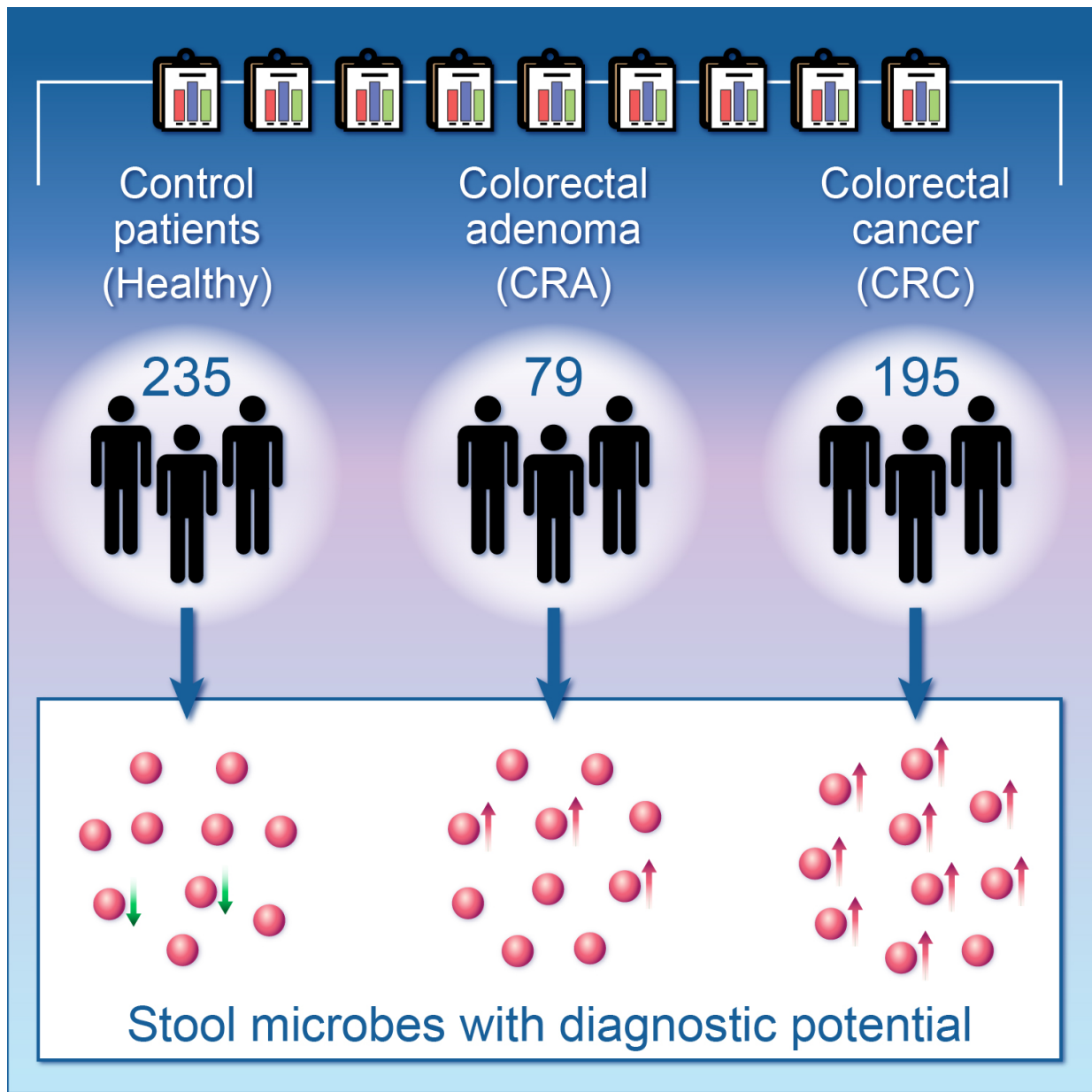# Research paves way for improved colorectal cancer test

April 19 2017, by Jeannette Jimenez

Researchers reanalyzed raw bacterial DNA sequence data from several studies and confirmed previously reported types of bacteria associated with colorectal cancer and identified other bacteria not previously associated with the disease. Credit: K. Prince/Texas Children's Hospital

The type of bacteria in your gut may help diagnose colorectal cancer. Researchers at Baylor College of Medicine and other institutions have identified specific types of bacteria that seem to be abundant in individuals with colorectal cancer. Using a combination of markers specific for these fecal microbes, scientists anticipate that a noninvasive, sensitive clinical diagnostic test potentially can be developed. The study is published in *Gut*.

