smaller reference DB (duo to available 48GB RAM)**
- **Only processed to taxonomy mapping**
  - ~35,000 Core-Hrs on Dell's Sanger
  - 30 TB data

Source: Weizhong Li, UCSD

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

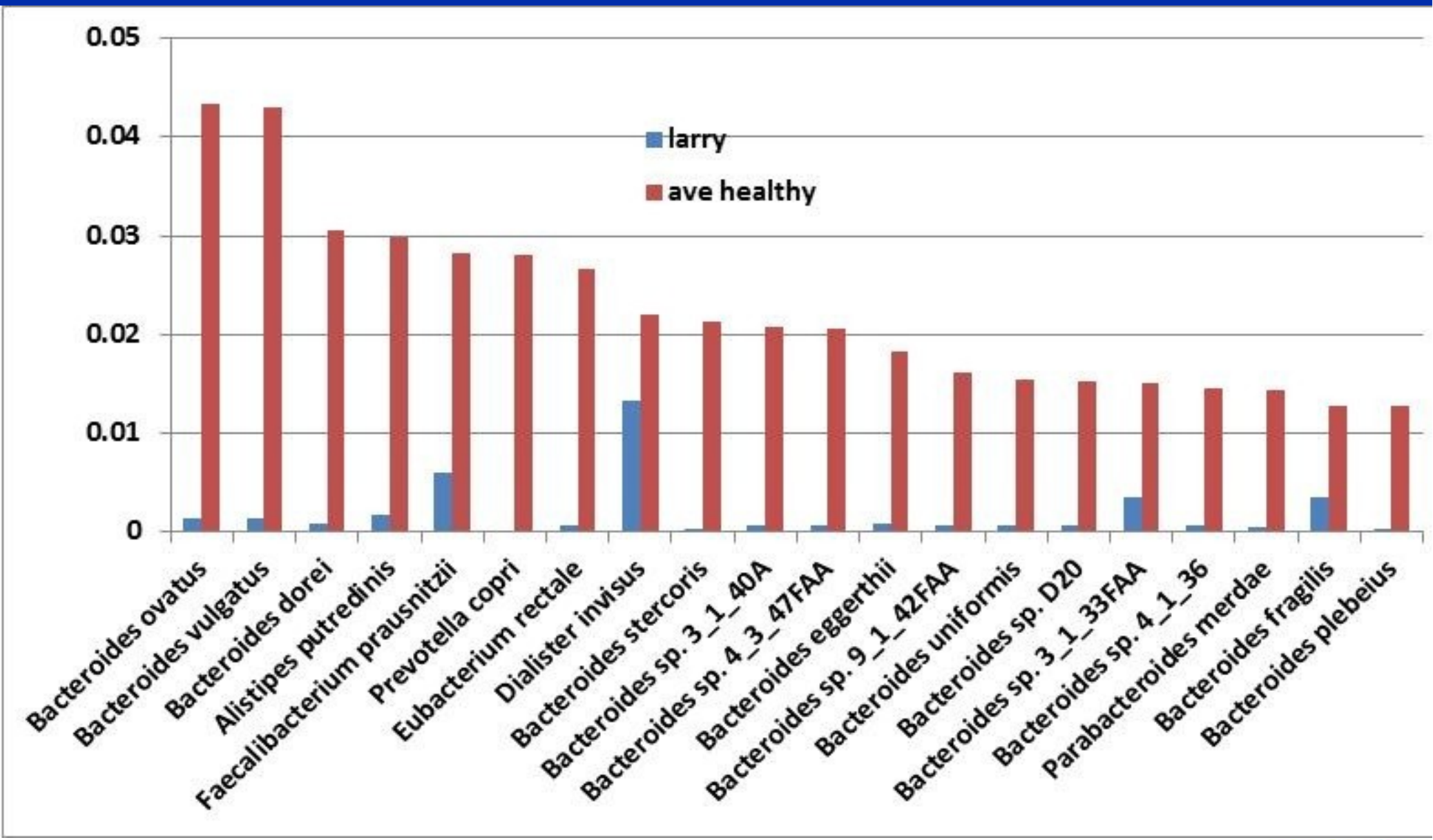
Lessons From Ecological Dynamics:
Invasive Species Dominate After Major Species Destroyed



