



(19)대한민국특허청(KR)  
(12) 공개특허공보(A)

(51) 。 Int. Cl. (11) 공개번호 10-2007-0065869  
C12Q 1/68 (2006.01) (43) 공개일자 2007년06월25일

(21) 출원번호	10-2007-7003939	(87) 국제공개번호	WO 2006/020005
(22) 출원일자	2007년02월20일	(43) 공개일자	2007년06월25일
심사청구일자	2007년05월15일		
번역문 제출일자	2007년02월20일		
(86) 국제출원번호	PCT/US2005/025138	(87) 국제공개번호	WO 2006/020005
국제출원일자	2005년07월15일	국제공개일자	2006년02월23일

(30) 우선권주장 60/589,627 2004년07월21일 미국(US)

(71) 출원인 더 리젠츠 오브 더 유니버시티 오브 캘리포니아  
미합중국, 94607-5200 캘리포니아, 오클랜드, 프랭클린 스트리트 1111, 5th 플로어

(72) 발명자 왕 데이비드  
미국 캘리포니아주 90211 비버리힐즈 사우스 클라크 드라이브 120

(74) 대리인 김진희  
강승옥

전체 청구항 수 : 총 6 항

(54) 타액 트랜스크립툼 진단법

(57) 요약

본 발명은 타액에서 유전자 산물을 측정하여 질병 상태의 진행을 진단, 확인 및 모니터링하는데 유용한 프로브 및 방법에 관한 것이다.

대표도

도 6

특허청구의 범위

청구항 1.

인간 타액 샘플을 수득하는 단계;

상기 인간 타액 샘플로부터 인간 mRNA를 수득하는 단계;

상기 mRNA를 증폭하여 핵산 증폭 산물을 제공하는 단계;

상기 핵산 증폭 산물을 분리하는 단계; 및

정상 개체와 인간 질병 상태를 나타내는 개체간에 차등 발현되는 mRNA를 확인하는 단계를 포함하는 인간 질병 상태용 마커(marker)의 확인방법.

## 청구항 2.

제1항에 있어서, 인간 질병 상태가 암, 자가면역 질환, 당뇨 및 신경계 질환중에서 선택되는 것인 방법.

## 청구항 3.

제1항에 있어서, 인간 mRNA를 수득하는 단계가 인간 타액 샘플을 RNase 저해제로 처리하는 단계를 포함하는 것인 방법.

## 청구항 4.

제3항에 있어서, RNase 저해제가 RNAlater™ 조성물을 포함하는 것인 방법.

## 청구항 5.

타액 샘플을 수득하는 단계; 및

상기 샘플을 RNAlater™ 조성물과 혼합하는 단계를 포함하는 타액 RNA의 보존방법.

## 청구항 6.

타액 수집용 컨테이너, 및 RNAlater™ 조성물을 포함하는 키트.

## 명세서

### 기술분야

### 발명의 분야

본 발명은 일반적으로 인간 질병 상태의 검출 및 진단, 및 그에 관련된 방법에 관한 것이다. 보다 구체적으로, 본 발명은 타액에서 유전자 산물을 측정하여 질병 상태의 진행을 진단, 확인 및 모니터링하는데 유용한 프로브 및 방법에 관한 것이다.

### 정부의 권리

35 U.S.C. §202(c)에 따라, 미국 국립보건원에서 지원을 해(허가번호 UO1 DE15018 및 RO1 DE15018) 이루어진 본 원에 개시된 발명에 대해 미국 정부가 부분적인 권리 행사를 할 수 있음이 인정된다.

### 배경기술

## 발명의 배경

타액은 혈청의 수동적 "한외여과물"이 아니지만(Rehak, N.N. et al. 2000 *Clin Chem Lab Med* 38:335-343), 효소, 호르몬, 항체 및 기타 분자의 특징적인 조성을 갖는다. 과거 10년간, 진단액으로서 타액의 사용은 진단학 및 다양한 증상의 위험이 있는 집단을 예측하는데 있어서 성공적으로 적용되어 왔다(Streckfus, C.F. & Bigler, L.R. 2002 *Oral Dis* 8:69-76). 카리에스(caries), 치주염, 구강암, 타액선 질병 및 전신성 질환, 예를 들어 간염 및 HIV를 모니터링하기 위한 진단 바이오마커(biomarker)가 타액에서 확인되었다(Lawrence, H.P. 2002 *J Can Dent Assoc* 68: 170-174).

인간 유전자 변형은 세포내적 및 세포외적 둘 다로 검출가능하다(Sidransky, D. 1997 *Science* 278:1054-1058). 핵산이 혈액, 소변 및 뇌척수액을 비롯한 대부분의 체액에서 확인되었으며, 질병을 위한 진단 바이오마커로 사용하기 위해 성공적으로 적용되어 왔다(Anker, P. et al. 1999 *Cancer Metastasis Rev* 18:65-73; Rieger-Christ, K.M. et al. 2003 *Cancer* 98:737-744; Wong, L.J. et al. 2003 *Cancer Res* 63:3866-3871). 최근의 관심사는 타액에서 핵산 마커의 검출을 개발하는데 맞추어져 있다. 지금까지, 타액내 대부분의 DNA 또는 RNA는 바이러스 또는 세균 기원인 것으로 밝혀졌다(Stamey, F.R. et al. 2003 *J Virol Methods* 108:189-193; Mercer, D.K. et al. 2001 *FEMS Microbiol Lett* 200:163-167). 구강암 환자의 타액에서 종양 세포 DNA 이질성을 입증하는 몇몇 보고가 있었다(Liao P.H. et al. 2000 *Oral Oncol* 36:272-276; El-Naggar, A.K. et al. 2001 *J Mol Diagn* 3:164-170). 본 발명자들은 타액에서 검출가능한 인간 mRNA에 대한 입증 사례를 확인하지 못했다.

미국에서 2004년도에 백삼십만명이 넘는 새로운 암 사례가 진단될 것으로 예상되었다(Cancer facts and figures 2004. Atlanta: American Cancer Society, 2004). 암으로 미국에서 올해에만 약 563,700명이 사망할 것으로 보이며, 이는 일분마다 한사람꼴로 사망하는 것이다. 암 치료에 진전이 있었음에도 불구하고, 이들 숫자는 과거 10년간 꾸준히 증가하였다. 또한, 구강암과 같은 일부 암의 경우에는, 과거 수십년간 전체 5년 생존율이 개선되지 않고 있으며, 약 30-50%의 낮은 비율을 유지하고 있다(Epstein, J.B. et al. 2002 *J Can Dent Assoc* 68: 617-621; Mao, L. et al. 2004 *Cancer Cell* 5:311-316). 예후가 개선되지 않는 결정적인 요인은 암의 상당 비율이 초기에 무증상 병소이며 진단된 상태에 이르기까지 진단 또는 치료되지 않는다는 데 있다. 암의 초기 검출이 이러한 질병으로부터 사망율을 감소시키는데 가장 효과적인 수단이다.

