Application of ‘Omics Technology to Infectious Diseases and the Human Microbiome

Karen Nelson
Marcus Jones

J. Craig Venter™ INSTITUTE

Historical…

- Genome of *Haemophilus influenzae* - 1995
- Reverse vaccinology - 2000
- Sargasso Sea Study - 2004
- First Human microbiome publication - 2006
- Diploid human genome - 2007
- Genome transplantation - 2007
- Global Ocean Survey, GOS - 2007
- Synthetic microbial genome - 2008
- >11,000 influenza genomes (75% of total worldwide and ongoing)
- Sequenced most major pathogens (*e.g.* TB, malaria, cholera, *T. parva*, *T. cruzi*)
- Vaccine development program – 2010
- Bacterial cell controlled by a synthetic chromosome – 2010

First Genome Sequenced 1995

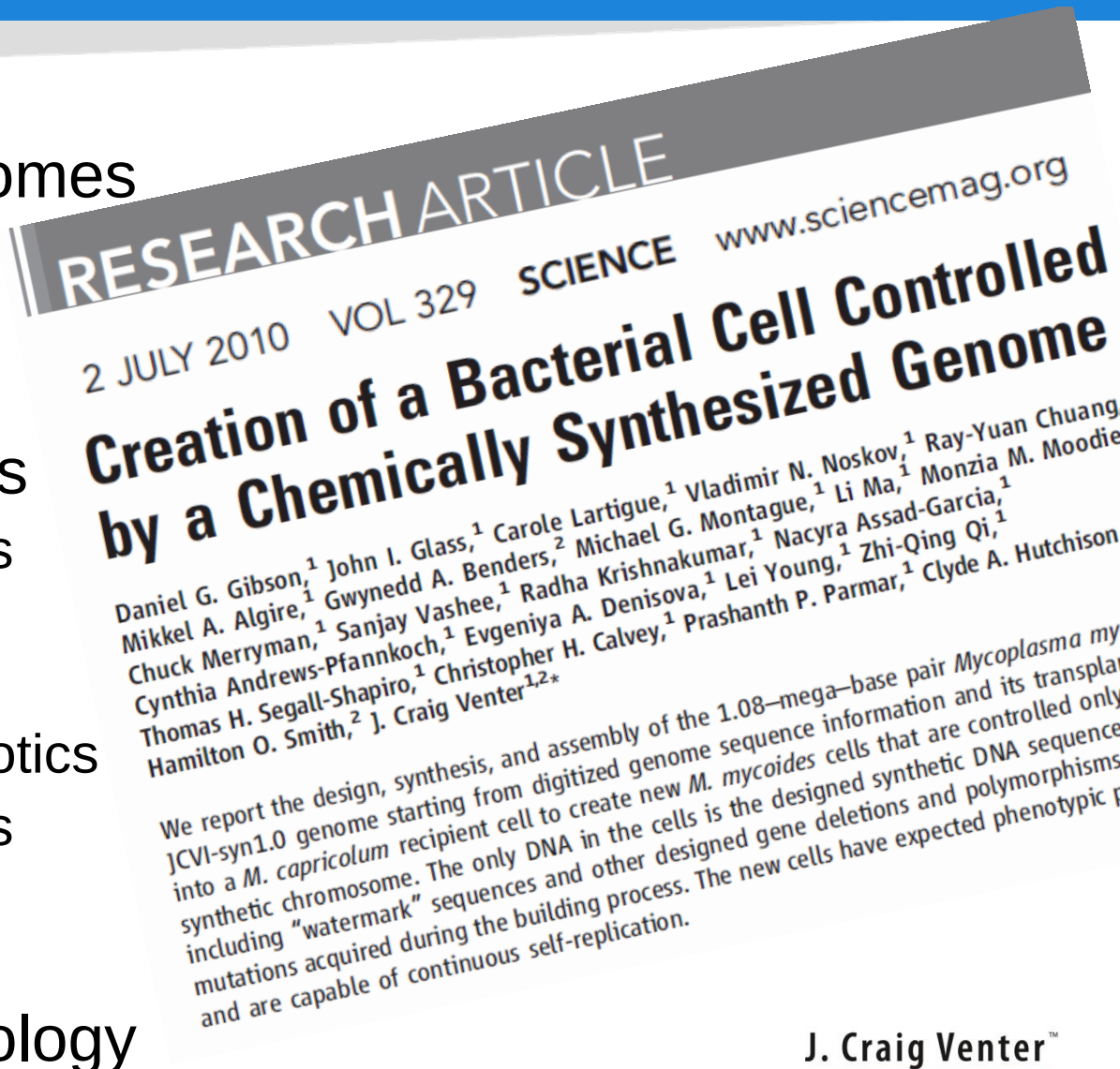
25 thousand sequences

6.25 x 10^8 pairwise comparisons

J. Craig Venter™
INSTITUTE

DNA synthesis Makes "Impossible" Genetic Manipulations Doable in Real Time

- We can synthesize genes and chromosomes cheaply and rapidly
- Enormous potential for new health and industrial applications
  - Production of biofuels
  - Small molecule therapeutics
  - New vaccines, antibiotics
  - Therapeutic microbes
  - Chloroplasts as plant factories
- Understand basic biology



RESEARCH ARTICLE

2 JULY 2010 VOL 329 SCIENCE www.sciencemag.org

Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome

Daniel G. Gibson,[1] John I. Glass,[1] Carole Lartigue,[1] Vladimir N. Noskov,[1] Ray-Yuan Chuang,[1] Mikkel A. Algire,[1] Gwynedd A. Benders,[2] Michael G. Montague,[1] Li Ma,[1] Monzia M. Moodie,[1] Chuck Merryman,[1] Sanjay Vashee,[1] Radha Krishnakumar,[1] Nacyra Assad-Garcia,[1] Cynthia Andrews-Pfannkoch,[1] Evgeniya A. Denisova,[1] Lei Young,[1] Zhi-Qing Qi,[1] Thomas H. Segall-Shapiro,[1] Christopher H. Calvey,[1] Prashanth P. Parmar,[1] Clyde A. Hutchison,[1] Hamilton O. Smith,[2] J. Craig Venter[1,2]*

We report the design, synthesis, and assembly of the 1.08–mega–base pair *Mycoplasma my* JCVI-syn1.0 genome starting from digitized genome sequence information and its transpla into a *M. capricolum* recipient cell to create new *M. mycoides* cells that are controlled only synthetic chromosome. The only DNA in the cells is the designed synthetic DNA sequence including "watermark" sequences and other designed gene deletions and polymorphisms mutations acquired during the building process. The new cells have expected phenotypic p and are capable of continuous self-replication.

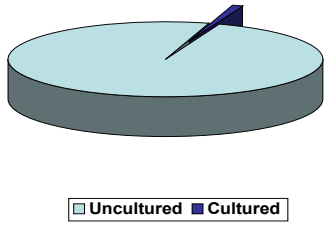
J. Craig Venter™ INSTITUTE

Metagenomics

- **We are capable of sequencing and analyzing the genomes of culturable species**
- **These species are estimated to represent less than 1% of total microbial diversit**

Culture dependent analysis:
Culture and obtain pure colonies
Complete genome sequencing of DNA
Organism has to be cultured in the laboratory

Culture-independent analysis
16S ribosomal RNA (rRNA) sequencing
Whole genome sequencing, assembly, annotation



- **Metagenomics:** sequence based analysis of complete microbial communities without need for culturing

Made possible by number of parallel developments:

- **Assembly and data analysis capabilities developed to being able to tease apart these large datasets**
- **Sequencing capabilities capable of achieving great depths of coverage at reduc cost**
- **Demonstrated proof of concept via Sargasso Sea study**
- **Global Ocean Sampling (GOS) largest protein dataset in existence**

Other "omics" technologies. Proteomics, metatranscriptomics, metabolomics

J. Craig Venter™ INSTITUTE

Changes in Sequencing Technologies

ABI 3730*xl* ***1-2 Mb/day***

Illumina GA IIx
50 Gb/12day run

ABI SOLiD
100Gb/12 day run

454 GS FLX +
0.6Gb/23hr run

Illumina HiSeq 2000 (2500†)
600 Gb/11day run

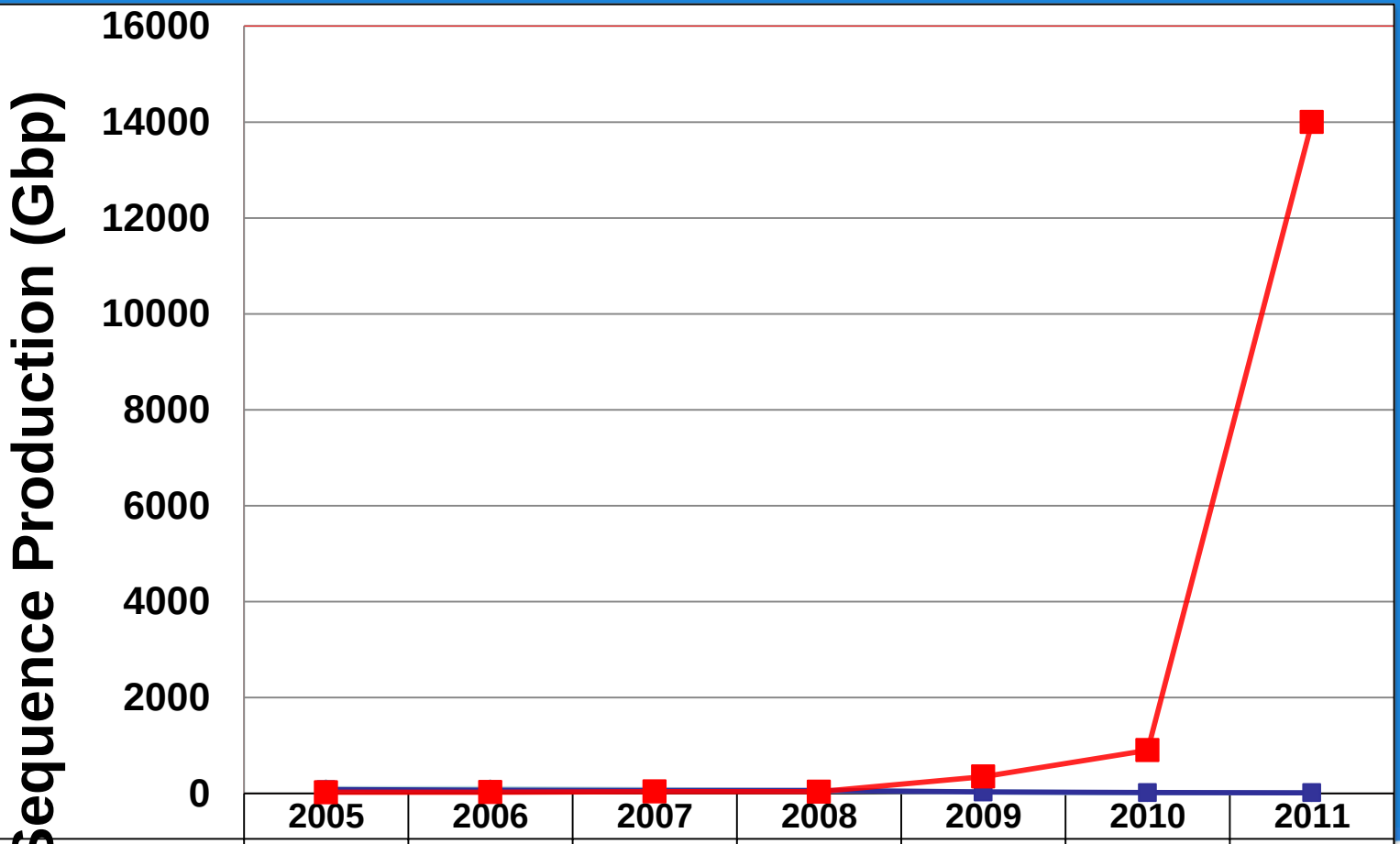
Ion Torrent
1Gb/2hr run

Ion Proton‡

† ***HiSeq 2500 upgrade****: up to 120Gb/27 hour run (available now for $50K)*
‡ ***Ion Proton****: up to 100Gb/4 hour run (available at the end of 2012)*

J. Craig Venter™ INS J. Craig Venter™ INSTITUTE

Changes in genomics sciences



	2005	2006	2007	2008	2009	2010	2011
# of Sequencers	100	100	78	55	29	9	9
FTE	77	70	64	60	30	17	13
Sequence data produced (Gbp)	24	27	38	36	350	900	14000

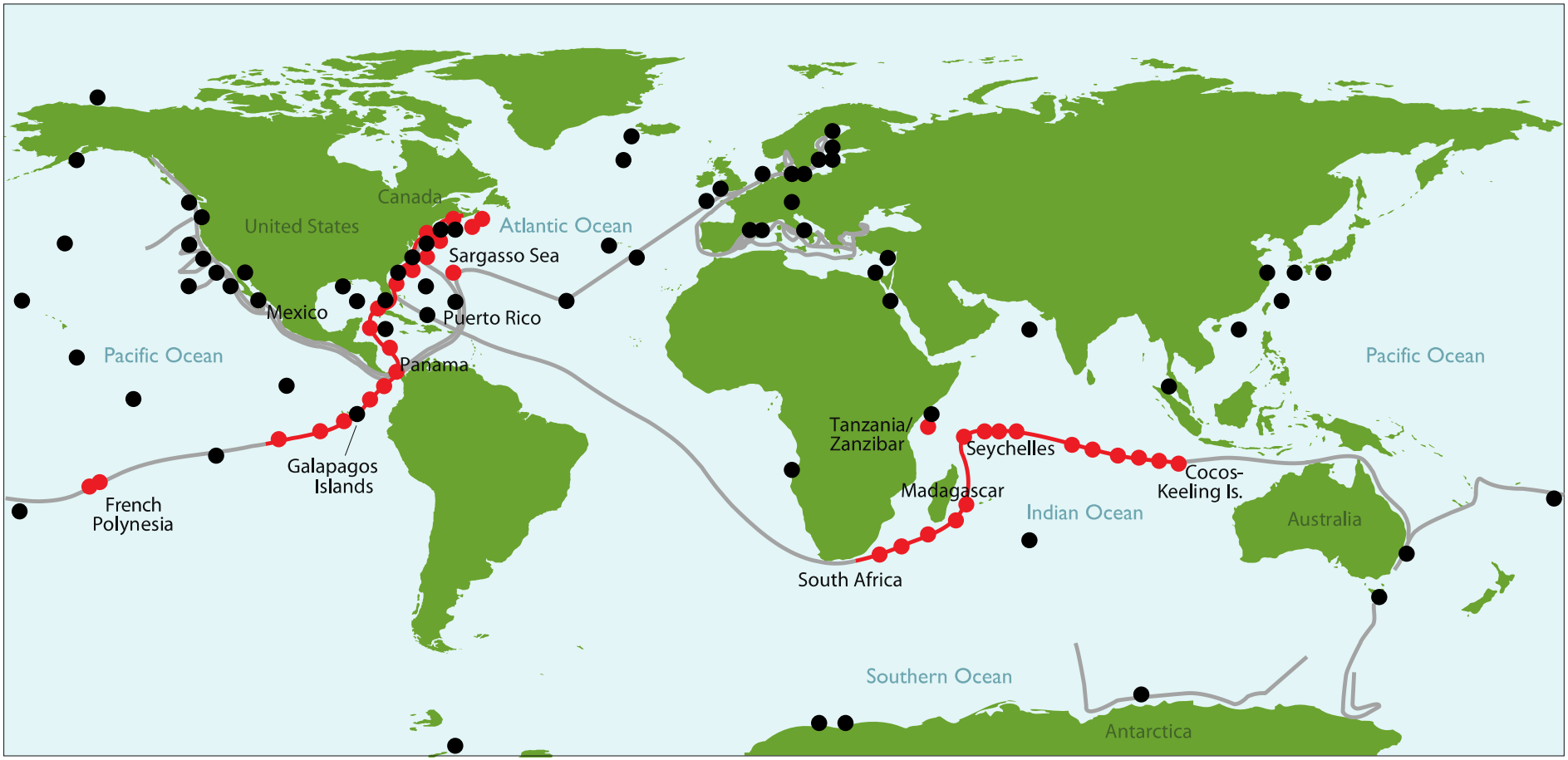
J. Craig Venter INSTITUTE

Sargasso Sea study

- Venter and colleagues at the JCVI
- Generated 1,987,936 DNA reads
- Approximately 1, 625 Mb of DNA
- 1.2 million new genes identified
- ~1,412 rRNA genes
- Estimated 1,800 species
- 12 complete genomes recovered
- Demonstration of the power of genomics

