

## 체액(Body Fluid)을 이용한 질병 진단학적 바이오마커의 발굴

인체의 대부분 (약 60%)은 체액으로 이루어져 있으며, 그 중 혈장, 혈청, 점액, 침, 소변, 담즙 등 외세포액 (extracellular fluid)이라 불리는 것은 총 체액 중 3분의 1을 차지하고 있다. 이러한 형태의 체액들은 체내를 이동하면서 조직세포에 영양분과 산소를 운반하고 노폐물 제거, 병원체의 박멸 및 체온조절 등의 다양한 기능을 수행하며 체내 환경과 건강 상태를 유지한다. 체액은 오믹스 기술의 발전을 통해 임상연구, 진단분야에서 매우 매력적인 대상으로 간주되어 왔다. 지속적으로 몸에서 생성되는 체액은 샘플링이 용이하여 다중 샘플링이 가능할 뿐만 아니라, 생검과 같은 침습적인 절차가 필요하지 않은 이점을 가지고 있기 때문이다. 또한 체액에 존재하는 단백질, DNA, RNA 등은 특정 질환 여부나 상태를 나타내는 지표물질인 바이오 마커로 사용될 수 있다. 정상인에게 없는 성분이 특정 질병에 걸린 환자 혈액이나 체액에서 발견되면 이를 토대로 질병 유무와 상태를 판단할 수 있다. 생물학적 체액은 이러한 특성을 통해 예후 및 진단 테스트 키트 개발 대상이 될 수 있는 잠재력을 가진다. 현재까지 주로 혈액이 바이오마커 식별에 가장 널리 사용되는 체액이며, 침, 소변, 뇌척수액 (CSF) 이 주로 사용되고 있다.[1,13] (그림1)

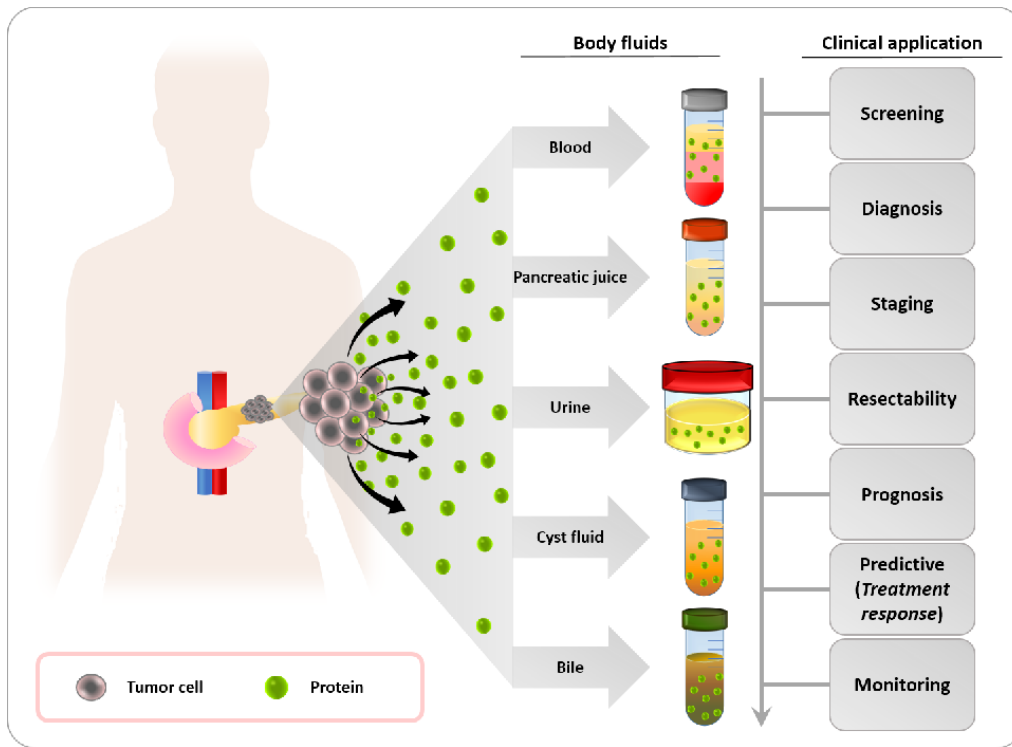


그림.1 Body fluids for the identification of potential protein biomarkers for pancreatic cancer. [2]