암 세포의 유전적 수차는 유전자 발현 패턴을 변경시키며, 암 표현형이 드러나기 훨씬 이전에 확인될 수 있다. 동일 기원의 정상 조직과 비교하여 암에서 독점적으로 또는 우선적으로 일어나는 변화가 분자 바이오마커로 이용될 수 있다(Sidransky, D. 2002 *Nat Rev Cancer* 2:210-219, 2002). 엄밀하게 확인된 바이오마커가 새로운 수단을 제시할 수 있으며, 암 초기 검출 및 암 위험 평가에 대한 주요 표적을 구성할 수 있다. 핵산을 기초로 한 각종 바이오마커가 암을 검출하는데 새로우면서도 강력한 수단인 것으로 입증되었다(Hollstein, M. et al. 1991 *Science* 253:49-53; Liu, T. et al. 2000 *Genes Chromosomes Cancer* 27:17-25; Groden, J. et al. 1991 *Cell* 66:589-600). 그러나, 대부분의 이들 마커는 후기 비침습성 및 전이 암으로부터의 암 세포주 또는 생검 재료에서 동정되었다. 발명자들은 바이오마커를 사용하여 암을 그의 최초 단계에 검출하는 능력에 여전히 한계가 있다. 더우기, 생검의 침습 특성은 고위험 군에서 암을 스크리닝하는데 부적합하도록 만든다. 이는 초기 검출을 개선할 새로운 진단 도구를 개발하는 것이 절실함을 제시하는 것이다. 초기 단계 또는 전암성 단계에서 암 발생을 예측할 수 있는 체액내 분자 마커의 동정이 상기 수단이 될 것이다.

## 발명의 요약

본 연구의 목적은 정상 대상으로부터 얻은 무세포 타액에서 트랜스크립토프 프로파일을 결정하는 것이다. 고품질도 올리고뉴클레오티드 마이크로어레이가 트랜스크립토프 프로파일링에 널리 사용되었다. 타액 트랜스크립토프 패턴이 타액 트랜스크립토프 진단법 응용에 기준 데이터베이스를 생성하는데 이용되었다.

타액은 다른 체액과 마찬가지로, 인간 건강 및 질병을 모니터링하는데 이용되고 있다. 이러한 연구는 정보 인간 mRNA가 무세포 타액에 존재함을 보여준다. 타액 mRNA는 구강 및 전신성 질병의 위험이 큰 집단 및 환자를 확인하기 위한 잠재적인 바이오마커를 제공한다. 고품질도 올리고뉴클레오티드 마이크로어레이가 타액 mRNA를 프로파일하는데 이용되었다. 무세포 타액에 수천개의 인간 mRNA가 존재함이 입증되었다. 정량적인 PCR(Q-PCR) 분석에 따라 본 마이크로어레이 연구로 확인된 mRNA의 존재가 증명되었다. 기준 데이터베이스는 정상 타액의 mRNA 프로파일을 기초로 설계되었다. 본 발명의 일 구체예로, 타액 트랜스크립토프 진단법(STD)은 질병 진단 뿐만 아니라 건강한 정상인을 조사하는데 사용된다.

다른 구체예로, 타액 트랜스크립토프 진단용 타액 RNA를 최적으로 보존하기 위한 실용적이면서도 사용하기에 친숙한 실온 프로토콜이 개발되었다.

## 발명의 상세한 설명

### 바람직한 구체예의 상세한 설명

본 발명은 인간 질병 상태를 초기에 검출, 진단 및 예측하는 것에 관한 것이다. 질병에 걸린 개체의 타액에서 얻은 RNA 분자로부터 발현된 폴리펩티드 또는 특정 서열의 발현 RNA 분자 형태의 질병 상태 마커가 개시된다. 이들 마커는 질병 상태의 지표이며, 정상 대상의 발현에 대해 차등적으로 발현되는 경우, 환자에서 질병 상태의 존재를 진단하게 된다. 이러한 마커는 당해 분야의 선행기술에 비해 상당한 이점을 제공한다. 이들은 타액 샘플에서 검출되기 때문에, 개체가 샘플을 취하기 전에 질병 상태(예: 종양)를 나타내는지 의심할 필요가 없고, 게다가 타액 샘플의 채취는 조직 생검 또는 채혈에 비해 환자에 훨씬 덜 침습적일 뿐만 아니라 덜 고통스럽다. 따라서, 개시된 검출 방법은 무증상 개체의 광범한 스크리닝에 적합하다.

### 실시예

#### 실시예 1

##### 마이크로어레이 기술을 이용한 무세포 타액의 RNA 프로파일링

##### 재료 및 방법

##### *정상적인 대상*

UCLA(캘리포니아주 로스앤젤레스의 캘리포니아 대학) 기관 감사 위원회에 의해 승인된 프로토콜에 따라, UCLA의 의학 센터에서, 이비인후과, 두경부 수술 분과로부터의 10명의 정상 공여자로부터 타액 샘플을 얻었다. 다음과 같은 포함 기준을 이용하였다: 연령 30세 이상; 악성 종양, 면역결핍, 자가면역 질환, 간염, HIV 감염 또는 흡연의 이력이 없음. 연구 집단은 6명의 남성과 4명의 여성으로 구성되었으며, 평균 연령은 42세였다(32세 내지 55세).

##### *타액 수집 및 처리*

간행된 프로토콜(Navazesh, M. 1993 *Ann N Y Acad Sci* 694:72-77)에 따라 아침 9시에서 10시 사이에 타액을 수집하였다. 대상에게 타액 수집에 앞서 적어도 한 시간 동안 식사, 음주, 흡연 또는 구강 위생 처리를 금하도록 요청하였다. 타액 샘플을 4 °C에서 15분간 2,600 × g으로 원심분리하였다. 타액 상등액을 세포상으로부터 분리하였다. 이어서, RNase 저해제(Suprase-In, Ambion Inc., Austin, TX, USA) 및 프로테아제 저해제(Aprotinin, Sigma, St.Louis, MO, USA)를 무세포 타액 상등액에 첨가하였다.

##### *무세포 타액으로부터 RNA 분리*

제조사로부터의 변형된 프로토콜을 이용하여 무세포 타액 상등액으로부터 RNA를 분리하였다(QIAamp Viral RNA kit, Qiagen, Valencia, CA, USA). AVL 완충액 (2,240 μl)과 충분히 혼합된 타액(560 μl)을 실온에서 10분간 인큐베이션하였다. 순수 에탄올(2,240 μl)을 첨가하고, 용액을 1분동안 6,000 × g으로 원심분리시켜 실리카 컬럼을 통과시켰다. 이어서, 컬럼을 두번 세척하고, 20,000 × g으로 2분동안 원심분리한 후, 2분동안 9,000 × g에서 30 μl의 RNase를 함유하지 않는 물로 용출시켰다. RNA의 분취액을 제조자의 설명에 따라 RNase를 함유하지 않는 DNase (DNase I-DNA-비함유, Ambion, Inc., Austin, TX, USA)로 처리하였다. 분리된 RNA의 품질을 세 가지 살림 유전자 전사물인 글리세르알데히드-3-포스페이트 데하이드로게나제(GAPDH), 액틴-β(ACTB) 및 리보솜 단백질 S9(RPS9)에 대한 RT-PCR로 조사하였다. PRIMER3 소프트웨어(genome.wi.mit.edu)를 이용하여 프라이머를 설계하고 다음과 같이 상용적으로 합성하였다(Fisher Scientific, Tustin, CA, USA): GAPDH를 위해 5'TCACCAGGGCTGCTTTAACTC3'(서열 번호 1) 및 5'ATGACAAGCTTCCCGTTCTCAG3'(서열 번호 2); ACTB를 위해 5'AGGATGCAGAAGGAGATCACTG3'(서열 번호 3) 및 5'ATACTCCTGCTTGCTGATCCAC3'(서열 번호 4); RPS9를 위해 5'GACCCTTCGAGAAATCTCGTCTC3'(서열 번호 5) 및 5'TCTCATCAAGCGTCAGCAGTTC3'(서열 번호 6). Ribogreen® RNA 정량화 키트(Molecular Probes, Eugene, OR, USA)를 이용하여 RNA의 양을 추정하였다.

