



**Results from the Sorcerer II
Global Ocean Sampling
Expedition**

RESEARCH ARTICLE

Environmental Genome Shotgun Sequencing of the Sargasso Sea

J. Craig Venter,^{1*} Karin Remington,¹ John F. Heidelberg,³
Aaron L. Halpern,² Doug Rusch,² Jonathan A. Eisen,³
Dongying Wu,³ Ian Paulsen,³ Karen E. Nelson,³ William Nelson,³
Derrick E. Fouts,³ Samuel Levy,² Anthony H. Knap,⁶
Michael W. Lomas,⁶ Ken Nealson,⁵ Owen White,³
Jeremy Peterson,³ Jeff Hoffman,¹ Rachel Parsons,⁶
Holly Baden-Tillson,¹ Cynthia Pfannkoch,¹ Yu-Hui Rogers,⁴
Hamilton O. Smith¹

We have applied "whole-genome shotgun sequencing" to microbial populations collected en masse on tangential flow and impact filters from seawater samples collected from the Sargasso Sea near Bermuda. A total of 1.045 billion base pairs of nonredundant sequence was generated, annotated, and analyzed to elucidate the gene content, diversity, and relative abundance of the organisms within these environmental samples. These data are estimated to derive from at least 1800 genomic species based on sequence relatedness, including 148 previously unknown bacterial phylotypes. We have identified over 1.2 million previously unknown genes represented in these samples, including more than 782 new rhodopsin-like photoreceptors. Variation in species present and stoichiometry suggests substantial oceanic microbial diversity.

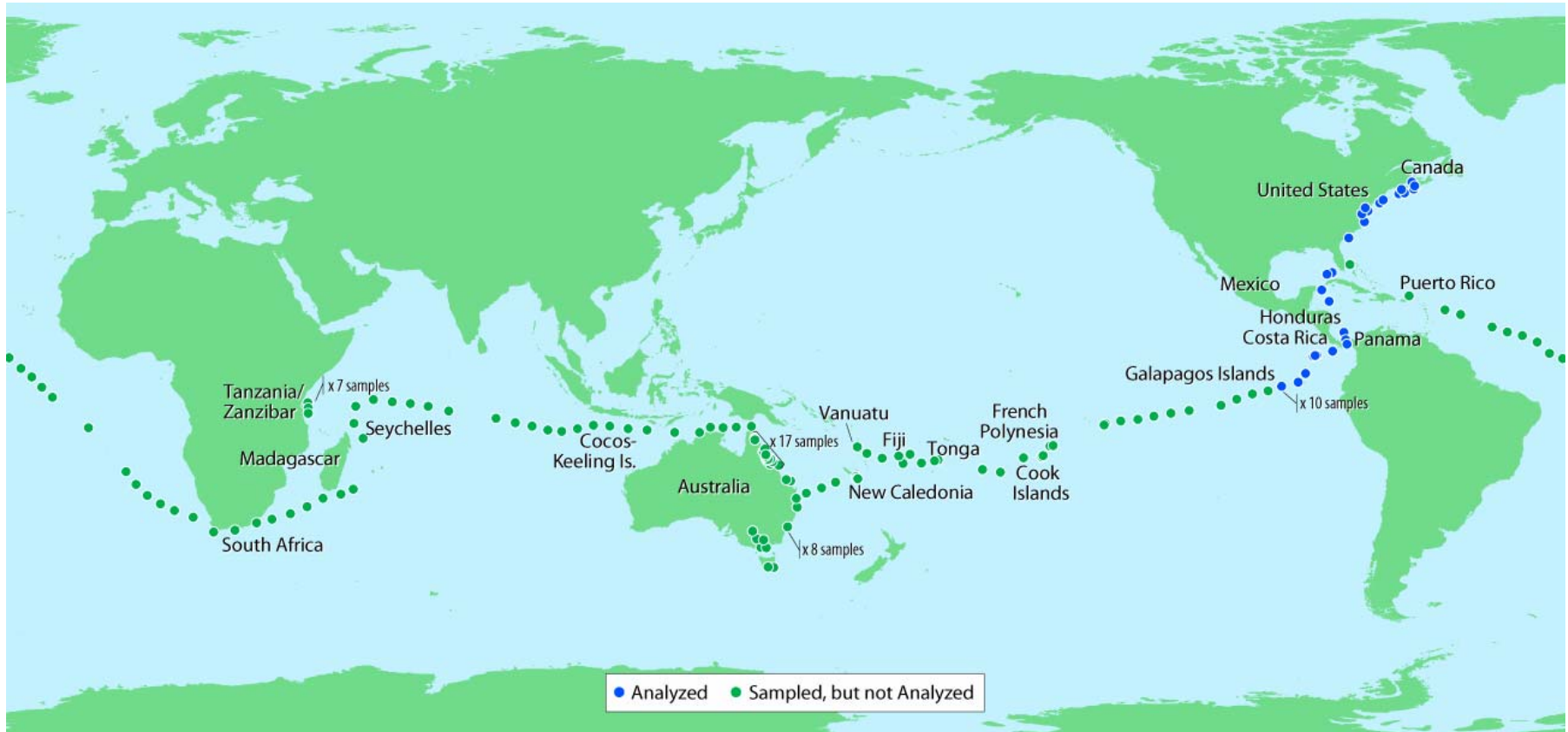
2 APRIL 2004 VOL 304 SCIENCE www.sciencemag.org