2022년 집계된 글로벌 체액 수집 및 진단 시장 규모는 301.9억 미국 달러로 평가되었고, 또한 2023년부터 2030년까지 연평균 성장률(CAGR) 6.5%로 시장의 지속적인 확장이 예상되고 있다. [12] 최근의 시장 조사에 따르면, 액체 검사 (liquid biopsy) 업계 규모는 2026년까지 58억 달러를 넘을 것으로 예상되고 있다. 이는 글로벌 시장에서 질병과 연관된 여러 가지 생체정보가 혈액이나 타액을 포함하는 일련의 체액에 존재하는 것으로 밝혀지면서, 비교적 채취가 간편한 체액을 이용하는 체외진단 기술이 급속도로 발전하고 있음을 시사한다. [8]

## Recent Trends in Biomarker Discovery Analysis using Body Fluids

체액은 게놈 (Genome), 전사체 (Transcriptome) 그리고 단백질체 (Proteomics)의 주요 세가지 영역에서 활용되고 있다. [3] 최근 게놈 분야에서는 혈장, 혈청 및 기타 다양한 체액에 존재하는 200bp 길이의 이중 가닥 DNA 조각인 Cell-free DNA가 주요 연구 대상이며, 최근 암 진단 분야에서 후성유전학적 상태를 반영한 메틸화 상태, 돌연변이 상태, 전좌 (translocation) 가 활발하게 연구되고 있다. 전사체 분야에서는 체액내 분해 안정성이 높은 작은 크기의 RNA를 포함하는 순환 세포 외 RNA (Circulating cell free RNA: cfRNA)가 분자 도구로 활용되고 있다. cell free RNA (cfRNA)에는 messenger RNA (mRNA), long-noncoding RNA (lncRNA), circular RNA (circRNA), miRNA, Piwi-interacting RNA (piRNA), transfer RNA (tRNA), 그리고 기타 여러가지의 non-coding RNA 분자가 포함되어 있으며, 특히 혈장 내 cfRNA는 불규칙한 단편화 상태로 대부분 poly-A tail이 없다는 특징이 있다. 이러한 cfRNA는 대사 질환, 암, 비만 연구 등 다양한 연구분야에서 활용된다.[4] (그림2)