##### *표적 cRNA 제조*

분리된 RNA를 공개된 방법(Ohyama, H. et al., 2000 *Biotechniques* 29:530-536)에 따라 선형 증폭시켰다. 요약하면, T7-올리고-(dT)24를 프라이머로 이용한 역전사를 실시하여 제 1 스트랜드 cDNA를 합성하였다. 시험관내 전사(IVT)의 첫번째 라운드를 T7 RNA 폴리머라제(Ambion, Inc., Austin, TX, USA)를 이용하여 실시하였다. BioArray<sup>TM</sup> 고수율 RNA 전사물 표지 시스템(Enzo Life Sciences, Farmingdale, NY, USA)을 cRNA 산물을 비오틴화하기 위해 두번째 라운드의 IVT에 사용하였다; 표지된 cRNA를 GeneChip<sup>®</sup> 샘플 정화 모듈(Affymetrix, Santa Clara, CA, USA)을 이용하여 정제하였다. cRNA의 양 및 품질을 분광학 및 겔 전기영동에 의해 결정하였다. 각 분리 및 증폭 단계로부터의 분취물 소량을 RT-PCR에 의한 품질 평가에 이용하였다. 단편화된 cRNA(Kelly, J.J. et al. 2002 *Anal Biochem* 311:103-118에 의해 개시된 대로 제조됨)의 품질을 또한 2100 바이오어널라이저(Agilent Technologies, Palo Alto, CA, USA)를 이용하여 모세관 전기영동에 의해 평가하였다.

#### *HG-U133A 마이크로어레이 분석*

약 19,000개의 유전자에 상당하는 22,215개의 인간 유전자 cDNA 프로브 세트를 함유하는 Affymetrix 인간 게놈 U133A 어레이(즉, 각 유전자는 하나보다 많은 프로브 세트로 나타내어질 수 있음)를 유전자 발현 프로파일링에 적용하였다. 어레이 데이터를 Microarray Suite(MAS) 소프트웨어(Affymetrix)를 이용하여 정규화하고 분석하였다. 각 프로브 세트에 대해 검출 p-값을 얻었다.  $p < 0.04$ 를 갖는 임의의 프로브 세트에 "존재(present)"를 할당하였으며, 이는 매칭되는 유전자 전사물이 신뢰성있게 검출됨을 의미한다(Affymetrix, 2001). 각 어레이상에 존재하는 프로브 세트의 총 수를 얻고 존재하는 유전자의 존재 퍼센트(P%)를 계산하였다. 유전자 온톨로지 수집 도구(netaffx.com)를 이용하여 선택된 유전자(10개의 모든 어레이에 존재,  $p < 0.01$ )에 대해 기능적 분류를 실시하였다.

#### *Q-PCR에 의한 정량적 유전자 발현 분석*

iCycler<sup>TM</sup> 열 순환기(Bio-Rad, Hercules, CA, USA)를 이용하여 Q-PCR을 실시하였다. 분리된 타액 RNA(증폭 없음) 2  $\mu$ l 분취액을 MuLV 역전사 효소(Applied Biosystems, Foster City, CA, USA)를 이용하여 cDNA로 역전사시켰다. 생성된 cDNA(3  $\mu$ l)를 iQ SYBR Green Supermix(Bio-Rad, Hercules, CA, USA)를 이용한 PCR 증폭에 이용하였다. Sigma-Genosys(Woodlands, TX, USA)에 의해 다음과 같이 프라이머를 합성하였다: 인터류킨 1-베타(IL1B)를 위해 5'GTGCTGAATGTGGACTCAATCC3'(서열 번호 7) 및 5'ACCCTAAGGCAGGCAGTTG3'(서열 번호 8); 스타티핀(SFN)을 위해 5'CCTGCGAAGAGCGAAACCTG3'(서열 번호 9) 및 5'TCAATACTGGACAGCACCTCC3'(서열 번호 10); 유비쿼터스 튜블린-알파(K-ALPHA-1)를 위해 5'AGCGTGCCTTTGTTCACTG3'(서열 번호 11) 및 5'CACACCAACCTCCTCATAATCC3'(서열 번호 12). 특이적 PCR 산물에 관계적인 조건으로 모든 반응을 삼중으로 실시하였다. 특정 템플레이트의 초기 cDNA 양을 LightCycler 소프트웨어 3.0(Bio-Rad, Hercules, CA, USA)을 이용하여 표준 곡선으로부터 외삽하였다. 표준 곡선에 의한 정량에 대한 상세한 과정은 앞서 개시되었다(Ginzinger, D. 2002 *Exp Hematol* 30:503-512).

#### 결과:

##### *RNA 분리 및 증폭*

560  $\mu$ l 무세포 타액 샘플로부터 평균  $60.5 \pm 13.1$  ng(n=10)의 전체 RNA를 수득하였다(표 1). RT-PCR 결과는 10개의 모든 타액 샘플이 살림 유전자: GAPDH, ACTB 및 RPS9를 코딩하는 mRNA를 함유하고 있음을 입증한다. 이들 유전자의 mRNA는 80 °C에서 6개월 이상동안 유의적인 분해없이 보존될 수 있다(도 1). 2 라운드의 T7 RNA 선형 증폭후, 비오틴화된 cRNA의 평균 수율은  $A_{260}/A_{280} = 2.067$ 로  $42.2 \pm 3.9$   $\mu$ g이었다(표 1).

[표 1]

10명의 정상 공여자로부터 수득한 무세포 타액의 유전자 발현 프로파일링

대상	성별	연령	RNA (ng) <sup>a</sup>	cRNA (μg) <sup>b</sup>	존재하는 프로브 <sup>c</sup>	프로브 P% <sup>d</sup>
1	F	53	60.4	44.3	3172	14.24
2	M	42	51.6	40.8	2591	11.62
3	M	55	43.2	34.8	2385	10.70
4	M	42	48.2	38.0	2701	12.12
5	M	46	60.6	42.7	3644	16.35
6	M	48	64.8	41.8	2972	13.34
7	F	40	75.0	44.3	2815	12.63
8	M	33	77.8	49.3	4159	18.66
9	F	32	48.8	41.4	2711	12.17
10	F	32	79.8	44.4	4282	19.22
평균 ± SD		42±8.3	60.5±13.1 2	42.2±3.94	3143±665.0	14.11±2.98

<sup>a</sup> 560 μl 무세포 타액 상등액내 전체 RNA의 양

<sup>b</sup> 2 라운드의 T7 증폭후 cRNA의 양

<sup>c</sup> HG U133A 마이크로어레이상의 존재 콜(call)을 나타내는 프로브의 수(검출 p<0.04).

<sup>d</sup> 존재 퍼센트(P%) = 존재 콜에 할당된 프로브의 수/총 프로브의 수(HG U133A 마이크로어레이에 대해 22,283).

cRNA는 단편화전에 200 bp 내지 4 kb 범위였으며; 단편화후 약 100 bp로 집중되었다. cRNA 프로브의 품질을 하이브리드화전에 모세관 전기영동으로 확인하였다. 원형 샘플 및 각 증폭 단계로부터의 산물에 대해 PCR/RT-PCR을 이용하여 ACTB mRNA가 검출가능하였다; 첫번째 cDNA, 첫번째 시험관내 전사(IVT), 두번째 cDNA 및 두번째 IVT(도 2).