J. Craig Venter

I N S T I T U T E

The Sorcerer II Expedition

Global Ocean Sampling Route



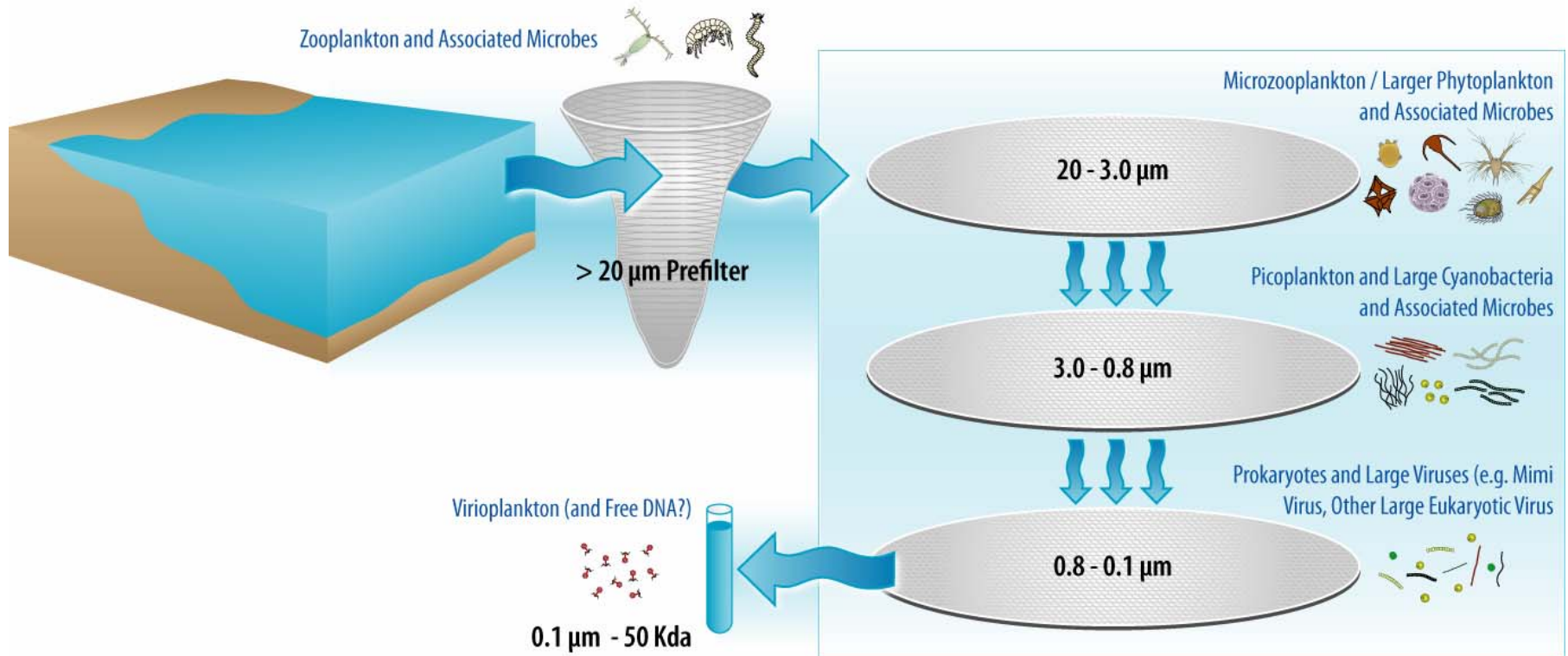
www.sorcerer2expedition.org

J. Craig Venter

I N S T I T U T E



Sample Fractionation



JTC Sequencer Lab