”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

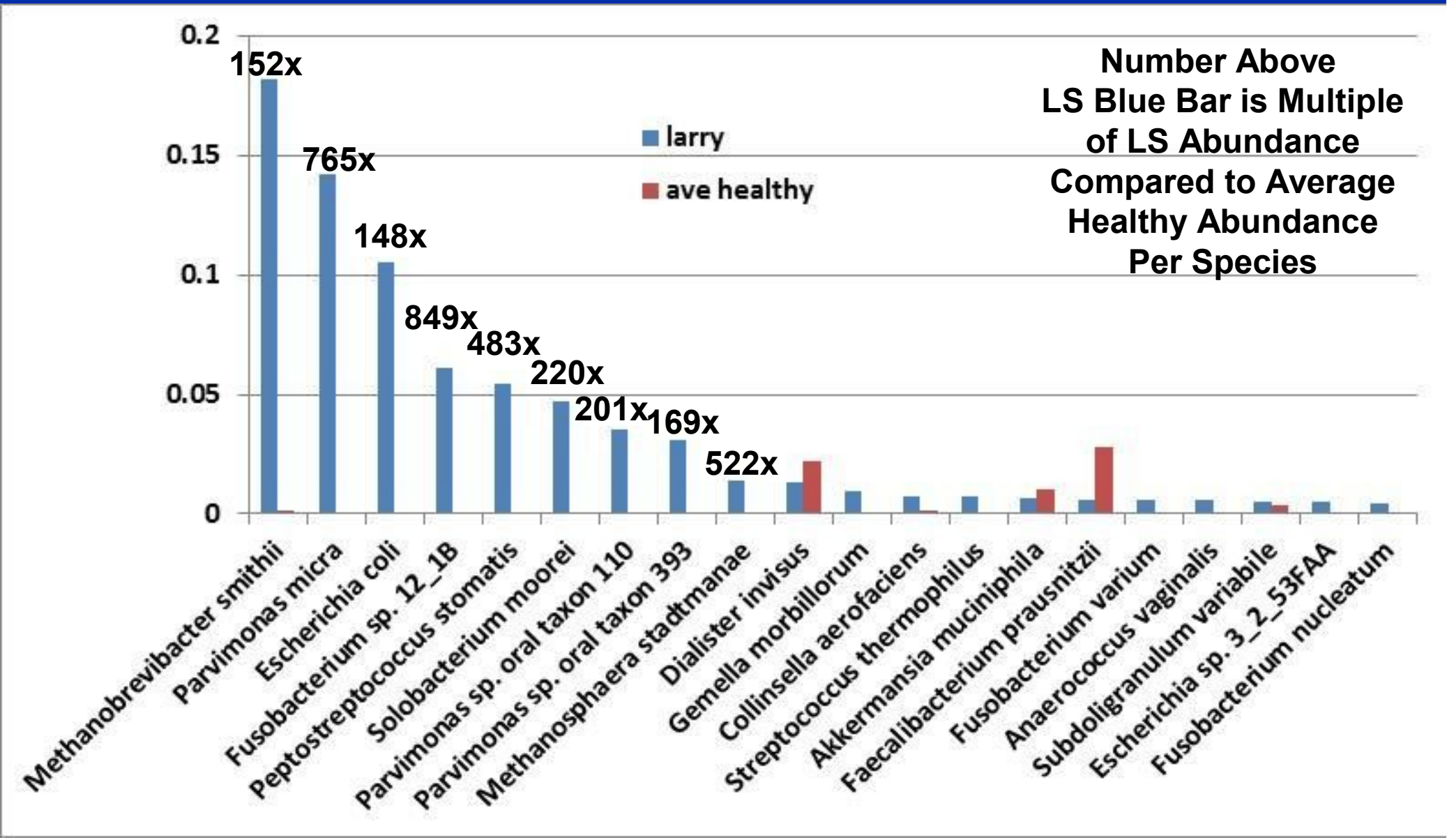
Almost All Abundant Species (≥1%) in Healthy Subjects Are *Severely Depleted* in Larry's Gut Microbiome