타액 mRNA의 마이크로어레이 프로파일링

22,283 cDNA 프로브를 함유하는 HG U133A 어레이를 이용하여 10명의 정상 대상의 타액 mRNA 프로파일을 얻었다. 존재 콜이 할당된 각 어레이에서 평균 3,143±665.0 프로브 세트(p<0.04)가 발견되었다(n=10). 이들 프로브 세트는 약 3,000개의 상이한 mRNA를 나타낸다. 평균 존재 콜 퍼센트는 14.11±2.98%(n=10)이었다. 10개의 어레이로부터의 데이터를 포함하는 기준 데이터베이스가 생성되었다. GAPDH, ACTB 및 RPS9를 나타내는 프로브 세트는 모든 10개의 어레이상에 존재 콜을 할당하였다. 검출 p<0.01로 10개의 모든 어레이상에 존재 콜이 할당된 185개 유전자에 상응하는 총 207개의 프로브 세트가 있었다. 이들 유전자들은 생물학적 과정에서 그들의 공지 역할 및 분자 기능을 기초로 분류되었다(표 2). 185개 유전자의 주요 기능은 세포 성장/유지(119 유전자), 분자 결합(118 유전자) 및 세포 구조 조성(95 유전자)에 관련된다. 이들을 "정상 타액 핵심 트랜스크립톰(NSCT)"으로 명명하였다.

[표 2]

10명의 정상 공여자로부터 얻은 무세포 타액내 185개 유전자의 생물학적 과정 및 분자 기능(Gene Ontology Mining Tool 을 이용하여 얻은 데이터)

생물학적 과정 <sup>a</sup>	유전자, n <sup>b</sup>	분자 기능 <sup>a</sup>	유전자, n <sup>b</sup>
세포 성장 및/또는 유지	119	결합	118
대사	93	핵산 결합	89
생합성	70	RNA 결합	73
단백질 대사	76	칼슘 이온 결합	12
뉴클레오티드 대사	10	기타 결합	23

기타 대사	18	구조 분자	95
세포 조직화 및 생물발생	2	리보솜 구성원	73
항상성	3	세포골격 구성원	17
세포 사이클	5	근육 구성원	2
세포 증식	11	폐물	15
수송	5	수송자	4
세포 운동성	8	효소	20
세포 소통	34	시그널 전달	10
외부 자극에 대한 반응	19	전사 조절자	7
세포 부착	3	해독 조절자	5
세포-세포 시그널링	5	효소 조절자	9
시그널 전달	17	세포 부착 분자	1
폐물	8	미지의 분자 기능	6
발생	18		
사멸	2		
미지의 생물학적 과정	11		

<sup>a</sup> 한 가지 유전자는 다수의 분자 기능을 가지거나 상이한 생물학적 과정에 관여할 수 있다.

<sup>b</sup> 다수의 유전자가 특정 그룹/서브그룹으로 분류될 수 있다.

*Q-PCR 확인 및 정량 분석*

실시간 정량적 PCR(Q-PCR)을 이용하여, 185개의 "정상 타액 핵심 트랜스크립톰" 유전자로부터 선택된 유전자를 정량하여 타액내 인간 mRNA의 존재를 확인하였다. 확인을 위해 IL1B, SFN 및 K-ALPHA-1를 무작위로 선택하고, 10개의 모든 어레이에 존재 여부를 할당하였다. Q-PCR 결과는 IL1B, SFN 및 K-ALPHA-1의 mRNA가 10개의 모든 비증폭된 원형 무세포 타액에서 검출가능함을 보여주었다. 이들 전사물(n=10)의 상대적인 양(카피 수)은 다음과 같다: IL1B에 대해  $8.68 \times 10^3 \pm 4.15 \times 10^3$ ; SFN에 대해  $1.29 \times 10^5 \pm 1.08 \times 10^5$ ; 및 K-ALPHA-1에 대해  $4.71 \times 10^6 \pm 8.37 \times 10^5$ . Q-PCR에 의해 측정된 이들 유전자의 상대적인 RNA 발현 수준은 마이크로어레이에 의해 측정된 것과 유사하였다.

타액은 이상적인 진단 매질로 작용하기 위한, 저렴하고 비침습성이며 접근가능한 체액의 요구를 만족시켜 준다. 타액내 특이적인 정보성 바이오마커는 질병을 진단하고 인간 건강을 모니터링하는 기능을 위해 가한층 요구된다(Bonassi, S. et al. 2001 *Mutat Res* 480-481:349-358; Streckfus, C.F. et al. 2002 *Oral Dis* 8:69-76; Sidransky, D. 2000 *Nat Reviews* 3:210-219). 타액 성분을 알아내는 것은 상기 매질을 이용하여 질병을 진단하는데 가능성있는 바이오마커를 확인하는데 필수적이다(Pusch, W. et al. 2003 *Pharmacogenomics* 4:463-476). 본 발명에 앞서, 정보 분자가 일반적으로 타액에 소량으로 존재한다는 비평적인 생각이 있었다. 그러나, 새로운 증폭 기술 및 고도의 민감성 분석으로, 이제 이것은 더 이상 제한이 될 수 없다(Xiang, C.C. et al 2003 *Nucleic Acids Res* 31:e53). 본 실시예에서, 무세포 타액 상등액으로부터 인간 RNA가 성공적으로 분리되었다. 타액 mRNA의 품질은 RT-PCR, Q-PCR 및 마이크로어레이 실험에 사용하기에 충분한 것으로 입증되었다.

타액이 고유적으로 미생물을 함유한다는 점에서 타액과 기타 체액(예: 혈액)간에는 상이한 차이가 있다(Sakki, T. & Knuuttila, M. 1996 *Eur J Oral Sci* 104:619-622). 또한, 일부 외부 물질(예: 음식물 조각)이 타액의 조성을 좀 더 복잡하게 만든다. 따라서, 바이오마커를 검출하기 위한 매질로서 전체 타액 대신 타액을 유체/상등액 상태로 사용하는 것이 좀 더 간편하고 보다 정확하다. 본 실시예에서는, 펠렛 및 타액 상등액을 분리하기 위한 조건이 무세포 유체 상에서 검출된 RNA에 기여할 수 있는 세포 원소의 기계적인 파괴가 방지되도록 최적화되었다(St. John, M.A.R. et al. 2004, in press). 이들 결과는 트랜스크립톰 분석에 무세포 타액을 사용하는 것이 적당하면서도 효율적임을 증명해 준다. 인간 mRNA가 무세포 타액 상등액에 존재한다는 것이 새로운 발견임에 반해, 핵산은 오래전부터 다른 무세포 체액에서 검출되어 질병 진단에 이용되어 왔다(Sidransky, D. 1997 *Science* 278:1054-1058). 예를 들어, 특이적인 종양 유전자, 종양 억제 유전자 및 초위성체(microsatellite) 변경이 환자 혈청에서 확인되었다(Anker, P. et al. 2003 *Int J Cancer* 103:149-152). 또한, 상이한 악성 종양이 있는 환자의 혈청으로부터 종양 mRNA가 분리되어 증폭되었다(Kopreski, M.S. et al. 1999 *Clin Cancer*

Res 5:1961-1965; Fleischhacker, M. et al 2001 *Ann NY Acad Sci* 945:179-188). 세포외적으로 검출된 이들 게놈 메신저는 질병용 바이오마커로서 제공될 수 있다고 널리 받아들여지고 있다(Sidransky, D. 1997 *Science* 278:1054-1058).