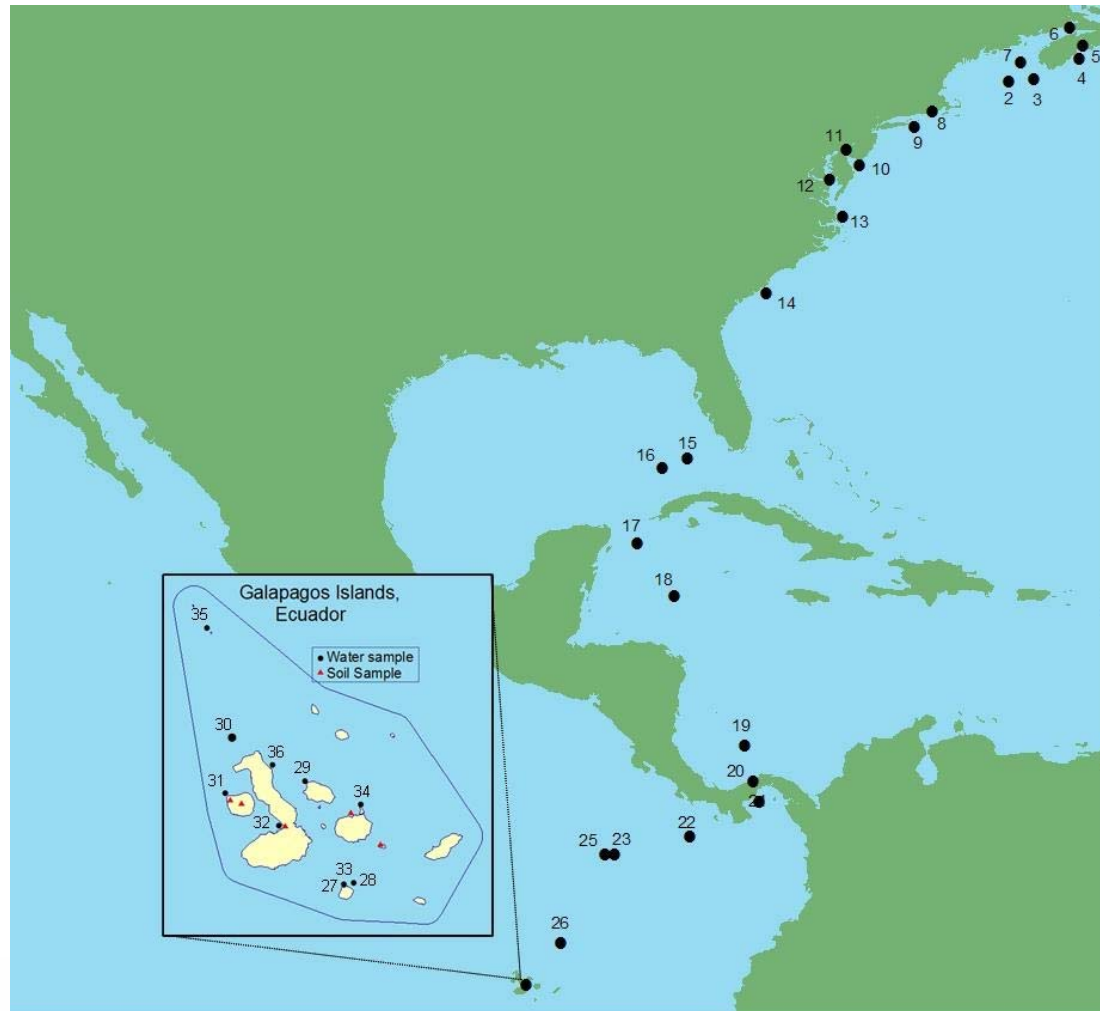
Capacity: 240,000 sequences/day or 80 million lanes/year at 24 runs per day



J. Craig Venter

I N S T I T U T E

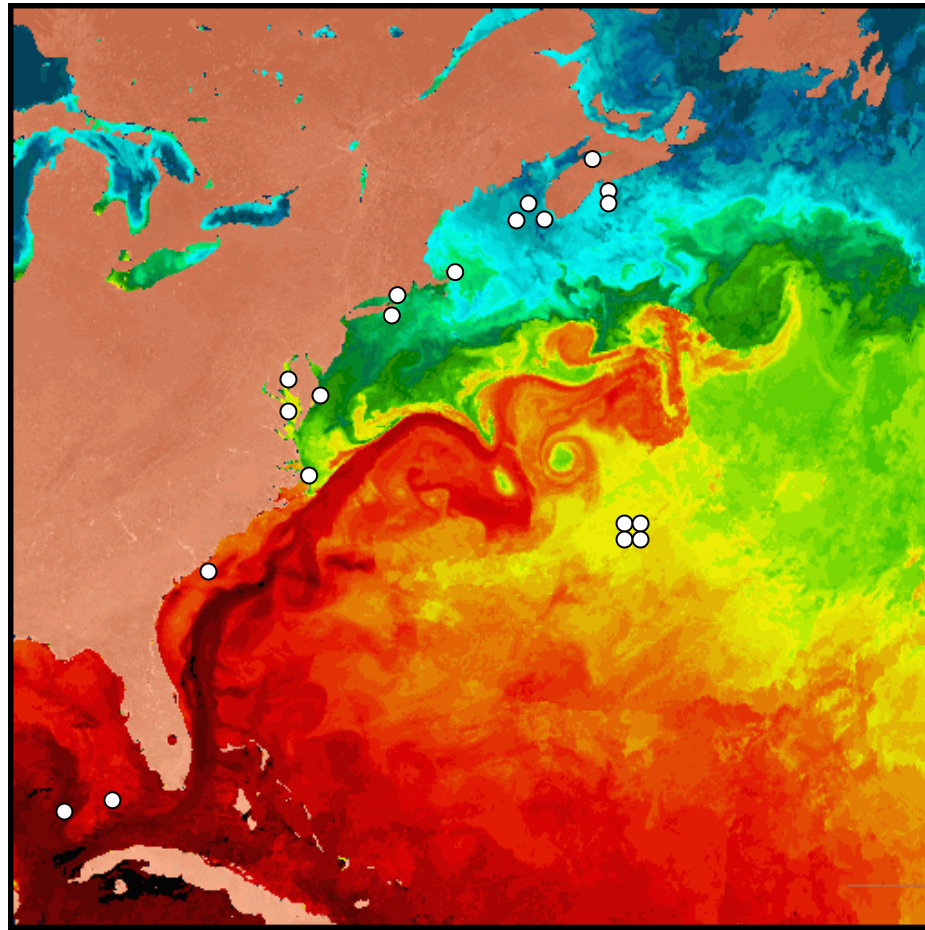
Preliminary Analysis: Nova Scotia through Galapagos