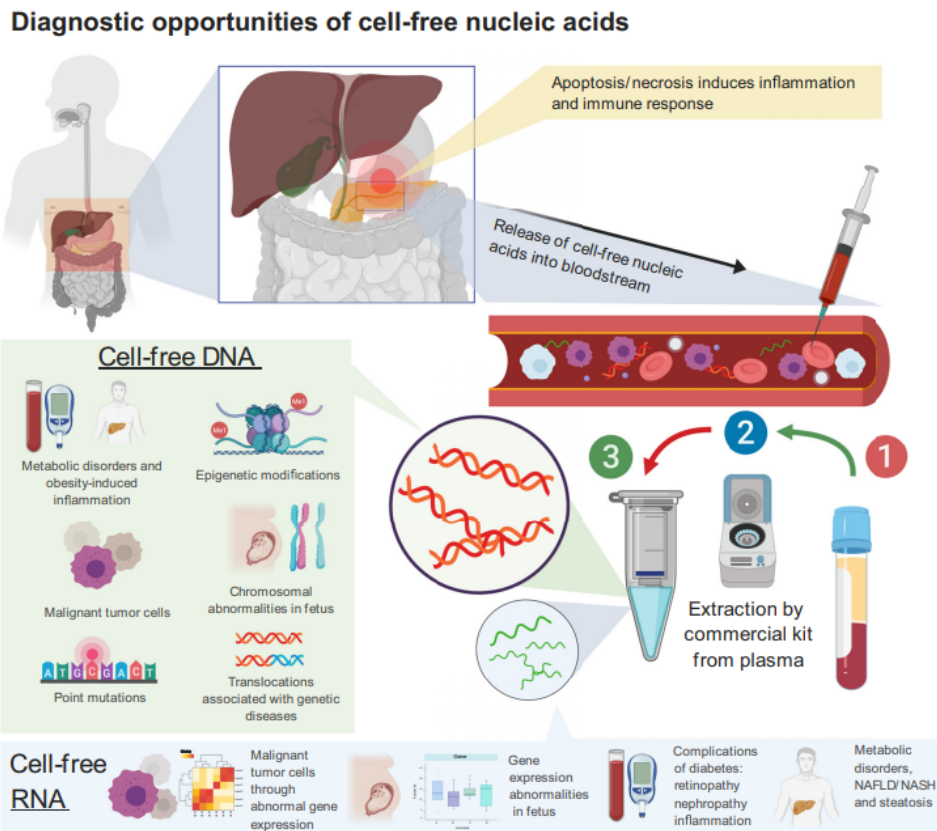


그림2. Diagnostic opportunities of cell-free DNA and RNA within prenatal screening and disease diagnostics. [4]

또한 인체 체액을 대상으로 한 단백질체 분석(HBFP, human body fluid proteome)은 숙주의 질병 발병, 감염 메커니즘 및 병리학적 증상과 관련된 대규모 단백질 정보를 제공하는 바이오 마커 발굴을 위해 전염병 및 신경학적 질병 연구에서 활용한다. [5]

## Proteomics-Driven Body Fluid-Based Biomarkers

현재 오믹스 기반기술인, Tandem mass spectrometry (LFQ-LC-MS/MS)는 고해상도 질량 분석기와 결합된 나노 고정밀 액체 크로마토그래피(HPLC) 시스템을 활용하며, 높은 처리량으로 대규모 단백질 특성화를 가능하게 하여 프로테오믹스 기반 체액 바이오 마커 연구 분야에 혁신을 가져왔다.[6]

이러한 오믹스 기반기술은 다른 접근 방식보다 체계적이고 편향적이지 않은(unbiased) 발견경로를 제공하여 인기

를 얻고 있으며, 샷건 (Shotgun) 접근법을 이용하여 발견단계에서 대규모 집단 내의 단백질 바이오 마커를 테스트하고 특성화 할 수 있다. 해당 접근법은 라벨링이 필요하지 않고 여러 환자의 샘플을 프로파일링 할 수 있으며, 특정 임상 샘플 그룹에 존재하거나 부재한 단백질 뿐만 아니라, 질병군 대비 정상군에서 차별적으로 발현된 단백질 (DEPs)을 식별할 수 있도록 도움을 준다.[1,7] 예를 들어, 다양한 체액에서 얻은 자궁경부암의 단백질체 분석은 혈액 (혈장 및 혈청), 소변, 자궁 점액, 자궁 질액, 생리 플루이드 등 다양한 체액에서 진행된다. Shotgun proteomics은 이러한 체액에서 단백질을 추출한 후 digestion, LC 분리 및 질량 분석이 수행된다. 이후 질량분석 raw data 가 통계적으로 분석되고, 특정 단백질의 유효성 검증 (MRM/PRM) 및 생체 표지자 발견이 이루어진다.[1] 선택된 후보 단백질들은 검증 단계에서 목표 지향적 Proteomics 접근법을 사용하여 검증될 수 있다. 이렇게 분석된 데이터셋은 문헌에서 얻은 정보로부터 누적 질병 발전 경로분석과 함께 컴퓨터 기반 분석에 활용된다. (그림 3)[1]

