#### ISSN: 1747-0862

# Enabling Circulating Cell-free mRNA Profiling to Empower Cancer Early Detection

### **Chen-Hsiung Y\***

Circulogene Theranostics, 3125 Independence Drive, Suite 301, Birmingham, AL, 35209, USA

### Abstract

DNA methylation occurs early in tumorigenesis and is highly pervasive across cancer types, making it suitable for early detection and intervention of cancer. Gene expression at the transcription level is reflected by epigenetic regulation of chromosome especially DNA methylation, therefore providing a parallel approach to achieve the same goal. Non-invasive and real-time gene expression profiling was performed on plasma circulating cell-free mRNA (cell free m-RNA) enriched from cancer patients using proprietary high sensitivity RT-qPCR assays. A plasma cell free m-RNA expression database covering 750 genes in 9 major cancer pathways was established with multiple layers of cancer type-specific characteristics: (i) the distribution of cell free m-RNA species across 9 major cancer pathways; (ii) the differential expression of target genes (high, medium or low expression); (iii) the global cell free m-RNA expression landscape in circulation; and (iv) the unique cell free m-RNA signatures to differentiate lung, breast, and pancreatic cancer. These novel blood-based metrics and biomarkers can be deployed for early detection and stratification of cancer.

Keywords: Gene expression • Pancreatic cancer • Biomarkers • Protein-coding mRNAs

# Introduction

The emergence of relatively non-invasive liquid biopsy as a complementary approach to surgical biopsies has fueled intensive research effort and investment. Circulating cell-free nucleic acids in this respect have revolutionized cancer diagnosis in recent decade, allowing non-invasive, real-time and longitudinal interrogation for genomic alterations using a single sample of blood [1]. Consequently, the use of cell-free nucleic acids as biomarkers could facilitate the early detection of diseases such as cancer and enable simple, specific monitoring of disease progression.

In addition to DNAs, protein-coding mRNAs from the tumor tissues are released into the blood, enriched over time and can reflect changes in tumorspecific gene expression. Plasma cfDNA methylation and mutation events are less dynamic and likely provide limited information on tissue homeostasis and disruption. In contrast, circulating cell-free mRNA (cell free m-RNA) profiling could provide richer molecular content compared to other non-invasive biomarkers and constitutes a unique non-invasive interrogation of tissue function in scenarios such as early detection of disease, early drug engagement and response in patients. Combined with advances in molecular diagnostics, systematic profiling of cell free m-RNA can improve our understanding of cancer pathology and identify novel biomarkers for early detection, without the need for invasive biopsy. The potential clinical utility of cell free m-RNA has been demonstrated in patients with various malignant cancers [2-4]. Understanding the mechanisms underlying the presence of mRNA transcripts in circulation is essential to interpret their clinical value.

# Methodology

Circulating cell free m-RNA abundance can be influenced by physiological state, level of nucleases in the blood, the half-life of individual cell free m-RNA

\*Address for Correspondence: Dr. Chen-Hsiung Yeh, Circulogene Theranostics, 3125 Independence Drive, Suite 301, Birmingham, AL, 35209, USA, E-mail: cyeh@circulogene.com

**Copyright:** © 2020 Chen-Hsiung Y. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Received 12 October 2020; Accepted 21 October 2020; Published 28 October 2020

and the clearance rate by immune system, liver and kidney [2]. Certain cell free m-RNA species are in complexed forms that protects them from degradation by RNases. This ensures their unique stability in the circulation, in contrast to complex-free RNA, which is rapidly degraded [5]. Therefore, key challenges in the cell free m-RNA testing include its extremely low abundance, susceptible to degradation, relatively unstable and poor extraction efficiency. To circumvent these limitations, a signal amplification step following cell free m-RNA extractions should be performed as we discovered in our studies. Circulating cell free m-RNA expression reflects dysregulation of cancer immunity, tumor cell growth, proliferation and stromal interaction, which makes cell free m-RNA expression signature a promising biomarker for early diagnostic, prognostic and therapeutic purposes [6,7].