J. Craig Venter

I N S T I T U T E

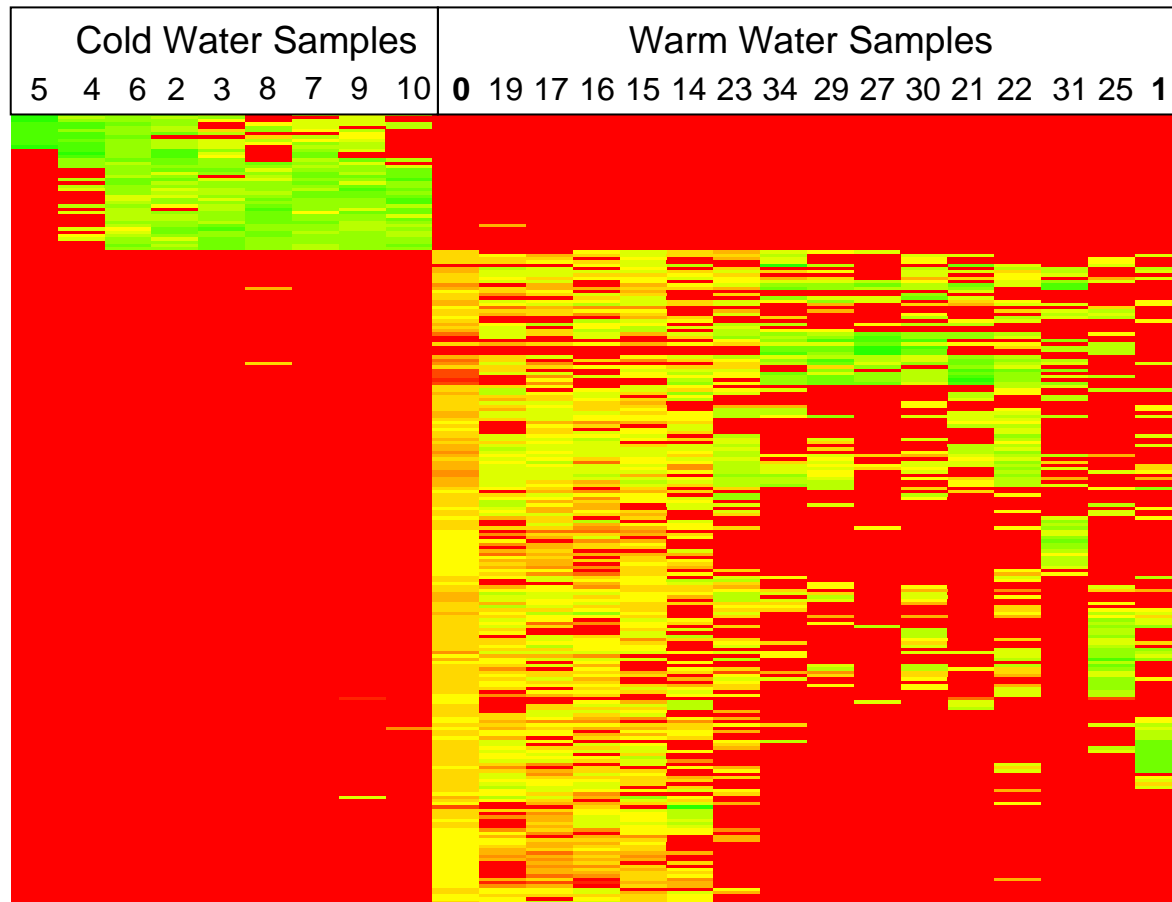
Water Temperature: *Nova Scotia through Galapagos Islands*