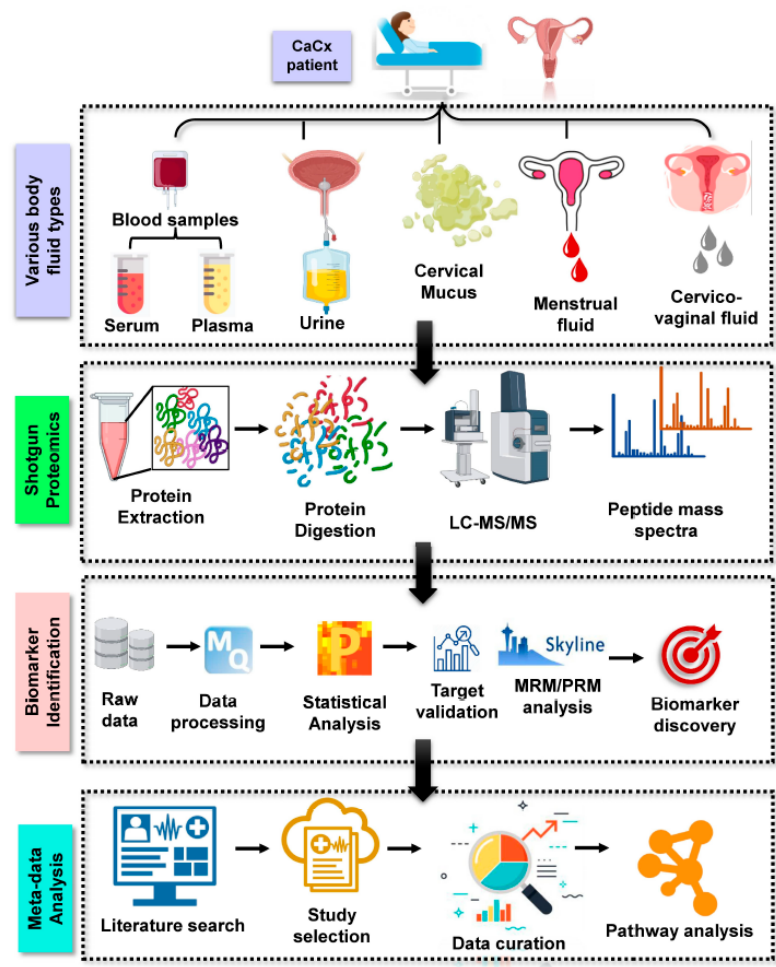


그림 3. Depicting the general method for proteomic analysis of cervical cancer derived from various body fluids. [1]

실질적으로 탐색적 소수의 검증된 바이오 마커들은 해당 질환과의 연관성에 대한 더 많은 증거가 수집됨에 따라 궁극적으로 검증 및 임상적으로 유용한 바이오 마커 범주로 이동할 수 있다. Lab scale에서 수십 명을 대상으로 한 결과를 일반화할 수는 없기 때문에, 특정 단백질이나 유전자가 실제로 질병에 영향을 미치는지 확인하기 위해서는 병원이 갖고 있는 환자 빅데이터를 이용해 수백~수천 명을 대상으로 한 임상이 필수적으로 수행되어야 한다. 진단 기준 확립까지는 많은 절차가 필요하며 대규모 시간과 노력이 투입되는 임상연구가 진행되어야 한다.

임상학적 체액 바이오마커는 주로 네 가지 유형으로 분류되어진다. 진단 바이오마커(Diagnostic biomarkers)는 질병의 초기 감지를 가능하게 하며, 예후 바이오마커(Prognostic biomarkers)는 질병의 예상 진행에 관한 정보를

제공한다. 치료 바이오마커(Therapeutic biomarkers)는 약물 또는 소분자 치료의 대상으로 활용될 수 있는 단백질을 말한다. 현재 환자의 안전성을 침해하지 않으면서 약물 효과를 예측하는 게 갈수록 중요해지고 있기 때문에 예측 바이오마커(Predictive biomarkers) 또한 대상 치료에 대한 환자의 반응을 예측하므로 치료 효과 가능성이 있는 환자의 하위 집단을 확립하는데 사용될 수 있다.

Proteomics 분석 시 유의해야할 점이 있다. 체액에는 농도 범위가 굉장히 넓은 다양한 단백질이 포함되어 있으므로, 종합적인 단백질체 분석을 위해 체액 내 고품량 단백질의 제거가 필수적으로 수행된다. 또한 단백질의 구조적 modification(당화, 절단)은 인체 체액 내 단백질에서 발생하는 일반적인 현상이며, 이러한 PTM은 질병 상태와 관련될 수 있으므로 적절한 Proteomics 플랫폼을 사용하여 특성화 및 정량화 할 필요성이 있다. 단백질의 양, 구조 및 기능에 대한 특성을 바이오 마커로 활용하여 초기 진단 또는 예측을 개선하는 데 매우 유용하게 사용된다.[5]

전기영동, 크로마토그래피, 단백질 어레이 등을 포함한 현대 Proteomics 방법은 젤 기반 기술에서 Shotgun proteomics로 알려진 젤 프리 질량 분석(Gel-free MS) 접근법으로 진화했다. 이러한 기술은 다양한 형태의 체액을 이용하여 꾸준히 바이오 마커 개발에 활용되고 있다.(그림4)[1]