J. Craig Venter™ INSTITUTE

Global Ocean Sampling Expedition

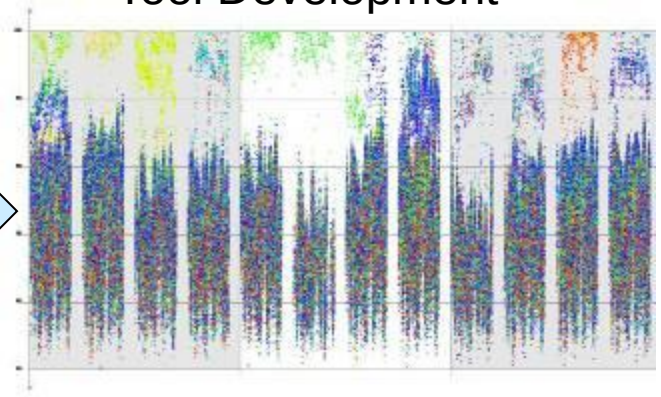


J. Craig Venter® INSTITUTE

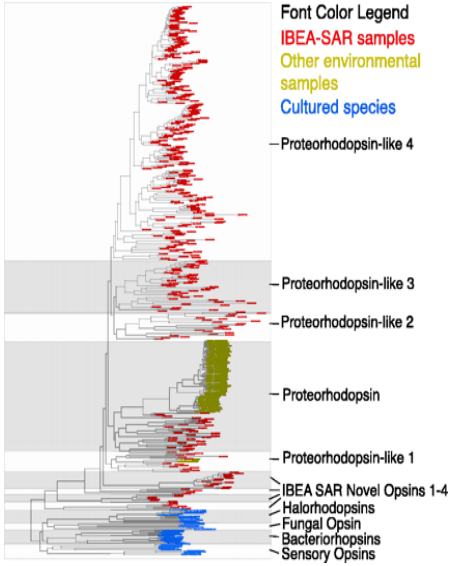
Global Ocean Sampling and Analysis

Sampling and Sequencing

Tool Development



Data Analysis



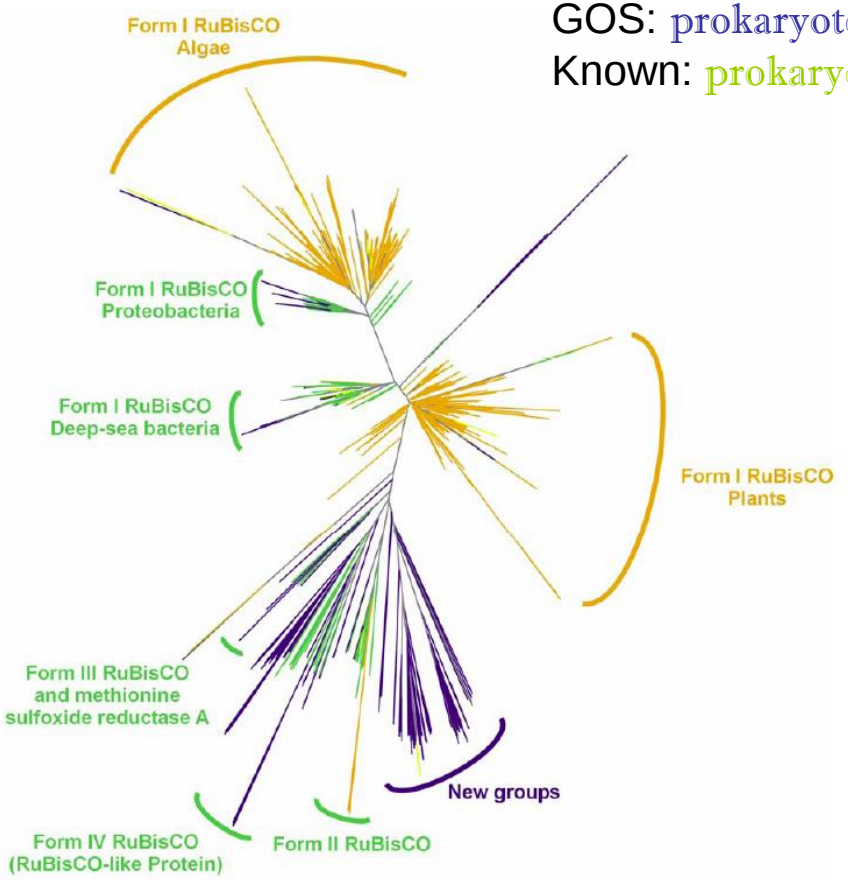
J. Craig Venter™ INSTITUTE

GOS increases size and diversity of known protein families

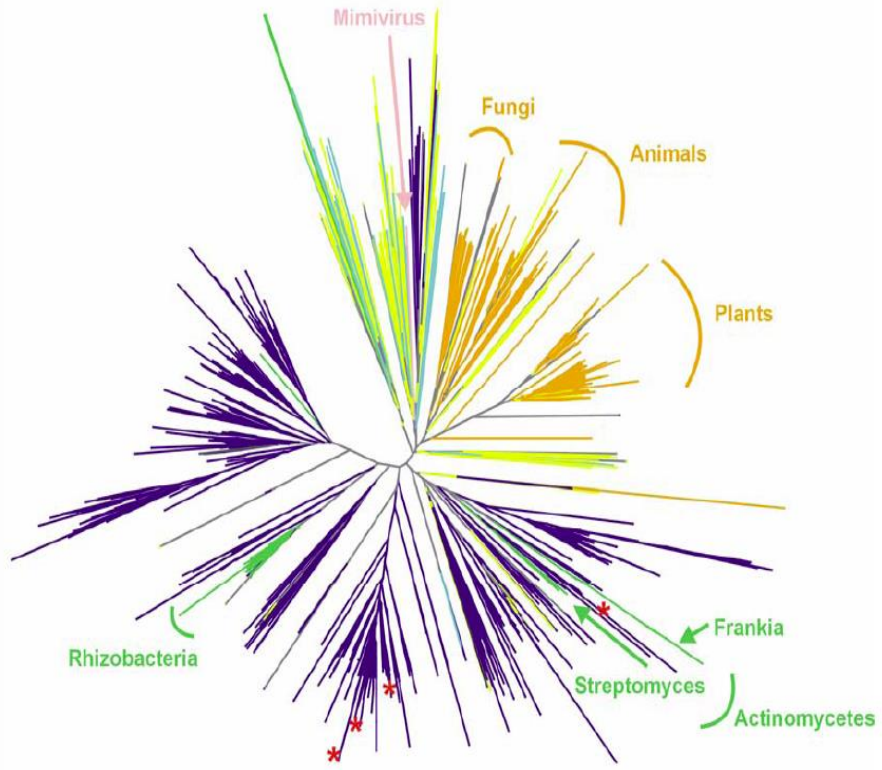
(Yooseph et al, 2007 *PLoS Biol*)

GOS: prokaryotes, eukaryotes

Known: prokaryotes, eukaryotes



RuBisCO



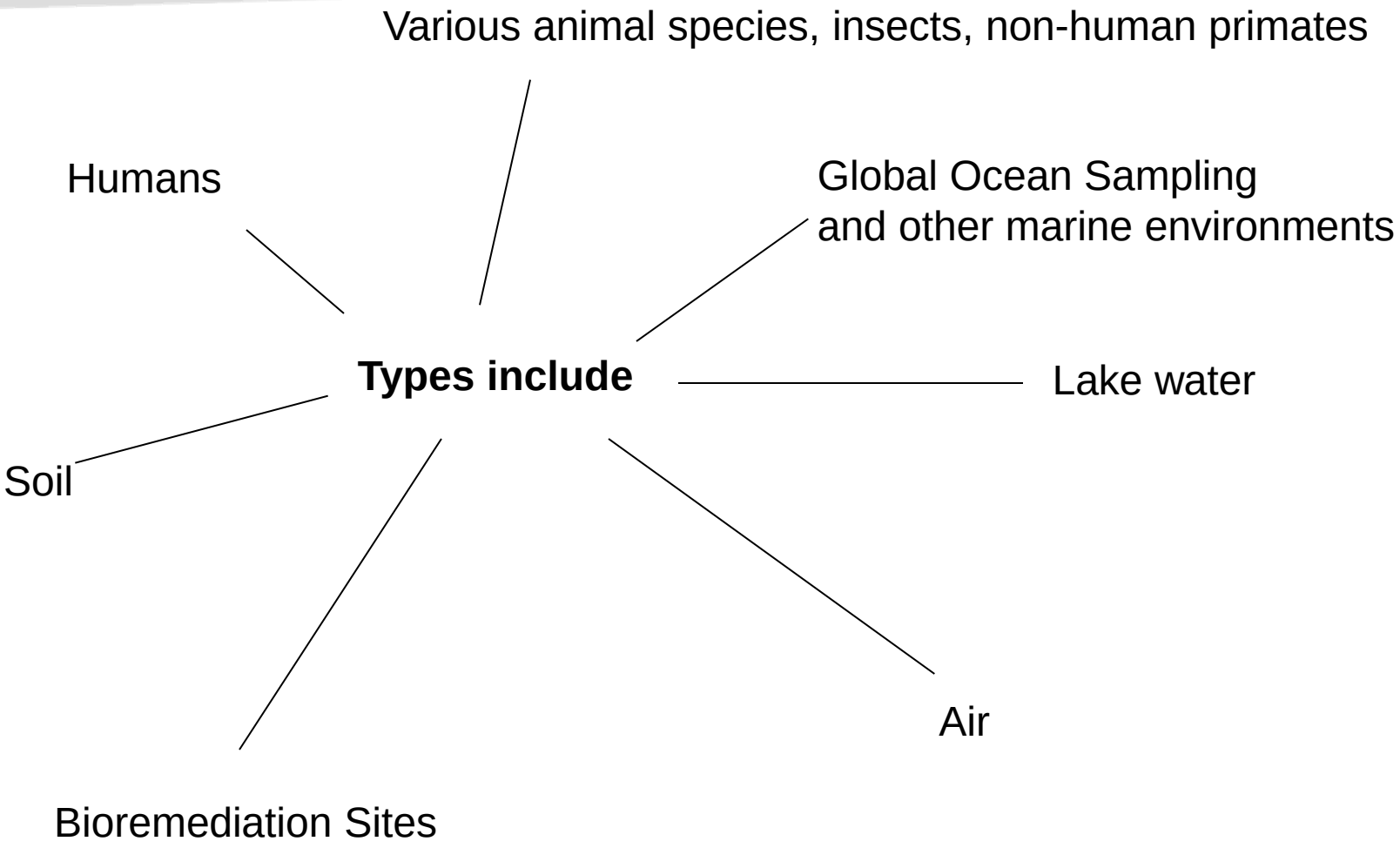
Glutamine synthetase (type II)

J. Craig Venter INSTITUTE

Spin off “omics” studies transcriptomics – metabolomics

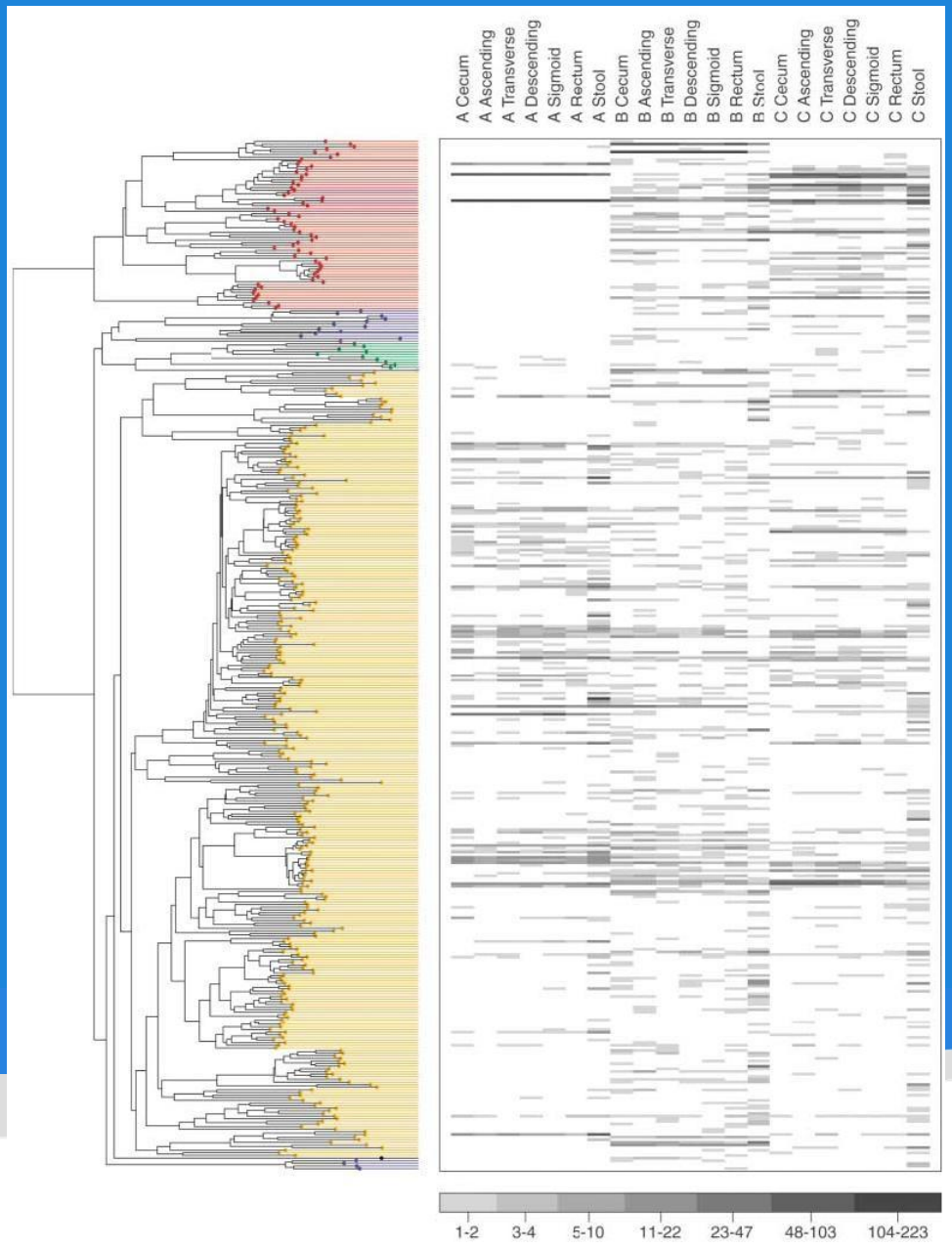
J. Craig Venter™
INSTITUTE

Metagenomic projects



J. Craig Venter™
INSTITUTE

Human Colon



Mucosal samples were obtained during colonoscopy from healthy-appearing sites within the six major subdivisions of the human colon: cecum

ascending colon

transverse colon

descending colon

sigmoid colon

rectum.

Fecal samples were collected from each subject 1 month following colonoscopy.

From 11,831 bacterial and 1524 archaeal 16S sequences, identified 395 phylotypes

J. Craig Venter™ INSTITUTE

Eckburg et at., 2005 Science

Stomach

1,833 full-length 16S sequences

Described 128 16S rDNA phylotypes

Derived from 23 human subjects

Bik, E.M. et al. (2006) PNAS 103, 732-737

PNAS

- First human metagenomic paper
- Investigated the gastrointestinal tract (via fecal samples) of two healthy adult
- 78 Mbp
- 2062 amplified 16S rDNA

RESEARCH ARTICLE

Metagenomic Analysis of the Human Distal Gut Microbiome

Steven R. Gill,[1*‡] Mihai Pop,[1†] Robert T. DeBoy,[1] Paul B. Eckburg,[2,3,4] Peter J. Turnbaugh,[5] Buck S. Samuel,[5] Jeffrey I. Gordon,[5] David A. Relman,[2,3,4] Claire M. Fraser-Liggett,[1,6] Karen E. Nelson[1]

The human intestinal microbiota is composed of 10^{13} to 10^{14} microorganisms whose collective genome ("microbiome") contains at least 100 times as many genes as our own genome. We analyzed ~78 million base pairs of unique DNA sequence and 2062 polymerase chain reaction–amplified 16S ribosomal DNA sequences obtained from the fecal DNAs of two healthy adults. Using metabolic function analyses of identified genes, we compared our human genome with the average content of previously sequenced microbial genomes. Our microbiome has significantly enriched metabolism of glycans, amino acids, and xenobiotics; methanogenesis; and 2-methyl-o-erythritol 4-phosphate pathway–mediated biosynthesis of vitamins and isoprenoids. Thus, humans are superorganisms whose metabolism represents an amalgamation of microbial and human attributes.

Our body surfaces are home to microbial communities whose aggregate

≥100 times as many genes as our 2.85–billion base pair (bp) human genome (*1*). Therefore, a

of single organisms, recent reports from Venter *et al.* (*9*) and Baker *et al.* (*10*) have demonstrated the utility of this approach for studying mixed microbial communities. Variations in the relative abundance of each member of the microbial community and their respective genome sizes determine the final depth of sequence coverage for any organism at a particular level of sequencing. This means that the genome sequences of abundant species will be well represented in a set of random shotgun reads, whereas lower abundance species may be represented by a small number of sequences. In fact, the size and depth of coverage (computed as the ratio between the total length of the reads placed into contigs and the total size of the contigs) of genome assemblies generated from a metagenomics project can provide information on relative species abundance.