우리가 아는 한, 타액내 인간 mRNA가 포괄적으로 프로파일된 것은 최초 보고라고 생각한다. 마이크로어레이 기술을 이용하여, 본 발명자들은 약 3,000개의 상이한 인간 mRNA가 각 정상 대상의 무세포 타액에 존재함을 발견하였다. 정상 집단으로부터 얻은 무세포 타액내의 타액 트랜스크립토펙 패턴이 건강을 모니터링하기 위한 데이터베이스로 제공될 것으로 구상된다. 본 발명자들이 본 발명에 따라 30,000개가 넘는 유전자로 구성된 인간 게놈을 알아내고(Venter J.C. et al. 2001 *Science* 291:1304-1351), 본 실시예에 사용된 HG U133A 마이크로어레이상의 프로브 세트는 약 19,000개까지만의 인간 유전자만을 나타내며, HG U133A 마이크로어레이에 의해서는 검출할 수 없었던 추가의 유전자 전사물이 무세포 타액에 존재할 것을 예상하고, 본 발명을 이용하여 검출할 수 있었음에 특히 주목하여 주기 바란다. 본 실시예에서 확인된 유전자 전사물, 특히 정상 타액 핵심 트랜스크립토펙(NSCT) mRNA는 정상 무세포 타액의 일반적인 트랜스크립토펙을 나타낸다. 본 발명자들은 여러 질병 증상을 갖고 있는 환자로부터 얻은 타액에서 상이한 정보성 진단 트랜스크립토펙을 확인할 수 있을 것으로 구상하였다. 따라서, 인간 타액 mRNA는 구강내에서 나타나는 구강 및 전신성 질병의 진단 바이오마커로 사용될 수 있을 것으로 구상된다.

본 발명의 일 구체예로, 타액 트랜스크립토펙 진단법은 정상 환자의 건강 상태를 모니터링하기 위해 사용된다. 다른 구체예로, 타액 트랜스크립토펙 진단법은 암(예: 전립선암, 결장암, 유방암, 폐암, 구강암 등) 및 전신성 질병, 예를 들어 자가면역 질환, 당뇨, 골다공증; 신경계 질환, 예를 들어 알츠하이머병, 파킨슨씨병 등의 초기 진단을 위한 질병용 마커를 검출하기 위해 사용된다.

## 실시예 2

### 구강암 검출용 타액 트랜스크립토펙 진단법

#### 목적:

구강액(타액)은 비침습성이며 접근이 용이하고 매우 효율적인 진단 매질의 요구를 충족한다. 인간 RNA의 대형 패널이 타액에서 신뢰적으로 검출될 수 있다는 본 발명의 발견은 새로운 임상 접근으로서 타액 트랜스크립토펙 진단법을 제시하였다. 본 실시예에서, 발명자들은 원리 증명 질병으로서 구강 편평세포암종(OSCC)을 이용하여 이와같은 새로운 접근법의 진단 가치를 평가하였다.

일차 종양에 존재하는 동일한 돌연변이가 병든 환자로부터 시험된 체액에서도 동정될 수 있는 것으로 나타났다(Sidransky, D. 1997 *Science* 278:1054-1059). 혈액, 소변 및 뇌척수액(CSF)의 암 관련 핵산이 암 진단용 바이오마커로 이용되고 있다(Anker, P. et al 1999 *Cancer Metastasis Rev* 18:65-73; Rieger-Christ, K.M. et al. 2003 *Cancer* 98:737-744; Wong, L.J. et al. 2003 *Cancer Res* 63:3866-3871). 보다 최근에, 혈청 또는 혈장내 mRNA 바이오마커가 암에 걸린 환자에서 RT-PCR에 기초한 검출 전략의 표적이 되고 있다(Kopreski, M.S. et al. 2001 *Ann NY Acad Sci* 945:172-178; Bunn, P.J., Jr. 2003 *J Clin Oncol* 21:3891-3893). 체액에서 이러한 바이오마커의 수가 증가할수록 유전적 변형을 집단 검사할 수 있도록 하는 보다 강력하고 비용 효율적인 방법을 이용할 수 있는 가능성이 높아진다. 대형 패널의 인간 mRNA가 타액에 존재한다는 마이크로어레이 기술에 따른 본 발명의 발견(실시예 1)은 질병 진단에서의 응용 뿐만 아니라 건강한 정상인을 조사하는데 새로운 임상적 접근인 타액 트랜스크립토펙 진단법을 제공한다. 이는 타액으로부터 RNA 특징을 이용하는데 매우 생산적이고 강력하며 재현성있는 접근법이다. 또한, 타액을 진단 유체로 사용하는 것은 저렴하고 비침습성이며 접근가능한 진단 방법의 요구를 만족시켜 준다(Lawrence, H.P. 2002 *J Can Dent Assoc* 68:170-174, 2002). 본 실시예에서, 발명자들은 상이한 mRNA 발현 패턴이 암 환자로부터의 타액에서 동정될 수 있고, 상이하게 발현된 전사물이 암 검출용 바이오마커로 제공될 수 있다는 가설을 검증하였다. 본 연구에서 원리 증명 질병은 구강 편평세포암종(OSCC)이었다. 원리는 구강암 세포가 타액 환경내에 들어 있고, 유전 이질성이 OSCC 환자 타액에서 검출된다는 것이다(El-Naggar, A.K et al. 2001 *J Mol Diagn* 3:164-170; Liao, P.H et al. 2000 *Oral Oncol* 36:272-276, 2000).

#### 실험 설계:

무자극 타액을 일차 T1/T2 OSCC 환자(n=32) 및 정상 대상(n=32)으로부터 매치되는 연령, 성별 및 흡연 이력에 따라 수집하였다. RNA 분리를 타액 상등액으로부터 실시하고, T7 RNA 폴리머라제를 이용한 2 라운드 선형 증폭을 실시하였다. 인간 게놈 U133A 마이크로어레이를 인간 타액 트랜스크립토펙의 프로파일링에 적용하였다. 상이한 유전자 발현 패턴을 10



명의 매치되는 암 환자 및 대조군에서 t 테스트 비교와 배수 변화 분석을 조합하여 분석하였다. 정량적 PCR(qPCR)을 이용하여 마이크로어레이에 의해 유의적인 차를 나타내는( $P < 0.01$ ) 선택된 유전자의 유효성을 확인하였다. 이들 타액 mRNA 바이오마커의 예측력을 수용체 작동 특성 곡선 및 분류 모델에 의해 분석하였다.

#### 결과:

마이크로어레이의 분석 결과는 암 환자와 대조군 사이에 타액내 유의적으로 상이한 발현 수준을 나타낸 1,679개의 유전자가 있음을 보여주었다( $P < 0.05$ ). OSCC 타액에서 적어도 3.5배 증가를 나타낸 7개의 암-관련 RNA 바이오마커( $P < 0.01$ )는 OSCC 환자( $n=32$ ) 및 대조군( $n=32$ )으로부터의 타액 샘플에 대한 qPCR에 의해 일관적으로 확인되었다. 이들 타액 RNA 바이오마커는 인터류킨 8(IL-8), 인터류킨 1-베타(IL1B), 이중 특이성 포스파타제 1(DUSP1), H3 히스톤, 3A 패밀리(H3F3A), 오르니틴 데카복실라제 안티자임 1(OAZ1), S100 칼슘 결합 단백질 PS(100P) 및 스피르디민/스피르민 N1-아세틸 트랜스퍼라제(SAT)의 전사물이다. 이들 바이오마커의 조합은 대조군으로부터 OSCC를 구별하는데 민감성(91%) 및 특이성(91%)을 양산하였다.

#### 결론:

구강암 검출을 위한 본 조사에서 타액 트랜스크립톰 진단법의 유용성이 성공적으로 입증되었다. 이러한 새로운 임상적 접근은 초기 암 검출을 위한 강력하고 매우 생산적이며 재현성있는 수단인 것으로 구상된다. 타액 트랜스크립톰 프로파일링은 다른 주요 질병을 평가하고 건강한 정상인을 검사하는데 적용할 수 있을 것으로 구상된다.

