Epigenomics AG, Berlin
Epigenomics, Inc., Seattle

Friday, July 18, 2003



Dockets Management Branch (HFA-305)
Food and Drug Administration
5630 Fishers Lane, rm. 1061
Rockville, MD 20852

Subject: Stakeholder Comments on Multiplex Text Guidance Document, Docket No. 03D-0120

**Introduction**

Epigenomics, Inc. would like to thank the FDA for the opportunity to comment on the Draft Guidance titled, "Multiplex Tests for Heritable DNA Markers, Mutations, and Expression Patterns," issued April 21, 2003, docket number Docket No. 03D-0120.

It is valuable to industry for the FDA to develop this guidance in advance of the first submissions of multiplex tests, and that the document takes a broad view of what constitutes a multiplex test, that is, more than just chip arrays.

Our comments on the guidance fall into two categories: 1) characterization of the range of targets for molecular tests, and the regulatory implications, and 2) specific comments about the document text.

Epigenomics' interest in this guidance comes from our work identifying DNA methylation markers that can be used as components in diagnostic tests, either as ASRs or as approved kit components. It is expected that a small panel of markers (2 to 5) will be used together to make a highly sensitive and specific test. Our initial focus is on markers for the early onset of cancer. If the Office of In-Vitro Devices would like to understand more about DNA Methylation technology, we could make an informal scientific presentation, at your convenience.

**Characterization of Molecular Markers**

The introduction to the draft guidance document, especially Section III. Genetics vs. Expression, highlights some key differences among molecular markers. We would like to add several other distinctions that we believe are important and relevant to the development, implementation, and interpretation of multiplexed molecular marker tests.

First, not all DNA-based tests are genetic. DNA methylation markers, which constitute our area of expertise, represent markers that are epigenetic in nature (not genetic), are reflective of gene expression, are not "fixed", yet are "DNA-based". As such, DNA methylation markers occupy a unique position intermediate to genetic and expression markers. On the one hand, they enjoy the benefits of a chemically stable DNA analyte. On the other hand, they vary among different somatic tissues, and can change over the course of a lifetime to reflect environmental and dietary exposures.

Second, perhaps a more important distinction between different types of markers is not "genetic" versus "expression", but "germline" versus "somatic". The reason that this distinction is important is that somatic events generally result in mosaicism in the biological sample to be analyzed, whether the event is genetic, epigenetic, or expression-based. For example, a p53 mutation that arises late in tumor development may be present only in a subset of the samples obtained from a

single tumor. This is a genetic event for which the detection is affected by tissue heterogeneity. Such tissue heterogeneity becomes an even greater problem with quantitative measures, as in the case of gene expression markers or quantitative genetic events such as Loss of Heterozygosity or gene amplification. Likewise, DNA methylation markers can behave as continuous measures due to variations in tissue composition. The sample heterogeneity of markers that reflect somatic events is an important issue that will need to be addressed for clinical specimens subjected to multiplexed molecular marker assays, whether these events are genetic or expression-based. This issue is important as it relates to the ability of various markers to be detected in a background of non-affected tissue or analyte (sensitive detection). Generally, such detection is more straightforward for qualitative, rather than quantitative changes.

Third, not all genetic markers are fixed in time. We refer to this as the "temporal stability" of the marker. Somatic mutations can accumulate over time, and in some cases can even be reversible (e.g. gene amplification). The only markers that are truly invariable are not "genetic" markers, but "germline" markers. Indeed, these are not DNA-based "changes", but DNA-based variants in the population. Among somatic events, there is a gradient of temporal stability with somatic mutations being most stable, followed by DNA methylation changes, and with expression markers showing the most temporal variability.

Fourth, the distinction between discrete measures and continuous measures is a useful one. However, not all DNA-based tests will generate discrete data. Some genetic tests yield quantitative continuous data (e.g. gene amplification, microsatellite length, etc.). In addition DNA methylation data can be discrete (methylation at a single CpG dinucleotide on an individual DNA molecule), or it can behave as a continuous variable (e.g. genomic 5-methylcytosine content or locus-specific methylation levels in heterogeneous DNA samples). Data structure of various DNA methylation measures can be quite complex, and will depend on assay technology.

