

Development and validation of a second generation microRNA-based assay for diagnosing tumor tissue origin



Ayelet Chajut¹, Shai Rosenwald¹, Tina B. Edmonston², Iris Barshack^{3,4}, Meora Feinmesser^{4,5}, Monica Huszar⁶, Wolf Mueller⁷, Franz Fogt⁸, Yael Spector¹, Ranit Aharonov¹ & Eti Meiri¹

¹Rosetta Genomics Ltd., Rehovot, Israel; ²Rosetta Genomics Inc., Philadelphia, PA; ³Department of Pathology, Sheba Medical Center, Tel-Hashomer, Israel; ⁴Sackler School of Medicine, Tel-Aviv University, Tel-Aviv, Israel; ⁵Department of Pathology, Rabin Medical Center, Beilinson Campus, Petah Tikva, Israel; ⁶Department of Pathology, Kaplan Medical Center, Rehovot, Israel; ⁷Department of Neuropathology at the Institute of Pathology, Ruprecht-Karls-University Heidelberg, Heidelberg, Germany; ⁸Department of Pathology, University of Pennsylvania Presbyterian Medical Center, Philadelphia, PA

Introduction

- Cancer of unknown primary (CUP) constitutes 3%-5% of all newly diagnosed cancer cases. It presents a major diagnostic challenge due to the therapeutic implications of the tissue origin of the cancer.
- MicroRNAs are a family of non-coding, regulatory RNA genes that are involved in development, differentiation and cell cycle and were shown to be involved in oncogenesis.
- MicroRNAs are highly tissue-specific and therefore ideal candidates for CUP diagnosis.
- We have previously presented a biologically-motivated classification assay, miRview™ mets, capable to identify the tissue of origin using a set of 48 microRNAs measured on a qRT-PCR platform. This assay can differentiate 25 different tumor types from 17 origins.
- We present here the development and validation of a second generation assay, miRview™ mets², that can identify 42 different tumor types using a custom array platform.

First generation PCR-based assay

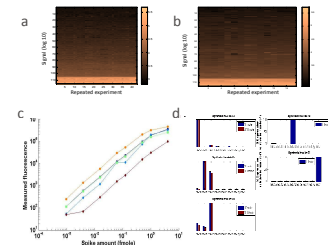
- A CLIA-approved qRT-PCR assay (miRview™ mets) measuring 48 probes, that identifies 25 different tumor types from 17 origins
- Based on two classification algorithms, a decision tree and KNN, as described previously [1]
- Validated on nearly 200 blinded samples [2]
- 86% accuracy for assigning tissue of origin. 2/3 of cases report single origin with 90% accuracy
- Independent validation by experts at M.D. Anderson Cancer Center showed 84% concordance with clinico-pathologic features in real CUP patients [3]
- Independent validation by experts at Heidelberg Medical Center showed 80% concordance with clinico-pathologic features in real CUP patients with CNS metastasis [4]

