Reviewer's report

Title: Noninvasive Fetal Trisomy Test (NIFTY) An Advanced Noninvasive Prenatal Diagnosis Methodology for Fetal Autosomal and Heterosomal Aneuploidies

Version: 3 Date: 23 March 2012

Reviewer: Joanna Holbrook

Reviewer's report:

Jiang et al Illumina sequenced the cell-free DNA of plasma from 903 pregnant women over a large range of gestational ages. The generated 2-4 million reads per sample and after QC were able to align an average of 1.7 million per sample, to the reference genome.

They develop the “NIFTY” approach to detect aneuploidy incorporating correction for GC content, fetal DNA content estimation and a separate procedure for gender detection.

They found:

• k-mer coverage had strong chromosome-specific relationship with GC content and the direction of the relationship depended on the chromosome average GC content.

• k-mer coverage is dependent on GC content of the chromosome – to the extent that GC-rich or GC-poor chromosomes have similar (higher?) k-mer coverage.

• Amount of cell free fetal DNA was weakly but positively correlated with gestational age, from an analysis of a subset group of 408 male fetal samples (+19 additional)

• Detection power of aneuploidy unsurprisingly increases with gestational week and number of unique reads. More surprisingly, sensitivity and specificity were better for male fetus’.

• By plotting kmer coverage against GC content, for chromosomes 13,18,21 in the 903 cases – aneuploid cases were clearly seen as outliers and could be detected with 100% sensitivity and specificity.

• For the X and Y chromosomes, high sensitivity and specificity was also achieved using the binary t-test methodology.

• When NIFTY is compared to the approaches of Quake et al and Lu et al in the study dataset: NIFTY performs best, Quake et al performs worst

Major Compulsory Revisions

The author must respond to these before a decision on publication can be reached. For example, additional necessary experiments or controls, statistical mistakes, errors in interpretation.

1. I can clearly see the outlier status for the T13, T18 and T21 cases in A, B and
C and understand how they would pass the t-test described in the methods but it’s hard to see how the XO, XXY and XYY cases are detected in D and E – please provide the values used for the t-tests and the logL values.

Minor Essential Revisions

The author can be trusted to make these. For example, missing labels on figures, the wrong use of a term, spelling mistakes.

1. State when the karyotyping used a gold standard for the predictive method – was carried out. Is this karyotyping on neonate blood, cord blood? Or from amniocentesis? Were there any controls for chimerism?

2. State accuracy of gender calling. I guess 100%?

3. Define tag number. Is this tag for chromosome?

Discretionary Revisions

These are recommendations for improvement which the author can choose to ignore. For example clarifications, data that would be useful but not essential.

• The relationship between cff-DNA and gestational week in figure S3, looks exponential to me rather than linear. Suggest data transformation and/or MIC analysis to better describe dependence.

• Please be more explicit in the interpretation of supplementary figures 1 and 2. Does high GC content tend to increase k-mer coverage in the context of GC-rich chromosome??

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:

I declare that I have no competing interests