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

Using microRNA Profiling to accurately identify tumour origin
Dalia Cohen

Our test is known as miRview™mets all over the world, except in the USA where it is known as Proto/oncom. Rosetta Gernomics is an Israeli based company and their CLIA laboratory in Philadelphia serves the whole world.

miRview™mets is the only assay that uses microRNA-based tests to identify primary origins of metastases and it can identify 25 different tumour types. It leverages proprietary, and highly sensitive, microRNA technology to measure the expression level of a total of 48 microRNA biomarkers. It uses a classifier to assign primary sites to the cancer samples based on microRNA expression.

MicroRNA is no longer the 'new kid on the block' although for a long time they were considered such. MicroRNA is a general term for a big family, which increases by the day, of non coding RNA. Their functions are mainly to emulate gene expression. The old dogma was that from DNA to transcription of mRNA and then translation to protein. Today we know that from DNA there is a step in which microRNA regulates the expression of the mRNA and this regulation is based mainly by inhibition of translation of the protein.

When considering the Biogenesis and processing of MicroRNA it is important to realise that the biology of microRNA is being transcribed from the DNA in the same way as mRNA and regulated like mRNA. It has been transcribed to a long chain with helpeen. This long transcript is being processed by Drosha to a small helpeen of about 60 betas and that is transported into the cytoplasm by a complex transport system. The Dicer then releases the stem loop, resulting in a double strand miRNA duplex, one strand of which becomes part of the RISC complex which is the silencing complex. Then seeds are hybridised on the three prong UTR of the mRNA and mostly this hybridisation, or binding, results in the ebishion of consolation.

At Rosetta we have profiled microRNA from over 10,000 samples and based on these many samples, when compared to protein and mRNA, microRNAs have been shown to have greater tissue specificity. We are using technologies that were developed in-house to sample each microRNA with high accuracy.

MicroRNAs are considered to be excellent biomarkers because they are highly tissue-specific. They are sensitive and are expressed across various pathological conditions, which makes sense if we think about their function. Their function is regulation of gene expression, and as of today, I do not know of any disease where microRNA has not been described as a component of the disease regulation. As of today in the world of microRNA the belief is that microRNA are regulating 80% of gene and ???????.

We have a straight forward algorithm, a very straightforward and transparent process which is easily explainable to the physician. MicroRNA have been shown to be more stable than mRNA in paraffin-embedded block samples (FFPE). This is basically our experience as well as it being recently described by Louis Hull in a direct comparison. We can use material between 2 to 11 years in a paraffin block and get the same quantity and quality of microRNA in tissue with our test.

The miRview™mets test:

- Identifies 25 different tumour types
- Leverages proprietary microRNA technology
- Utilizes a proprietary classifier
- In the majority of cases the test report s an origin accurate in nearly 90%, with specificity of 99%

- In the remaining cases two possible origins are reported:
  - Most likely
  - Second most likely
- Utilizes a relatively small number of molecular markers with a simple classifier

The test started with a Discovery Phase which was published by Rosenfeld et al in Nature Biotechnology. In this study expression amounts of over 600 microRNA were measured on array. More than 300 tumour samples were collected, including 131 metastatic tumours from 22 tissue types. A decision-tree was developed to identify the tissue of origin using 48 microRNAs. The algorithm was combined with a KNN algorithm to assure high accuracy of prediction. 83 samples (about a quarter of the data) were kept aside for a blind test set.

One of the questions that we asked from the beginning was 'How close are metastases compared to the primary tumour?'. There was a lot of work done on tumour of origin and mRNA but we asked ourselves how similar, or dissimilar, those metastases were from the origin. On the slide here are two examples. The first example is a perfect match of a primary colon cancer compared to a metastatic one. Basically the two profiles of microRNA and mRNA are identical. However when we consider the example of a primary stomach cancer compared to metastases in the lymph node here most of the microRNA were similar to each primary and the metastases, however there were specific ???? which is basically due to the environment. a-miR 150, which has been shown to be a miR which is highly ??? in lymphocytes compared to miR 143, that has been shown in epithelial cell as well as a muscle related miR 133, which is a very specific muscle microRNA. For every primary tumour, therefore, we look on the expression of microRNA comparing to each other the metastases and the primary and took these into account when the tree was developed.

We have heard already about a tree that was based on tissue microarray, but here we have a tissue decision tree which is based on microRNA. This decision tree is biologically motivated, has far fewer features, creates context, is very easy to incorporate external data into and is very transparent. I will now give some examples of how the tree is helpful in identifying the origin of HER metastases in the following slides:

1. Starting with node number 3, this node can basically divide into three from epithelial origin to non-epithelial origin, based on 2 microRNA. 205 and 200c fits very nicely the separation between epithelial and non-epithelial of the tree.
2. When we found the tumours 205 and 200 we were gratified later to see some publications showing :
   a. the family of 200 and 2065 regulate the epithelial to mesenchymal transition by targeting a specific transcription factor.
   b. The inhibition of epithelial tumours in transition in cancer cell migration

   All of which enhanced our faith that we are looking on a very tumour specific microRNA.
3. Continuing to look at the tree ,before I showed you node number 3 divided between epithelial to non-epithelial , here the tree is divided between GI origin and other epithelial, again looking very closely using, this time 3microRNA 205, 145 and 194 and this gave a sensitivity of 93% and specificity of 97.
4. Going down the tree, this time looking on node 13, using two miRs – miR 21 and let7e again we can differentiate between gastrointestinal and a lung carcinoma.