J. Craig Venter INSTITUTE

it²

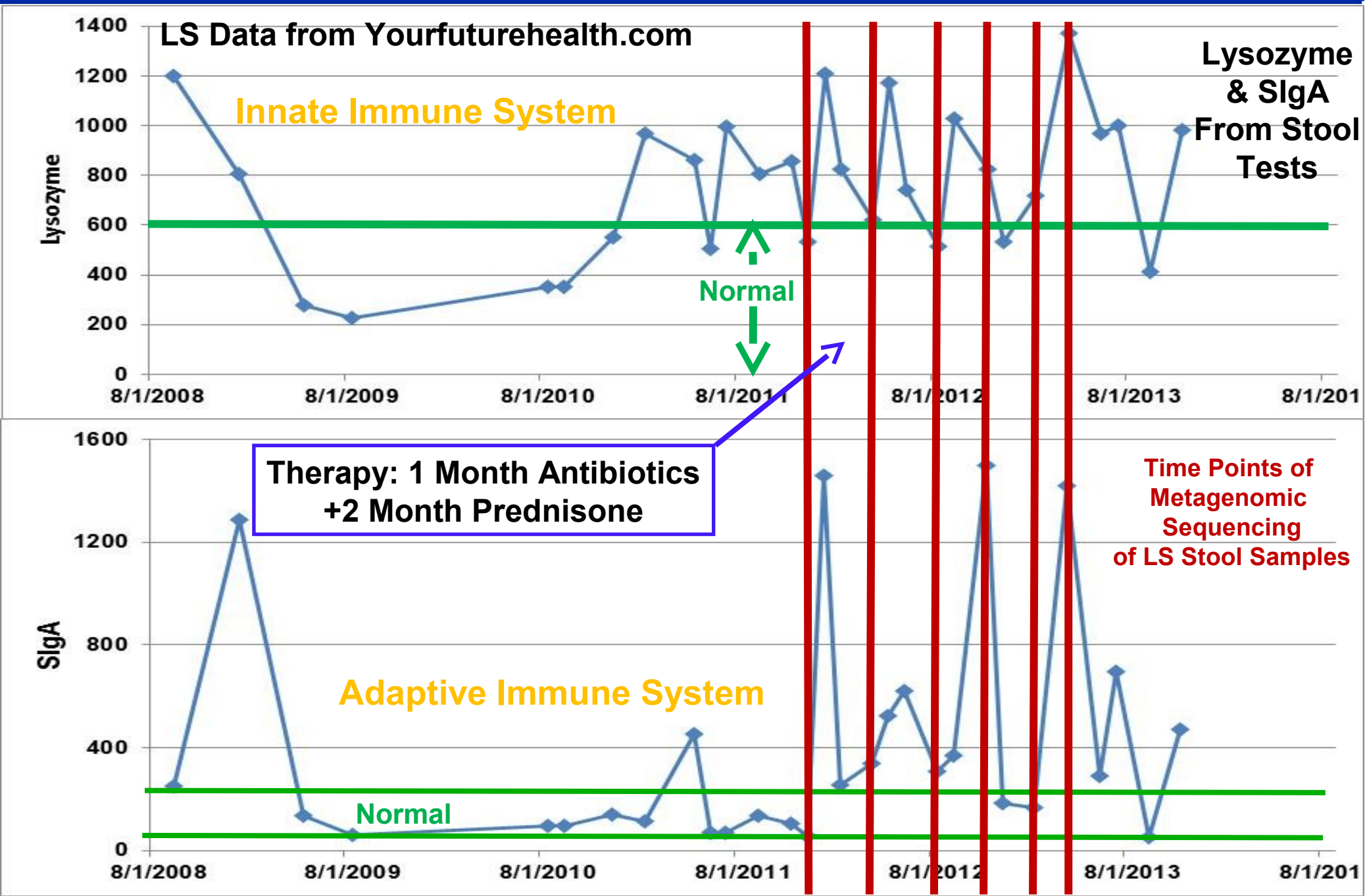
Top 20 Most Abundant Microbial Species In LS vs. Average Healthy Subject