Body Fluid Type	Methodology and Protocol	Cohort	Key Findings	Extra Comments (Merits/Demerits)	References
Plasma	2D-DIGE (silver staining); MS/MS (MALDI-TOF) to identify DEPs, and further validation by ELISA and statistical analysis by ANOVA.	Healthy: 40 individuals; CSCC and CIN patients: 80 individuals.	Cytokeratin 19 is upregulated in both the CIN 3 and CSCC IV conditions and tetranectin downregulated in CSCC.	Identification of DEPs along different stages of cervical cancer progression helps in understanding and prognosis of cancer.	[38]
Serum	Weak cation method, exchange chromatography fractionation in conjunction with MALDI-TOF spectrometry; liquid chromatography-electrospray ionization tandem mass spectrometry, and enzyme-linked immunosorbent assay (ELISA).	Healthy: 50 individuals; patients before surgery: 39; patients after surgery: 28.	The three peaks (m/z: 2435.63, 2575.3, and 2761.79 Da) may serve as predictive serum biomarkers for cervical cancer (CC).	Each patient group has obvious variation as the combined effect of age, stage, and tumor type reduces the power of marker detection.	[47]
Mucous	SELDI-TOF (surface-enhanced laser desorption and ionization-time of flight mass spectrometry).	Samples were collected from women attending urban hospital colposcopy clinics who were enrolled as a part of the study of cervical neoplasia. Samples were collected at the time of colposcopy by absorption into two Weck-Cel® sponges from 2-6 women matched for ages and races.	Annexin, tropomyosin, 14-3-3 sigma, calreticulin, and anterior gradient protein were identified.	The short sample size and inaccuracy of sample collecting techniques lowered the number of proteins discovered	[28]
Menstrual fluid	Genomic DNA was extracted from the menstrual blood collected on a napkin using a QIAamp DNA Mini Kit. Two rounds of PCR reaction using My11 and My09 primers for HPV detection. Fischer's exact test to examine the association between the distribution of genotypes or alleles for the TAP polymorphisms.	Control: 137 individuals; CIN3, CIN1, CIN2: 265 patients.	TAP1 1333V and TAP1 D637G were detected in the menstrual blood samples. The genotypes AA, AG, and GG were detected at each polymorphic site in the patients and the risk of developing high-grade cervical neoplasia was reduced for the AG and GG phenotypes as compared to the AA genotype. The risk of developing high-grade CIN was reduced in the patients that had a G allele than in those with an A allele.	The findings in the study have high specificity, sensitivity, and positive predictive value for the HPV virus and have received positive responses from over 5000 women.	[52]
Cervicovaginal fluid	Label-free quantification method based on LC-MS/MS method followed by ELISA. The PLS-DA model for further statistical analysis.	Development set—healthy: 10 individuals; LSIL: 10 individuals; HSIL: 10 individuals; cancer: 10 individuals Validation set—healthy: 14; LSIL: 8; HSIL: 6; cancer: 5.	ACTN4, VTN, ANXA1, ANXA2, CAP1, MUC5B and PKM2 from the 27 differentially expressed proteins have been indicated as promising biomarkers for cervical cancer.	The comparatively high number of samples gives better and more accurate results and reduces the chances of false biomarker discovery. The samples were also better classified into further four subgroups providing a comparison basis amongst the four groups.	[53]
Urine	Label-free quantification-UPLC-MS/MS analysis of pooled samples protein-protein interaction (STRING), pathway enrichment analysis, and molecular functions from KEGG and GO. Sensitivity as potential biomarkers tested by Western blotting and statistical analysis like logistic regression, ROC and AUC.	Healthy: 13 individuals; cervical cancer: 24 individuals.	Five Proteins with molecular weight >100 kDa were identified as potential biomarkers—LRG1, MMRN1 (upregulated), S100A1, CT344, SERPIN 33 (downregulated).	Rather than conventional gel-based MS analysis, non-gel based LPQ-MS analysis could aid in finding the low molecular weight potential biomarkers present in trace amounts in urine.	[55]

그림 4. Biomarkers identified from different body fluids of cervical cancer patients. [1]