Circulating cell free m-RNA in plasma is usually made up of degraded small fragments with size smaller than 200 nucleotides, very low concentration (average lower than 10 ng/mL), and with different terminal modification, these properties make it difficult to investigate [8-10]. The current molecular techniques employed for the detection and characterization of cell free m-RNA include microarrays, RT-qPCR and next-generation sequencing (NGS; RNA-Seg) [11-14]. Microarrays had been widely used to define circulating microRNA expression. However, due to their limited sensitivity microarrays can only screen the most abundant RNA in biofluids. On the contrary, both RT-qPCR and NGS can detect low abundant cell free m-RNA and remain currently the methods of choice. NGS has been used for cell free m-RNA studies, but some intrinsic problems were not solved, including labor-intensive, time-consuming, requirement of large volume of blood, no standard sequencing method for all RNA fractions, high cost for large scale RNA library preparation, as well as low mapping rate and thus low sensitivity [15-17]. RT-gPCR is a more convenient, sensitive and cost-effective approach, with pre-loaded custom plates, further enhancing its capability as an automate and high throughput platform. Developing a simple, highly reliable, cost-efficient and non-invasive diagnostic cell free m-RNA test system to screen and identify early stages without the use of a tissue biopsy would significantly reduce both the mortality and the economic burden associated with cancer.

Tissue biopsy and associated approaches which are highly dependent on skills of an operator and the availability of costly equipment could hardly fit into a model of point-of-care diagnostics. The absence of clear alternatives prompts the development and validation of functionalized gene signatures where each individual gene would, ideally, reflect certain pathophysiological process contributing to specific cancer progression in a given individual and predict its outcome. Although multiple research reports have demonstrated amazing promises of circulating cell free m-RNA for diagnostic application, this

field is still in its infancy. It is imperative and of paramount interest to harness highly sensitive cell free m-RNA detection technologies and establishing unique expression signatures as early-stage fingerprints of oncogenesis in biological fluids. The fundamental advantage of circulating cell free m-RNA over protein biomarkers is that, unlike proteins, nucleic acids can be detected by a PCR which has the detection threshold of a single molecule. A cell free m-RNA-based expression signature could, if necessary, be augmented by other blood-based biomarkers including cancer-specific cfDNA. The cell free m-RNA expression profiling can be considered as a compendium of transcripts collected from all organs. Some of these circulating transcripts correspond to "true" tissue-specific or cancer type-specific genes, strongly supporting interrogation of these biomolecules to dynamically monitor early pathological changes of tissues and organs. In contrast to poorly functional annotation of non-coding RNA, the coding cell free m-RNA expression profiling provides direct access to both genetic information and functional information pertaining to the tissue of origin and its physiology. Previous studies have reported transcripts in circulation encoding functional information of the liver, brain, immune system, or fetal development [7,11,12]. Therefore, cell free m-RNA expression pattern has the capability of integrating functional and genetic information of tissues, highlighting this analyte's unique potential as a noninvasive biomarker.

Our comprehensive cell free m-RNA profiling data here provides circulating transcript snapshots of gene expression signatures in patients with lung, pancreatic or breast cancer. The cell free m-RNA expression signature will allow non-invasive delineation of cancer type, early detection, and progression monitoring. Our data further provide promising proof of concept of using cell free m-RNA profiling to monitor early onset cancer activity, which could lead to improved therapeutic management of cancer patients, and eventually alleviate the need for invasive biopsies.

### **Results and Discussion**

While microarrays and RNA-Seq were previously benchmarked for

differential expression and prediction model development, the complexity of quantifying low-abundance cell free m-RNA is compounded by the presence of high and variable levels of globin mRNA and ribosomal RNA (rRNA). Although rRNA depletion and globin reduction have been shown to mitigate some of these issues, they require a large amount of total cfRNA pool and may induce biases in the quantification of gene expression [18]. To address these limitations, we applied targeted expression profiling methods based on multiplex RT-qPCR amplification followed by quantitative analysis of cell free m-RNA abundance by delta Ct, the difference of Ct values between reference gene (18S) and target gene.

Circulating cell free m-RNA was extracted from pooled plasma cohorts of patients with lung, pancreatic or breast cancer. Total 750 cancerassociated genes were profiled and categorized into 9 major cancer signaling pathways: immune response (IR), transcription factors (TF), DNA repair (DR), oncogenesis (ONC), tumor metastasis (TM), TP53 signaling (TS), MAP kinases (MK), cell surface markers (CSM) and DNA methylation (DM). The distribution of detected cell free m-RNA species in these 9 categories from the lung cancer cohort was demonstrated in Figure 1. We identified eighteen genes belonged to cell surface markers (21%), eleven genes involved in DNA repair (13%), fourteen genes are MAP kinases (17%), eighteen genes involved in TP53 signaling (21%), four genes are transcription factors (5%), five genes involved in immune response (6%), seven genes correlated with oncogenesis (8%), six genes associated with tumor metastasis (7%) and two gene related to DNA methylation (2%).