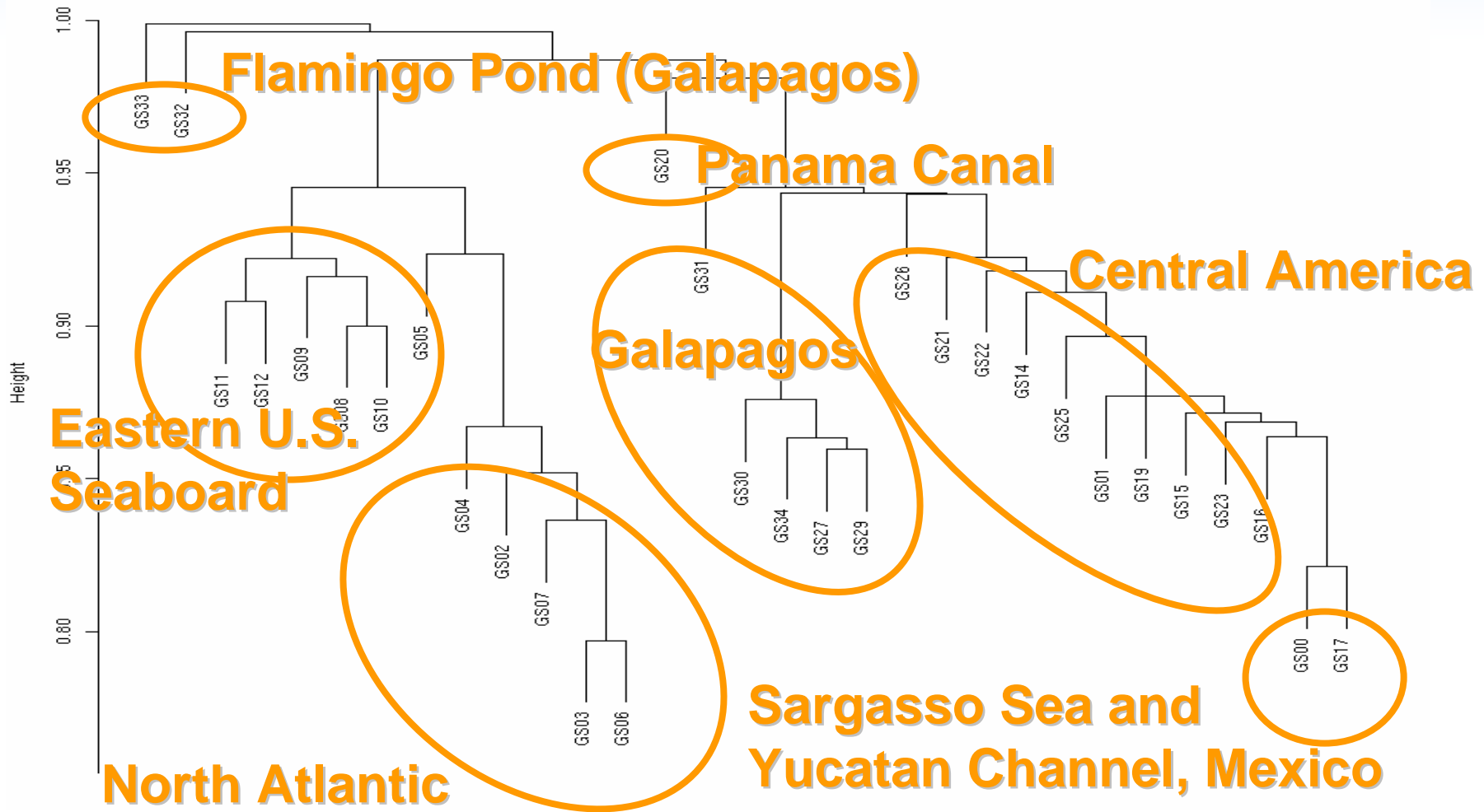
J. Craig Venter

I N S T I T U T E

Clustering of GOS sites: *Reads incorporated into significant scaffolds*



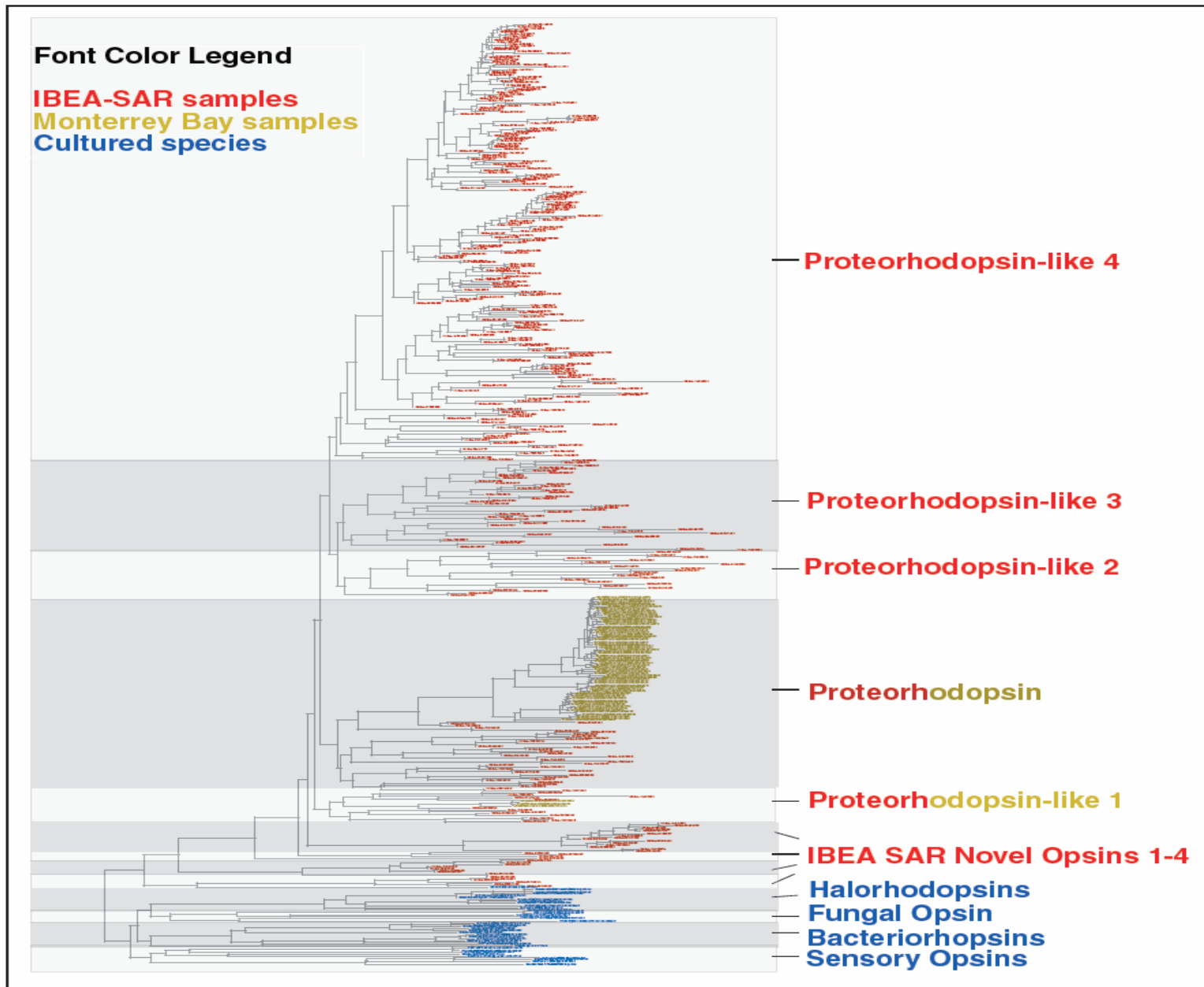
The Biodiversity of Each New Region is Different