## Cell free RNA Biomarkers Discovery from Body Fluids

혈청 및 다른 체액에서 발견되는 순환 cell free RNA에는 다양한 종류가 있다. 이 RNA는 빠르고 민감하며 안정성을 가지고 있으며 비교적 저비용 효율적인 진단의 새로운 형태로 큰 관심을 받고 있다. 특히 조직 특이적 유전

자 발현을 반영하는 것과 더불어, cfRNA는 Pathogenesis 과정의 중요한 스냅샷을 제공하며, 초기 질병 단계에서 질병을 검출하고 예후를 예측하고 모니터링하는 데 이상적인 분자이다.[4]

순환 cell free RNA는 생물학적 역할과 RNA 종류에 따라 그 기원이 다르다. 순환 mRNA는 일반적으로 세포 괴사 (necrosis)와 세포 자멸 (apoptosis)에서 기원하며, 분해로부터 보호하는 엑소좀과 apoptotic body와 관련이 있다. 그러나 특히 17~23bp 크기의 noncoding RNA 인 miRNA는 주로 장애 및 질병과 관련된 활성 분비에서 기원한다고 알려져 있으며, 미세 소포 (microvesicle)에 실려 전달된다. [8] 앞서 말한대로 miRNA 및 small ncRNA는 놀라운 안정성을 보이며, 장기간 보관된 혈청 샘플이나 심지어 미량의 농도에서도 프로파일링 할 수 있다.

RNA-RNA antisense 결합 기작을 통한 작용하는 miRNA는 유전자 발현 조절자로서 작용하여 병원성 조건을 예측하는 생체 바이오마커 뿐만 아니라 잠재적인 약물 표적으로 작용한다. 종합적으로, cfRNA는 당뇨 환자의 병태 생리학적 과정과 암 환자의 종양 상태 및 치료 효과를 반복적으로 스크리닝하고 모니터링하는데 사용되고, 민감하고 특이적이며 비침습적인 방법으로 주목받고 있으며, cfRNA 중심의 최신 연구는 이러한 관심을 뒷받침하고 있다.[4] (그림5)

Molecule	Phenotype/test	Tissue(s)	Technology	References
miRNA	NAFLD and NAFLD-related disease	Liver biopsy, blood, rodent liver, serum, plasma	TaqMan microRNA array, NGS, RT-qPCR, biochemical markers, ultrasound	(3, 87-92)
	Steatosis	Serum, plasma	NMR, ultrasound, TaqMan microRNA array, NGS	(93)
	Obesity	Blood, MIR-34a KO mice	RT-qPCR, ELISA, biochemical markers	(94,95)
	T1D	Human and murine islets	Biochemical markers, RT-qPCR	(96)
	T2D	ZDF rats, blood, murine liver	RT-qPCR, NGS, LC-MS/MS	(97-99)
mRNA	Diabetic retinopathy	Serum	RT-qPCR	(100)
	Oral cancer	Saliva	Microarray, RT-qPCR	(101)
	Lung cancer	Serum and bronchial lavage	RT-qPCR	(102,103)
	Prostate cancer	Plasma, PBMC, serum	RT-qPCR	(31)
	Breast cancer	Serum	RT-qPCR, PCR-based fluorescence microsatellite assay	(102, 104,105)
	Renal cell carcinoma	Urine	RT-qPCR	(5)
	Bladder cancer	Urine	Microarray	(106)
	Rectal cancer	Plasma	RT-qPCR	(107)
	Fetal genetic abnormalities	Plasma, placental tissue	RT-qPCR, microarray	(108-111)
	Sex determination	Plasma	RT-qPCR	(112)
lncRNA	Pre-eclampsia	Plasma	RT-qPCR	(113)
	Diabetic retinopathy	Whole blood	RT-qPCR	(114,115)
	Diabetic consequences	Cell culture, murine model, blood	WB, ELISA, RT-qPCR, NGS, cloning, cell assays, microscopy	(116-118)
	Diabetic retinopathy	Cell culture, murine model, blood	ELISA, RT-qPCR, microarray, WB, cloning, microscopy	(119,120)

cfRNA, cell-free RNA; ELISA, enzyme-linked immunosorbent assay; KO, knock out; LC-MS/MS, liquid chromatography tandem mass spectrometry; NAFLD, non-alcoholic fatty liver disease; NGS, next-generation sequencing; NMR, nuclear magnetic resonance; PBMC, peripheral blood mononuclear cells; RT-qPCR, real-time quantitative polymerase chain reaction; T1D, type 1 diabetes; T2D, type 2 diabetes, WB, Western blot; ZDF rat, Zucker diabetic fatty rat.