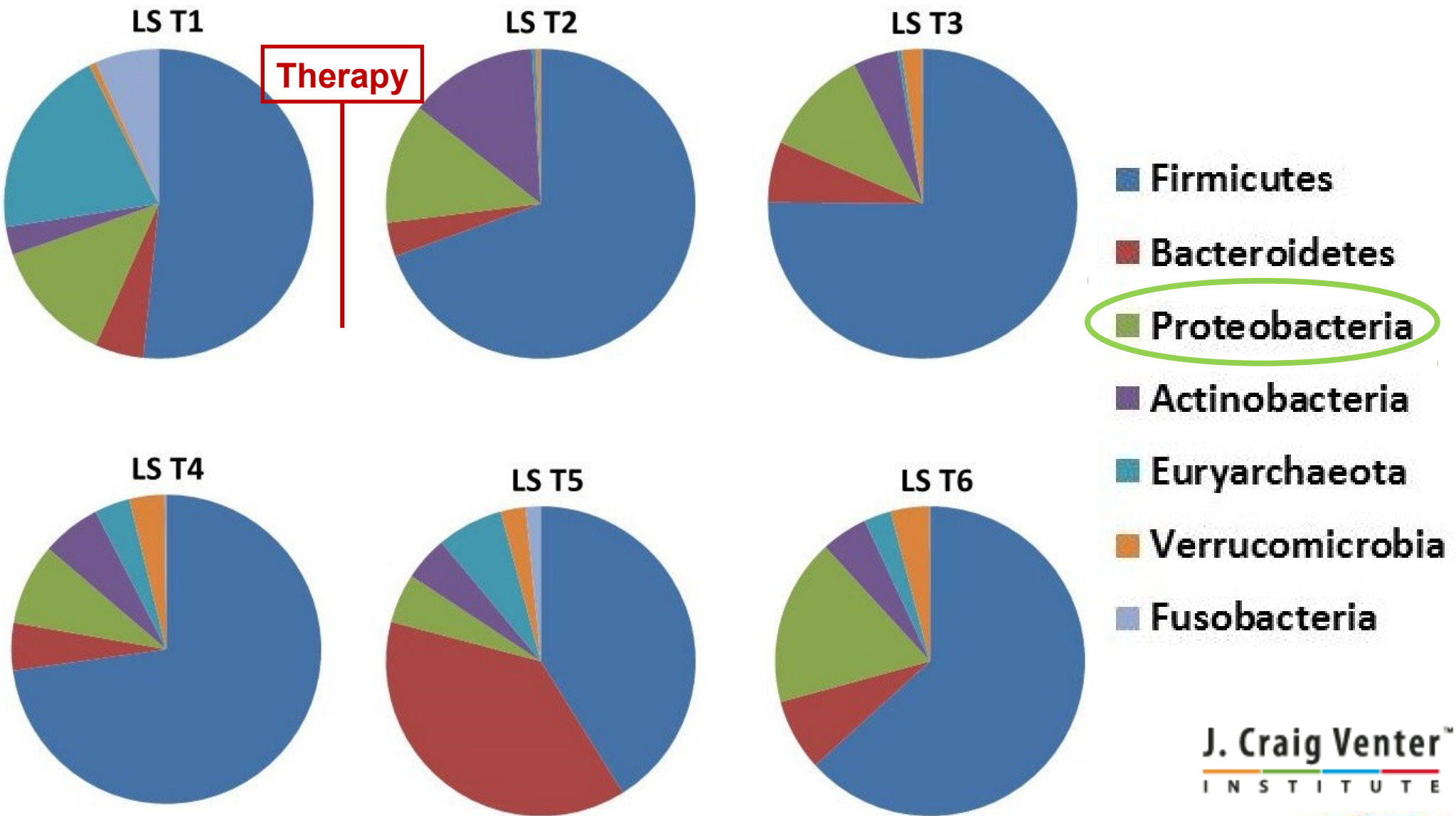
J. Craig Venter INSTITUTE

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



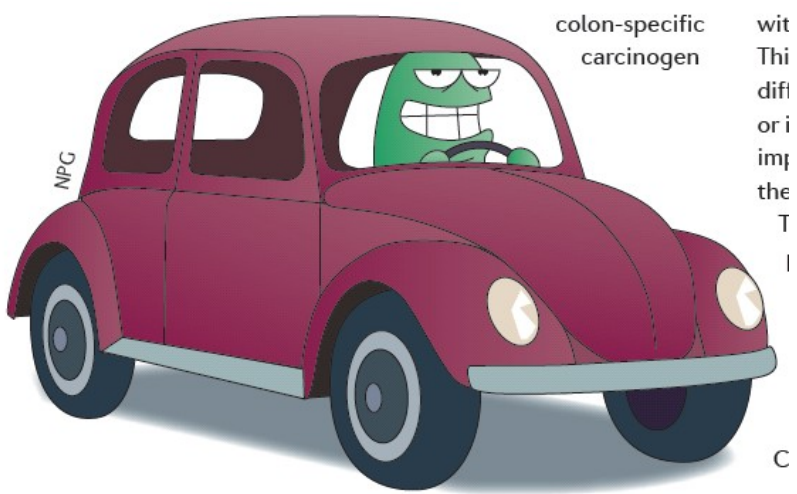
J. Craig Venter INSTITUTE

it²

Six Metagenomic Time Samples Over 16 Months