For quantification of cell free m-RNA expression levels, genes with delta Ct values between 0-15 was classified as "high expression" (blue); delta Ct values between 15-20 was interpreted as "medium expression" (green); and delta Ct values of 20-30 was called "low expression" (red). The genes with delta Ct values >30 were not color coded. Figure 2 illustrated a global cell free m-RNA expression and functional landscape we established in non-small cell lung cancer (NSCLC). Among the 9 key caner pathways, the circulating cell-free transcriptome composition of TP53 signaling, MAP kinases and cell surface markers were particularly dominant in NSCLC cohort.



Figure 1. Distribution of detected cell free m-RNA species in NSCLC.

IR	π	DR	Onc	TM	TS	MK	CSM
L 10.	RPLPO	AP EXH	ABL1	ά.PC	ABL1	∆.R:∆F	ALC:01
L 19	HMBS	μ.TM	0.KT1	BRING1	AKT1	0.TF 2	CSAR1
11.2	TEP	Δ.TR	0.KT2	CASPS	GKT2	BRAF	00160
La	PG K1	BARD1	APC	CCL7	ak ta	CCND1 CCNE1	00169
La La	UBC PP-IA	BRC41	BAX BCAR1	0042	ATM	CONE1	CD-19
LA	VR			COHI	BOH	CDKN1D	CD-14.
11.0	UR URNT	CCNO CHEK1	BCL2	COKNEA	B AN	COKNIC	0010
ILA	ATE 1	CHEK2	BCL3L11	CEAC40/1	80.2	CHUK	00:2
LS	ATE 2	DOLDENA	BID	CTRP1	BRCA4	COL141	00.209
11.10	4TF 9	ERCC1	BRAF	CTNN41	CASP1	CREBI	00.02
11.124	0.TF-4	EDCC0	CASPS	стак	CASP10	CREBAR	0024
11.129	CERPA	ERCCS	0.4 00 0	CXCL12	CASP 10	DUGDI	00.044
11-10	CERPS	ERCOL	CCND-1	CXCR4	CASPS	DUSPHO	00.047
LS	CERPO	ERCCS	CCND2	DUPK1	C-19.04	DUSPHE	00.16
L 17 A	08591	ERCCO	CONDS	DCC	CASPS	DUSP2	00.97
L 19	CREARP	ERCCO	CONT -	EPHB2	CASPS	DUSP4	00.98
CCL9	CTNNET	FANCA.	00042	ER992	CASP7	DUSPS	00.00
CCL19	DR1	FANCE	COHI	ETV8	CASPS	DUSPS	0000
2012	EXE 1	FANCD2	CDK2	F AT1	CARPS	DUSP7	004
COLS	E2F 0	FANCE	CDKK	F GF2	CCND1	DUSPS	20-60
CCR2	E9 R1	FANCE	CDKN14	FOFRI	CONST	DUSPS	02401.0
CCR4	ELK1	FANCO	CDKN1B	FNI	0042	FOS	006
CCRS	E9R1	FENI	CDKN24	FXYDS	CDHA	6 R 80	006
CCR7	ET SH	9.000450	COKN2B	ONRH1	CD4N14	HRAS	CD 65
282902	ET 22	9.000459	CO L141	HOF	CHIRCH	IGET	0066
0112902	FOS	9100459	CRK	HPST	CHEK2	KRAS	<b>CD</b> 7
111200	FO X82	G TF2H1	CTNNB1	HRAS	CSNK1D	MADD	CD70
SP1	F0 X01	G TF2HS	evea	HTA TP 2	DDR-1	M &P 2K1	CD72
C SF2	Giá Tát	HUSH	DVL1	IG F1	E2F1	MARKERI	CD74
CAFS	90.T02	IGE1	E2F1	L18	E 2F2	M dP 2K2	00786