Colorectal cancer is the second-leading cause of cancer-associated death in the United States. Individuals diagnosed early have a 90 percent chance of survival; however, more than 30 percent of individuals 50 years old or older, the high-risk group, indicate never having been screened for the disease. Currently, recommended screening methods include an invasive procedure - colonoscopy - and noninvasive tests, such as home-based fecal occult blood tests, and Cologuard for detecting colorectal cancer DNA markers and blood in stools. Each of these methods has its shortcomings, so to meet the need for a sensitive, noninvasive diagnostic test for colorectal cancer, the team of researchers looked at gut microbes as indicators of the disease.

"A number of studies have shown an association between fecal microbes and colorectal cancer; however, there is limited agreement in the types of microbes reported," said first author Dr. Manasi Shah, who was a graduate student at the University of Texas School of Public Health during the course of this project. "I was interested in finding a microbial marker for the disease. One way to do this is by carrying out a single-

institution study, but this takes a long time for sample collection, involves sequencing the microbes' DNA and is expensive. I noticed that some of the published studies provided the means for accessing the raw microbial DNA sequencing data of the samples. How great it would be, I thought, if I could leverage existing raw data across multiple cohorts and come up with a generalizable marker for the disease." Shah realized that to take this approach, in addition to her biostatistical training, she would need to learn the bioinformatics tools necessary for analyzing next generation sequencing data. She approached senior author Dr. Emily Hollister, assistant professor of pathology at Baylor and Texas Children's Hospital and director of microbial ecology for the Texas Children's Microbiome Center, proposed her idea and expressed interest in learning the tools required to reprocess microbial sequence data from its original format.

"Manasi had the interest, and we had the expertise," Hollister said. "In our center, we had been planning to compare a series of different statistical tools to analyze large amounts of microbiome data. Manasi's proposal fit very well with our goals."

Researchers reanalyzed raw bacterial DNA sequence data from several studies and confirmed previously reported types of bacteria associated with colorectal cancer and identified other bacteria not previously associated with the disease.

Easier said than done

"In our experience, collecting the raw data from the published studies was an uphill task," Shah said. "Some studies shared all the sample-associated microbial DNA sequences and clinical data, others only shared partial data or did not share any data at all. After much effort, I was able to gather data from nine of 12 published studies. This highlights the need for an initiative to encourage investigators to share their data

upon publication, which will help wider dissemination and reproducibility in the field."

The researchers also had to overcome the difficulties posed by the diversity of technological approaches used by different laboratories analyzing the samples.

"This was an incredibly large, complex multinational study," said co-author Todd DeSantis, co-founder and vice president of informatics at Second Genome Inc. "We saw many differences between medical centers in the way each collected and stored stool samples and in the methods used to process the bacterial DNA in stools. These differences can be problematic for identifying the bacterial strains that proliferate in cancer patients, but our Second Genome KnowlegeBase Team, led by co-author Thomas Weinmaier, found ways to enhance our software platform to address these differences along the way. The findings that emerged from this challenging data set helped validate our platform, and in the process we were able to deliver high-quality insights to advance our collaboration with Dr. Hollister."

After reanalyzing large amounts of raw bacterial DNA sequence data from several studies uniformly using a variety of statistical tools, the scientists confirmed previously reported types of bacteria associated with colorectal cancer and identified other bacteria not previously associated with the disease.

"The fact that even when we combined several different studies we could correctly classify a sample as a colorectal cancer case or control with 80 percent accuracy solely based on microbial abundances was very encouraging," Shah said."This is a promising first step to develop a noninvasive test that might be used in the detection of colorectal cancer, supplementing colonoscopy or fecal occult blood tests," Hollister said.

"The same strategy could be applied for developing diagnostic tests or therapeutics for other diseases such as inflammatory bowel disease, nonalcoholic steatohepatitis (a form of non-alcoholic fatty liver disease), type 2 diabetes and Alzheimer's disease, among others for which the microbiome is currently being investigated," Shah said.

 More information: Manasi S Shah et al, Leveraging sequence-based faecal microbial community survey data to identify a composite biomarker for colorectal cancer, *Gut* (2017). DOI: 10.1136/gutjnl-2016-313189

Provided by Baylor College of Medicine