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

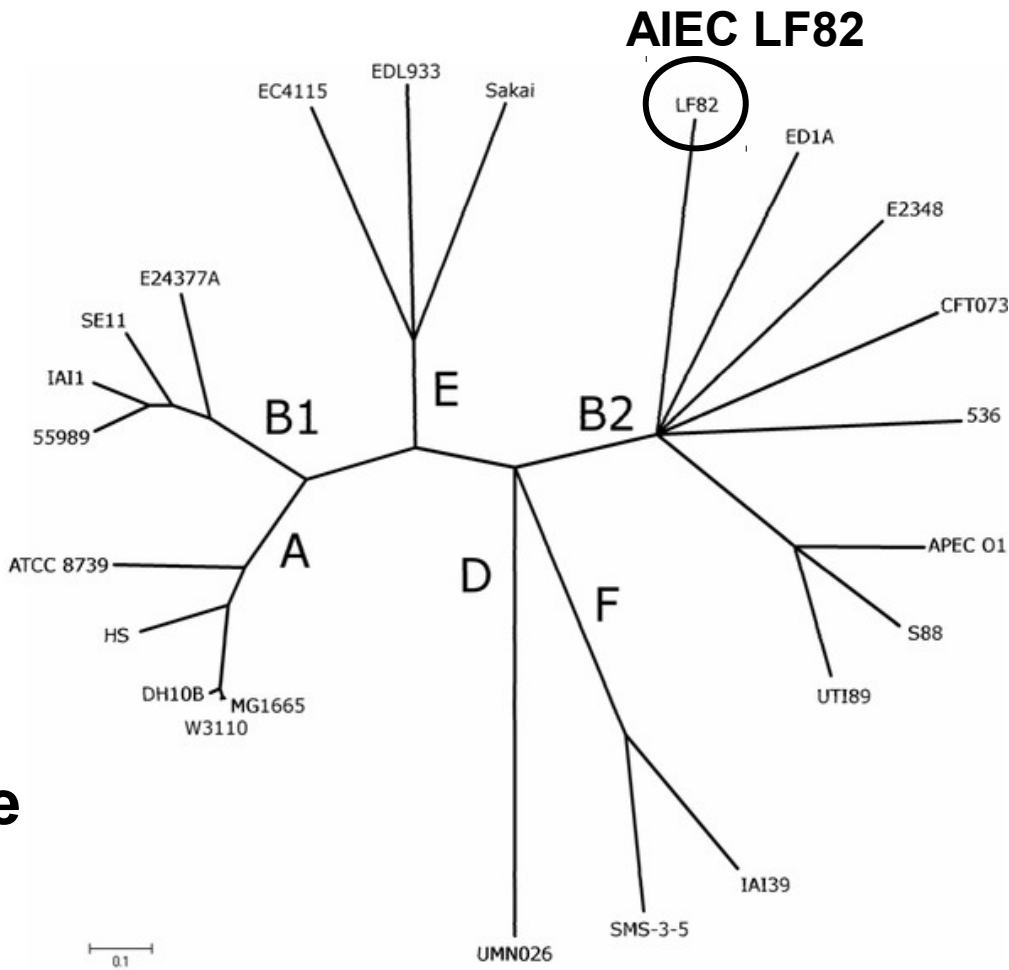
VOLUME 12 | OCTOBER 2012
NATURE REVIEWS | CANCER

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)

