**Overcoming the Unknown:**
**New Approaches to the Diagnosis and Treatment of Carcinomas of Unknown Primary.**
**London 15th October 2009**

# Diagnosis of metastatic cancer by molecular classification

**Mark Eriander**

I work for bioTheranostics, which is a wholly owned company by bioMerieux located in San Diego. We develop molecular diagnostic tests that are focussed on the oncology area and I am going to talk to you today about the molecular profiling that has been done in cancer classification.

To give you a little background about us before I begin a more general talk:

- We have a clinical lab in San Diego and we do real world testing there.
- At this point we have tested over 2,000 patient samples, primarily metastatic cancers or different types of biopsy where the origin of the cancer is uncertain or unknown. This gives a flavour, through real world testing, on what a lot of the issues really are. Beyond the single diagnosis of CUP there are a lot of metastatic cancers which are uncertain rather than unknown.
- We also have in the United States over a thousand oncologists that have used our tests.

The molecular component is definitely the 'new kid on the block' when it comes to integration of information about metastatic cancer origin and what we want to do is to provide a new piece of independent information that allows the oncologist to make a better decision as far as treating their patient. This is our challenge.

If you step back and talk about the molecular classification of cancer I think a lot of you know that there are many different avenues of research that are going on today – translational research to be able to move molecular into the clinical setting, where it can actually have practical application. We have many signatures that have to deal with risk, drug response, different disease subtypes and tumour microenvironment. Today I am going to talk to you about tumour origin of metastatic disease. I am going to try and give a more general overview of where we are today as far as the use of molecular profiling of metastatic cancer and classification of cancers is concerned.

An example of the emergence of genomic testing is the story of Steve:

- May 2008 – Steve, an active, otherwise healthy electrician began losing weight. For months he complained of severe stomach pain and cramping.
- Months later, after urgent care visits, Steve went to Sand Diego's Scripps Memorial Hospital where CT identified inoperable cancer of unknown origin between his liver and kidneys.
- Doctors said that Steve's cancer had been metastisized and its origin could not be identified and he had 4-6 months to live.

This is an example of a success story for molecular profiling and could be from Bio Theranostics or from another company who offer these kinds of services. In essence it is a typical example of an unknown or uncertain primary but in this particular case it turned out to be a germ cell tumour, which I will discuss more fully later in my talk.

These are the types of success stories that, as molecular biologists are moving into the clinical area, we obviously are very excited about.

How did it all begin? It began back in 1999 at the Broad Institute. Eric Lander and his group showed for the first time that you could actually use gene expression profiling to classify two different types of leukaemias. Their work showed that you could use gene work as a classifier. The specific expression pattern of the two genes allowed you to classify these different types of leukaemia . This was really a way of moving away from subjectivity and really into quantitation. Looking at the differential gene expression of specific genes allowed them to set up what they

called a class predictor. Once they had developed their predictor they could then test this in an independent set to validate this prediction or out rhythm.

Since this time there have been a lot of publications and papers showing that in fact you can take this beyond these two cancer types. For example Sridah Ramaswamy, who is now at MGH, demonstrated at the time that you could take fourteen different tumour types and use the same sort of approach. You divide the tumours into two groups. One is the training set where you develop your predictor and then you take an independent test to demonstrate the performance of that predictor in a blind trial.

The general approach that all of this work takes, wether by ourselves or other companies that offer these kinds of services, is that you would first take a tissue biopsy. It is key that you have a pathologist up front looking at that biopsy to check firstly 'Does it have a tumour on it?', and secondly 'If there is a tumour does it need to be macro-dissected or micro-dissected?' Many times we receive core needle biopsies from the liver and half of the core will be normal patastites and the other half will be malignancy. I see molecular profiling is based on looking at residual markers of the development of the cell lineage of those cell types, so it is very important that you get the right cells. In the clinical setting versus the research setting we have to learn to abide by the rules of pathologists. They like to use formalin; they like to fix it this way because it gives them maximum morphology. We need to be able to have a methodology, if we are going to get into gene expression, of being able to get RNA out such that, although degraded, it can still be measured robustly and quantitatively to be able to make predictions of origin. This process is called demodification.