#### 환자 및 방법

##### 환자 선별

UCLA 의학 센터; 남가주 대학(USC) 의학 센터(Los Angeles, CA); 및 캘리포니아 샌프란시스코 대학(UCSF) 의학 센터(San Francisco, CA)로부터 구강 편평세포암종(OSCC) 환자를 모집하였다. 일차 T1 또는 T2 OSCC의 기록을 가진 32명의 환자를 이 연구에 포함시켰다. 모든 환자들은 최근 일차 질병으로 진단받았으며, 화학요법, 방사선치료, 수술 또는 대체요법의 형태의 어떤 이전 치료도 받지 않았다. 비교할만한 흡연 이력을 가진 동일한 수의 연령 및 성별이 매치되는 개체를 대조 그룹으로 선택하였다(St. John, M.A.R et al. 2004 IL-6 and IL-8: Potential Biomarkers for Oral Cavity and Oropharyngeal SCCA. Archives of Otolaryngology-Head & Neck Surgery, in press). 두 대상 그룹 사이에서, 평균 연령: OSCC 환자,  $49.8 \pm 7.6$ 세; 정상 대상,  $49.1 \pm 5.9$ 세(스튜던트 t 테스트  $P > 0.80$ ); 성별( $P > 0.90$ ); 또는 흡연 이력( $P > 0.75$ ) 면에서 유의적인 차이가 없었다. 어떠한 개체도 이전에 악성 종양, 면역결핍, 자가면역 질환, 간염 또는 HIV 감염의 병력을 가지지 않았다. 모든 대상들은 실험에 타액을 공여하는데 동의를 한다고 기관 감사 위원회에 서명하였다.

##### 타액 수집 및 RNA 분리

간행된 프로토콜(Navazesh, M. 1993 *Ann N Y Acad Sci* 694:72-77)에 따라 아침 9시에서 10시 사이에 무자극 타액 샘플을 수집하였다. 대상에게 타액 수집에 앞서 적어도 한 시간 동안 식사, 음주, 흡연 또는 구강 위생 처리를 금하도록 요청하였다. 타액 샘플을  $4^{\circ}\text{C}$ 에서 15분간  $2,600 \times g$ 으로 원심분리하였다. 상등액을 펠렛으로부터 제거하고, RNase 저해제(Suprase-In, Ambion Inc., Austin, TX)로 처리하였다. QIAamp Viral RNA 키트(Qiagen, Valencia, CA)를 사용하여  $560 \mu\text{l}$ 의 타액 상등액으로부터 RNA를 분리하였다. 분리한 RNA 분취물을 제조자의 지시에 따라 RNase를 함유하지 않는 DNase(DNaseI-DNA-비함유, Ambion Inc., Austin, TX)로 처리하였다. 분리한 RNA의 품질을 세 개의 세포 관리 유전자 전사물인 글리세르알데히드-3-포스페이트 데하이드로게나제(GAPDH), 액틴- $\beta$ (ACTB) 및 리보솜 단백질 S9(RPS9)에 대한 RT-PCR로 조사하였다. 세 개의 모든 유전자에 대해 PCR 산물을 나타내는 샘플만을 후속 분석에 사용하였다.

##### 마이크로어레이 분석

10명의 OSCC 환자(남성 7명, 여성 3명, 연령= $52 \pm 9.0$ ) 및 10명의 성별 및 연령이 매치되는 정상 공여자(연령= $49 \pm 5.6$ )로부터의 타액을 마이크로어레이 연구에 이용하였다. 타액으로부터 분리한 RNA를 RiboAmp<sup>TM</sup> RNA 증폭 키트(Arcturus, Mountain View, CA)로 선형 증폭시켰다. 예정량( $0.1 \mu\text{g}$ )의 대조 RNA를 다섯개의 독립적인 실험에 20개의 샘플을 병행 이용하여 RNA 증폭 효율을 측정하였다. 실시예 1에 기재된 프로토콜에 따라, Affymetrix 인간 게놈 U133A 어레이(HG U133A, Affymetrix, Santa, Clara, CA)를 유전자 발현 분석에 적용하였다.

어레이를 스캔하고, Microarray Suit 5.0 소프트웨어(Affimetrix, Santa Clara, CA)로 형광 강도를 측정 후, 정규화 및 모델에 기초한 분석(Li, C. & Wong, W.H. 2001 *PNAS USA* 98:31-36)을 위해 DNA-칩 분석기 소프트웨어에 입력하였다. S-플러스 6.0(Insightful, Seattle, WA)을 모든 통계적 시험에 이용하였다. 상이하게 발현된 전사물을 결정하기 위하여 세 가지 기준을 이용하였다. 먼저, 모든 샘플에서 "부재" 콜로 할당된 어레이상의 프로브 세트를 제외시켰다. 둘째로, OSCC(n=10)와 대조군(n=10)에서 평균 유전자 발현 시그널 강도의 비교를 위해 양측(two-tailed) 스튜던트 t 테스트를 이용하였다. 통계적 유의성을 위해 0.05의 임계적 알파 수준을 정의하였다. 셋째, 통계적으로 유의한 차이를 나타낸 유전자 전사물에 대해 배수비를 계산하였다( $P < 0.05$ ). 적어도 2배 변화를 나타내는 유전자 전사물만이 추가 분석을 위해 포함될 것이다.

### 정량적인 PCR 확인

마이크로어레이 분석으로 동정된 상이하게 발현된 전사물의 서브세트를 확인하기 위하여 qPCR을 실시하였다. MuLV 역전사효소(Applied Biosystems, Foster City, CA) 및 프라이머로서 랜덤 헥사머(ABI, Foster City, CA)를 이용하여, 원형 및 비증폭 타액 RNA로부터 cDNA를 합성하였다. iCycler<sup>TM</sup> PCR 시스템(Bio-Rad, Hercules, CA, USA), iQ SYBR Green Supermix(Bio-Rad, Hercules, CA)에서 qPCR 반응을 실시하였다. PRIMER3 소프트웨어를 이용하여 프라이머 세트를 설계하였다. 모든 반응을 특정 산물에 관례적인 조건으로 삼중 실시하였다. 특정 템플레이트의 cDNA/RNA의 초기 양을 앞서 개시된 바와 같이 표준 곡선으로부터 외삽하였다(Ginzinger, D.G. 2002 *Exp Hematol* 30:503-512). 이러한 확인은 마이크로어레이 연구를 위해 이전에 이용된 20개를 포함한 모든 샘플(n=64)을 시험하여 완결하였다. 윌콕슨 순위 합 검정(Wilcoxon Signed Rank test)을 통계적 분석에 이용하였다.

### 수용체 작동 특성(ROC) 곡선 분석 및 예측 모델

RT-qPCR 결과를 이용하여, ROC 곡선 분석(Grunkemeier, G.L. & Jin, R. 2001 *Ann Thorac Surg* 72:323-326)을 S-플러스 6.0으로 작성하여 각 바이오마커의 예측력을 평가하였다. 최대 상응하는 민감성 및 특이성을 제공하는 것을 조사하여 각 바이오마커에 대해 최적화된 컷포인트를 결정하였다. 그 후, 최적화된 민감성 및 특이성 값 세트에 기초해 ROC 곡선을 플로팅하였다. 곡선 아래 면적을 ROC 곡선의 수치 적분을 통해 계산하였다. ROC 곡선 아래 면적이 가장 큰 바이오마커가 OSCC를 검출하기 위한 가장 강력한 예측력을 가지는 것으로 확인되었다.