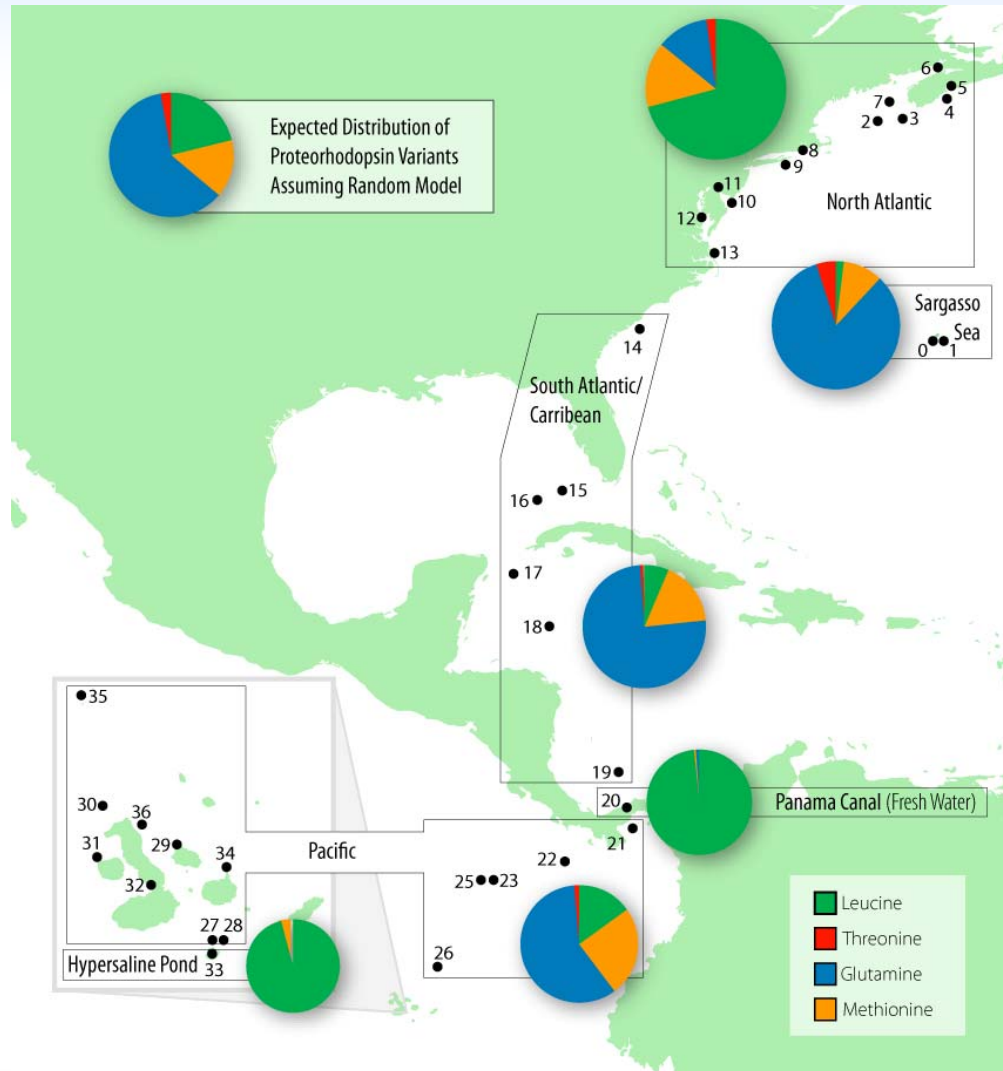
J. Craig Venter

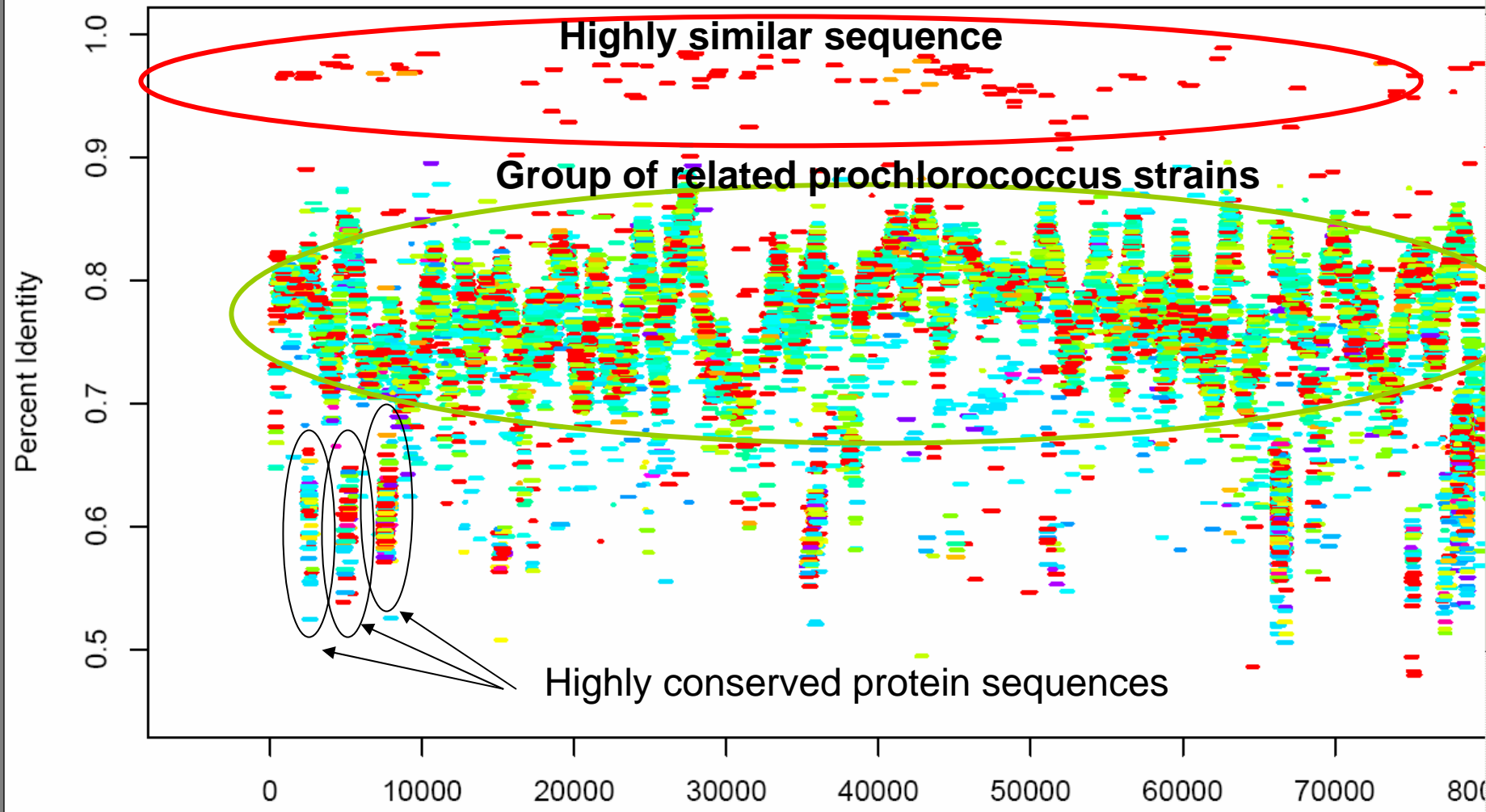
I N S T I T U T E

Phylogenetic Tree of Halorhodopsin-like Genes in Sargasso Sea