You can then extract your RNA and use this in the two different types of gene expression , or RNA platforms, that are used today. One is microarray based and the other is quantitative RT-PCR. Once these platforms have been used to measure the expression of a particular gene a pattern will have been produced. An algorithm is used to attempt to find the best match of that pattern to a set of patterns of truth. Truth being many different tumour types where you already know the truth of the tumour and you have a pattern associated with that truth. Once you align the results with those of the greatest similarity then you can make a prediction.

This is, in essence, the process of how molecular profiling is used today to be able to make predictions of origin or classification of cancer type.

In the different platforms in use today you have Microarray and real time of qRT-PCR (Reverse Transcriptates) .The advantage of microarray are of course that you can look at many genes. You can look at thousands of genes which allows you to have potentially a greater discriminatory ability between cancer types. Real time PCR, as we all know, is the gold standard as it has increased sensitivity, it is more reliable and it is also amenable to micro dissection, because in micro dissection you get a very small number of cell types. The disadvantages of microarray technology are in essence that it is more of a research tool because you are looking at so many thousands of genes you take a bit of a hit as far as dynamic range. That is to say, it is basically limited in its sensitivity. It is also more variable and complicated to analyse after the analysis. For qualitative PCR, although it has many advantages, the one thing you have to keep in mind is that because you are using a smaller number of gene set you do not have the redundancy built in, so in other words, if you incorrectly measure one of those genes you may be in trouble. It is very important that you have very robust methodologies and SOP's within you clinical lab and also ways to do QC.

Molecular classification of cancer is really closing the diagnostic gap. Genomic technology provides what we consider to be an independent source of information that we think is key. It brings an independent piece of information to the table to help the oncologist make a better decision. The basic premise for IHC and gene expression is that tumours retain a sufficient gene expression signature that reflects the origin, despite the fact that there is dedifferentiation and metastases. This is a very key assumption, and one that has been shown to be true.

The strength and value of the molecular approach is higher resolution, which we believe, compliments the current diagnostic modalities. The bottom line for all of this is that basically for targeted therapies of metastatic disease these are tumour specific and so this classification aligns with the oncologist making a decision about treatment.

There are today various commercial entities that offer molecular cancer classification schemes:

- Veridex
- Pathworks
- Prometheus which is actually licensed from Rosetta Gernomics
- Bio Theranostices

*Veridex:*

This was a 10 gene classifier that allowed you to classify six different cancer types. Although it is not commercially available it actually paved the way, and there were some very key studies that were done such as those talked about earlier by Dr Varadhachary. The bottom line is that overall accuracy is 78% and 73.3% in metastatic cancers.

This was using a methodology whereby you are looking at six different cancer types where they, in essence, used markers that were specific to different cancer types. Much like IHC is done today this was the first generation of molecular profiling. As we move forward we do not use this type of methodology, in fact what we do instead, is that we use collectively all the genes to be able to classify all the different cancer types. That actually gives you more power because you are not relying on specific markers.

Dr Greco and Dr Varadhachary demonstrated by using this method the first data to show that you could take CUP and using molecular profiling you emerge with a new paradigm of subset identification. This was a key step from my vantage point of moving forward from using more clinical ways of classifying CUP patients into good and poor prognosis into using molecular profiling and also IHC to identify subsets that you could then treat more effectively. The bottom line of the data which they published in JCO was that for those patients who received specific colon cancer regimens (all these patients were predicted to have colon primary origin) they had a much better outcome than those patients who were predicted to have colon but were treated by standard chemotherapy. This was a very significant movement forward.

*Pathworks*:

Pathwork's has a Tissue of Origin Test where they use a microarray platform and look at 15 different tumour types. They have also moved from the frozen tissue as they realised that in order to have a more practical system in the clinic they need to move to a form of fixed paraffin embedded sample. They have gone on to show that their assay works in FFPE as well.