이어서, 다중변수 분류 모델을 작성하여 암 예측을 위한 타액 마커의 최상의 조합을 결정하였다. 먼저, 의존적 변수로서 질병(OSCC) 및 비-질병(정상)의 이원 결과를 이용하여, 연령, 성별 및 흡연 이력의 제어를 위한 로지스틱 회귀 모델을 작성하였다. 후진 단계적 회귀(Renger, R. & Meadows, L.M. 1994 *Acad Med* 69:738)를 이용하여 최상의 최종 모델을 확인하였다. 로지스틱 회귀 모델을 검증하기 위하여 리브-원 아웃 교차 확인(leave-one out cross validation)을 이용하였다. 교차 확인 전략은 먼저 한 가지 관찰을 제거하고 이어서 모든 마커를 이용하여 나머지 경우로부터 로지스틱 회귀 모델을 맞춘다. 모델을 개선하지 않는 변수를 제거하도록 하기 위해 이들 각 모델에서 단계적인 모델 선택을 이용한다. 이어서, 남겨진 경우의 마커 값을 이용하여 그 관찰에 예측되는 부류를 계산하였다. 교차 확인 에러율은 부정확하게 예측된 샘플의 수를 샘플의 수로 나눈 것이다. 민감성 및 특이성 계산에 가능한 컷포인트로서 모델로부터 적합화된 확률을 이용하여, ROC 곡선을 유사한 절차에 의해 로지스틱 모델에 대해 계산하였다.

그 다음으로, 예측자로서 확인된 mRNA 바이오마커를 사용하여 S-플러스 6.0에 의해 트리-베이스(tree-based) 분류 모델, 분류 및 회귀 트리(CART)를 작성하였다. CART는 이원 회귀성 분할에 의한 분류 모델에 적합하며, 이때 각 단계는 암 대 정상 그룹의 최상의 스플릿(split)을 야기하는 예측자 변수의 조사를 포함한다(Lemon, S. C. et al. 2003 *Ann Behav Med* 26:172-181). CART는 S-플러스에 대한 디폴트 세팅으로 결정된 스플리팅(splitting) 기준을 가지는 엔트로피 함수를 이용하였다. 이어서, 이 방식에 따라, 전체 샘플(n=64)을 포함하는 모 그룹을 암 그룹과 정상 그룹으로 나누었다. 상이한 분류를 갖는 서브-브랜치를 야기하지 않는 모든 스플릿을 제거하기 위하여 초기 트리를 정리하였다.

### 결과

타액 상등액 560  $\mu$ l로부터 평균  $54.2 \pm 20.1$  ng(n=64)의 전체 RNA가 얻어졌다. OSCC와 연령 및 성별이 매치되는 대조군 간에 총 RNA 양에는 유의적인 차가 없었다(t 테스트,  $P=0.29$ , n=64). RT-PCR 결과는 모든 타액 샘플(n=64)이 세 가지 유전자(GAPDH, ACTB 및 RPS9)로부터의 전사물을 함유함을 입증하였으며, 이들은 인간 타액 RNA의 품질 대조군으로

사용되었다(실시에 1 참조). 2 라운드의 RNA 증폭후 지속적인 증폭 양(658±47.2, n=5)를 얻을 수 있었다. 비오티닐화된 cRNA의 양은 평균 39.3±6.0 µg(n=20)이었다. OSCC와 대조군 사이에 생산된 cRNA 양에는 유의적인 차이가 없었다(t 테스트, P=0.31, n=20).

HG U133A 마이크로어레이를 이용하여 암 환자와 매치되는 정상 대상 간의 타액 RNA 프로파일 차이를 확인하였다. 이전에 개시된 기준에 의해 포함된 10,316 전사물중에서, 0.05 미만의 P 값을 갖는 1,679개의 전사물이 확인되었다. 이들 전사물중에서, OSCC 그룹에서 836개는 상향 조절되고 843개는 하향 조절되었다. 이들 관찰된 전사물은 P<0.05를 사용한 양성 오류를 고려할 때 단독으로 기여하였다고는 볼 수 없다( $\chi^2$  테스트, P<0.0001). 10개의 모든 OSCC 타액 시료에서 3 배 이상의 조절 변화 및 0.01 미만의 P 값의 보다 엄중한 컷오프의 미리 정해진 기준을 이용하여, 표 3에 제시된 17개의 전사물을 확인하였다. 이들 17개의 타액 mRNA는 OSCC 타액에서 모두 상향 조절된 반면, 동일한 필터링 기준으로 하향 조절된 mRNA는 발견되지 않았다. 이들 유전자의 생물학적 기능이 표 3에 주어졌다.

**【표 3】**  
**마이크로어레이에 의해 확인된 OSCC에서 상향 조절된 타액 mRNA(>3배, P<0.01)**

유전자 기호	유전자 명칭	유전자은행 기탁 번호	위치	유전자 기능
<i>B2M</i>	$\beta$ -2-마이크로글로불린	NM_004048	15q21-q22.2	항아포프토시스; 항원 제시
<i>DUSP1</i>	이중특이성 포스파타제 1	NM_004417	5q34	단백질 변형; 신호 전달, 산화성 스트레스
<i>FTH1</i>	페리틴, 중 폴리펩티드 1	NM_002032	11q13	철이온 수송; 세포 증식
<i>GOS2</i>	추정 림프구 GO/G1 스위치 유전자	NM_015714	1q32.2-q41	세포 성장 및/또는 유지; 세포 사이클의 조절
<i>GADD45B</i>	성장 중지 및 DNA-손상-유도성 $\beta$	NM_015675	19p13.3	키나제 캐스케이드; 아포프토시스
<i>H3F3A</i>	H3 히스톤, 패밀리 3A	BE869922	1q41	DNA 결합 활성
<i>HSPC016</i>	가설 단백질 HSPC016	BG167522	3p21.31	미지
<i>IER3</i>	즉시 초기 반응 3	NM_003897	6p21.3	배발생; 형태발생; 아포프토시스; 세포 성장 및 유지
<i>IL1B</i>	인터루킨 1 $\beta$	M15330	2q14	시그널 전달; 증식; 염증; 아포프토시스
<i>IL8</i>	인터루킨 8	NM_000584	4q13-q21	혈관신생; 복제; 칼슘-매개 시그널링 경로; 세포 부착; 화학 주성; 세포 사이클 중지; 면역 반응
<i>MAP2K3</i>	미토젠-활성화된 단백질 키나제 키나제 3	AA780381	17q11.2	시그널 전달; 단백질 변형
<i>OAZ1</i>	오르니틴 데카복실라제 안티자임 1	D87914	19p13.3	폴리아민 생합성
<i>PRG1</i>	프로테오글리칸 1, 분비성 입자	NM_002727	10q22.1	프로테오글리칸
<i>RGS2</i>	G-단백질 시그널링 2의 조절자, 24 kDa	NM_002923	1q31	종양 발생; g-단백질 시그널 전달
<i>S100P</i>	S100 칼슘 결합 단백질 P	NM_005980	4p16	단백질 결합; 칼슘이온 결합
<i>SAT</i>	스퍼미딘/스퍼민 N1-아세틸트랜스퍼라제 EST, 매우 유사한 페리틴 경쇄	NM_002970 BG537190	Xp22.1	효소, 트랜스퍼라제 활성 철이온 항상성, 페리틴 복합체

인간 계놈 U133A 마이크로어레이를 이용하여 10명의 암환자 및 매치되는 10명의 정상 대상으로부터의 타액에서 RNA 발현 패턴 차이를 확인하였다. 모든 OSCC 타액 시료에서 3배 이상의 조절 변화 및 0.01 미만의 P 값의 컷오프의 기준을 이용하여, OSCC 타액에서 상당한 상향 조절을 나타내는 17개의 mRNA를 확인하였다.