A total of 65,059 and 74,462 high-quality sequence reads were generated from random DNA libraries created with fecal specimens of two healthy humans (subjects 7 and 8). These two subjects, ages 28 and 37, female and male,

Gill et al, *Science* 200

J. Craig Venter INSTITUTE

Human Microbiome
Metagenomics, Health and Disease

J. Craig Venter™
INSTITUTE

Human Microbiome

~10^{12} Human cells

~10^{13} Bacterial cells

>600 oral bacterial species

J. Craig Venter™ INSTITUTE

Human Microbiome

- Collective of the human microbiome exceeds the number of human cells by at least an order of magnitude.
- Many of these microbial interactions endow or enhance human physiology including processes related to development, nutrition, immunity and resistance to pathogens.
- The majority of the human microbiome remains unknown.
- Many relationships between the human host and microbiome remain to be determined

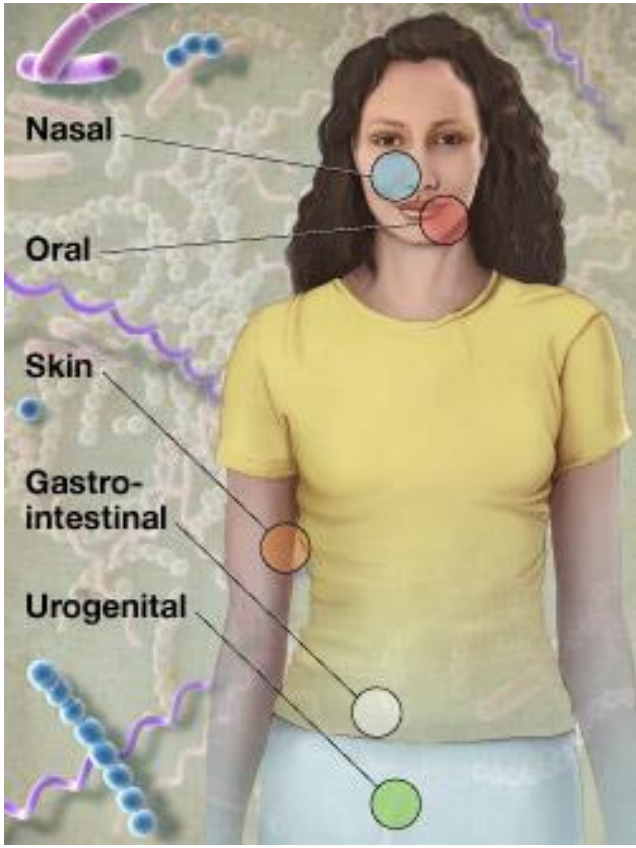


image courtesy of the NIH HMP website
http://nihroadmap.nih.gov/hmp/

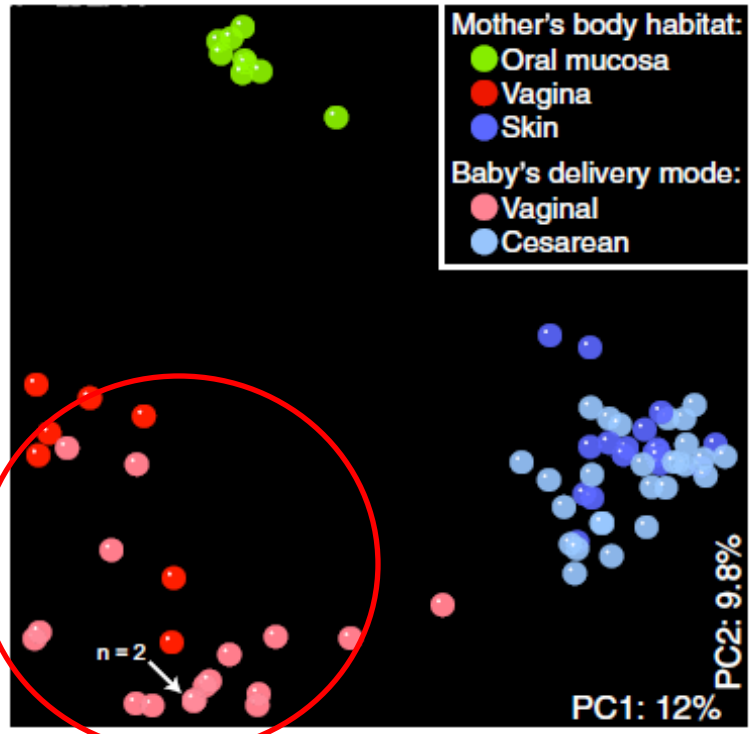
J. Craig Venter™
INSTITUTE

The Human Microbiome

Significant role: Example in the Gastrointestinal tract

- **They foster development of the mucosal wall.**
- **The development and maturation of the immune system is dependent on the presence of some members of the intestinal microbiota. Link to human health and disease.**
- **Essential for the metabolism of certain compounds as well as xenobiotics.**
- **Protection against epithelial cell injury.**
- **Regulation of host fat storage.**
- **Stimulation of intestinal angiogenesis.**

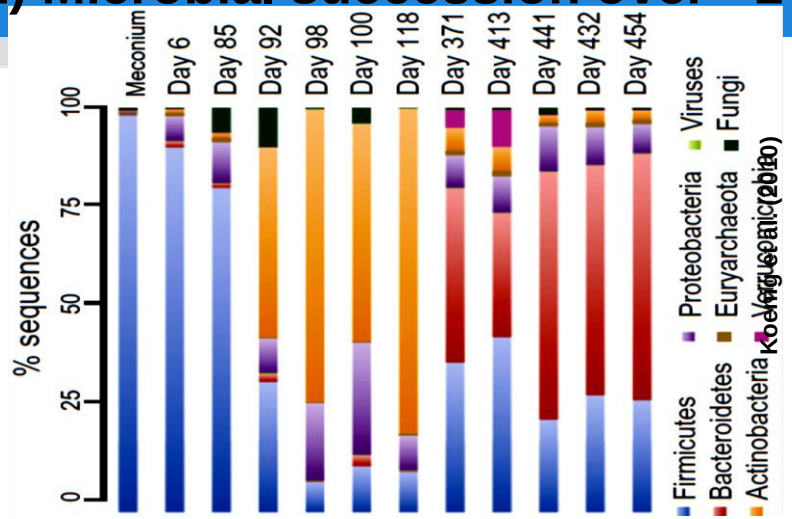
Microbiota are acquired *anew* each generation.



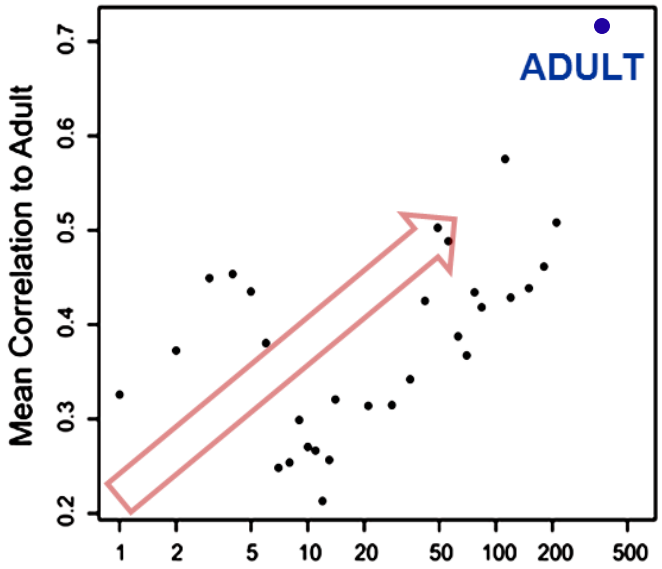
Dominguez-Bello et al. (2010).

1) Infants obtain microbes from mother or environment.

2) Microbial succession over ~1-2 yrs.



Koenig et al. (2010)



Palmer et al. (2007)

3) Microbiome becomes "adult-like" in ~1-2 yrs.

NIH Roadmap Human Microbiome Project

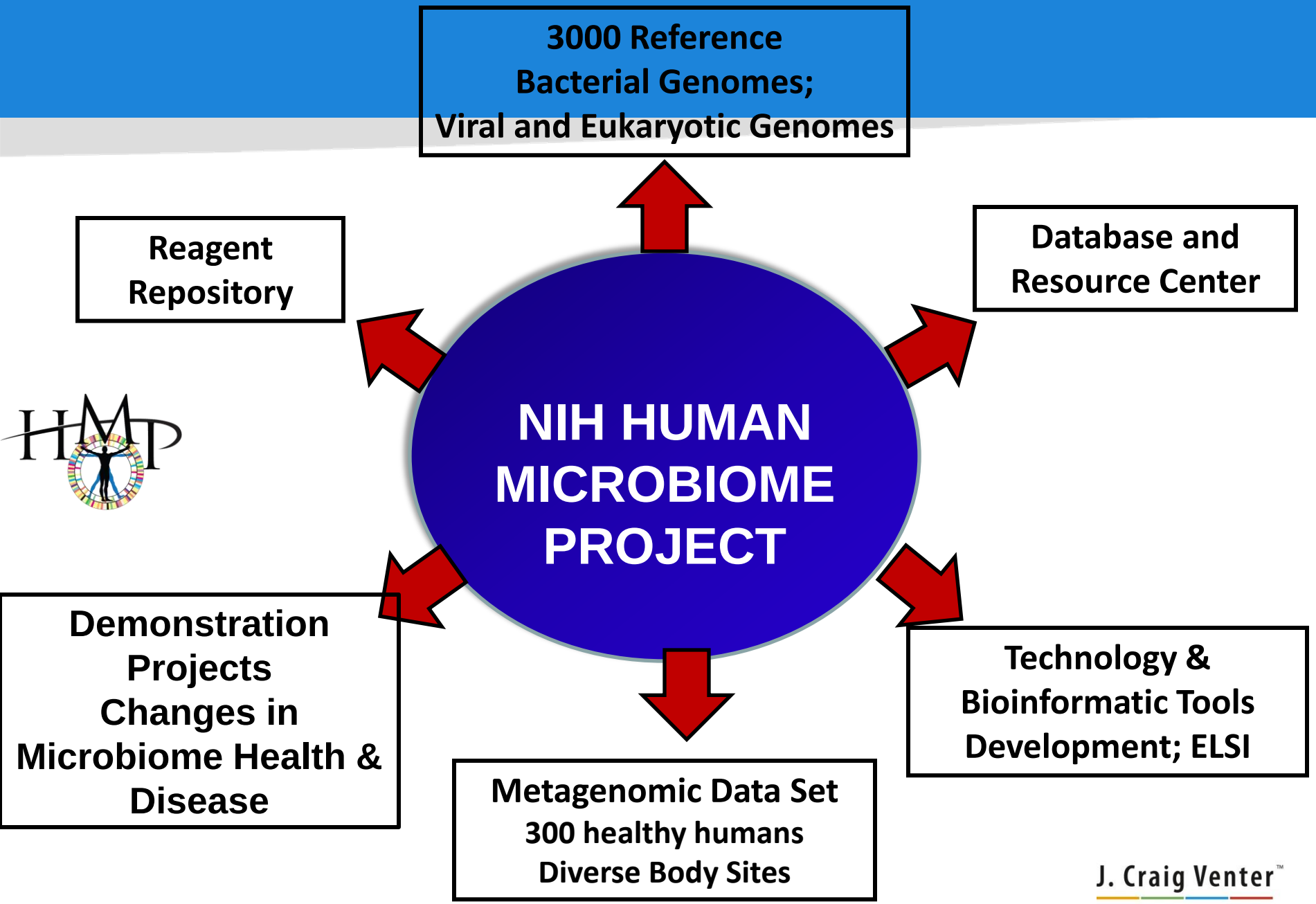
- **Budget > $175 million 2007-2013**
- **Goal: Characterize the microbes that inhabit the human body and examine whether changes in the microbiome can be related to health and disease**
- **Feasibility project designed to determine the value of microbial metagenomics to biomedical research**
- **Community Resource Project-generate reagents and data sets; rapidly placed in public domain**
- **Continuous Scientific Community Input**

  External Scientific Advisory Group, Workshops.

- http://nihroadmap.nih.gov/hmp
- http://www.human-microbiome.org/#

J. Craig Venter INSTITUTE

Slide courtesy Maria Giovanni-NIAID



Slide courtesy Maria Giovanni-NIAID

“Healthy Cohort” Body Sites

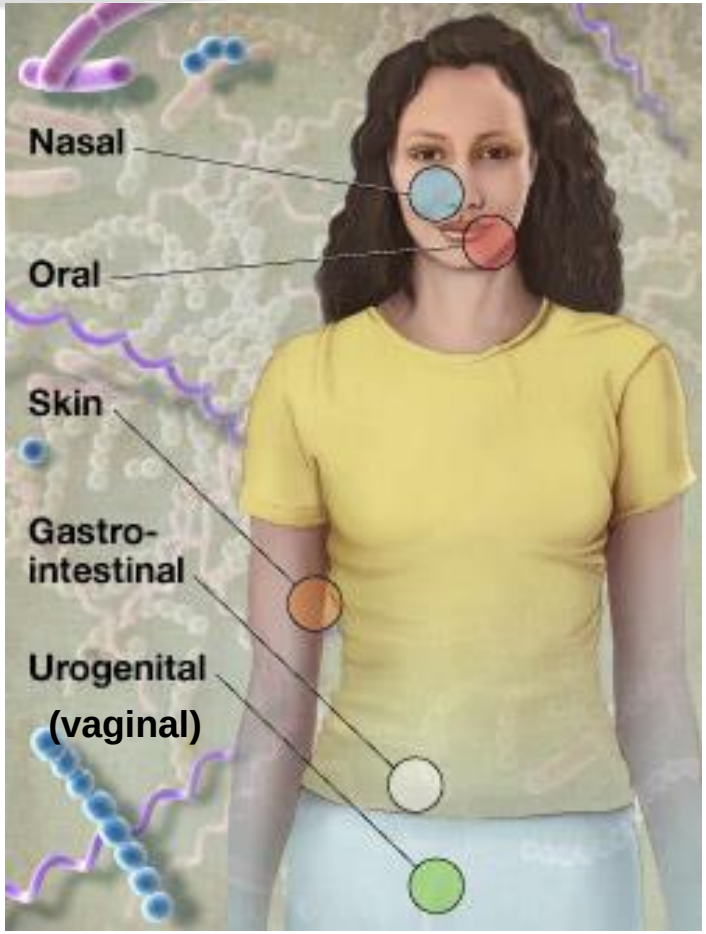
Oral
- Saliva
- Tongue dorsum
- Hard palate
- Buccal mucosa
- Keratinized (attached) gingiva
- Palatine tonsils
- Throat
- Supragingival plaque
- Subgingival plaque

Ski
- Retroauricular crease, both ears (2)
- Antecubital fossa (inner elbow), both arms (2)

Nasal
- Anterior right and left nares (pooled)