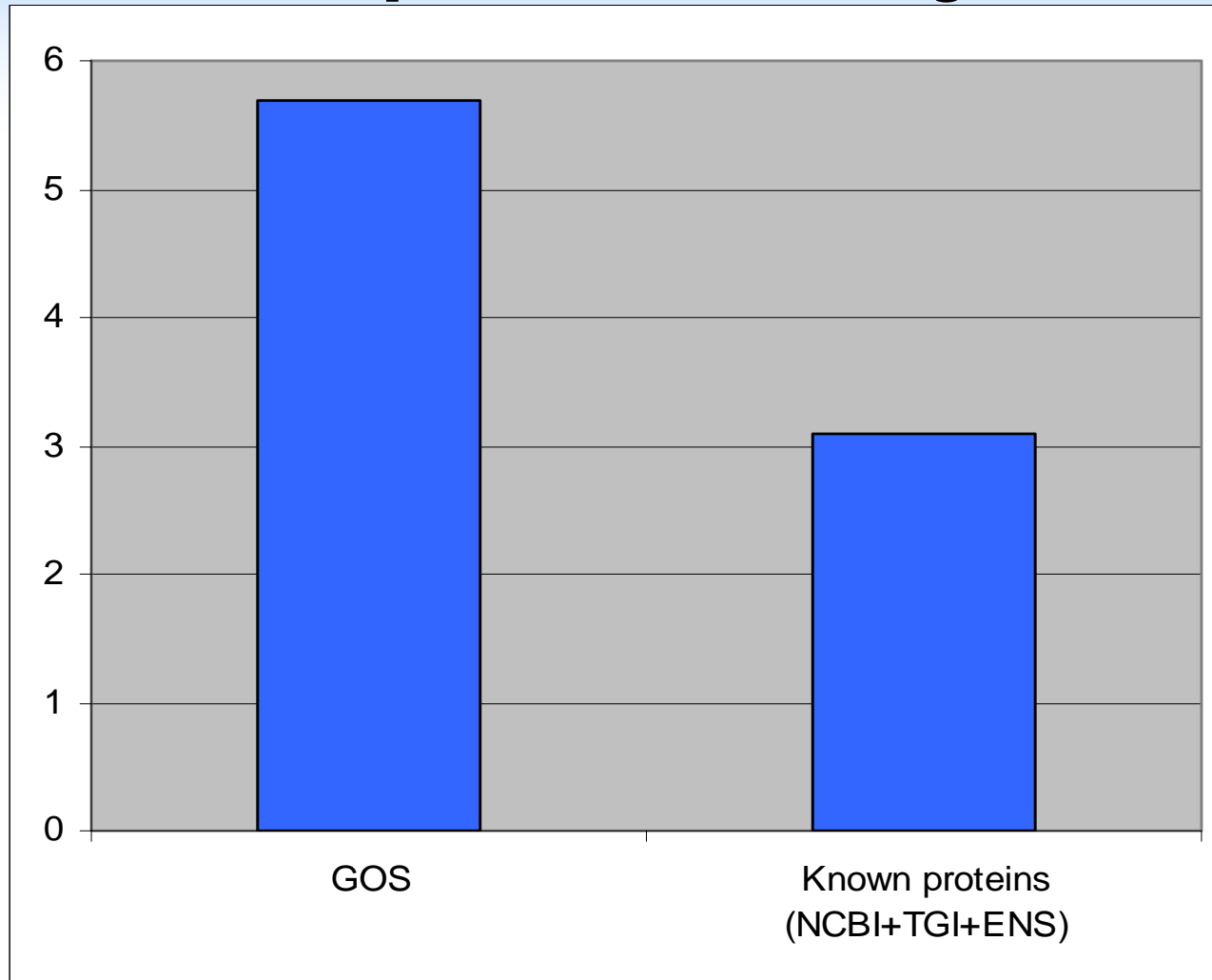
Proteorhodopsins Vary by Region





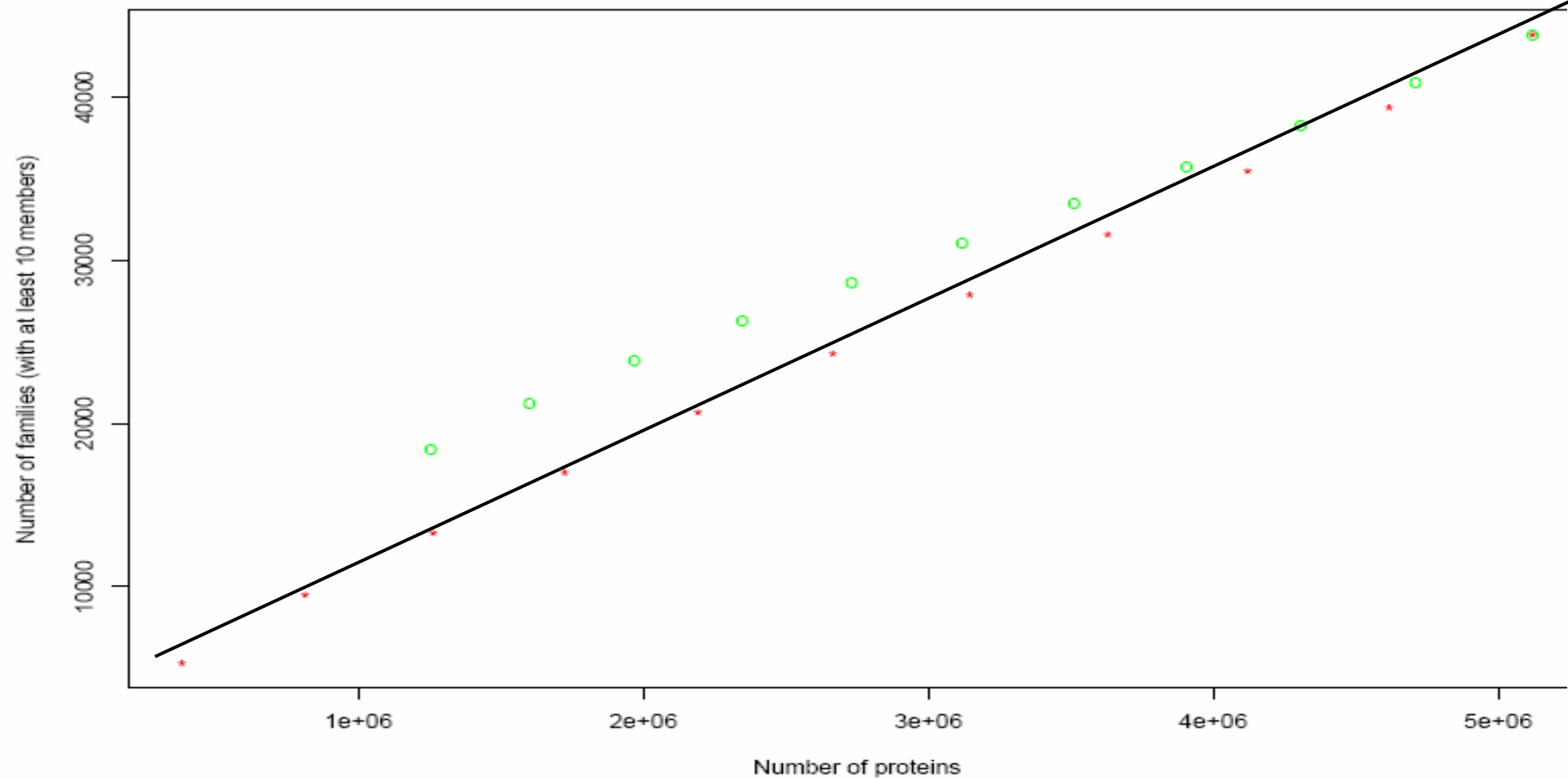
← Sequenced Prochlorococcus genome →

GOS protein analysis



GOS nearly doubles the number of proteins

Number of Protein Families on Earth



J. Craig Venter

I N S T I T U T E