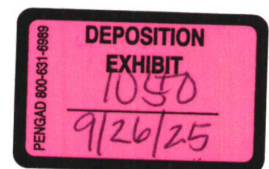


From: [Greg Porreca](#)
To: [Waller, Patrick](#); [Young, Dan](#)
Subject: new app
Date: Thursday, May 14, 2009 1:31:11 PM
Attachments: [Method-to-increase-sensitivity.doc](#)

Gregory Porreca, Ph.D.
Director of Technology
Good Start Genetics

8 St Mary's Street, Suite 621
Boston, MA 02215
P: 617-358-6004
F: 617-358-5677



ML-IPR-000022

Method to increase the sensitivity of sample isolation for DNA sequencing

05-13-2009

Greg Porreca and Bob Wassman

Good Start Genetics

Currently, the ability of DNA sequencing to detect mutations is limited by the ability of the upstream sample isolation (amplification, enrichment, capture, etc) methods to reliably isolate the locus of interest. If one wishes to make heterozygote basecalls for a diploid genome (e.g. a human sample presented for molecular diagnostic sequencing), it is imperative that the isolation method produces near- or perfectly-uniform amounts of the two alleles to be sequenced.

Typically, isolation methods fall into three 'buckets': 1) single- or several- target amplification (e.g. uniplex PCR, 'multiplex' PCR), 2) multi-target hybridization enrichment (e.g. Agilent SureSelect 'hybrid capture', Roche/Nimblegen 'sequence capture'), and 3) multi-target circularization selection (e.g. molecular inversion probes, padlock probes, 'SeleCirc'). Each of these methods can result in a pool of isolated product that does not adequately represent the input abundance distribution. For example, the two alleles at a heterozygous position can become skewed far from their input 50:50 ratio to something that results in a missed basecall during downstream sequencing. For example, if the ratio was skewed from 50:50 to 10:90, and the sample was sequenced to 10X average coverage, there is a high probability that one of the two alleles would not be observed once in the ten sequencing reads. This would reduce the sensitivity of the sequencing method by converting a heterozygous position to homozygous (where potentially the 'mutant' allele was the one not observed).

We have discovered that the classes of isolation methods above result in different modes of bias in the product. We therefore present a way to increase the sensitivity of the downstream sequencing by using a combination of multiple isolation methods (one or more from at least two of the classes above) for a sample. This is particularly important in molecular diagnostics where high sensitivity is required to minimize the chances of 'missing' a disease-associated mutation. For example, given a nominal false-negative error rate of 1×10^{-3} for sequencing following circularization selection, and a false-negative error rate of 1×10^{-3} for sequencing following hybridization enrichment, one can achieve a final false-negative rate of 1×10^{-6} by performing both techniques on the sample (assuming failures in each method are fully independent). For a recessive disease with carrier frequency of 0.1, caused by a single fully-penetrant mutant allele, the number of missed carrier diagnoses would decrease from 1000 per million patients tested to 1 per million patients tested. Furthermore, if the testing was used in the context of prenatal carrier screening, the number of affected children born as a result of missing the carrier call in one parent would decrease from 25 per million to 25 per billion born.

Additionally, approaches can be combined to effectively increase sequencing coverage in regions containing disease-associated alleles. Since heterozygote error

ML-IPR-000023

rate is largely tied to both deviations from 50:50 allele representation, and in the case of next-generation DNA sequencing deviations from average abundance (such that less abundant isolated targets are more likely to be undersampled at one or both alleles), selectively increasing coverage in these regions will also selectively increase sensitivity. Similarly, the algorithms used to determine base identity can use varying levels of stringency depending on whether the given position has any known disease-associated alleles. Stringency can be reduced in such positions, and effectively increase sensitivity for mutant allele detection at the cost of decreased specificity.

ML-IPR-000024