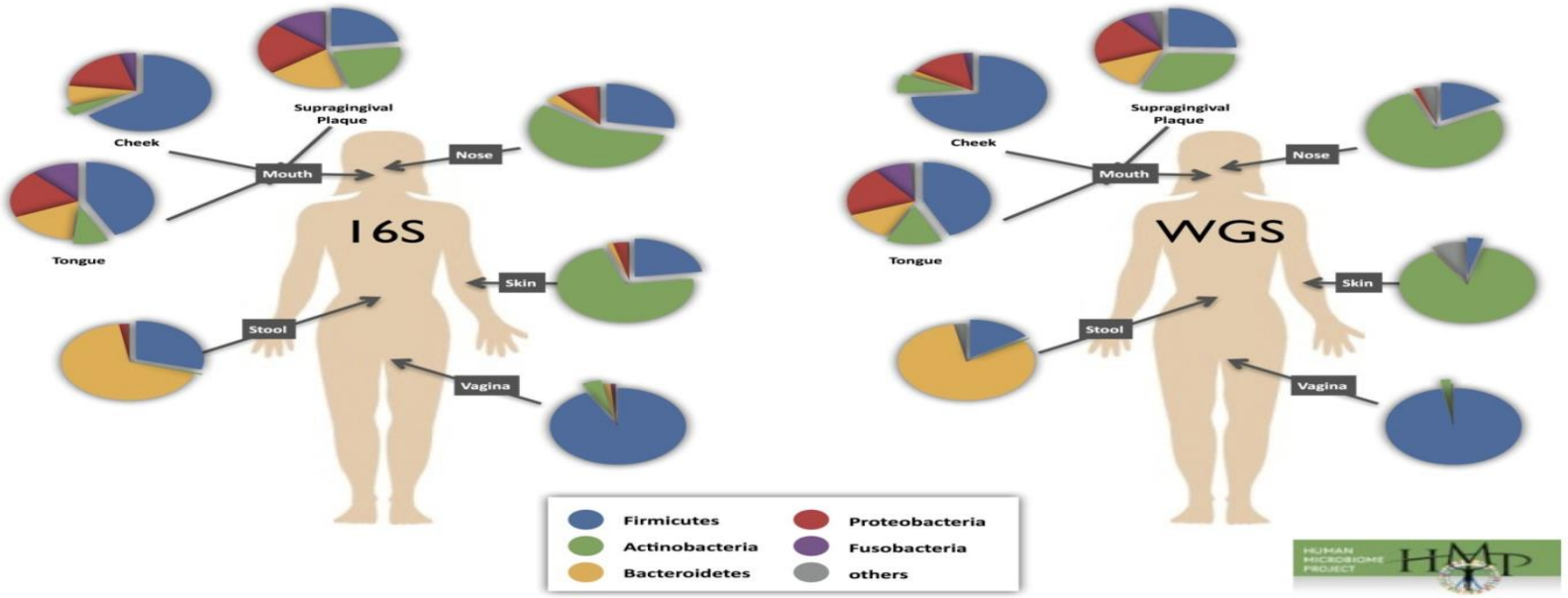
Gut
- Stool

Vaginal
- Posterior fornix, vagina
- Midpoint, vagina
- Vaginal introitus



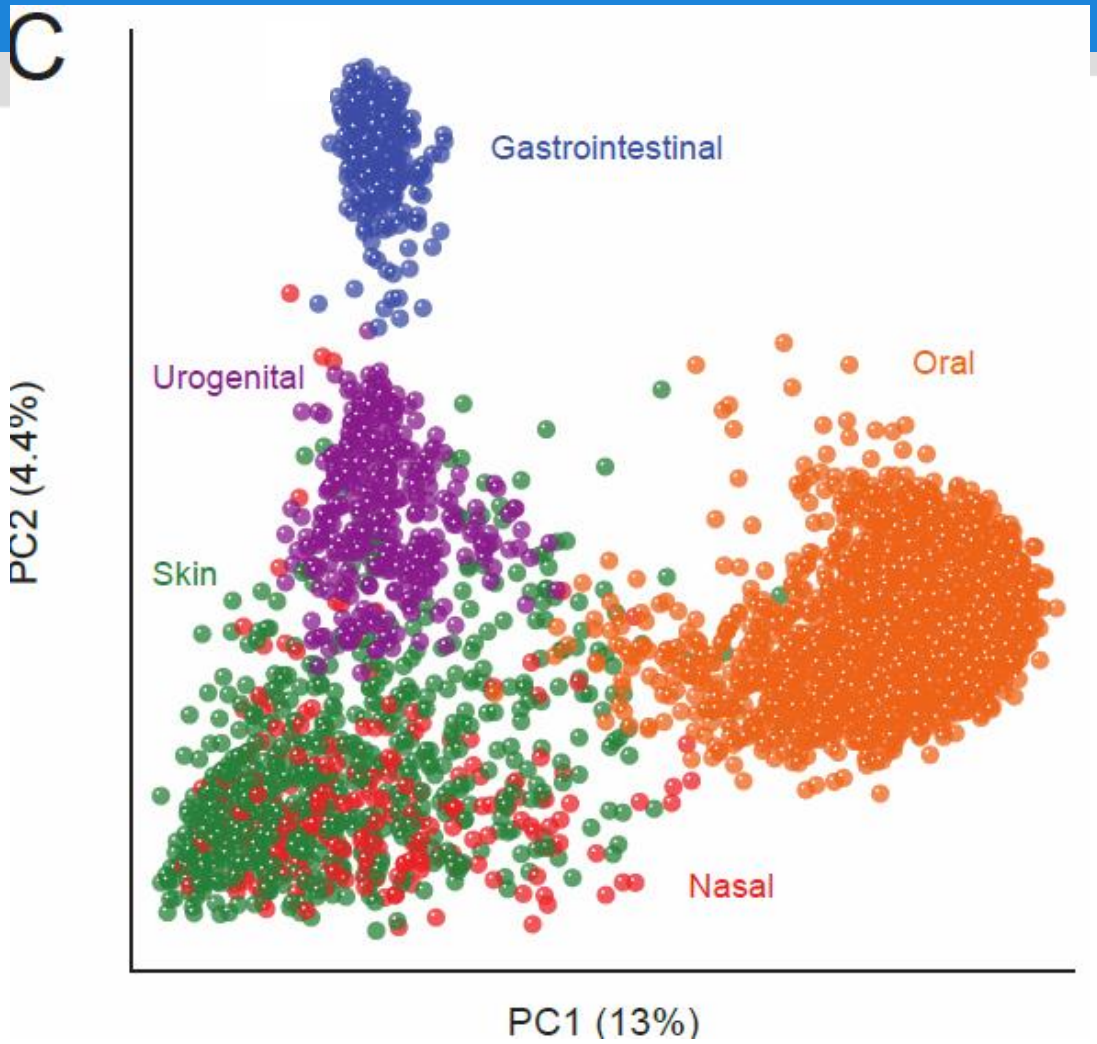
Slide courtesy of NHGRI

J. Craig Venter INSTITUTE



Supplementary Figure 8. Phylum abundances per body site. For each of the body sites studied by both 16S rRNA gene sequencing (A) and whole-genome shotgun sequencing (B) the five most abundant phyla are shown. The small remaining fraction of the data is collapsed and labeled as other phyla (grey).

In adults, each part of the body supports a distinct microbial community.



With no apparent relationship with gender, age, weight, ethnicity or race.

J. Craig Venter™ INSTITUTE

"Structure, Function and Diversity of the Human Microbiome in an Adult Reference Population" The Human Microbiome Consortium. **HMP Consortium (2012)**

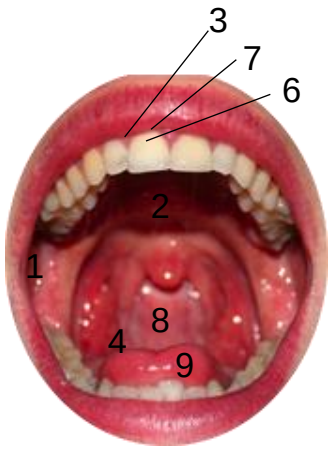
Some results from HMP:

HMP estimates for global microbiome:

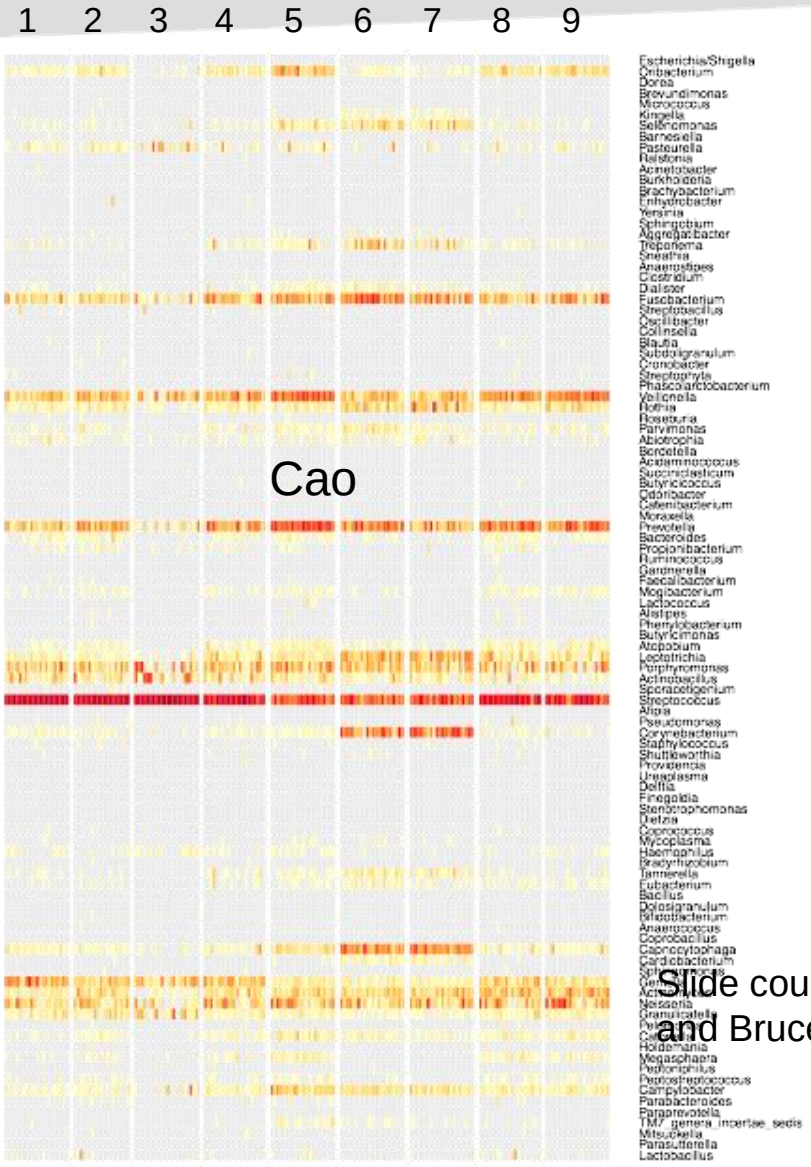
~ 10,000 microbial species

~ 8 million microbial genes

Sub-body sites have distinct communities



Soft	1	Cheek
	2	Palate
	3	Gums
	4	Tonsils
	5	Saliva
	6	Subgingival Plaque
	7	Supragingival Plaque
Hard	8	Throat
	9	Tongue



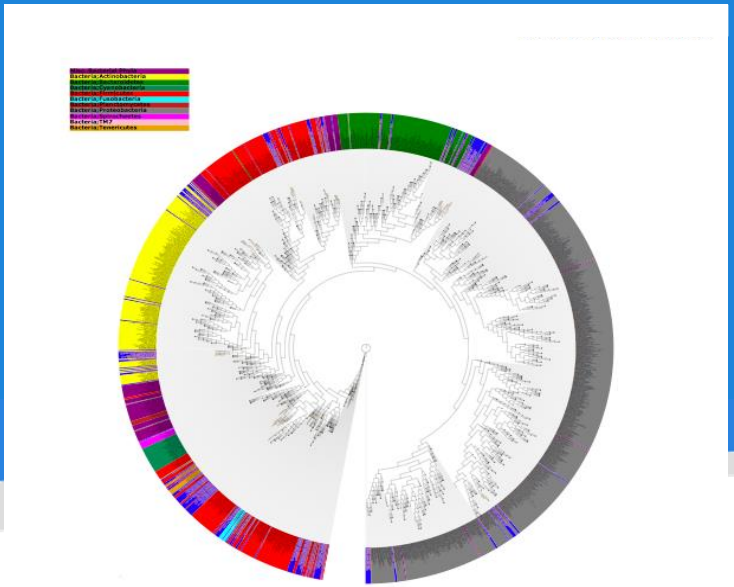
Slide courtesy of HMP Consortium and Bruce Birren, Broad Institute

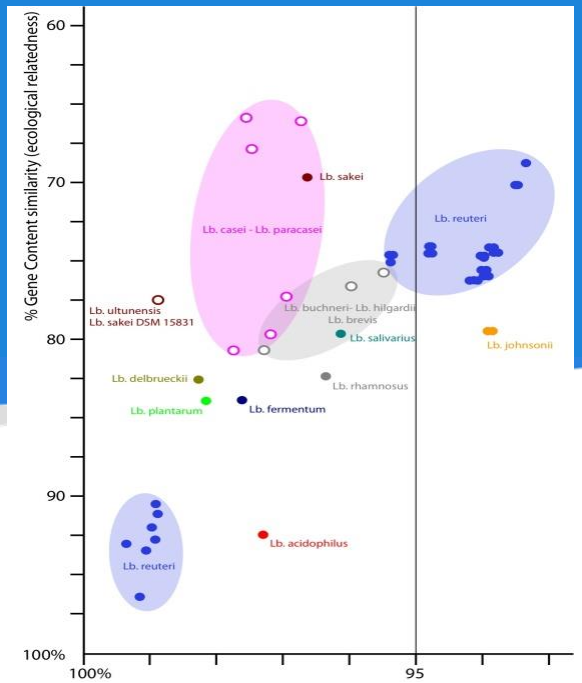
J. Craig Venter INSTITUTE

- **Reference Strains: Generate complete genomes from > 3000 prokaryotes.**
- **Build our understanding of those recognized through 16S profiles**
- **Provide for interpretation of metagenomics and other "omics" data**
- **Sequence reference phage, viruses and eukaryotes**

A Catalog of Reference Genomes from the Human Microbiome

178 genomes
~550,000 genes
Nelson et al., Science
May 21, 2010





Reference Genomes of the Human Microbiome Project

In order to facilitate the phylogenetic and functional analysis of the metagenomic sequences produced from human body sites, the HMP plans to sequence, or collect from publicly available sources, a total of 1000 reference genomes. The organisms included in this collection have all been isolated from a human body site. The information gained from the Reference Genomes will allow 16S RNA sequences and metagenomic sequence from uncharacterized microbiome organisms to be grouped phylogenetically with related organisms from the reference set providing information about the taxonomy of the unknown strains. Likewise, functional characterization of proteins in the reference organisms will aid in the functional annotation of related proteins contained in the sequence fragments derived from metagenomic samples.

Choosing Reference Organisms:

The HMP has developed a detailed set of guidelines for inclusion of a strain in the reference genome group. If you have suggestions for additional strains to include on the list or if you have a strain that you would like to contribute please use our feedback form to let us know.

- Guidelines for inclusion of a strain
- Feedback form - help us by recommending strains to include in the HMP reference genome collection
- Current breakdown of strains according to body site
- Phylogenetic Analysis - Below are phylogenetic trees of HMP organisms in the context of a wide sampling of sequenced and/or culturable bacteria:
  - All HMP reference genomes
  - Reference genomes isolated from airways
  - Reference genomes isolated from the gastrointestinal tract
  - Reference genomes isolated from the oral cavity
  - Reference genomes isolated from the skin
  - Reference genomes isolated from the vagina

HMP Catalog

For a full list of the HMP reference genomes please visit the HMP Project Catalog where you can search for strains by many features and characteristics, including body site and taxonomy. The collection of strains in the HMP Project Catalog represents projects at all stages including those that are planned (project status "targeted") as well as those that have reached completion (project status "complete"). Also included in the set are strains that are being sequenced by members of the International Human Microbiome Consortium (link to further down section of page). More information on this effort can be found further down this page.

Most of the HMP Reference Genomes will be sequenced only to the "standard draft" stage, a minimum standard for a draft genome that has been established by the HMP sequencing centers. Draft genome sequence does not include every base of the genome, rather they are assemblies of several large contiguous pieces of sequence (contigs) with subsequent gaps in sequence knowledge. About 15% of the reference strains will be taken closer to a "finished" or fully complete state. There are several finishing levels that genomes can be taken to, each with an associated cost. The same guidelines mentioned above for choosing which strains to include on the list are applied to decide which of the strains should be promoted to a higher state of finishing. A standardized set of Finishing Categories is currently under development by a multi-center, international group of researchers. Once finalized, they will be posted on this site and each strain will be assigned to one of the categories.

The Human Microbiome: Altering the future of medicine

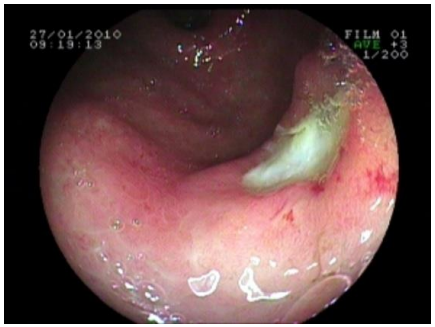
- Microbiome influenced by many factors including environment and host genetics
  - Complex bio-feedback mechanism: host <-> microbiome
- This population can be studied and altered to benefit the host
  - Normal flora of healthy individuals can potentially be mined to identify new probiotics
  - Population changes/shifts can be used as indicators of deterioration/improvement of health
  - Can be used for disease surveillance
- Need for integration of multiple “omics” approaches to understand the complexity of the microbiome and its broader implications

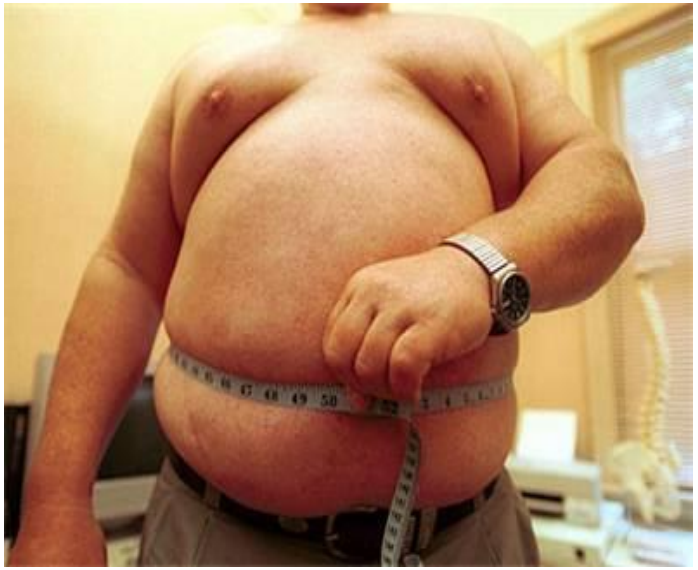
Disease related microbiome studies at JCVI

- Progression of esophageal cancer (NYU)
- Bacterial vaginosis and pre-term delivery (Illinois/Mayo; NIAID)
- Nasopharynx microbiome and vaccination in children (Gates)
- Skin microbiome, acne and psoriasis (NYU)
- Oral diseases including periodontitis (NYU)
- Colon cancer (Howard University)
- Type 1 Diabetes pilot (TEDDY)

Can we use as a biomarker for:

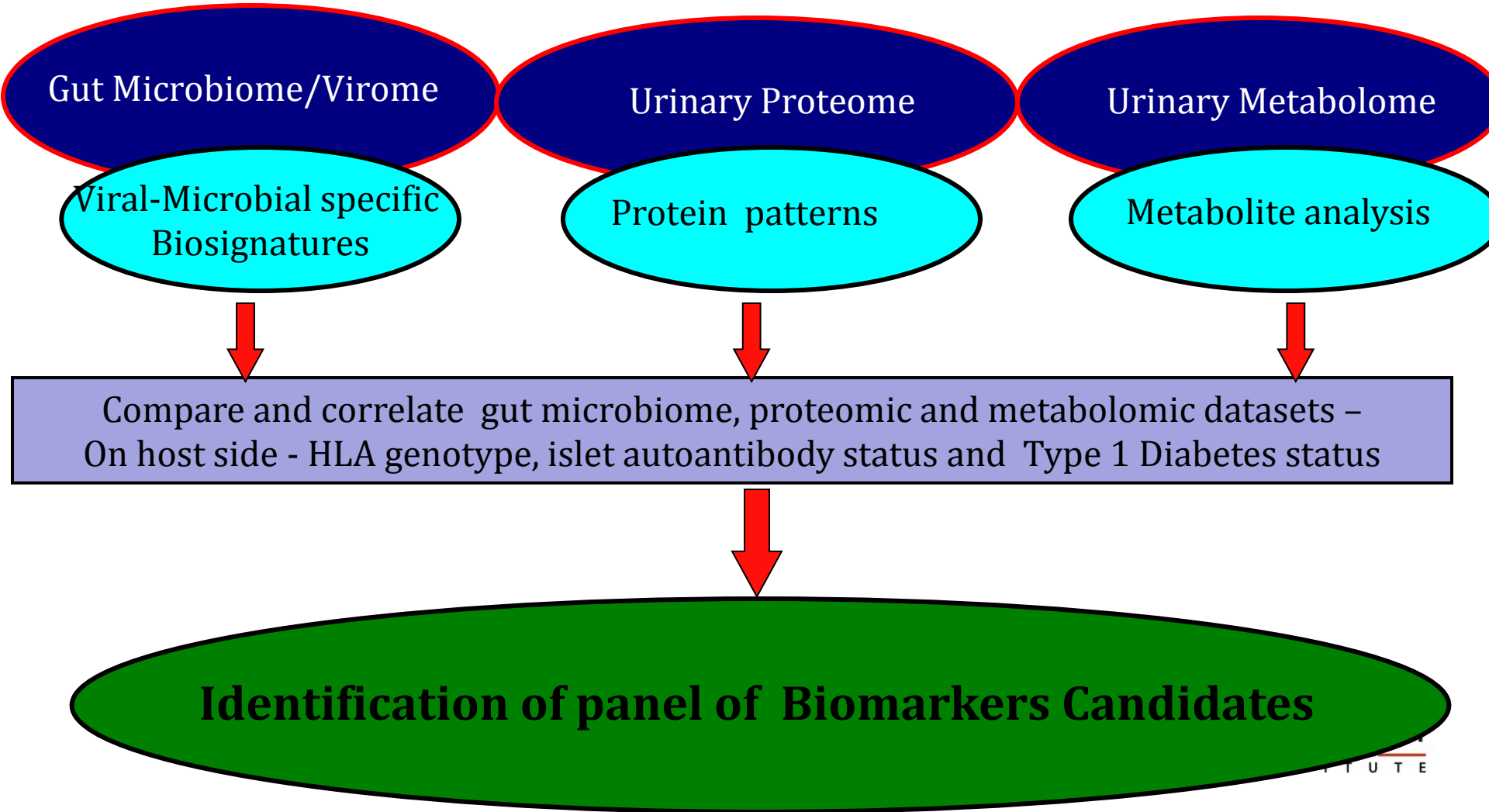
- Development of new predictive biomarkers so that preventive strategies based on pre- and probiotics can be developed.
- New therapeutic strategies
- Increase our understanding of the etiologies of complex diseases and health





Integrated "omics" approaches

NIDDK funded - Type 1 Diabetes Study



Still need:

- Technology development
- Informatics and data handling
- Education
- Well defined studies

J. Craig Venter™ INSTITUTE

Transcriptomic and Proteomic Analyses of the Microbiome and Infectious Diseases

J. Craig Venter™
INSTITUTE

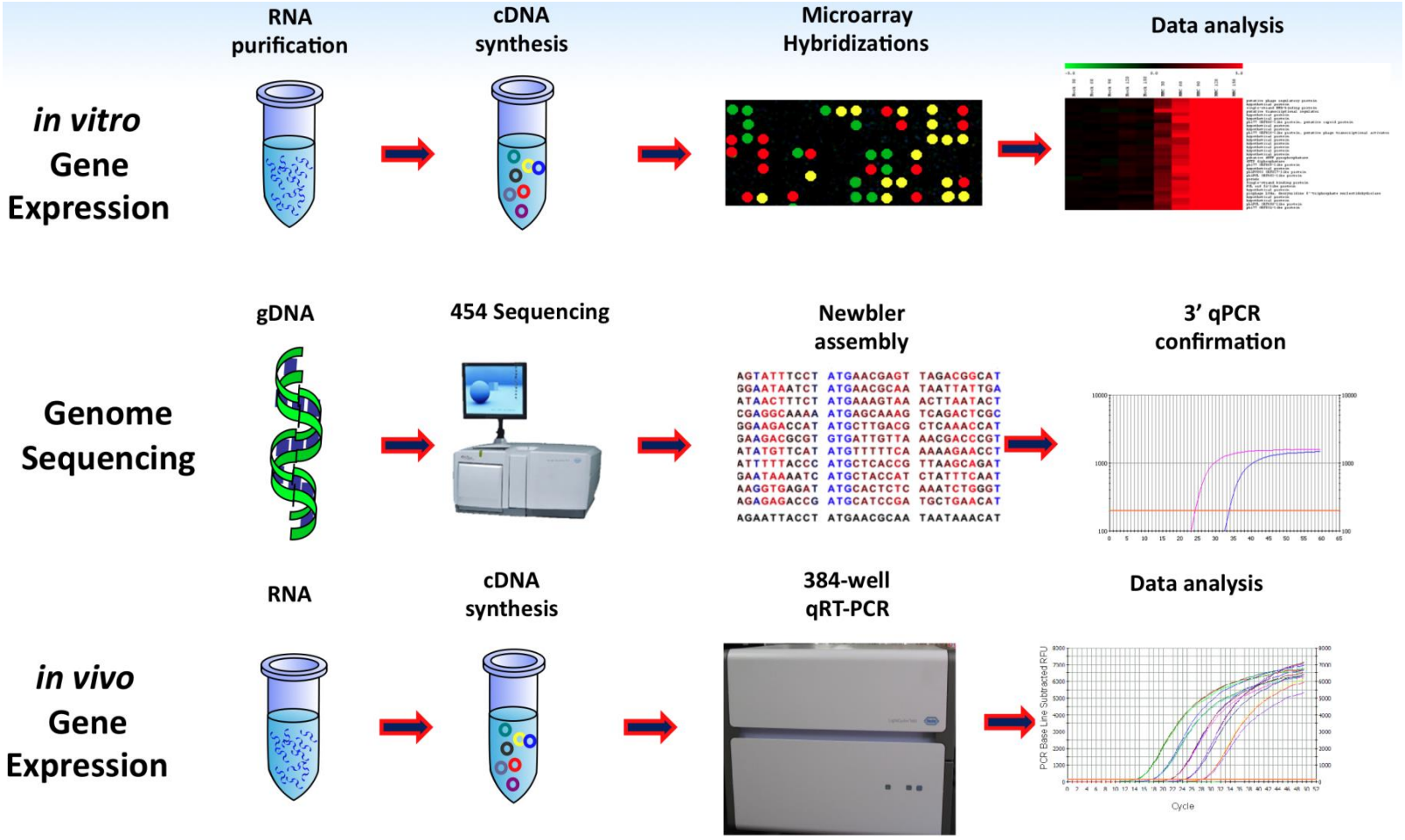
“Omics” Technologies

- Metagenomic Analysis
- Transcriptomics
- Proteomics
- Glycomics
- Lipodomics
- Metabolomics

J. Craig Venter™ INSTITUTE

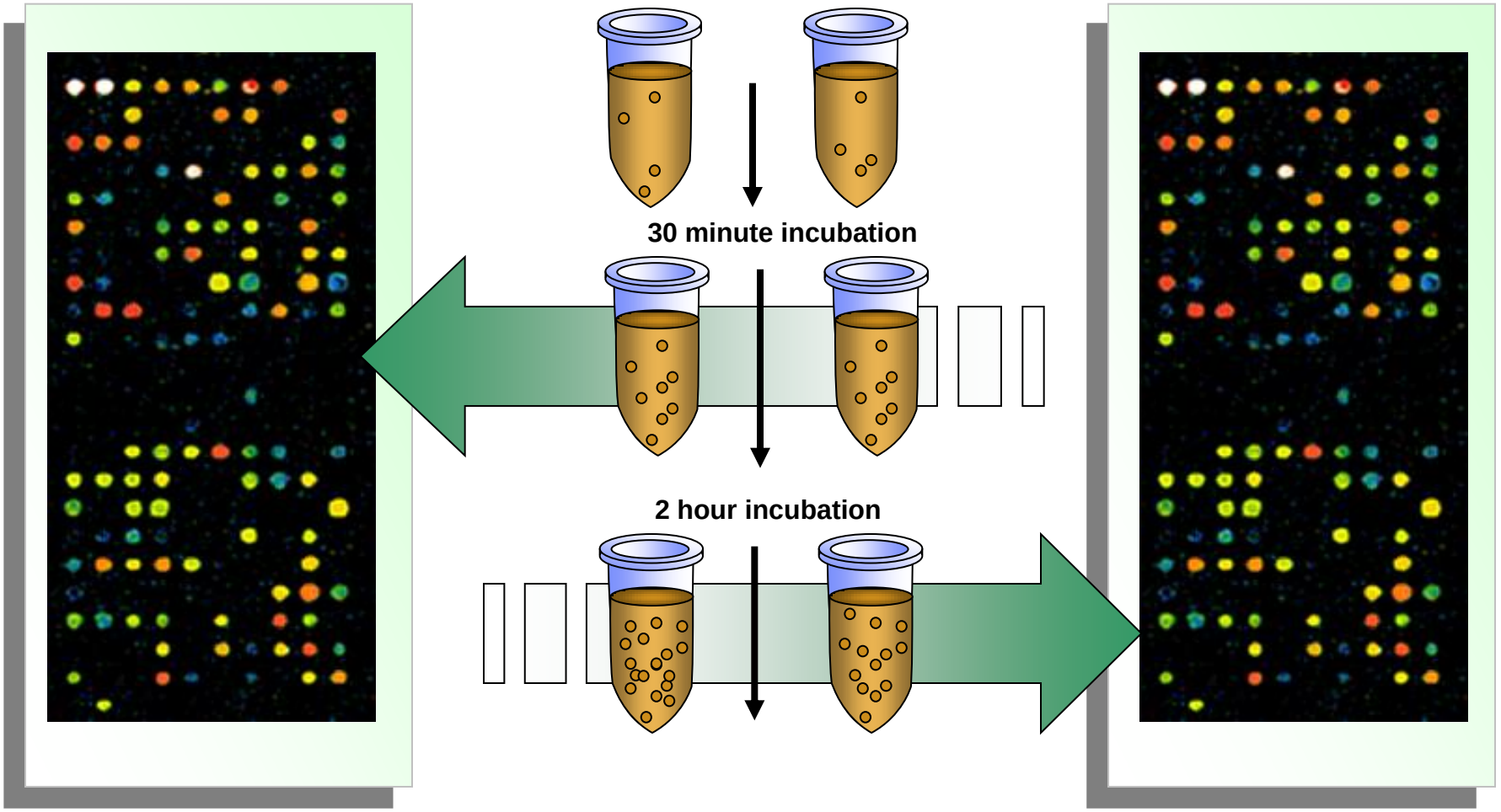
What is Transcriptomics?