e tate	GAT AG	LIGH	E GER	L1B	E 2F3	MAPSKS	00798
NFKB2	OTF 28	LIGS	ELK1	7.990	E 2F4	MAP2K4	0940
KRKR	ATF 2F1	LIGK	ERG62	Kiast	E 2F5	MAP 2K5	0069
2095	HAND1	MAPK10	FADD	KSSIR	E 2F0	M 4P 2K 6	0066
604	HAND2	MAPK11	FAG	KR49	E P900	M 4P 2K7	0064
0.000	HD4C1	MAPK12	FAGLO	LANGI	FAS	MAPSK1	0099
81:00	HIF1A	MAP K14	FGF2	LYPDS	FOS	MAPSK10	0066
L3R4	HNF44	MAPKE	FNI	MCAM	9400-64	MAPSK11	CHST10
0028	HOM: 5	MAPKS	FOS	MET	64.00-69	MAP9K2	00 L141
9803	HOFT	MBD4	FYN	NAS ATS	9100-59	MAPSKS	00 L142
0642	ID-1	M D-M2	F 201	MUP1	GEKSE	MAP 9K4	CR2
PTPRC	IRF1	MGNT	G R 92	MMIP10	HD4C1	MAPSKS	CSF1R
	JUN	MRE11A	a skse	MMP14	HD4C10	MAPSKe	CTLAS
0903	JUNE	MSH2	HOF	MAL P2	HD4C11	MAPSK7	DPP4
0000	JUND	MSHS	HRAS	MMPS	HD4C2	MAPSKA	ENG .
CTLA4	MAX	MSHE	JGF1	MM P7	HOUCS	MAPSKS	Fag
010102	MEF2A	NEN	IGF1R	MM P9	HD4C4	MOP4K1	FCERIA
HLA-DRA HLA-DRB1	MEF2B MEF2C	NTHL1 D GG1	ITGA 29	MT 44	HDUCS	MAP4K2	FC:GR14
HLA-DRS1	MEF2C MYB	0.001	ITGE 1	NT 22 NT 291	HDJC9 HDJC7	MAPK10	PLAKTA
TEX21 THE ROF 18		PCNA	ITSE S	MY C	HDUC7 HDUC9	MAPK10 MAPK11	HLA-DRA
	MYES	PNKP	JUN	NC4011	HDUCS	MAPK12	C10/2
1092	MYOD1	POLA1	KDR	NF2	HIETA	MAPK19	L128.91
	NFATS	POLA	KIT	NNE1	HPK2	MAPK14	L12Kar
CL2L1	NEATC1	POLD1	KRAS	NR449	KAT29	MAPK15	LIRA
B dox	NFATC2	POLG	LEFI	PECANH	MA.PK10	MAPKS	TGA 1
CANH	NFATCS	POLH	MAP 2K1	PNN	MA PK11	MAPKA	TG0.2
ALP.	NEATER	POLK	MAPSKS	Paca	MA PK12	MARKA	TGA 9
10.5	NEKBI	POLO	MAPK1	PTEN	MA PK19	MAPK7	KLR91
	NEYB	POLR18	MAPKIA	PTGG2	MA PK14	MAPKA	KLRC1
PTGG2	NRSCI	POLRIC	MAPKS	RB1	MA PKG	MAPKEP1	KLRD1
RP2	Pú X9	POLRSA	МАРКА	ROLI	MA PKS	MAPKEPS	KRT18
CYP10.2	PO USAF 1	POLRIB	MILLOC	RBL2	NDM2	марка	KR TS
	PP-1 Rd.	POLR2C	M D4/2	RET	Plast	MAPKAPKS	
FNIG	PP & RG	PRKDC	MYC	RHOC	PRKDC	MAPKAPKS	Makat
PRF1	RBI	Panaa	NEKAI	910034	P 90/42	MAPKAPKS	MYHND
9.21/19	REL	P90810	NFKB2	SERPINES	P SM 45	M 67 24	MYHE
S NEY	RELA	Panas	NEKSIA	SERPINES	P SM 46	MEF 2C	NYOCD
49	RELB	Panas	NRAS	9 ET	P-91/82	MINNET	NCANH
	SM 3D-1	PANDO	P1KSCA	20.000	PSMBA	MRNK2	NO 99
19 751	SM 4D-4	Pance	PIKSR1	S MA DE	PSMBS	MYC	NT SE
	SUIVE	RAD1	PTEN	anc.a	P3V02	NLK	PECANI
9 MA D7	90.009	R4.047	PTK2	SSTR2	PSIICS	NRAS	RETN
		RA.0228	PTK2B	STRC .	PSIIC4	PAR	91004.6
	92P-1			EPC-AM	PSNCS	P336	SELP