Tissue origins and samples

ID	Tumor organ of origin	Tumor type (as reported)	# training	# test validation
1	Adrenal	Adrenocortical Carcinoma	19	11
2	Adrenal	Pheochromocytoma	15	13
3	AmesSkin	Squamous Cell Carcinoma, Ames or Skin	28	15
4	Biliary tract	Cholangiocarcinoma or Adenocarcinoma of Extrahepatic Biliary Tract	51	15
5	Bladder/TCC	Urothelial Carcinoma	60	15
6	Brain	Astrocytic tumor (primary)	14	15
7	Brain	Oligodendroglioma (primary)	12	9
8	Breast	Adenocarcinoma of the Breast	57	15
9	Cervix	Squamous Cell Carcinoma, Uterine Cervix	29	14
10	Colon/Rectum	Colorectal adenocarcinoma	45	15
11	GI	Carcinoid, Gastrointestinal Tract	30	9
12	GI	Gastrointestinal Stromal Tumor (GIST)	19	10
13	Kidney	Renal Cell Carcinoma, chromophobe	23	15
14	Kidney	Renal Cell Carcinoma, clear cell	38	15
15	Kidney	Renal Cell Carcinoma, papillary	24	15
16	Liver	Hepatocellular Carcinoma	24	15
17	Lung	Lung, large cell or adenocarcinoma	37	15
18	Lung	Lung, small cell carcinoma	21	13
19	Lung	Carcinoid, Lung	25	15
20	Lung/H&N/Esophagus	Squamous Cell Carcinoma, Lung, Head&Neck, or Esophagus	140	15
21	Lung/pleura	Pleural Mesothelioma	32	12
22	Lymphoma	Lymphoma	108	26
23	Ovary	Ovarian Carcinoma	61	15
24	Ovary	Ovarian Primitive Germ Cell Tumor	5	2
25	Pancreas	Pancreatic Adenocarcinoma	25	16
26	Pancreas	Pancreatic Islet Cell Tumor	11	2
27	Prostate	Prostatic Adenocarcinoma	28	20
28	Sarcoma	Ewing Sarcoma	10	2
29	Sarcoma	Chondrosarcoma	11	3
30	Sarcoma	Malignant Fibrous Histiocytoma (MFH) or Fibrosarcoma	22	9
31	Sarcoma	Osteosarcoma	11	10
32	Sarcoma	Rhabdomyosarcoma	9	2
33	Sarcoma	Synovial Sarcoma	11	6
34	Sarcoma	Liposarcoma	18	10
35	Skin	Melanoma	28	15
36	Stomach/Esophagus	Gastric or Esophageal adenocarcinoma	47	15
37	Testis	Non-Seminomatous Testicular Germ Cell Tumor	18	15
38	Testis	Seminomatous Testicular Germ Cell Tumor	27	15
39	Thymus	Thymoma/Thymic Carcinoma	29	11
40	Thyroid	Thyroid Carcinoma, follicular	16	8
41	Thyroid	Thyroid Carcinoma, papillary	25	15
42	Thyroid	Thyroid medullary	19	6
		Total	1282	509

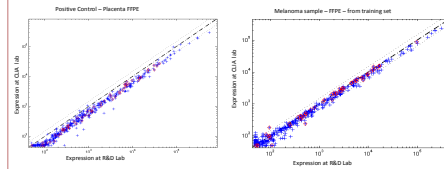
Array platform

- a) Microarray reproducibility of RNA extracted from FFPE sample (15 times) or (b) placenta RNA (42 times) were labeled and hybridized to the microarray. An overall mean correlation coefficient of 0.99 was achieved.
- c) Microarray sensitivity & dynamic range: five synthetic short RNAs were labeled and hybridized to the array. Lowest sensitivity was 0.1 fmole with a linear dynamic range of about three orders of magnitude.
- d) Microarray specificity: 1 and 0.1 fmole of the synthetic RNAs of hsalet-7a, b, c, d & i were labeled and hybridized individually to the microarray. Specificity of above 10-fold in the signal level was achieved



Inter Lab concordance

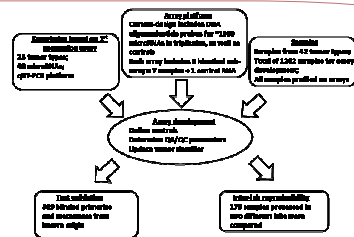
- 179 RNA samples were processed in two different laboratories:
 - Rosetta Genomics' R&D lab in Rehovot, Israel
 - Rosetta Genomics' CLIA lab in Philadelphia, PA
- A Pearson Correlation on the expression of the 64 assay microRNAs of above 0.95 was achieved in 160 (89%) samples
- The two labs agreed on the diagnosis in 175 (98%) of the cases, demonstrating the precision and robustness of the test.



Two samples as an example: Blue crosses are microRNAs on the array, those marked in red are the microRNAs in the assay

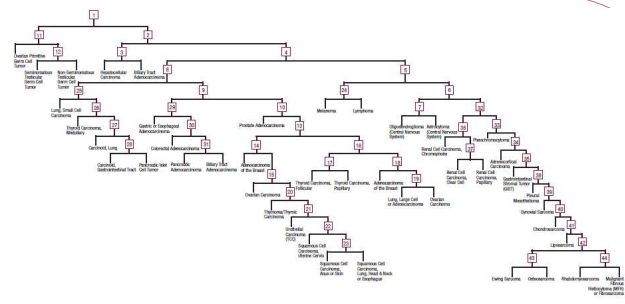
Development and validation of the array-based assay - overview