Transcriptomics Technologies



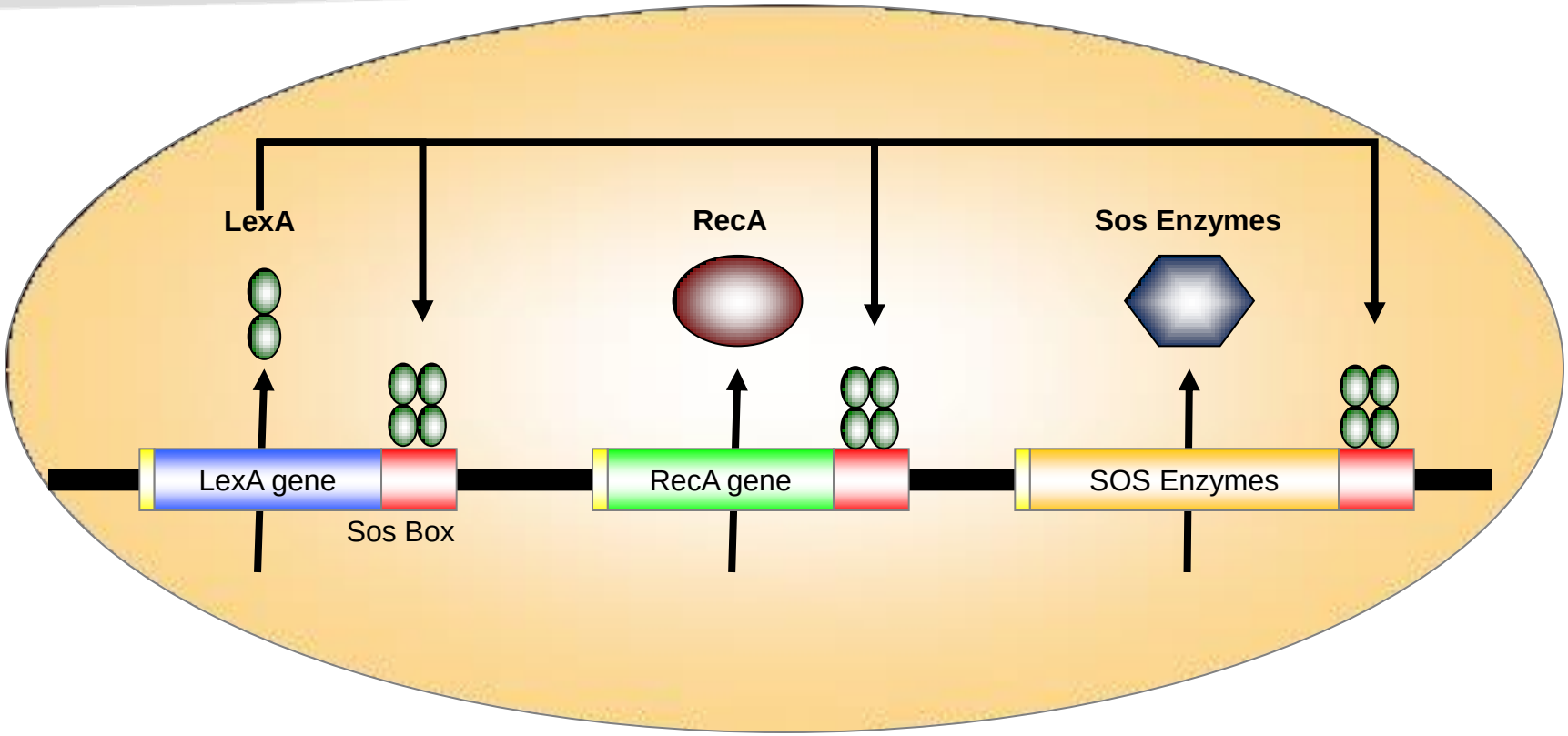
J. Craig Venter INSTITUTE

Characterization of *in vitro* Samples

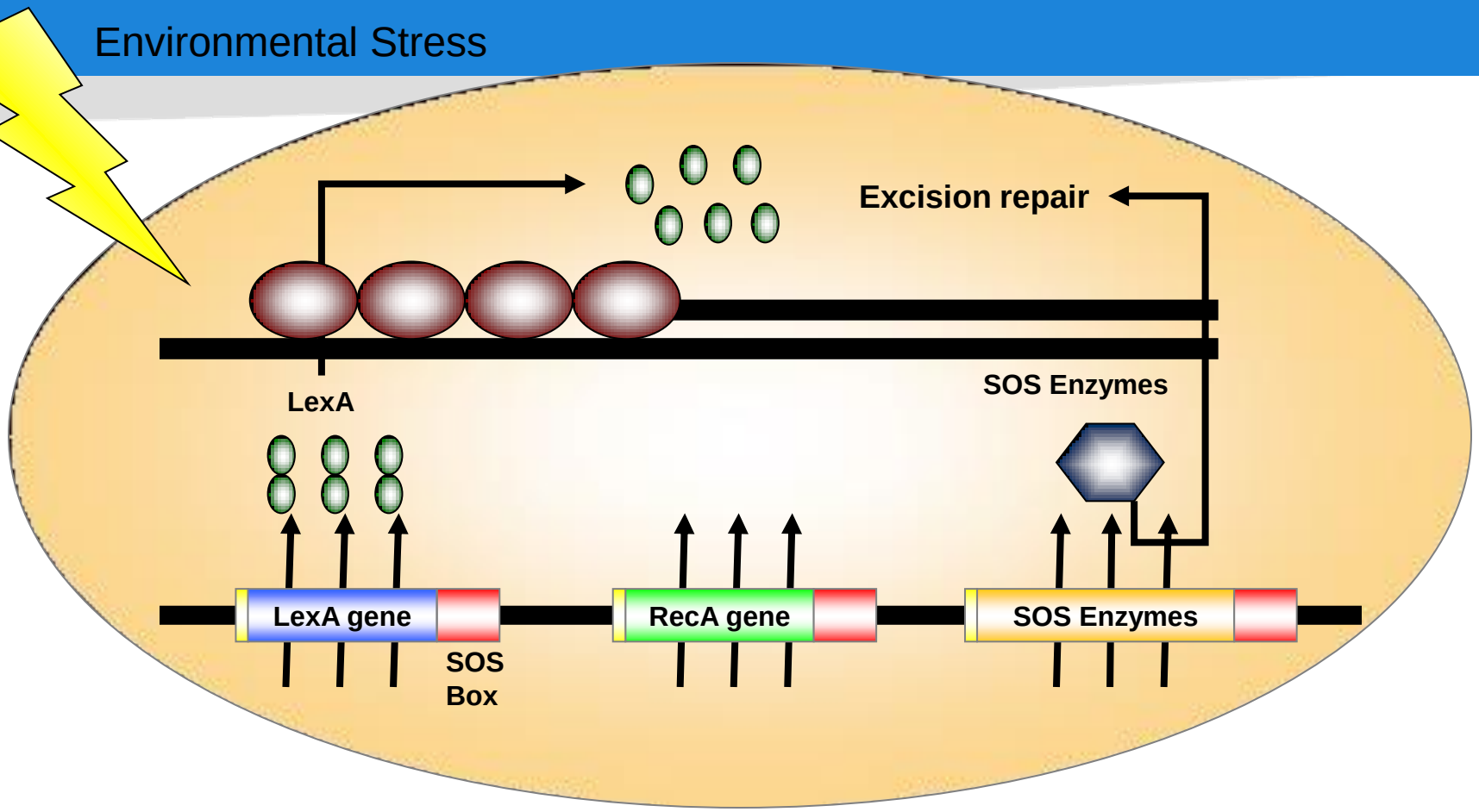


J. Craig Venter™
INSTITUTE

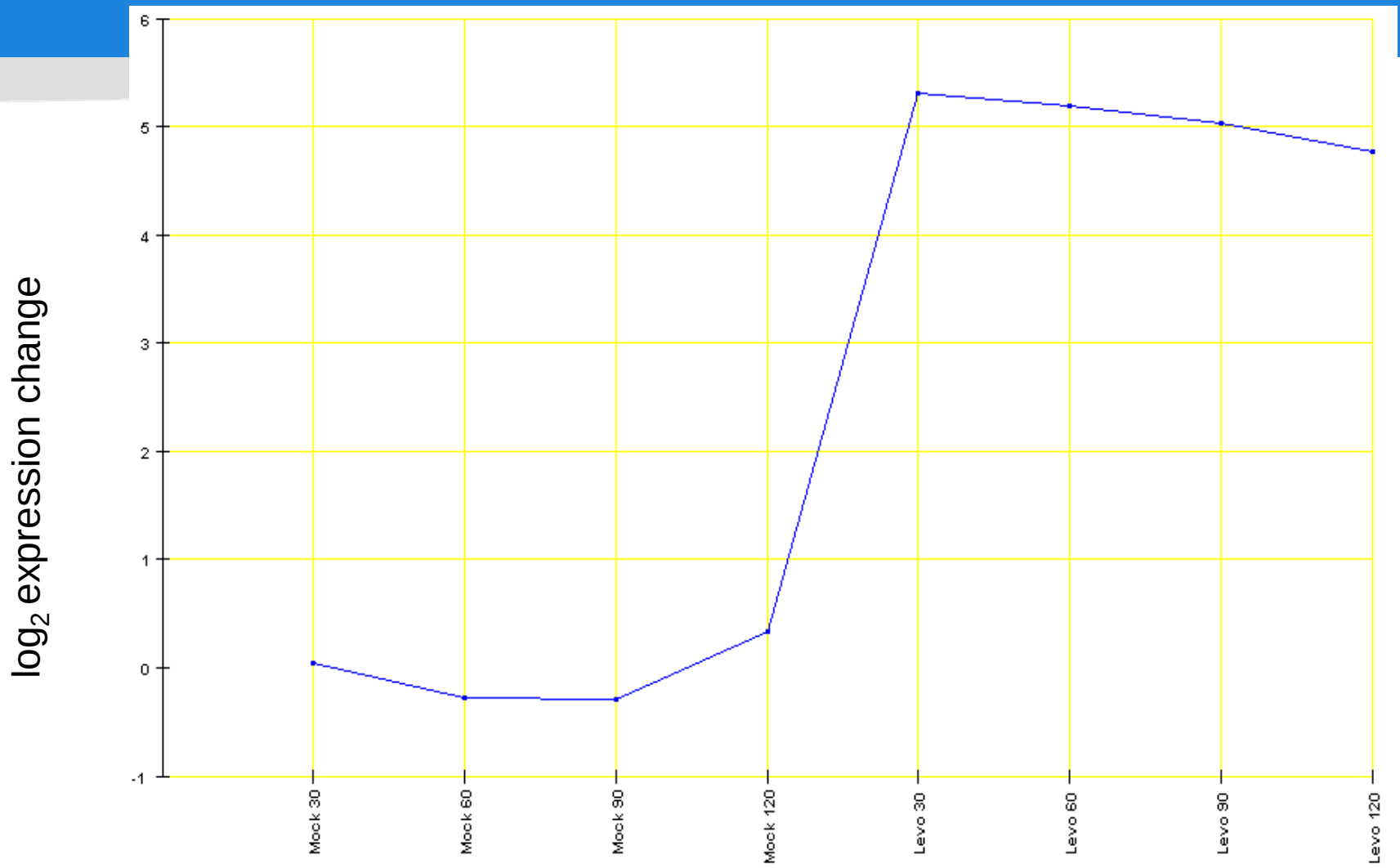
SOS



Environmental Stress

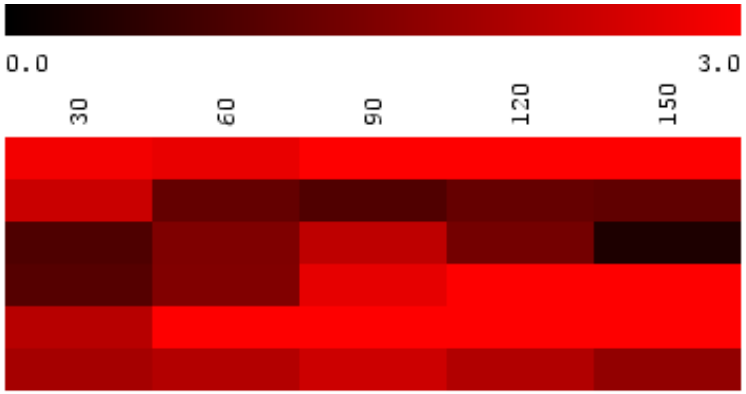


Effect of levofloxacin on *B. anthracis* γ-polymerase



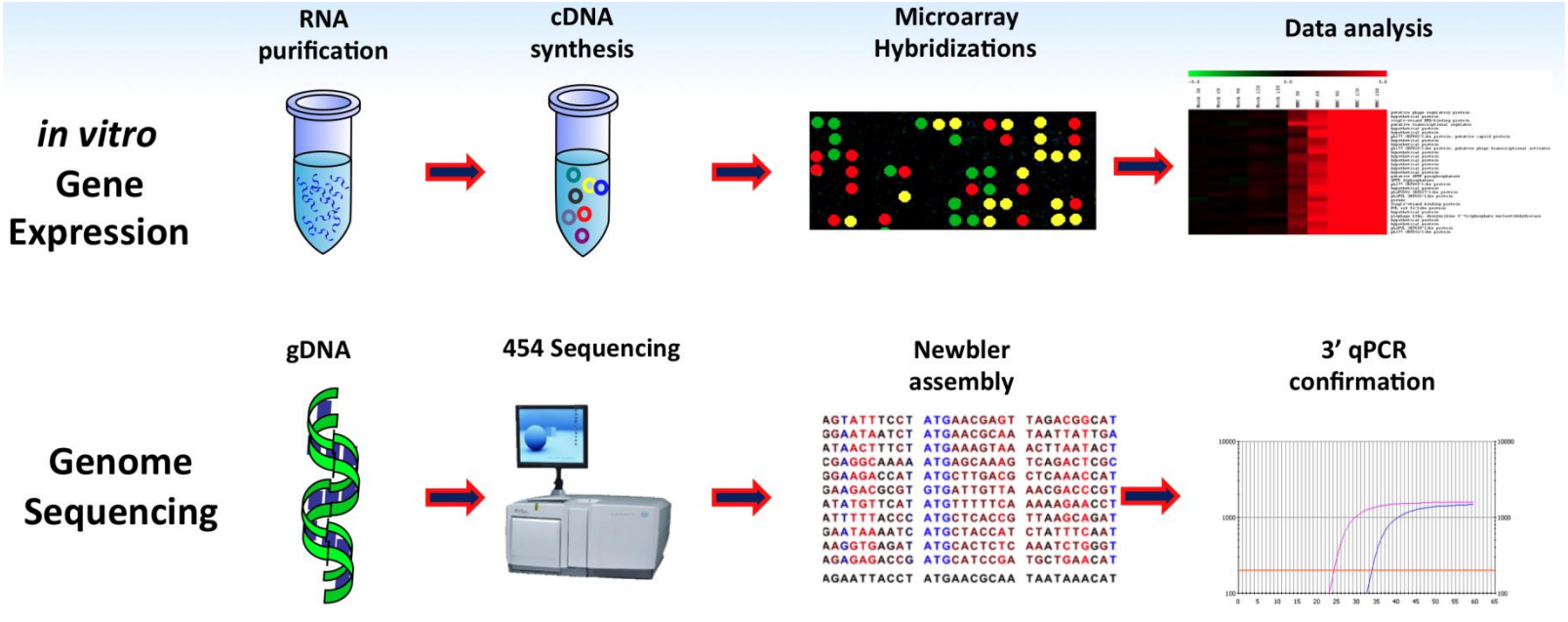
J. Craig Venter™ INSTITUTE

SOS response in *B. anthracis*



"DNA-damage-inducible protein P, putative"
LexA repressor
"recA protein, group I intron-containing"
"excinuclease ABC, A subunit"
"excinuclease ABC, B subunit"
"excinuclease ABC, C subunit"