a kd	99-1 99-3	RADEO	RáC1		_		
aki Finit Ca	99-9 97.471	RADSO RADS1	R.4.F1	TC:F 20	-	PAKS	STEG AL1
940 F NH C 9	999 97 474 97 472	RADSO RADS1 RADS2	R451	TCF 20 TQ FB 1	P SM De P TEN	Р-1. КВ Р-1. КВ	STEG AL1 EP-C-AM
aki Fini Eg Tinif	99-9 97.471	RADSO RADS1	R.4.F1	TC:F 20	PSMD4	PAKS PTK38	
AKI ENH ESI ENIF LTD.	999 97 474 97 472	Rá D50 Rá D51 Rá D52 Rá D93 RP 40	RAF1 RS1 RSLA RHCA	TOF 20 TO FR 1 TO FR 2 TOW	PSAIDA PTEN RS1 SFN	PAKE PTK38 RAC1	EPICIAN TEK TNERGES
akt FNH C:a TNIF LTA ACSE	999 97 471 97 472 97 479	RADSO RADS1 RADS2	R451	TOF 20 TO FO 1 TO FOR 2	PSNDA PTEN RS1	PAKS PTK38	EP-C-AM TEK
AKI ENH ESA TRAF 	99-9 97 474 97 472 97 479 97 474	Rá D50 Rá D51 Rá D52 Rá D93 RP 42	RAF1 RS1 RSLA RHCA	TOF 20 TO FR 1 TO FR 2 TOW	PSAIDA PTEN RS1 SFN	PAKE PTK38 RAC1	EPICIAN TEK TNERGES
2 KI F NH C 9 TN F L T 0. A C 5 V T 9 F 0. C D 94 A 9 TR 1	99-9 91 474 91 472 91 479 91 474 91 4754	RADSO RADS1 RADS2 RADS0 RPAS SMUS1 TRS	RAF1 RB1 RELA RHOA BHC1 BMAD4 BOB1	TOF 20 TO FO 1 TO FOR 2 TUNP 1 TIMP 2 TIMP 4	PSHD4 PTEN RSH SFN SFN SIRT1 SIRT2 SIRT3	PAK PTK28 RAC1 RAF1 RELA RP26KA1	EPICAM TEX TNFROFA TNFROFA TNFROFA TPOAR VCAM1
240 5 NH C29 TNF LTA ACE VEG FA CD94 A.GTR:1 A.GTR:2	9229 937 474 937 479 937 479 937 479 937 4754 937 4754	Rá DS0 Rá DS1 Rá DS2 Rá DS2 RP AS RP AS	RAF1 RB1 RELA RHOA BHC1 BMAD4	TOF 20 TO FO 1 TO FOR 2 TO 0 TIMP 1 TIMP 2	PSHD4 PTEN RSH SFN SFN SIRT1 SIRT2 SIRT3	PARE PTKSB RAC1 RAF1 RELA	EPICIAN TEX TNFRIGFA TNFRIGFA TPGILB1
240 5 NH C29 TNF LTA ACE VEG FA CD94 A.GTR:1 A.GTR:2	99-9 97 574 97 575 97 575 97 575 97 5755 97 575 97 575 97 576	RADSO RADS1 RADS2 RADS0 RPAS SMUS1 TRS	RAF1 RS1 RS1A RHOA SHC1 SMLD4 SOS1 SP91 SRC	TOF 20 TO FO 1 TO FOR 2 TUNP 1 TIMP 2 TIMP 4	PSHD4 PTEN RSH SFN SFN SIRT1 SIRT2 SIRT3	PAK PTK28 RAC1 RAF1 RELA RP26KA1	EPICAM TEX TNFROFA TNFROFA TNFROFA TPOAR VCAM1
2 KU F NH C 2 T NF L T A A CE V EQ F A C D 3 K A CE V EQ F A A GTR 1 A GTR 2 E D NH L JF	9299 937 674 937 679 937 679 937 679 937 6756 937 6756 937 6759 937 679 739 937 679	R.4.050 R.4.051 R.4.052 R.4.094 R.P.40 R.P.40 S.MU-6.1 TR65 T.R.2(1	RAF1 RS1 RS1A RHOA SHC1 SMAD4 SOS1 SPP1	TCF 20 TG FB-1 TG FBR2 TUMP TIMP 1 TIMP 2 TIMP 2 TIMP 2 34	PSHIDA PTEN RSH SFN SRT1 SRT2 SRT3 SRT3 SRT4	PARE PTK39 RAC1 RAF1 RELA RP36KA1 RP36KA2	EPICAM TEX TNFROFA TNFROFA TNFROFA TPOAR VCAM1