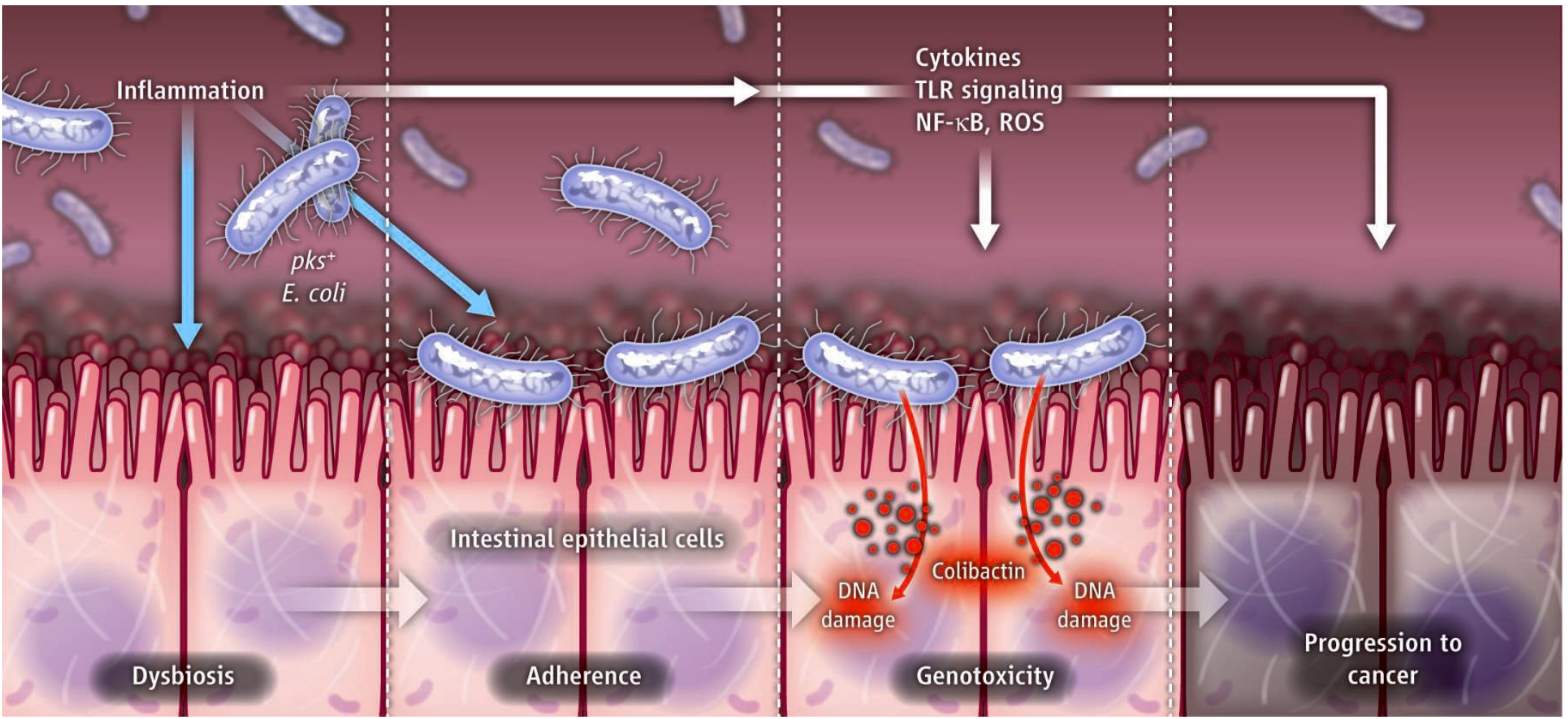


E. coli/Shigella Phylogenetic Tree
Miquel, et al.
PLOS ONE, v. 5, p. 1-16 (2010)

*Family Containing E. coli

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

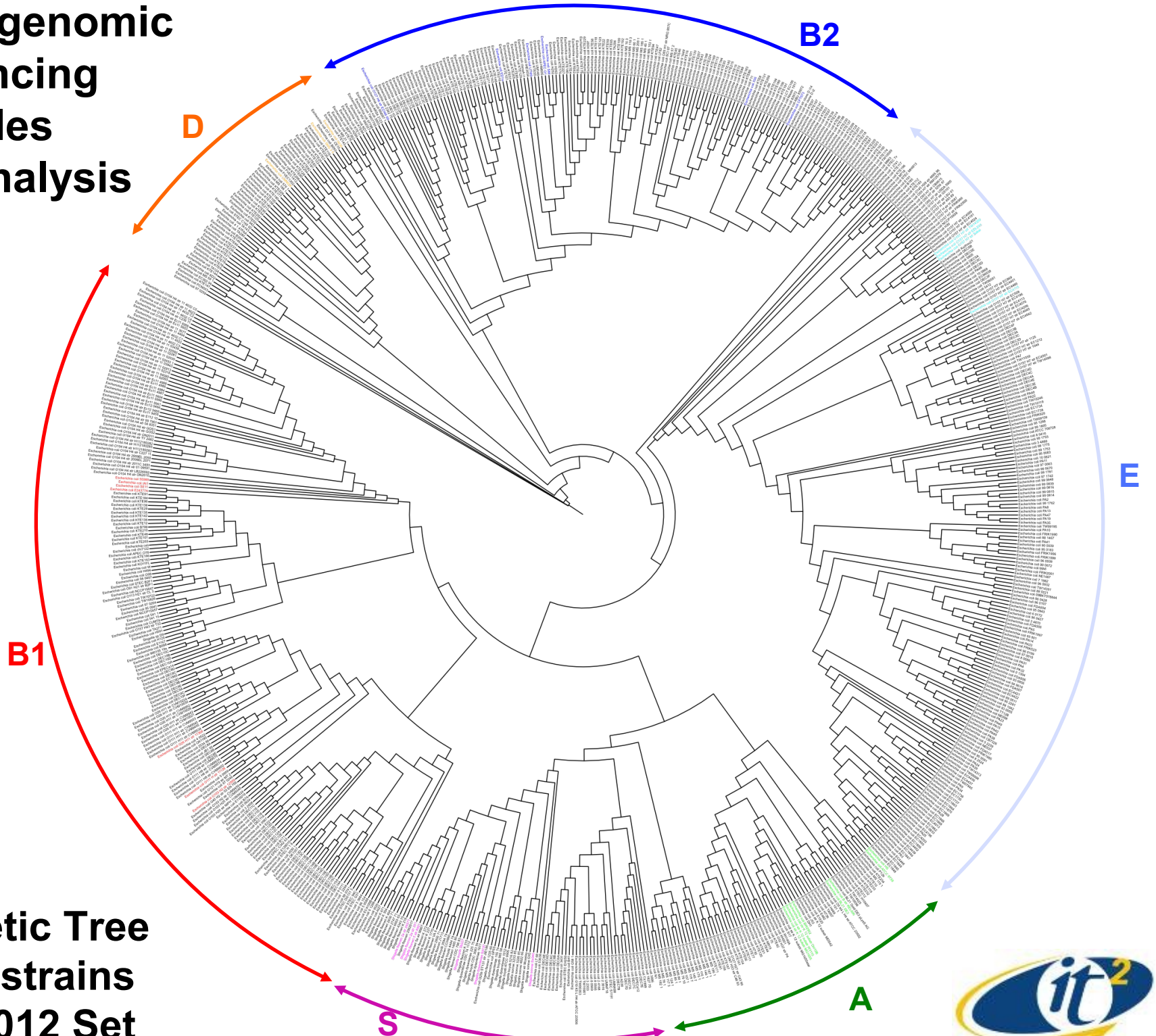