J. Craig Venter™ INSTITUTE



J. Craig Venter INSTITUTE

Mapping Promoters:
Chip-chip Analysis of *in vitro/ ex vivo* Samples

Cross-link protein to DNA

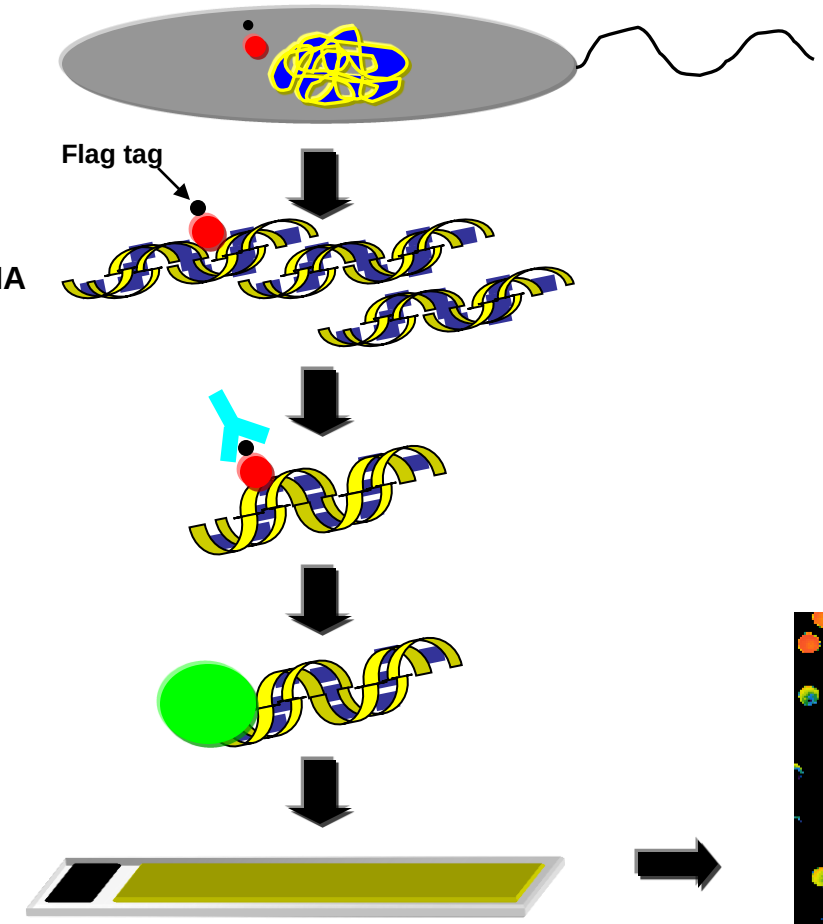
Sonicate to fragment to DNA

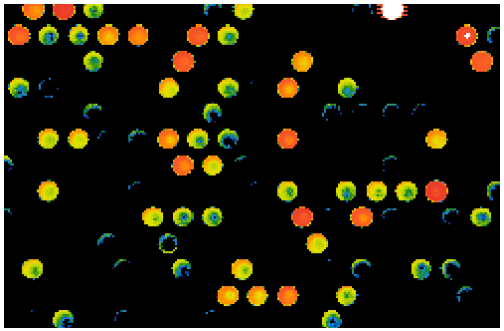
Flag tag

Immunoprecipitate

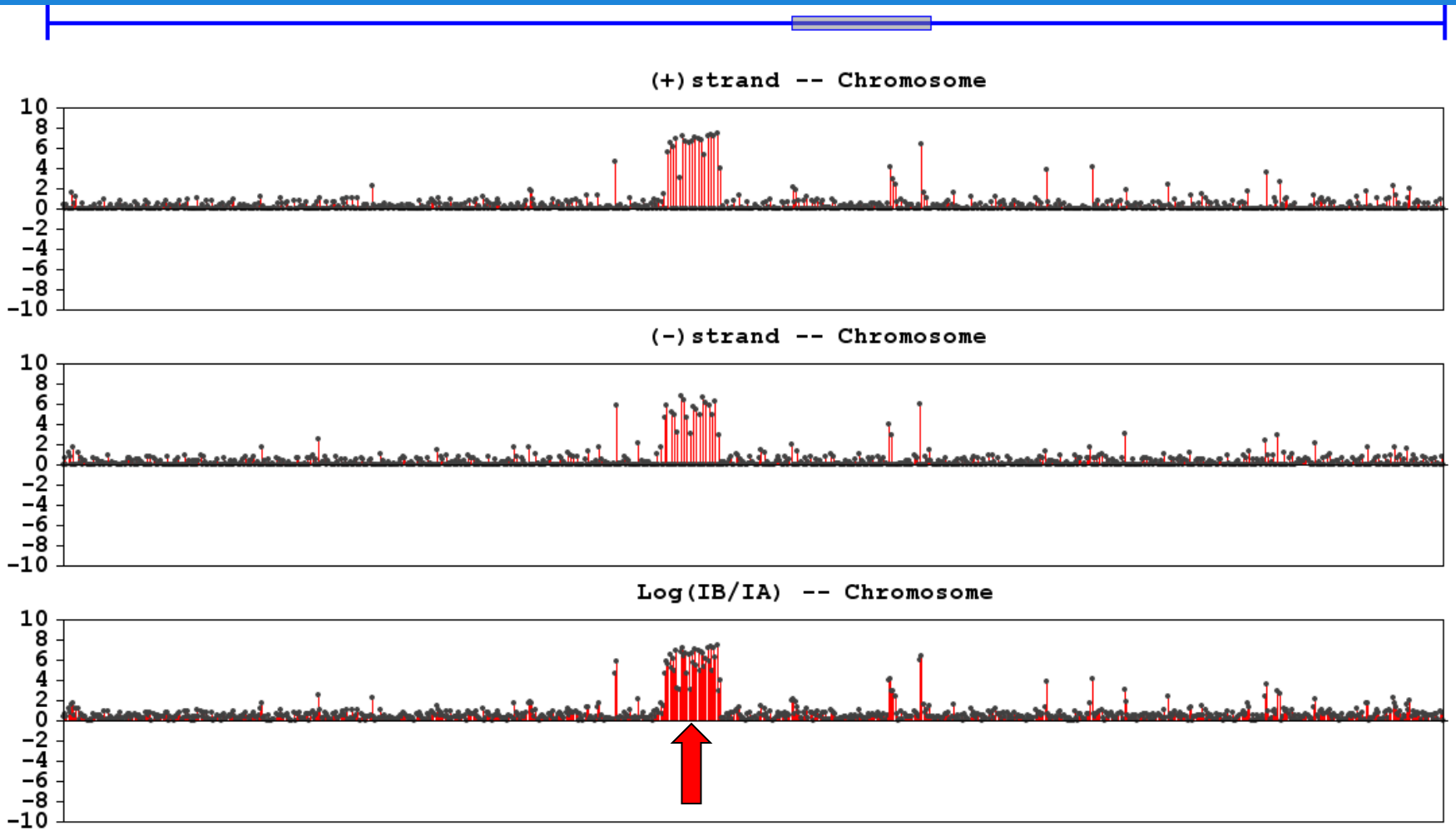
Purify and label DNA

Hybridize to microarray



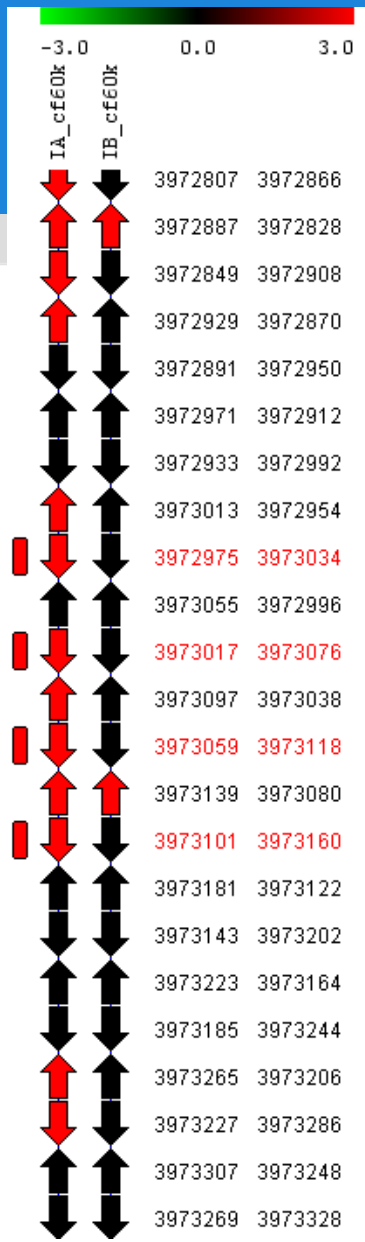


J. Craig Venter INSTITUTE



RpoE
Biding
Site Oligos

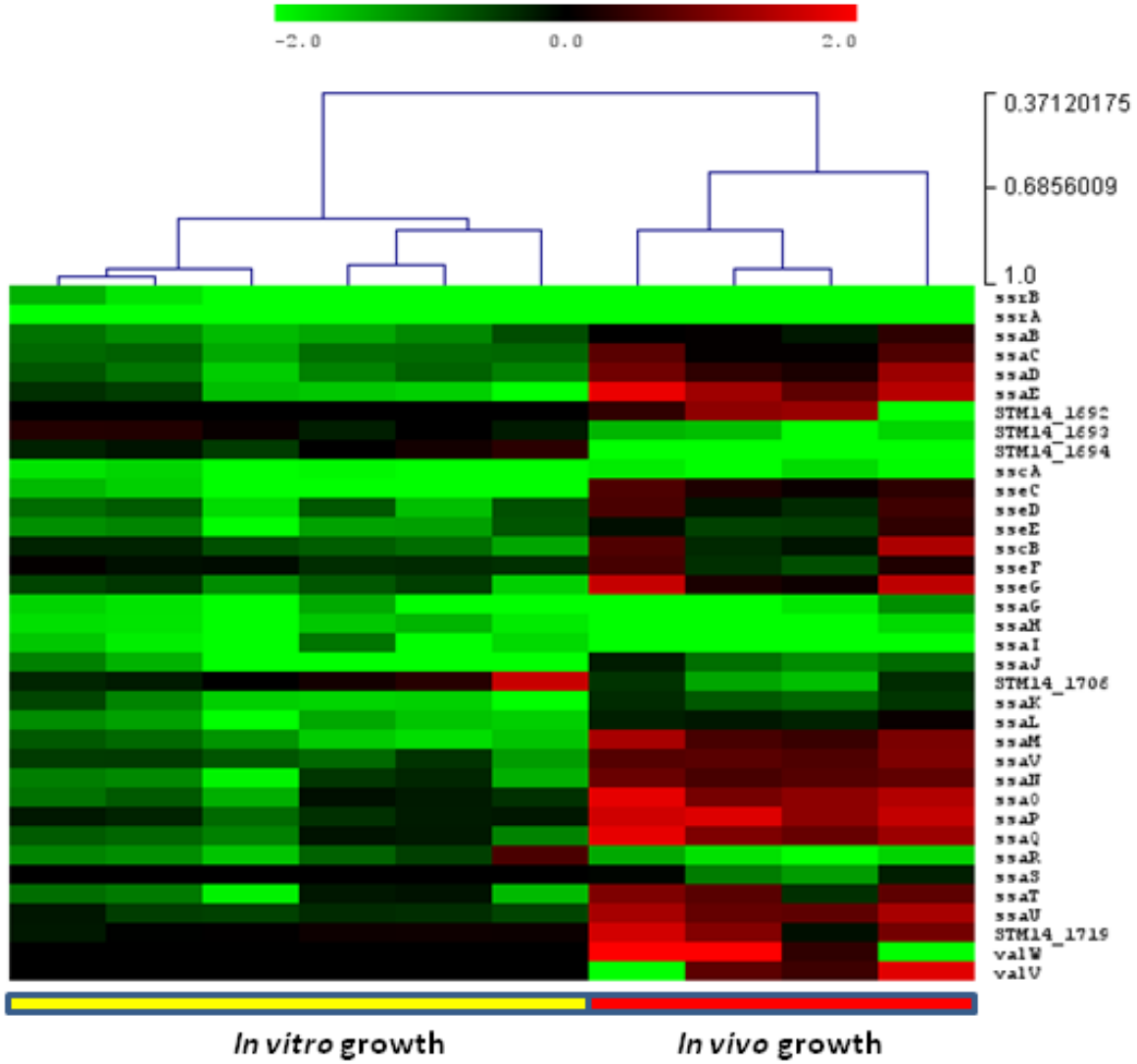
J. Craig Venter™
INSTITUTE



3972807	3972866	**Salm14028S_chr_3972807_3972866_F**	STM14_4530 [3972800 3972913 R] (60)
3972887	3972828	**Salm14028S_chr_3972828_3972887_R**	STM14_4530 [3972800 3972913 R] (60)
3972849	3972908	**Salm14028S_chr_3972849_3972908_F**	STM14_4530 [3972800 3972913 R] (60)
3972929	3972870	**Salm14028S_chr_3972870_3972929_R**	STM14_4530 [3972800 3972913 R] (44 match feat right)
3972891	3972950	**Salm14028S_chr_3972891_3972950_F**	STM14_4530 [3972800 3972913 R] (23 match feat right)
3972971	3972912	**Salm14028S_chr_3972912_3972971_R**	STM14_4530 [3972800 3972913 R] (2 match feat right)
3972933	3972992	**Salm14028S_chr_3972933_3972992_F**	INTERGENIC
3973013	3972954	**Salm14028S_chr_3972954_3973013_R**	INTERGENIC
3972975	3973034	**Salm14028S_chr_3972975_3973034_F**	INTERGENIC
3973055	3972996	**Salm14028S_chr_3972996_3973055_R**	INTERGENIC
3973017	3973076	**Salm14028S_chr_3973017_3973076_F**	INTERGENIC
3973097	3973038	**Salm14028S_chr_3973038_3973097_R**	INTERGENIC
3973059	3973118	**Salm14028S_chr_3973059_3973118_F**	INTERGENIC
3973139	3973080	**Salm14028S_chr_3973080_3973139_R**	INTERGENIC
3973101	3973160	**Salm14028S_chr_3973101_3973160_F**	INTERGENIC
3973181	3973122	**Salm14028S_chr_3973122_3973181_R**	INTERGENIC
3973143	3973202	**Salm14028S_chr_3973143_3973202_F**	INTERGENIC
3973223	3973164	**Salm14028S_chr_3973164_3973223_R**	INTERGENIC
3973185	3973244	**Salm14028S_chr_3973185_3973244_F**	INTERGENIC
3973265	3973206	**Salm14028S_chr_3973206_3973265_R**	INTERGENIC
3973227	3973286	**Salm14028S_chr_3973227_3973286_F**	INTERGENIC
3973307	3973248	**Salm14028S_chr_3973248_3973307_R**	INTERGENIC
3973269	3973328	**Salm14028S_chr_3973269_3973328_F**	INTERGENIC

J. Craig Venter™ INSTITUTE

RNASeq Data



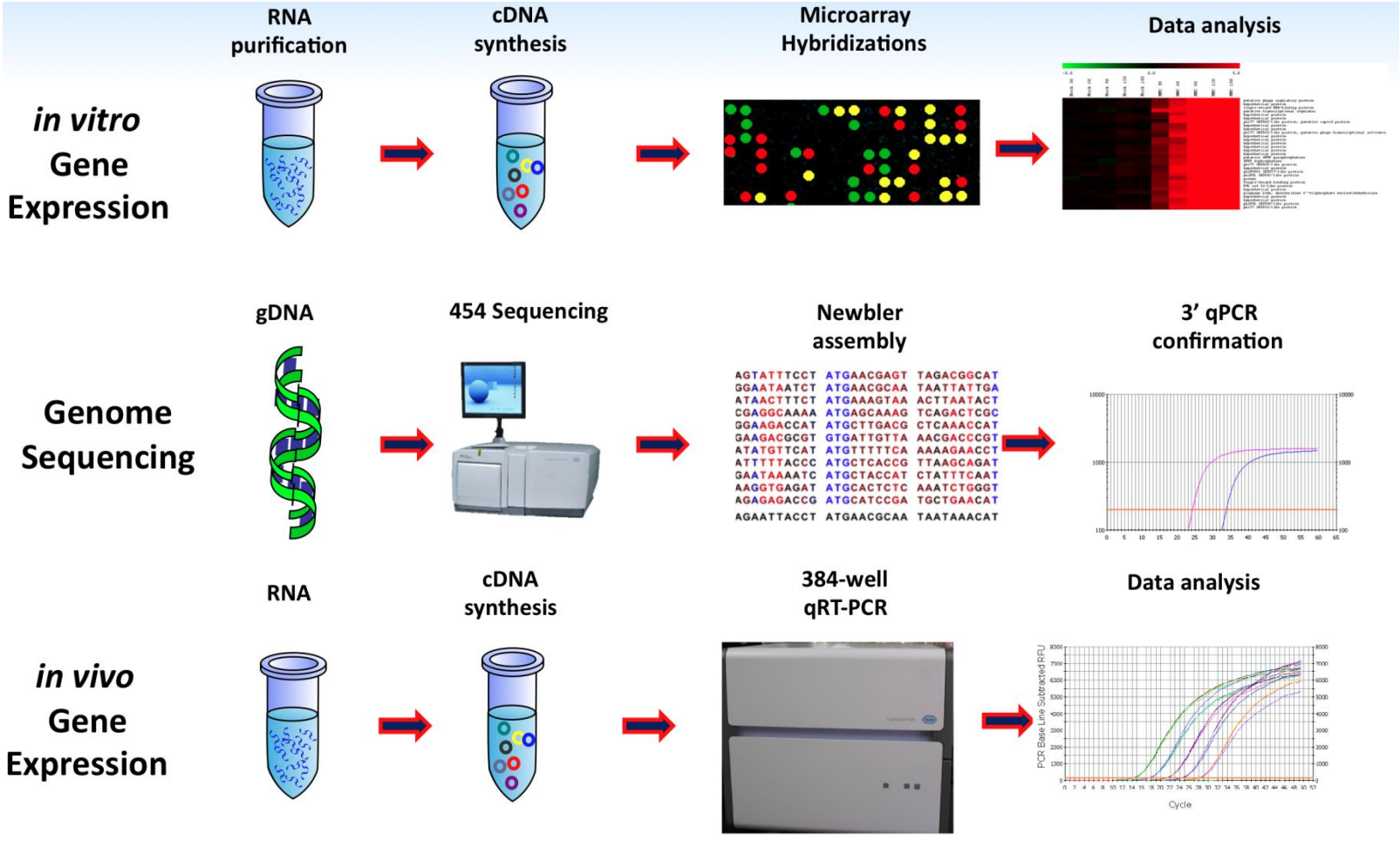
J. Craig Venter™ INSTITUTE

Linezolid Lineage Clinical Isolates

locus	start	end	symbol	WT	SNP	3577	5612	5892	7210	function
SA2212	2483696	2482581	N/A	G	C	negative	positive	positive	positive	hypothetical protein
SA1577	1815520	1808960	N/A	T	G	negative	positive	positive	positive	hypothetical protein
SA1118	1268775	1270448	N/A	T	C	negative	positive	positive	positive	hypothetical protein
SA1924	2173700	2172273	N/A	A	G	negative	positive	positive	positive	hypothetical protein
SA1669	1906904	1905519	fumC	A	G	negative	positive	positive	positive	fumarate hydratase
SA0500	579620	583171	rpoB	T	A	negative	negative	positive	positive	DNA-directed RNA polymerase subunit beta
SA0264	318934	319926	N/A	C	T	negative	negative	positive	positive	hypothetical protein

