

Error Detecting/Correcting Barcodes for Accurate, Simultaneous High-Throughput Sequencing

Published date: Feb. 1, 2012

Technology description

Summary

Background DNA sequencing is critical for research efforts ranging from understanding biological processes to forensic sciences. Conventional DNA sequencing methods are extremely costly, but highthroughput sequencing reduces costs per sequence by orders of magnitude. Pyrosequencing, a form of high-throughput sequencing, thus has the potential to revolutionize many sequencing efforts, including the assessment of microbial community diversity throughout our bodies and our planet, by generating hundreds of thousands of sequences in each run. However, these analyses have been limited by the expense of each individual run. One solution to this limitation is barcoding, in which a unique tag is added to each primer before PCR amplification. Because each sample is amplified with a known tagged primer, sequencing can be performed on an equimolar mixture of PCR-amplified DNA from each sample, and each sequence can be assigned to the correct sample based on its sample-specific barcode. This technique has been used successfully to sequence up to twenty-five samples in a single pyrosequencing run. However, prior barcoding methods are limited both in the number of unique barcodes they use and in their ability to detect sequencing errors that produce mistaken sample assignments.

-- for example, to screen large populations of patients, or to do long timeseries. Scaling from dozens to hundreds/thousands of samples completely transforms the type of research questions that can be asked. For example, Knight's team routinely splits a single 454 run across 100-300 samples, which has allowed them to find diversity in human-associated microbial communities that was completely unsuspected (leading to papers accepted in Nature Methods, PNAS, Nature, etc.). For example, they showed that the left and right hands of the same person share only 18% of their bacterial species on average, that lean and obese individuals have markedly different bacterial communities in the gut, and that people from different continents have completely different gut microbial communities -- these findings would not have been possible had only a few subjects been examined, as was typical in earlier studies.

Application area

Technology Rob Knight and Micah Hamady of the University of Colorado have developed an improved method for designing barcodes for use in high throughput sequencing, e.g. pyrosequencing.

Advantages

These barcodes, uniquely, can both detect and correct sequencing errors, thus allowing high-confidence sequencing of hundreds of samples in a single pyrosequencing run. This multiplexing greatly reduces the cost of entry for end users because, for many applications, a few hundred or a few thousand sequences are sufficient to see the most important patterns

Institution

[University of Colorado](#)

联系我们



叶先生

电话 : 021-65679356

手机 : 13414935137

邮箱 : yeyingsheng@zf-ym.com