240 FNH CS TNF LTA ACE VIES FA CDS4 AGTR1 AGTR2 EDN1	97 9 97 471 97 479 97 479 97 479 97 479 97 4754 97 4755 97 475 97	R4050 R4051 R4052 R4094 R940 R940 R940 SMU41 TR55 TR51 TR51 TR52	RAF1 RS1 RS1A RHOA SHC1 SMLD4 SOS1 SP91 SRC	TCF 20 TG FB4 TG FB42 TUMP1 TUMP1 TUMP2 TUMP4 TUPR224 TUPR224 TUF2F10	PSHIDE PTDN RSH SEN SET SET SET SET SET SET SET SET SET SET	PARE PTK39 RAC1 RAF1 RELA RP36K21 RP36K22 RP36K23	EPICAUL TEX TNFR:9F4 TNFR:9F9 TP9:0.94 VC:0011
240 FNR 29 TNF LTA ACE VEQ FA 2024 AGTR1 AGTR2 EDN1 LIF	929 97 471 97 473 97 473 97 473 97 475 97 475 97 97 97 97 97 97 97 97 97 97 97 97 97	RADS0 RADS1 RADS2 RADS0 RP40 RP40 SMUS 1 TR54 TR54 TR54 TR54 X452	RAF1 RS1A RHOA SHC1 SMAD4 SCS1 SPP1 SRC TCFS TGFS1 TGFSR1	TCF 20 TG FB1 TG FB2 TUMP1 TIMP1 TIMP2 TIMP4 TMF2510 TP53 TP53 T94R	PSNID: PTDN RSH SFN BRT1 BRT2 BRT3 BRT4 SRT4 SRT4 SRT4 TNFRSF10	PARE PTK19 RAC1 RAF1 RF14 RF24KA1 RF24KA1 RF24KA1 RF24KA1 RF24KA1 RF24KA1 RF24KA1	EPICAUL TEX TNFR:9F4 TNFR:9F9 TP9:0.94 VC:0011
240 FNR CS TNF LTA ACS VEG FA ACS LEG ACTR2 EDNH LIF LIF LV96 MIF (FATC2	925 97 471 97 475 97 475 97 475 97 475 97 475 97 475 97 475 97 475 10 10 10 10 10 10 10 10 10 10 10 10 10	RADS0 RADS1 RADS2 RADS2 RPAS SMUS 1 TRS3 TRS1 TRS1 TRS2 KAS2 KAS2 KPA	RAF1 RS1A RS1A RHCA SHC1 SMAD4 SCS1 SPP1 SRC TCFS TCFS TGFS1	TCF 20 TG FB4 TG FB42 TLMP4 TLMP4 TLMP4 TLMP4 TLMP4 TLMP4 TMF52 TD TP53 TP96	PSNID2 PTDN RSH SFN SRT1 SRT2 SRT2 SRT2 SRT2 SRT5 SRT5 SRT5	PARE PTK19 RAC1 RAF1 RELA RP3640.1 RP3640.2 RP3640.5	EPICAUL TEX TNFR:9F4 TNFR:9F9 TP9:0.94 VC:0011
4M FNR 29 TNF 20 50 50 50 50 50 50 50 50 50 50 50 50 50	929 97 471 97 473 97 473 97 473 97 475 97 475 97 97 97 97 97 97 97 97 97 97 97 97 97	RADSO RADS1 RADS2 RADS3 RP33 RP33 SMUS1 TRS3 TRS41 TRS42 KAS2 KP4 KP4	RAF1 RS1A RHOA SHC1 SMAD4 SCS1 SPP1 SRC TCFS TGFS1 TGFSR1	TCF 20 TG FE1 TG FER2 TUMP1 TIMP2 TIMP4 TMF24 TMF2510 TP20 TP240 T24R	PSNID: PTDN RSH SFN BRT1 BRT2 BRT3 BRT4 SRT4 SRT4 SRT4 TNFRSF10	PARE PTK19 RAC1 RAF1 RF14 RF24KA1 RF24KA1 RF24KA1 RF24KA1 RF24KA1 RF24KA1 RF24KA1	EPICAUL TEX TNFR:9F4 TNFR:9F9 TP9:0.94 VC:0011