Many of the issues raised here are relevant to molecular markers in general, as opposed to being specific for multiplex issues. However, given the unique position of DNA methylation markers as DNA-based, epigenetic (non-genetic) molecular markers reflective of gene expression, we thought that it would be useful to elaborate on some of these other distinctions. We have generated the accompanying Table to help organize our own thoughts on molecular markers (see Attachment 1). Perhaps you will find this overview of some use in developing strategies for evaluating multiplexed molecular markers.

Given a molecular marker, two characteristics which we think are important to IVD validation are:

1) How the target varies over time
2) How the target varies over the sample

**Specific Comments**

The following comments on are directed at specific content from various sections of the draft Multiplex document. We understand that providing guidance is a complex task and appreciate the effort that has gone into the document so far.

The document contains some material that is not specific to multi-marker panels. It describes procedures that should be standard for single or multi-marker panels. It is hard to tell which requirements are specific for multiplexed assays and which are about molecular tests in general. It might be useful to have two guidance documents, one to address molecular marker based tests, and one to address multiplexed tests.

The words "target, pattern, marker and mutation" are used throughout the document. It is unclear what is meant by "pattern". A general definition of "pattern" might be, a combination of multiple markers, however, how one combines those markers into a pattern is unclear. For instance can the marker be combined using multivariate techniques. If the word "pattern" means something different, then the document seems to require that each of the markers be validated independently as single markers. This would seem unnecessary to us, because the big advantage of multiplexed assays should be that multiple markers together, the "panel," has higher performance than the individual markers.

The document discusses cut-off points. If we consider multiple markers, it may be clearer to use the expression "cut-off hyperplanes". The fact that techniques such as support vector machines (SVM) are mentioned in the document, and which we use regularly, reinforces this point.

We are pleased that the document explicitly allows cross-validation techniques (leave-one out, jackknife, etc.) when no independent test set is available. This is a statistically sound technique and one that is in regular use in our company. It is helpful that the document acknowledges this technique.

Multiple testing correction methods are not discussed, but perhaps they should be. During multi-marker development, it is a tendency to perform multiple analyses on the data, to find a question which the markers discriminate. Statistical results of these studies must be corrected for the multiple questions asked to avoid identifying spurious correlations.

The ideas presented in the document on proving equivalence between two tests is quite useful, especially the consideration give to a lack of gold standard. Epigenomics expects that in the long run molecular markers will be more informative than previous diagnostic techniques, so the issue of gold standard is very important. We note again, that this is really about molecular markers in general, and not specifically about multi-marker tests.

**Closing**

Finally, it may represent a biased view, but from Epigenomics perspective, a more informative titled for the document would be: "Multiplexed Molecular Marker Tests; Draft Guidance for Industry and FDA Reviewers."

Thanks for your attention to our comments. We look forward to reviewing the next draft of this valuable document.

Thomas Skillman 7/18/03

Thomas Skillman
Director of Regulatory Affairs
Epigenomics, Inc.
1000 Seneca St. Suite 300
Seattle, WA 98101
tom.skillman@epigenomics.com
206-883-2914

Attachment 1

| MOLECULAR MARKERS | | | | | |
|---|---|---|---|---|---|
| | **Germline Polymorphism** | **Somatic Mutation** | **DNA Methylation** | **RNA Expression** | **Protein Expression** |
| **Type of Marker** | Genetic | Genetic | Epigenetic | Epigenetic/Expression | Epigenetic/Expression |
| **Analyte** | DNA | DNA | DNA | RNA | Protein |
| **Stability of Analyte** | High | High | High | Low | Moderate |
| **Type of Change** | Qualitative | Usually Qualitative | Qualitative and Quantitative | Usually Quantitative | Usually Quantitative |
| **Type of Measure** | Discrete | Discrete | Discrete and Continuous | Continuous | Continuous |
| **Temporal Stability** | Fixed | Stable | Moderate | Variable | Variable |
| **Tissue Heterogeneity** | Homogeneous | Heterogeneous | Heterogeneous | Heterogeneous | Heterogeneous |
| **Sensitive Detection in Heterogeneous Samples** | Not Applicable | Yes | Yes | Difficult | Difficult |
| **Informational Content** | Moderate | Low | High | Very High | Very High |
| **Technical Issues** | Difficult to assay large numbers of selected SNPs | Tissue heterogeneity; uniqueness of each event | Challenging assays | Unstable analyte; Temporal and tissue variability | Anonymous protein fragments |
| **Multiplex Issues** | Haplotype linkage | Rare occurrence of many multiple markers | Technical difficulty of multiplexing many markers | Noise, False Clusters | Noise, False Clusters |