To summarize what we called the discovery phase. Most of what I have just presented is presented in the paper by Rosenfeld et al mentioned before.

- The classifier is combined from the answer of the decision tree and the KNN classifier
- When the two give the same answer, a single answer is returned, otherwise two answers are given
- We have a total accuracy of 86%
- Two thirds of the samples return with a single answer
- The single answer predictions have 90% accuracy

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

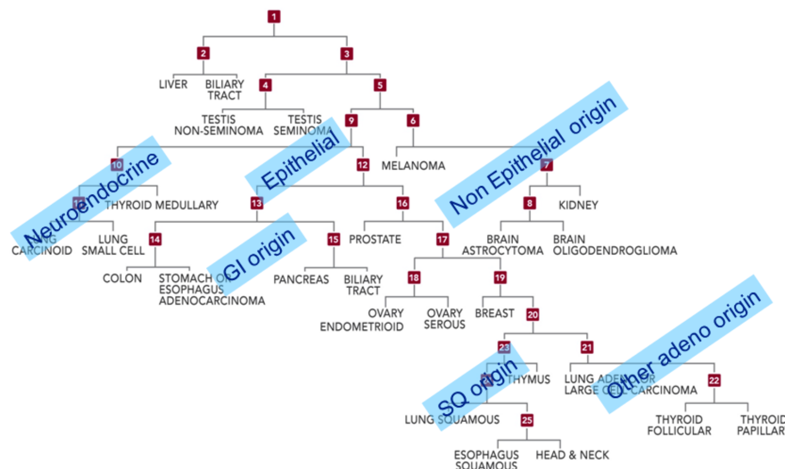
- Most of the tissue classes have 100% accuracy
- Importantly with blinded metastatic samples we got 85% accuracy for single answer predictions.

For the new miRview™mets assay development was based on qRT-PCR and you saw before Mark very nicely described from the chip to pure qRT-PCR, so I am only going to describe the pure qRT-PCR when we are doing our assay. In order to develop the assay hundreds of additional samples were added, along with additional tissue origins and histological types were added to original tissue that we worked with in the discovery phase. We started with over 100 microRNAs that, in the tree structure, was modified to fit the new classes and platform. We have an automatic QA for RT-PCR and in the CLIA Lab, with the CLIA assay, includes 48 microRNAs measured in duplicate and we always have positive and negative controls for each batch of samples.

This slide is a description of the tissues that we have, which shows the different tissues and the different types in the tissue that miRview™mets have in the test.

We looked at whether or not we have similarity between primaries and metastatic tumours and found they were very similar to each other. The example here again shows a perfect match between a lung adenocarcinoma and metastases. In 10 samples from metastatic tumours and 4 from the primary we have a very similar, almost identical, expression.

This slide shows the Decision Tree of today with the additional tissue type. This tree can differentiate the different tumour types from the neuroendocrine , epithelial, GI origin, non epithelial, squamous origin and other adeno origin.



This next slide shows examples of work carried out on Node number 12 using only 2 microRNA and I think that this is where the tissue specificity of the microRNA have been described so you can see here in this specific case miR192 in combination with the expression of miR106 can differentiate very nicely between the digestive and the non digestive drugs.

In the assay we added 240 new blinded samples, 44% of which were metastatic tumours, with a sensitivity of 84% and specificity of 97%. 67% of the samples were classified with a single answer which gave a sensitivity of 90% and specificity of 99%.

I now present an example of success in test re-classification. In this specific case the pathologist identified lung Ca metastatic to the brain, however, when we did our assay the test classified a hepatocellular carcinoma. There was a very high expression of miR-122 which is known to the literature and by us, to be highly specific to the liver. A blinded re-examination was done the results of which were negative to lung specific markers and positive for several markers, including

HEPA 1, to liver specific markers. The tree of our test using a combination of miR 200c and miR 122 showed how we distinguish between samples.

We are in the process of doing further validation on our test. A large blinded study is underway in conjunction with M D Anderson which is using 100 samples from CUP patients and our results will be evaluated in light of clinical and pathological patient characteristics. We have another study in Heidelberg in Germany, the first phase of which was 104 metastases to the brain of known origin and we have now started the second phase with 60 metastases to the brain which were initially identified as CUP. The classification will be done in Heidelberg with respect to possible diagnoses. These will also be blinded.

To summarise the test we are able with the miRview™mets to identify the primary site of metastatic tumours for patients with cancer of unknown primary. The studies have shown a relatively small number (48) of microRNA which can identify the tissue of origin with high accuracy. A qRT-PCR assay was developed and achieved high accuracy of classification on validation samples.

I would like to end my presentation by thanking the Rosetta Genomics Team and our many collaborators.