Escherichia coli Strain NC101



Science
AAAS

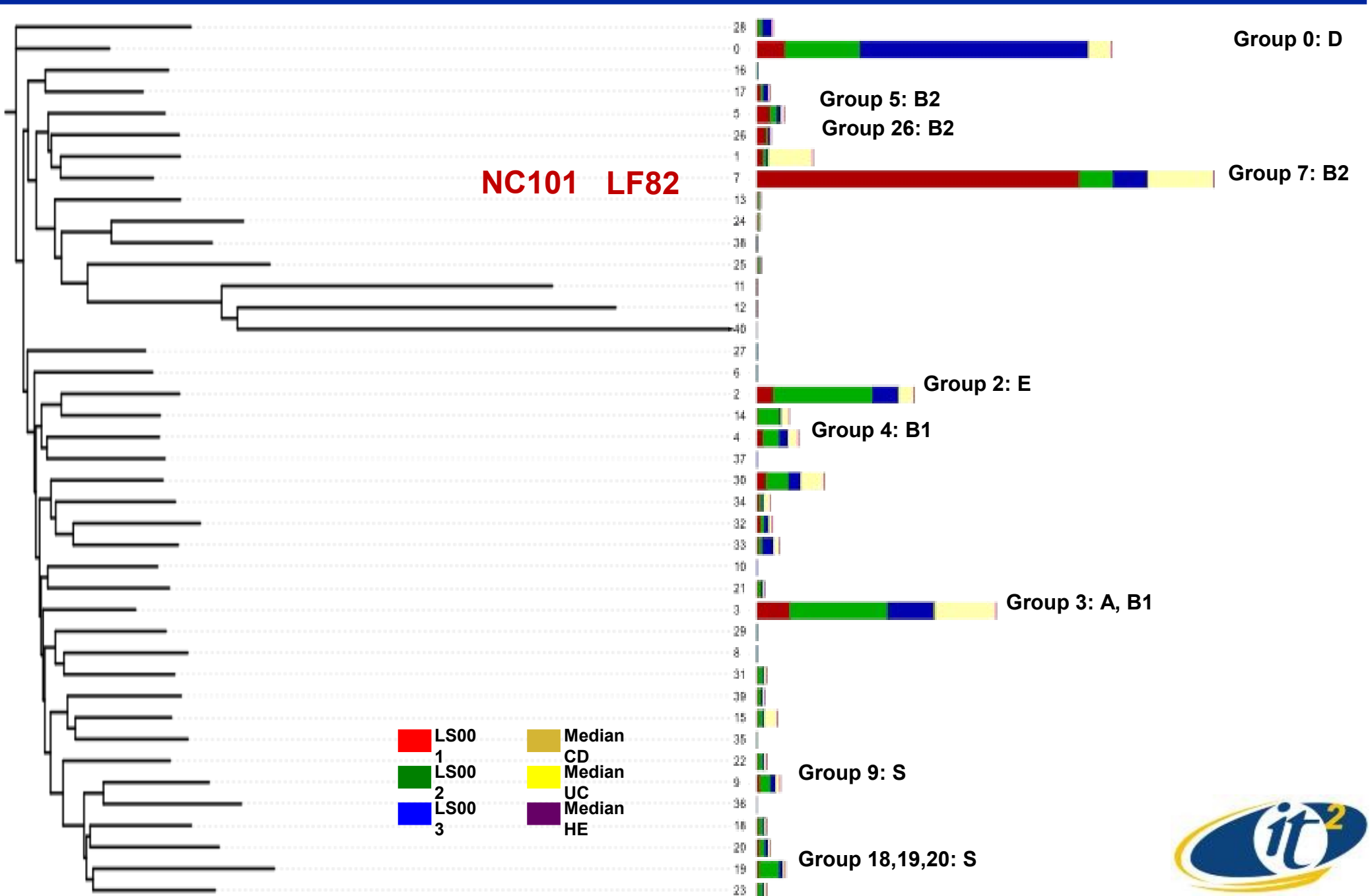
Bacteria Deliver a Genotoxic Hit
Robert F. Schwabe and Timothy C. Wang
Science **338**, 52 (2012);
DOI: 10.1126/science.1229905