Pathworks report on the fifteen type of tumour. When a test sample comes in of a particular metastatic cancer, usually unknown or uncertain primary, they give it a similarity score from low to high. They then compare the similarity score against their database of truth.

Pathworks also published a paper where they went on to look at using their system for carcinoma of unknown primary. They worked on 21 samples which had been diagnosed as CUP and they went on to profile all of them and see what their predictions were. In this particular study there was no truth, that is to say, they were CUPs and no one knew what the primaries were. They were able to evaluate their predictions versus the initial path reports to see if there was any concordance or discordance. What they found was that about 33% were consistent with clinical pathology data, 43% were inconsistent and 24% were indeterminate. This is not to say that this is not a good test. Obviously they

are a competitor of ours but on the other hand the fact of the matter is that there was no truth. Which was right, which was not right? One does not know in this particular experiment because there is no truth. However one thing that it does highlight is shown on this slide. Here are a couple of patients where they predicted colorectal origin showing the therapies used, and the IHC and you can see that that would have changed therapy in both cases if the system had been used.

The next generation from a technology point of view is to consider RT-PCR, which is a much more quantitative measurer of gene expression. I will talk about two tests, both of which use either whole genome or a micro RNA expression profiling of tumour specimens and from that they build reference databases to classify using discriminatory genes or micro RNAs . You then select a subset that is much more compact and smaller in number, where you are really in a clinical setting where you are required to have a very robust test and here we are using RT-PCR or, in the case of micro RNA they use a different method but it is similar in the sense that that it is real time PCR.

*Rosetta genomics/Prometheus:*

Rosetta genomics license their process to Prometheus who are based in San Diego. They have a test which is based on 48 miRNAs which quantifies 22 different tumour types. It is compatible with FFPE and has excellent overall accuracy. No performance or technical specifics are at present available. It will be commercially available later this year.

They can classify 22 tumour types and you will hear more about this later this morning.

*Bio Theranostics:*

Our test was developed using whole genome gene expression profiling of over 600 tumour samples. We also included metastatic. We started by profiling 22,000 genes and using BioPharmetic tools brought that down to 87 genes of content, and 5 genes which were used for normalisation of input of RNA. This test is offered in our field lab today.

The sample requirements for us is that we need about 300 cells or greater, which is actually a very small number. This is why we are able to be successful with core needle biopsies which are being used more and more today because of morbidity issues and other things. We are routinely using samples that are coming from FNAs. We are compatible with FFPE and one of the issues for us, and anyone who is working with RNA, is that there are two ways that you can decal a bone biopsy; one of them uses HCL and the other formic acid. Formic acid is good, HCL not so good, so we are trying to educate the pathologists more about this because they could use either one. Maybe a pathologist can tell me that one is better, but I am holding out that formic is best as far as morphology is concerned.

How do we carry out our process? We start, and this is absolutely key, with a Board certified anatomical pathologist to make sure that we are dealing with a malignancy and also where the malignancy is on the section. The second thing that I have talked about that is extremely important is micro-dissection. This slide shows a base of cancer surrounded by stromal cells and that you can actually take those evasive cells out and isolate only the invasive cancer that is of interest. Once this has been done you can extract the RNA, generate cDNA, perform real time PCR, compare against the algorithm and produce results which are returned to the Medical Director, who signs out the case.

One of the things that I want to point out is that the earliest generations of gene expression profiling used specific or tissue markers. Unfortunately Mother Nature did not give us PSAs for every single tumour type or every tissue of origin so we found, through a lot of iteration from competition analysis, that in fact a much more powerful way to classify cancers is not to rely on, and look for, those markers but instead to identify a set

of genes that collectively, as in a collective discretion profile, will allow you to be able to classify cancers. This was a key point and a key fork in the road for us in understanding a better way to use gene expression for classification of cancer. It was important to us to use a data driven process. When we profiled the 22,000 gene expressions across 600 tumour types we did not use any preconceived knowledge. We did not say 'Well I happen to know from reading the literature that this particular marker is specific'. Instead we used a data driven process where we demanded the identity of the best set of collective set of genes that would allow us to classify those cancers. This was a very key step as well in the development of this test.