그림5. Examples of studies on circulating cell-free RNA as potential diagnostic tool. [4]

특히 exosome와 같은 세포 외 소포체 (EV) 내에서 분비되는 miRNA는 막 구조에 둘러싸여 뛰어난 안정성을 가지면서 서로 다른 조직 간의 인근 및 내분비 통신을 중재하며, 이를 통해 원격 세포의 유전자 발현과 기능을 조절한다고 알려져 있다. 특히 이러한 프로세스가 손상될 때 조직 기능 장애, 노화 및 질병으로 이어질 수 있기 때문에, Exosomal small ncRNA는 위암 환자를 조기 스크리닝하고 예후를 평가하기 위한 혁신적인 비침습적 검사 분자도구로 활용된다. 현재 액체 생체검사에 관한 연구는 순환 DNA, 순환 종양 세포, exosome 및 miRNA의 검출을 포함하고 있다. [9,10] (그림6)



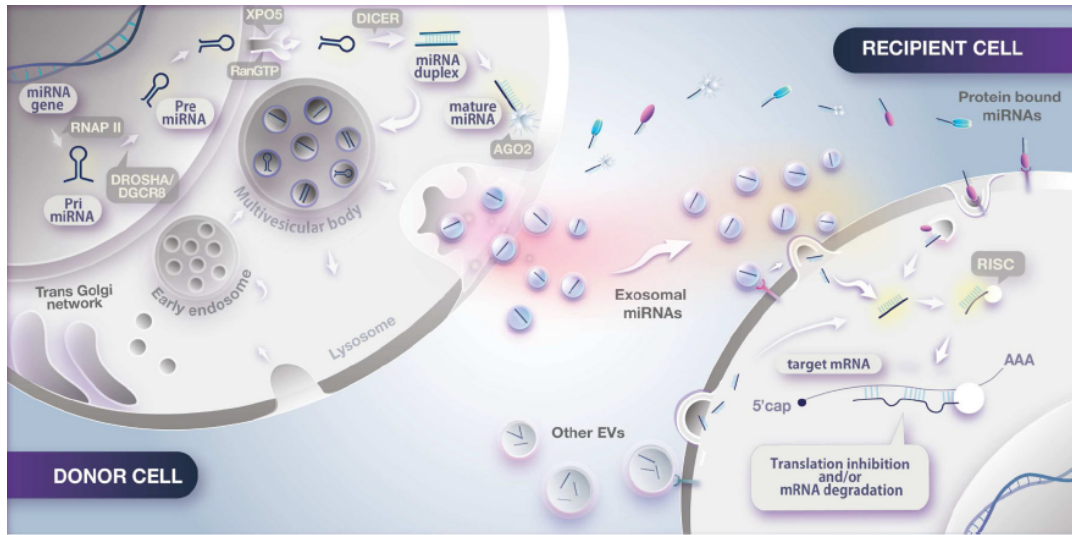


그림6. Intercellular communication via miRNAs. [10]

현재 체액 내에서 miRNA 분석을 위해 여러 기술이 사용되며, miRNA 마이크로어레이, qRT-PCR (quantitative Real-Time Polymerase Chain Reaction) 그리고 NGS시퀀싱(Next-Generation Sequencing)이 포함된다. 생물정보학적 분석을 위한 NGS 방식을 통해 샘플 내에서 알려지거나 새로운 전사체(transcripts)를 감지하고 상대적 양적 정량을 수행한다. 또한 qRT-PCR (quantitative Real-Time Polymerase Chain Reaction) 접근 방식을 사용하여 절대 정량을 측정하는 것이 일반적이다. 최근에는 매우 낮은 RNA input 양을 가지고 RNA-seq을 수행할 수 있도록 여러 프로토콜이 개발되고 수행된다.[11]

### E-biogen's Body Fluids Analysis Service

이바이오젠에서는 체액 타입의 시료를 대상으로 특정 질환 및 환경의 바이오마커 발굴 스크리닝을 할 수 있는 분석 서비스를 제공하고 있으며, 이바이오젠 홈페이지(www.e-biogen.com)를 방문하면 다양한 서비스 정보를 확인할 수 있다.