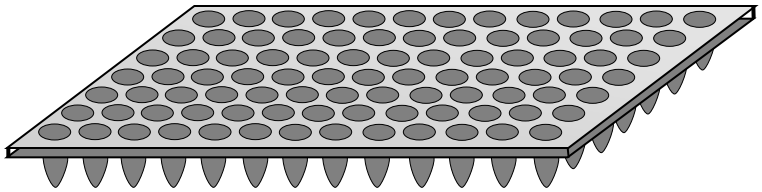
Two-step Resistance

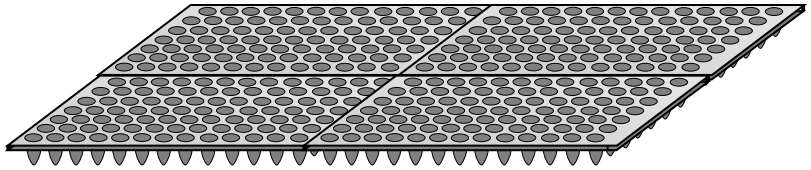
J. Craig Venter™ INSTITUTE



J. Craig Venter INSTITUTE

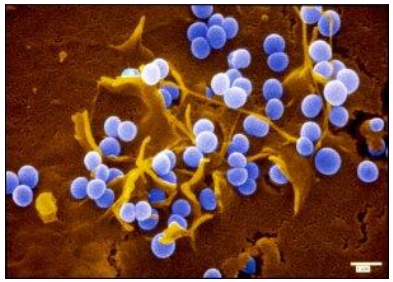
qRT-PCR Validation using Roche 480 LC

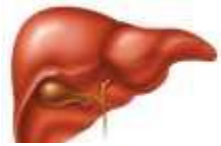


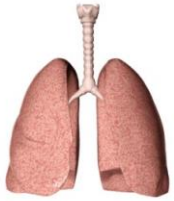


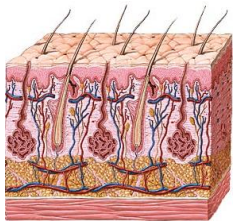
J. Craig Venter INSTITUTE

Characterization of Pathogen Gene Expression During Infection



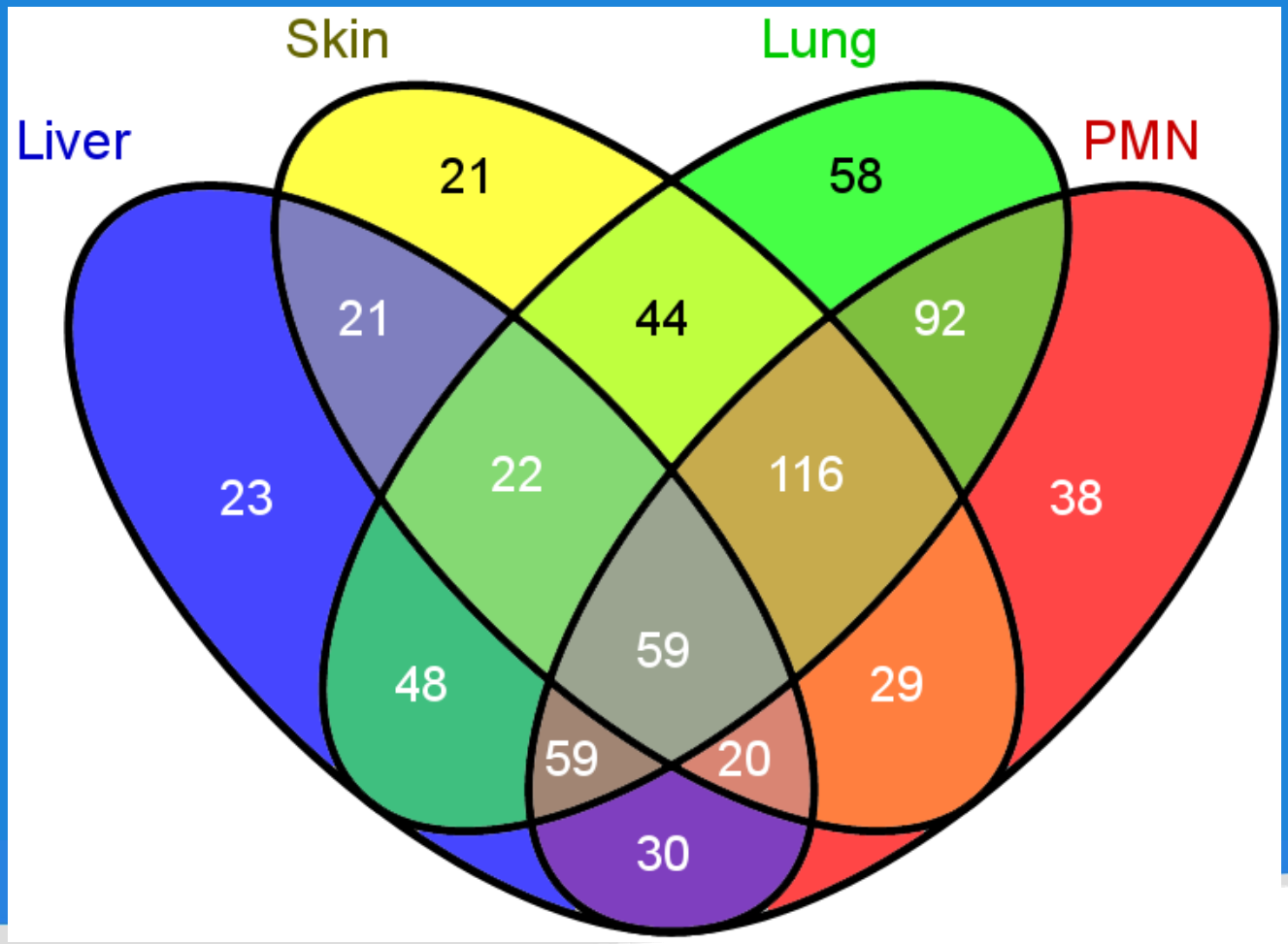






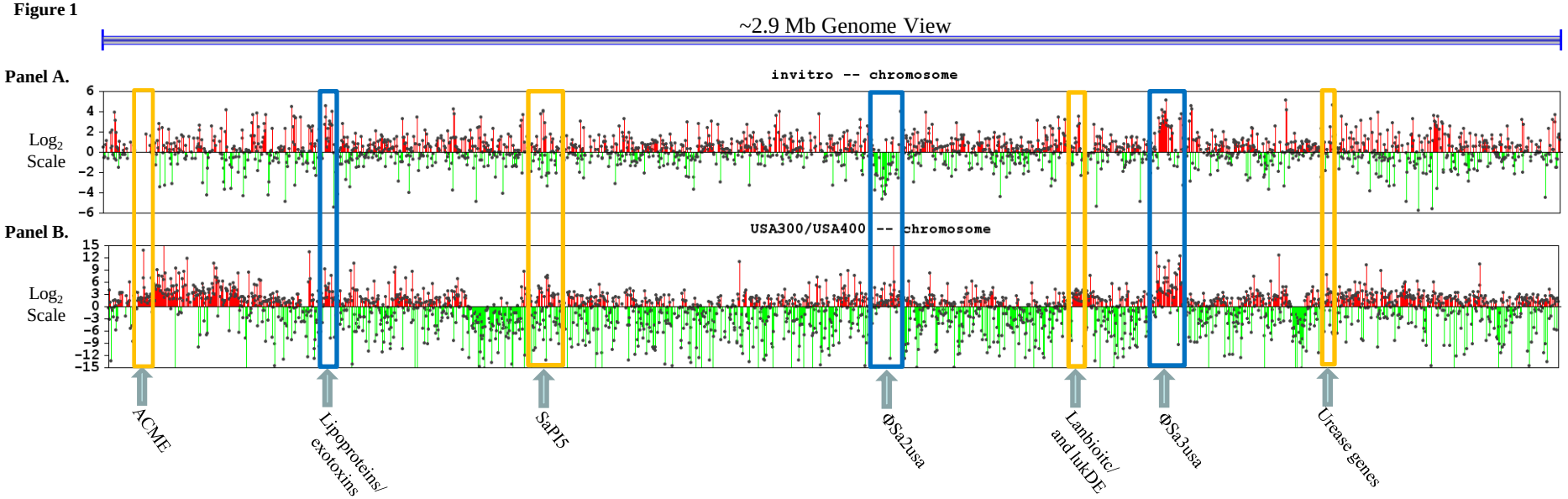
J. Craig Venter INSTITUTE

Hypothetical genes differentially expressed *in vivo*

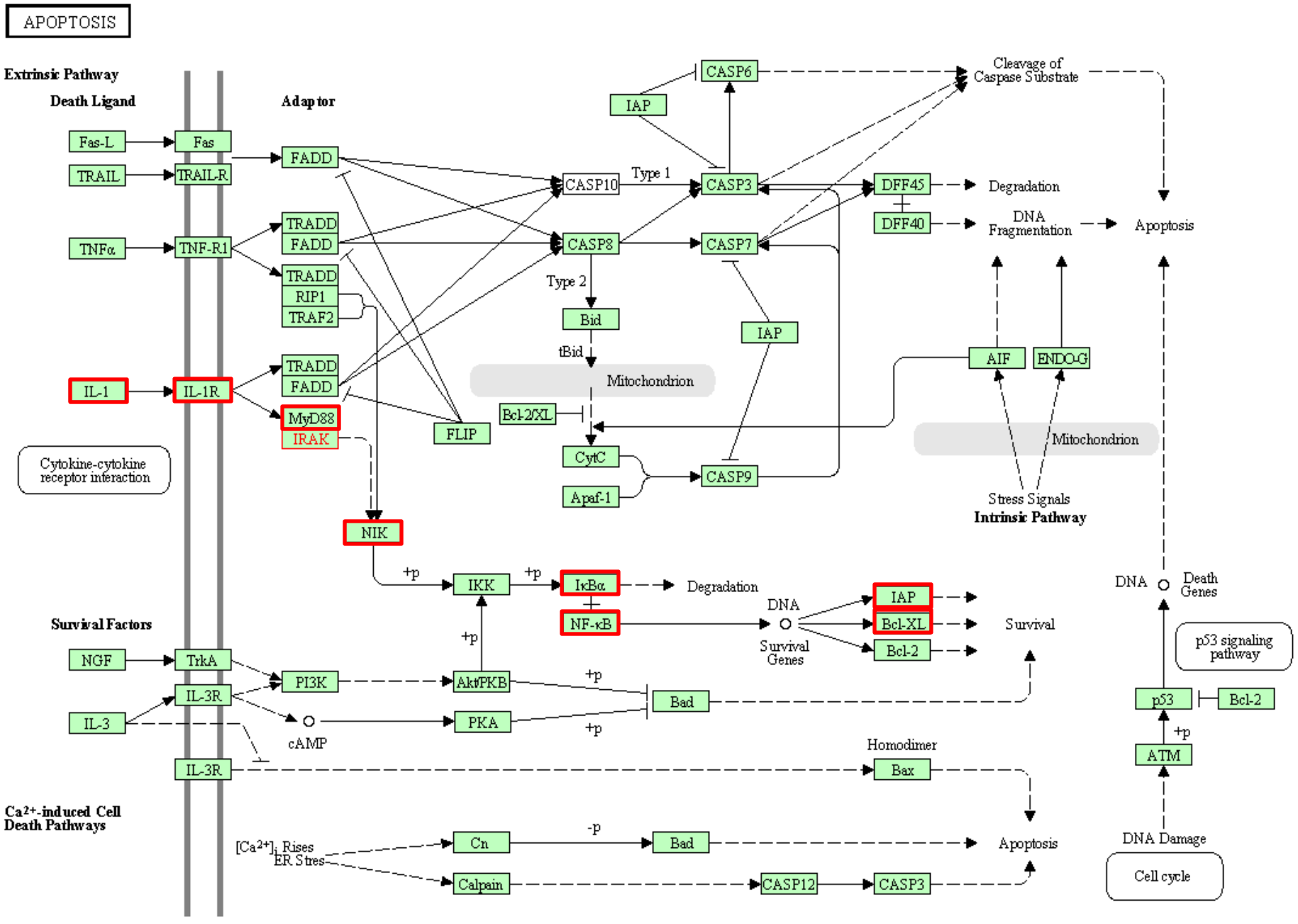


J. Craig Venter INSTITUTE

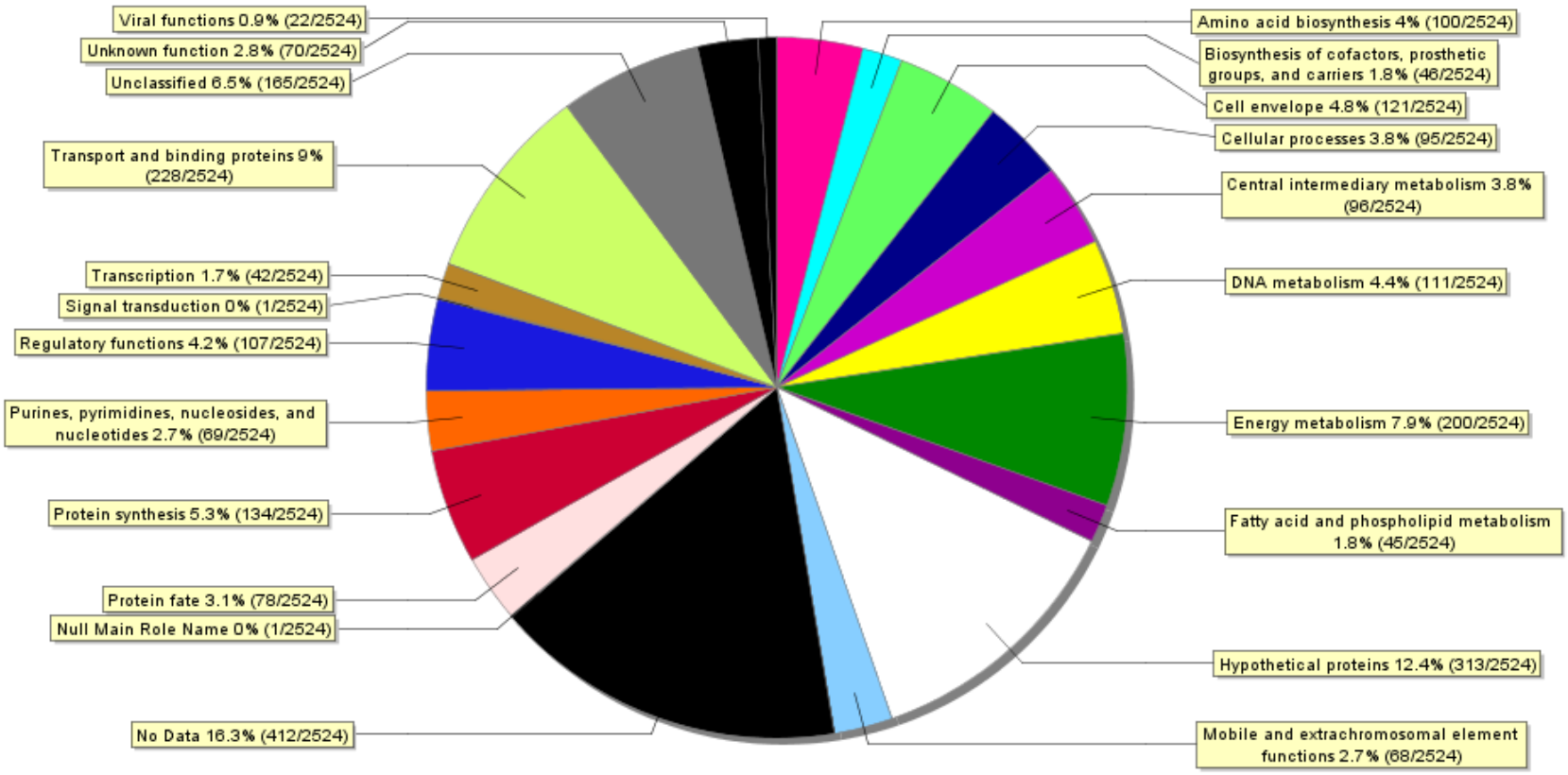
Figure 1



J. Craig Venter INSTITUTE



Role Category Analysis



J. Craig Venter INSTITUTE

What is Proteomics?

Additional Public Health Concerns

- UTIs are the most common cause of hospital-acquired infections accounting for approximately 40% of the total

- Many of these UTIs are caused by the ESKAPE pathogens. There is an increasing shortage of effective antibiotics againsts pathogens with multiple resistances
  - Enterococcus faecium
  - Staphylococcus aureus
  - Klebsiella pneumoniae
  - Acenitobacter baumanii
  - Pseudomonas aeruginosa
  - Enterobacter species

- Carbapeneme (Kp, Ec); MDR (Pa); penicillins and vancomycin (Ef, Sa)

- Large number of immune-compromised patients: HIV/AIDS, transplant and cancer patients

More informative Methods for UTI and ASB Diagnosis ?

- Vaginal and urinary tract microbiome profiling (sensitive detection of protective bacteria, ESKAPE pathogens, anaerobes missed in urine cultures): metagenomics

- Protein profiling to identify the bacteria and survey antimicrobial and immune responses: proteomics

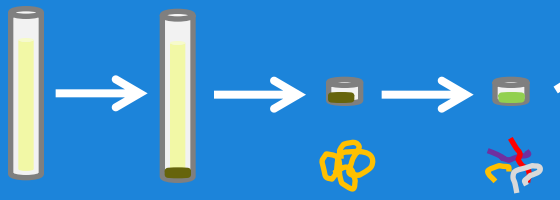
Fouts et al., J Transl Medicine (2012) 10, 174:
"Integrated next generation sequencing of 16S rDNA and metaproteomics diferentiate the healthy urine microbiome from asymptomatic bacteriuria in neuropathic bladder associated with spinal cord injury"

J. Craig Venter™ INSTITUTE

Proteomics: Analysis Stages

Sample acquisition

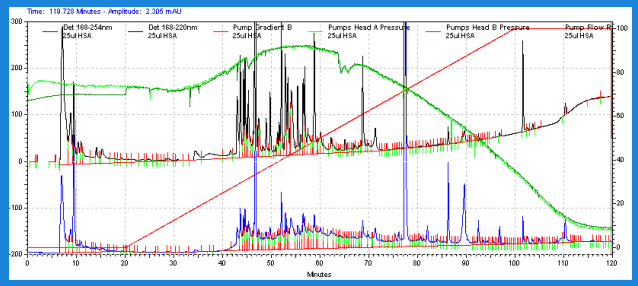
Pellet isolation



Protein denaturation

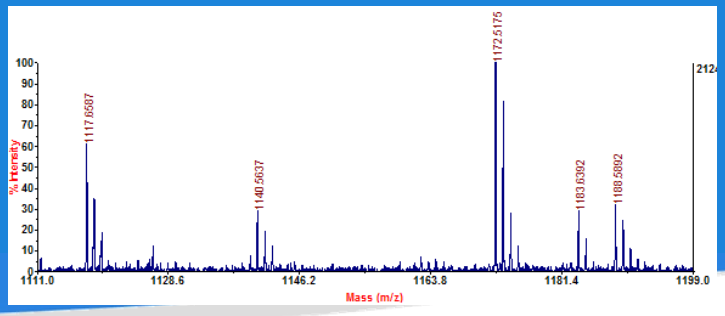
Protein digestion

Peptide separation



Electrospray

LC-MS/MS



Mascot data analysis

Quantitative methods

- Annexin A1 n=14 Tax=Eutheria RepID=ANXA1_HUMAN
- Alkyl hydroperoxide reductase subunit C [Klebsiella pneumoniae 342]

J. Craig Venter INSTITUTE

Database searches

- human protein sequence database UniRef90
- uropathogenic *E. coli*
- *Klebsiella pneumoniae*
- *Proteus mirabilis*
- *Pseudomonas aeruginosa*
- *Enterococcus faecalis*
- *Enterobacter hormachei*
- *Lactobacilus jensenii*
- *Morganella morganii*
- *Corynebacterium urealyticum*
- *Peptoniphilus asaccharolyticus*
- *Streptococcus pneumoniae*
- *Prevotella intermedia*
- *Staphylococcus epidermidis ….*

The database search space comprises ~80,000 distinct proteins

Patent application: Pieper et al., January 2013

J. Craig Venter™ INSTITUTE

Acknowledgements

- All JCVI faculty, staff and collaborators
- Funding Agencies: NIH
  - NIAID
    - Genome Sequencing Center (**NIH-HHSN272200900007C**)
    - System Biology for EnteroPathogens (**NIH-HHSN27220070058C**)
    - Pathogen Functional Genomics Resource Center (**N01-AI-15447**)
  - NIDCR
  - NIDDK

J. Craig Venter™ INSTITUTE