Deep Metagenomic Sequencing Enables Strain Analysis

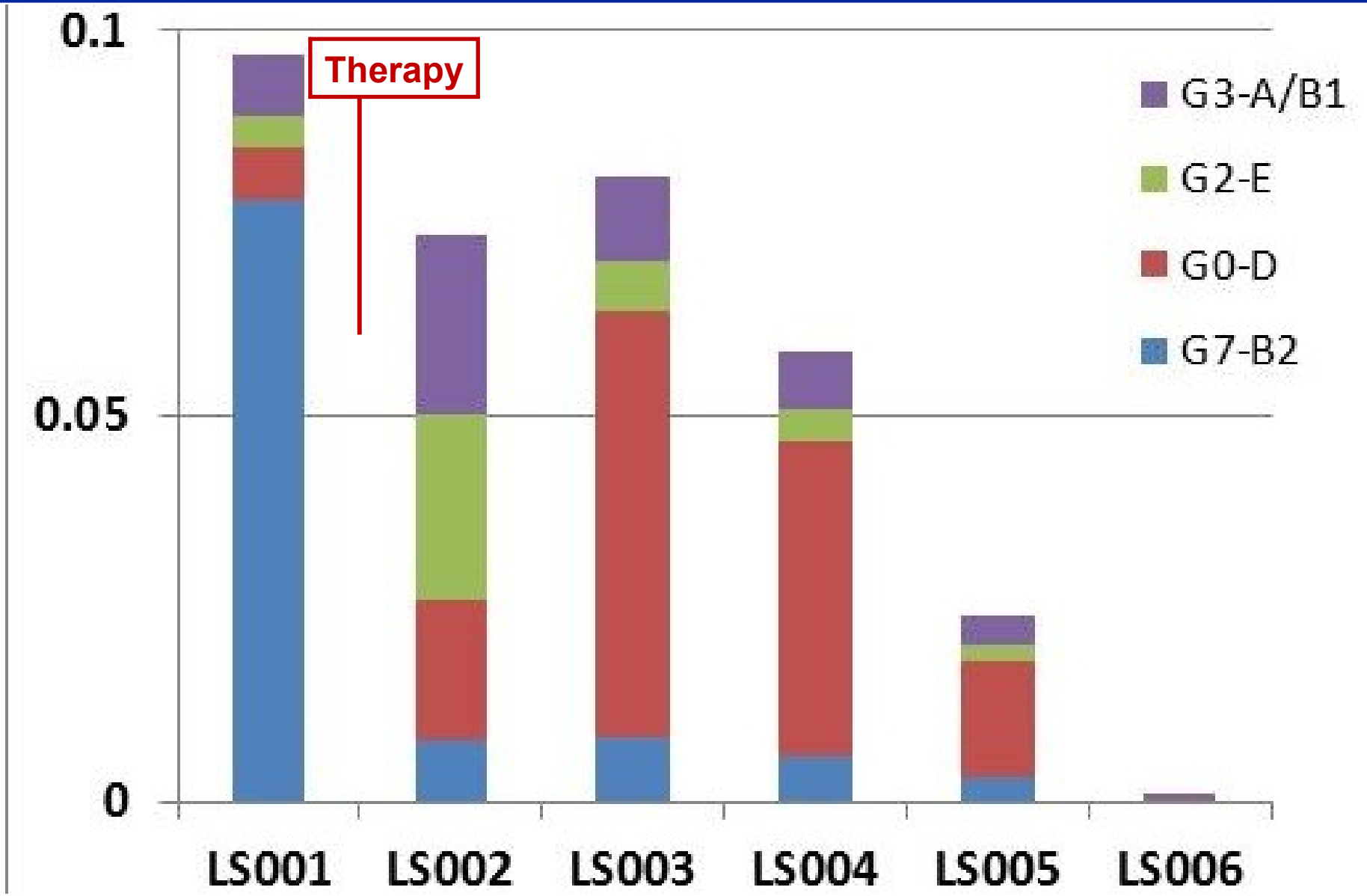


Phylogenetic Tree
778 Ecoli strains
=6x our 2012 Set

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

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