High expression Medium expression Low expression

Figure 2. Fingerprint of cell free m-RNA functional clusters detected in NSCLC.



Figure 3. Cancer type-specific cell free m-RNA expression profiles can be used to identify novel biomarkers for early detection and stratification of cancer.



Figure 4. Blood-based cell free m-RNA fingerprint - a non-invasive, real-time and longitudinal liquid biopsy for early-stage cancer detection.

From the representative cell free m-RNA expression heatmaps shown in Figure 3, differentially expressed cancer type-specific genes can be easily identified in the same cluster, for example, ERCC2, MDM2, POLR2B, PSMB10 in DNA repair cluster are highly expressed and are pancreatic cancer-specific genes; while FANCG is breast cancer-specific gene; and POLH, RPA2 are strongly expressed as lung cancer-specific genes. Among cell surface markers, C5AR1, CD24, CD28, CD40LG, CD96, KRT18, SELP are highly expressed as pancreatic cancer-specific genes; whereas CD7, CD8A, FAS are breast cancer-specific genes; and CD79A and MS4A1 are strongly expressed as lung cancer-specific genes. Together, we have obtained highly distinct cell free m-RNA expression profiles and functional clusters specific for NSCLC, breast cancer, and pancreatic cancer. Pancreatic cancer revealed the largest heterogeneity of gene expression as a wide spectrum of cell free m-RNA molecules are produced by its transcriptional machinery. In contrast, NSCLC has relatively low cell free m-RNA heterogeneity and fewer specific genes that contribute to the total cell free m-RNA composition.

Multiple pathways are responsible for transducing mechanical and growth stimuli into changes in gene expression during cancer development. In this project we have established an unprecedented functional cell free m-RNA database which will lead to (i) classification of cell free m-RNA species by their functions, (ii) identification of differentially expressed cell free m-RNA in a particular cancer type, (iii) illustration of a comprehensive landscape of cell free m-RNA in circulation, and (iv) establishment of specific cell free m-RNA expression signatures for different cancer types. The cell free m-RNA expression profiles identified in this study represent gene fingerprints in circulation for specific cancer types, thus offering the exciting possibility of detecting early stage cancer (Figure 4).

# Conclusion

We have comprehensively explored cell free m-RNA expression profiles and signatures of different cancer types, thereby establishing a plasma-based functional transcriptomic databank, including differential gene expression, classification, functional clustering and cancer type-specific signatures. We believe that our work has research, clinical, and diagnostic value, and provides greater dimensionality to the current landscape of cell free m-RNA research and makes a relevant jump into understanding and devising strategies to tackle early cancer detection.

# **Future Perspectives**

Prior studies have laid a solid foundation for tissue-based gene expression signatures in cancer diagnosis and prognosis. However, gene expression signatures for early cancer detection were hampered by tissue unavailability and limited bioinformatics tools, and in many ways spurred renewed interest in developing non-invasive blood-based high-dimensional analytical algorithms. With our cell free m-RNA gene expression database, it is now possible to derive highly accurate and sensitive cancer type-specific signatures that are amenable to large-scale clinical validation. Our profiling model comprised of 750 cancer-associated genes has the potential to be translated into widely applicable assays. While the clinical validity of cell free m-RNA signatures will ultimately be proven by cross-validation in prospective studies, the accurate extraction of information from genomic and/or epigenomic studies is of vital importance for guiding such studies.

The establishment of cancer type-specific cell free m-RNA expression signatures is necessarily an ongoing and dynamic process, in which, with the inclusion of more early-stage patient samples with consistent clinical information, an early detection signature will be continuously refined. Due to biological and technical limitations, cell free m-RNA-based expression signatures may not be able to achieve 100% accuracy, yet the application of advanced feature selection algorithms and the combination of genetic and clinical data will enable their robust performance with dramatically reduced complexity.