From this the interesting question arose of 'What did we find?'. 'What were the genes that came out of this screening?':

1. A lot of transcriptional factors that are important for developmental biology. This makes a lot of sense. We are really talking about looking at lineage of cell types and the fact that they still maintain or retain some of that.
2. Some high tissue specific genes.

We have 39 different types of tumour that we classify. At bioTheranostics we believe that you want to have as big a universe as possible. Even though often we will say that there are only six really important cancers from a real world point of view you are getting all kinds of metastatic cancers coming in. There are all kinds of possibilities and the greater your universe of classification the better shot you have at getting it right and, of course, that is the bottom line.

A report feature that we have includes a graph with the 39 different types on the y axis and a probability on the x axis. This report, therefore, shows that the sample that you have sent in is most similar to a particular cancer type. The other key feature of our report is that we also do a rule out, so everything below a certain point has a 95% confidence of ruling it out as the origin.

We did a blinded study with Dr Greco with latent primaries. There were 28 different patients who had biopsies that were actually able to provide some tissue and of those we were able to get 20 samples that would yield enough for RNA. The reason we got into this study with Dr Greco was that there was the opportunity to do work with him where he had had CUP patients who later on, through clinical manifestation, were able to identify the primary origin of that cancer. This was a way of dealing around that chicken egg loop of logic that you can never get out of. Tony Greco told me on the phone that if I could get even half of them right he would consider this to be not too bad a test. We were able to get 75% correct and this was reported at ASCO a couple of months ago.

The clinical implication for us is 'What do we have?' . What we are building here at a molecular level are cancer classifiers and very succinctly, as written by one of my colleagues, this is a gene expression-based assay that accurately predicts tumour tissue type on the basis of the expression levels of key discriminatory genes. The clinical indications that we see in the real world testing are:

- patients whose oncologists or pathologists send in samples of a typical clinical presentation. This is a key type of sample that we get.
- Many times we get cases in where there is a differential and many different choices. As Dr Owen was saying: if you have a CK7 basically positive and CK20 positive you have lots of possibilities.
- Further to diagnose cancer of unknown primary.
- The fourth type of utility we see is that many times pathologists or oncologists already have a pretty good idea of what they think it is but they like to have a non subjective independent consult in this case of a molecular profile.

Earlier I mentioned a patient who presented with a metastatic legion and this slide shows the pathology report from Scripps Memorial and it was basically a superclavicular lymph node classified as an undifferentiated carcinoma. They then did a pretty big work up from an IHC point of view to see what it was. They looked at markers that were epithelial in origin, and looked at melanoma markers, lung, lymphoma type markers and they also several germ cell markers. We predicted a seminoma. After the prediction they went back and looked at Oct ¾ as well, which was IHC negative. The oncologist treated it as a germ cell and, as of today, the patient has no detectable cancer. So this was a success story about using molecular profiling. This was very exciting for us as it really showed that you could use gene expression profiling as an adjunct to bring new information to the table that helped this particular patient.

In summary I have shown you the different technologies out there today and the number of tumour types that they classify. Also the tissue requirements where, if you have frozen tissue, you need a lot of tissue. We can, however, work right now with about 300 cells and a very small number of sections.

Classification of the origin of metastatic cancers is basically:

- Using gene expression profiling has transitioned us into using a new tool that will continue to be validated and be demanded to be validated by the research and clinical communities.
- I have shown you that there are various platforms.
- Overall molecular profiling is giving you prediction accuracies of around 80 to 90%.
- We and other investigators believe that molecular diagnostics is really going to reduce the level of what we consider to be diagnostic uncertainty. That is the goal.
    - We believe that it also increases the predictive accuracy by complementing current IHC and imaging methods that are being used today, and
    - We believe also that by using molecular tools we should be able to identify subsets of CUP that would potentially benefit from treatment based on identification of tissue origin.

Thank you very much for your attention and I hope that this presentation was useful.