- Assay development was based on knowledge collected for the development of the first generation assay, as well as profiling of samples from current and additional tissue origins
- The assay was developed on a custom-designed array platform on a total of 1282 primary and metastatic samples from 42 tumor types
- For tumor classification, we used the main components of the classification process used in the first generation assay, with several improvements
- The assay performance was established by a blinded test validation of 509 samples and an inter-lab reproducibility study of 179 samples



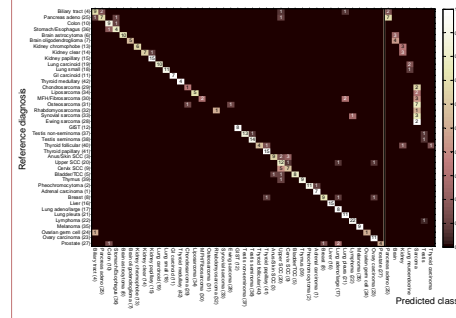
Tumor Classification

- The test was designed to identify 42 tumor types
- The test relies on two classifiers to determine the tissue of origin: a binary decision-tree classifier and a k-nearest neighbors (KNN) classifier [1, 2]
- Both classifiers assign a tissue of origin based on the normalized expression of 64 microRNAs
- Each of the two classifiers (tree and KNN) predicts one of the 42 tumor types or one of the following 7 tumor classes: sarcoma; renal cell carcinoma; Lung, small cell carcinoma or carcinoid; Testicular Germ Cell Tumor, seminomatous or non-seminomatous; Astrocytic or oligodendroglial tumor (primary); Thyroid Carcinoma, follicular or papillary; Adenocarcinoma of Biliary Tract or Pancreas.
- The two predictions are then combined into a single predicted tissue origin or two different predictions, based on whether the two classifiers agree or on their confidence measures
- When two predictions are reported, they are ranked by the positive predictive value of each answer



Performance

- The test performance was assessed using an independent set of 509 validation samples which included primary as well as metastatic tumor samples, preserved as FFPE blocks, whose original clinical diagnosis ("reference diagnosis") was one of the tumor types on which the classifier was trained
- The samples were processed according to the SOP by personnel who were blinded to the original reference diagnosis, and classifications were automatically generated by dedicated software
- 489 samples (96%), including 146 metastatic tumor samples were processed successfully and produced tissue-of-origin predictions
- 418 samples (85%), the reference diagnosis was predicted by at least one of the two classifiers, resulting in an overall sensitivity (positive agreement) of 85%. Specificity (negative agreement) was above 99%
- 403 samples (82%) produced a single tissue-of-origin prediction
- For these single-prediction cases, the sensitivity (positive agreement) was 90% (361/403).



Confusion matrix on the validation set for the 403 single-answer cases (numbers in parenthesis represent the tumor type ID as in the table)

Summary

- Previous studies have demonstrated the high tissue-specificity of microRNA expression, and have demonstrated their significant potential to classify human malignancies. An assay utilizing microRNA expression in a biopsy for identifying the tissue of origin has been in clinical use (miRview™ mets).
- A second generation assay was developed which can identify 42 different tumor types by utilizing the expression of 64 microRNAs as measured on a custom-designed microarray.
- A blinded validation set of 509 samples was studied and overall provided an accurate result in 85% of the samples; in 82% of cases a single tissue-of-origin prediction was given with an accuracy of 90%.
- This assay (miRview™ mets²) provides an important new tool for diagnosing tumor tissue origin.

References

1. MicroRNAs accurately identify cancer tissue origin. Rosenfeld, Aharonov et al. *Nature Biotechnology* 26:462-469, 2008
2. Validation of a microRNA-based qRT-PCR test for accurate identification of tumor tissue origin. Rosenwald, Gilad et al. *Modern Pathology* 23: 814-823, 2010
3. Prospective gene signature study using microRNA to predict the tissue of origin (ToO) in patients with cancer of unknown primary site (CUP). Varadhachary, Edmonston et al. *ASCO-2010 J Clin Oncol* 28:15s, (suppl); abstr 4151, 2010
4. Accurate classification of metastatic brain tumors using a novel microRNA-based test. Muller, Spector et al. *The Oncologist* 16:165-74, 2011