# References

- Schwarzenbach, Heidi, Dave SB Hoon and Klaus Pantel. "Cell-Free Nucleic Acids as Biomarkers in Cancer Patients". Nat Rev Cancer 11(2011): 426-437.
- Fleischhacker, Michael and Bernd Schmidt. "Circulating nucleic acids (CNAs) and cancer–A survey". *Biochim Biophys Acta* 1775(2007): 181-232.
- Bettegowda, Chetan, Mark Sausen, Rebecca J Leary and Isaac Kinde, et al. "Detection of Circulating Tumor DNA in Early- and Late-Stage Human Malignancies". Sci Transl Med 6(2014): 224ra24.
- Dominguez-Vigil, Irma G, Ana K Moreno-Martinez, Julia Y Wang and Michael H Roehrl, et al. "The Dawn of the Liquid Biopsy in the Fight against Cancer". Oncotarget 9(2018): 2912-2922.

- Yuan, Tiezheng, Xiaoyi Huang, Mark Woodcock and Meijun Du, et al. "Plasma Extracellular RNA Profiles in Healthy and Cancer Patients". Sci Rep 6(2016): 19413.
- Rapisuwon, Suthee, Eveline E Vietsch and Anton Wellstein. "Circulating Biomarkers to Monitor Cancer Progression and Treatment". *Comput Struct Biotechnol J* 14(2016): 211-222.
- Tsang, Jason C. H., Joaquim Si-Long Vong, Liona C Y Poon and Jiang Peiyong, et al. "Integrative Single-Cell and Cell-Free Plasma RNA Transcriptomics Elucidates Placental Cellular Dynamics". Proc Natl Acad Sci USA 114(2017): E7786-E7795.
- Spornraft, Melanie, Benedikt Kirchner, Bettina Haase and Vladimir Benes, et al. "Optimization of Extraction of Circulating RNAs from Plasma-Enabling Small RNA Sequencing". *PLoS One* 9(2014): e107259.
- Wei, Zhiyun, Arsen O Batagov, Sergio Schinelli and Jintu Wang, et al. "Coding and Noncoding Landscape of Extracellular RNA Released by Human Glioma Stem Cells". Nat Commun 8(2017): 1145.
- Tsui, Nancy B, Peiyong Jiang, Yuen Fei Wong and Tak Y Leung, et al. "Maternal Plasma RNA Sequencing for Genome-Wide Transcriptomic Profiling and Identification of Pregnancy-Associated Transcripts". *Clin Chem* 60(2014): 954-962.
- Pan, Wenying, Thuy T M Ngo, Joan Camunas-Soler and Chun-Xiao Song, et al. "Simultaneously Monitoring Immune Response and Microbial Infections during Pregnancy through Plasma cfRNA Sequencing". *Clin Chem* 63(2017): 1695-1704.
- Enache, Liviu S., Elena L. Enache, Christophe Ramière and Olivier Diaz, et al. "Circulating RNA Molecules as Biomarkers in Liver Disease". Int J Mol Sci 15(2014): 17644-17666.
- Cortez, Maria A., Carlos Bueso-Ramos, Jana Ferdin and Gabriel Lopez-Berestein, et al. "MicroRNAs in Body Fluids-The Mix of Hormones and Biomarkers". Nat Rev Clin Oncol 8(2011): 467-477.
- Etheridge, Alton, Inyoul Lee, Leroy Hood, David Galas, et al. "Extracellular microRNA: A New Source of Biomarkers". *Mutat Res* 717(2011): 85-90.
- 15. Max, Klaas E. A., Karl Bertram, Kemal Marc Akat and Kimberly A Bogardus, et al. "Human Plasma and Serum Extracellular Small RNA Reference Profiles and their Clinical Utility". *Proc Natl Acad Sci USA* 115(2018): E5334-E5343.
- Yeri, Ashish, Amanda Courtright, Rebecca Reiman and Elizabeth Carlson, et al. "Total Extracellular Small RNA Profiles from Plasma, Saliva, and Urine of Healthy Subjects". Sci Rep 7(2017): 44061.
- 17. Veytsman, Boris and Ancha Baranova. "High-Throughput Approaches to Biomarker Discovery and the Challenges of Subsequent Validation". *Gen Methods Biomark Res Their Appl* (2015): 3-6.
- Liu, Jinny, Elizabeth Walter, David Stenger and Dzung Thach. "Effects of Globin mRNA Reduction Methods on Gene Expression Profiles from Whole Blood". J Mol Diagn 8(2006): 551-558.

How to cite this article: Yeh, Chen-Hsiung. "Enabling Circulating Cell-free mRNA Profiling to Empower Cancer Early Detection." *J Mol Genet Med* 14(2020): S3 doi: 10.37421/jmgm.2020.14.S3