Service	Antibody array		*LC-MS Proteomics , PTM	**Small RNA-Seq, Exosomal small RNA-Seq	**Arraystar microarray
	RayBiotech	Full moon BioSystems			
Application	Protein profiling	Protein, phosphorylation profiling	Protein , PTM (glycosylation, phosphorylation) profiling	miRNA profiling	LncRNA, circular , piRNA, Super Enhancer LncRNA profiling
Sample Type	Serum/plasma, Exosome, other body fluids	Serum/plasma	Serum/plasma, Exosome CSF, urine, saliva, other body fluids	Whole blood, Serum/plasma , Exosome, Urine	Whole blood, Serum/plasma, Exosome
Experiment method	High-throughput ELISA Antibody based Array		Agilent 1260 & 1290 System Thermo Q-Exactive MS	NextSeq 500/550, SE 75bp	RNA microarray designed specific probes on chip
Data type	Semi-quantitative/ quantitative	Semi-quantitative	Semi-quantitative	Semi-quantitative	Semi-quantitative
Turnaround time	~ 4 weeks after Protein QC		~ 6 weeks after Protein QC	~ 4 weeks after RNA QC	~ 6 weeks after RNA QC
Link	<a href="#">Arrays   RayBiotech</a>	<a href="#">Arrays - Full Moon BioSystems</a>	<a href="#">ebiogen Inc. (e-biogen.com)</a>	<a href="#">ebiogen Inc. (e-biogen.com)</a>	<a href="#">Featured ncRNA Microarrays - piRNA Microarray - Super-enhancer LncRNA Array</a>

\* Body fluid는 시료 특성에 따라 단백질 양 및 처리방법이 상이할 수 있어, 자세한 상담과 관련 문헌 검토가 필요합니다.

\*\* 이바이오젠에서는 요구되는 RNA 양과 품질 평가를 위한 Quality Control을 통해 서비스 최종 진행 여부를 판단합니다.

< 참고문헌 >

1. Insights on Proteomics-Driven Body Fluid-Based Biomarkers of Cervical Cancer, *Proteomes*, 2022, 10, 13.
2. Proteomic biomarkers in body fluids associated with pancreatic cancer, *Oncotarget*, 2018, Vol. 9, (No. 23), pp: 16573-16587
3. High-throughput miRNA sequencing and identification of biomarkers for forensically relevant biological fluids, *Electrophoresis*, 2016, Volume 37, Issue 21, Pages 2780-2788
4. Cell-free DNA and RNA—measurement and applications in clinical diagnostics with focus on metabolic disorders, *Physiol Genomics* 53: 33–46, 2021.
5. Human body fluid proteome analysis, *Proteomics*. Author manuscript; available in PMC 2010 July 26. NIH-PA
6. Review of Liquid Chromatography-Mass Spectrometry-Based Proteomic Analyses of Body Fluids to Diagnose Infectious Diseases, *Int. J. Mol. Sci.* 2022, 23, 2187.
7. Body fluid biomarkers in multiple sclerosis, *Lancet Neurol* 2014; 13: 113–26
8. Current challenges and best practices for cell-free long RNA biomarker discovery, Cabús et al. *Biomarker Research* (2022) 10:62
9. Exosome-derived noncoding RNAs in gastric cancer: functions and clinical applications, Tang et al. *Mol Cancer* (2021) 20:99
10. Extracellular miRNAs: From Biomarkers to Mediators of Physiology and Disease, *Cell Metab.* 2019 October 01; 30(4): 656–673
11. MicroRNAs in body fluids—the mix of hormones and biomarkers, *Nat Rev Clin Oncol.* ; 8(8): 467–477.
12. Body Fluid Collection And Diagnostics Market Size, Share & Trends Analysis Report By Sample Type (Blood, Saliva, Urine, Cerebrospinal Fluid), By Products, By Technology, By Application, By Region, And Segment Forecasts, 2023 – 2030 Report ID: GVR-4-68040-019-8
13. HBFP: a new repository for human body fluid proteome, *Database*, Volume 2021, 2021, baab065