정량적 PCR을 실시하여 32명의 OSCC 환자와 32명의 매치되는 대조군으로부터의 타액을 포함하는 확대된 샘플 크기에서 마이크로어레이 결과를 확인하였다. 타액 mRNA 바이오마커의 9개 후보인 *DUSP1*, *GADD45B*, *H3F3A*, *IL1B*, *IL8*, *OAZ1*, *RGS2*, *S100P* 및 *SAT*를 그들의 암 관련성에 기초하여 선택하였다(표 3). 표 4는 qPCR에 의해 결정된 OSCC 환자로부터의 타액에서 상기 9개의 후보의 정량적 변화를 제시한다. 결과로부터 9개의 후보 mRNA중 7개의 전사물(78%)인 *DUSP1*, *H3F3A*, *IL1B*, *IL8*, *OAZ1*, *S100P* 및 *SAT*가 OSCC 환자의 타액에서 유의적으로 증가되었음을 확인하였다(윌콕슨 순위합 검정,  $P < 0.05$ ). qPCR에 의해 *RGS2*( $P=0.149$ ) 및 *GADD45B*( $P=0.116$ )의 양에서는 통계적으로 유의적인 차이가 검출되지 않았다. 확인된 7개의 유전자는 증가량에 따라 다음과 같은 세 등급으로 분류될 수 있다: *IL8*을 포함한 고도로 상향조절되는 mRNA(24.3배); *H3F3A*(5.61배), *IL1B*(5.48) 및 *S100P*(4.88배)를 포함한 적절히 상향 조절되는 mRNA; 및 *DUSP1*(2.60배), *OAZ1*(2.82배) 및 *SAT*(2.98배)를 포함한 낮은 수준으로 상향 조절되는 mRNA. ROC 곡선 아래 면적, 역치값, 및 OSCC에 대한 7개의 잠재적 타액 mRNA 바이오마커 각각에 대한 해당 민감성 및 특이성에 대한 상세한 통계가 표 5에 기술되었다. 데이터는 OSCC의 존재를 예측하기 위해 7개의 잠재적인 바이오마커중에서 *IL8* mRNA의 성능이 최고임을 보여주었다. *IL8*을 위한 계산된 ROC 곡선 아래 면적은 0.85였다. 역치값  $3.19E-18$  mol/L로, 타액내 *IL8* mRNA는 정상으로부터 OSCC를 구별하는데 88%의 민감성 및 81%의 특이성을 제시한다.

[표 4]

타액에서 선택된 9개의 전사물에 대한 정량적 PCR 확인 (n = 64)<sup>a</sup>

유전자 기호	프라이머 서열 (5' → 3')	확인	P 값	평균 배수 증가
DUSP1	F: CCTACCAGTATTATTTCCCGAAG (서열 번호 13)	예	0.039	2.60
	R: TTGTGAAGGCAGACACCTACAC (서열 번호 14)			
H3F3A	F: AAAGCACCCAGGAAGCAAC (서열 번호 15)	예	0.011	5.61
	R: GCGAATCAGAAGTTCAGTGGAC (서열 번호 16)			
IL1B	F: GTGCTGAATGTGGACTCAATCC (서열 번호 17)	예	0.005	5.48
	R: ACCCTAAGGCAGGCAGTTG (서열 번호 18)			
IL8	F: GAGGGTTGTGGAGAAGTTTTTG (서열 번호 19)	예	0.000	24.3
	R: CTGGCATCTCACTGATTCTTG (서열 번호 20)			
OAZ1	F: AGAGAGAGTCTTCGGGAGAGG (서열 번호 21)	예	0.009	2.82
	R: AGATGAGCGAGTCTACGGTTC (서열 번호 22)			
S100P	F: GAGTTCATCGTTCGTGGCTG (서열 번호 23)	예	0.003	4.88
	R: CTCCAGGCATCATTGAGTCC (서열 번호 24)			
SAT	F: CCAGTGAAGAGGGTTGGAGAC (서열 번호 25)	예	0.005	2.98
	R: TGGAGGTTGTCATCTACAGCAG (서열 번호 26)			
GADD45B	F: TGATGAATGTGGACCCAGAC (서열 번호 27)	아니오	0.116	
	R: GAGCGTGAAGTGGATTGTC (서열 번호 28)			
RGS2	F: CCTGCCATAAAGACTGACCTTG (서열 번호 29)	아니오	0.149	
	R: GCTTCCTGATTCACTACCCAAC (서열 번호 30)			

qPCR을 실시하여 32명의 OSCC 환자와 32명의 매치되는 대조 대상으로부터의 타액을 포함하는 확대된 샘플 크기에서 마이크로어레이 결과를 확인하였다. 표 3에 제시된 17개의 후보로부터 9개의 잠재적인 타액 mRNA 바이오마커가 선택되었다. 이중 7개가 qPCR로 확인되었다( $P < 0.05$ ). 샘플은 OSCC 환자로 부터의 32개의 타액 및 매치되는 정상 대상으로 부터의 32개의 타액을 포함한다.

윌콕슨 순위합 검정:  $P < 0.05$ 이면, 확인됨(예);  $P \geq 0.05$ 이면, 확인안됨(아니오).

**[표 5]**

OSCC 관련 타액 mRNA 바이오마커의 수용체 작동 특성(ROC) 곡선 분석

바이오마커	ROC 곡선 아래 면적	역치/컷오프 (M)	민감성 (%)	특이성 (%)	선택 기준
DUSP1	0.65	8.35E-17	59	75	(34)
H3F3A	0.68	1.58E-15	53	81	(54)
IL1B	0.70	4.34E-16	63	72	(44)
IL8	0.85	3.19E-18	88	81	(55)
OAZ1	0.69	7.42E-17	100	38	(37)
S100P	0.71	2.11E-15	72	63	(40)
SAT	0.70	1.56E-15	81	56	(35)

qPCR 결과를 이용하여, ROC 곡선을 분석하여 각 바이오마커의 예측력을 평가하였다. 최대 해당 민감성 및 특이성을 제공하는 최적의 컷포인트를 결정하였다. ROC 곡선 아래의 가장 큰 면적을 갖는 바이오마커가 OSCC를 검출하기 위한 가장 강한 예측력을 갖는 것으로 확인되었다.

질병 식별을 위한 타액 mRNA의 유용성을 입증하기 위하여, 두개의 분류/예측 모델을 조사하였다. 7개의 확인된 바이오마커중 IL1B, OAZ1, SAT 및 IL8 4개에 기초하여 로지스틱 회귀 모델을 작성하였으며, 이들은 조합시 최상의 예측을 제공한다(표 6). 계수 값은 이들 4개 마커에 대하여 양성이었으며, 이는 타액에서 이들 농도의 동시적인 상승이 OSCC 대상으로부터 샘플이 얻어질 확률을 증가시켰음을 나타내는 것이다. 로지스틱 회귀 모델에 기초한 리브-윈-아웃 교차-확인 에러율은 19%이었다(12/64). 리브-윈-아웃 분석으로 생성된 모델(64개)중 하나를 제외한 모두는 표 6에 특정된 전체 데이터 모델에서 유의적인 것으로 밝혀진 4개 마커의 동일 세트를 이용하였다. 로지스틱 회귀 모델에 대해 ROC 곡선을 계산하였다. 50%의 컷오프 확률을 이용하여, 91%의 민감성 및 91%의 특이성을 얻었다. ROC 곡선 아래의 계산된 면적은 로지스틱 회귀 모델에 대해 0.95였다(도 3).

**[표 6]**

로지스틱 회귀 모델에 의해 선택된 OSCC에 대한 타액 mRNA 바이오마커

바이오마커	계수 값	표준 에러	P 값
절편	-4.79	1.51	0.001
IL1B	5.10E+19	2.68E+19	0.062
OAZ1	2.18E+20	1.08E+20	0.048
SAT	2.63E+19	1.10E+19	0.020
IL-8	1.36E+17	4.75E+16	0.006

조합하여 최상의 예측을 제공하는 7개의 확인된 바이오마커중 4개(IL1B, OAZ1, SAT 및 IL-8)에 대해 로지스틱 회귀 모델을 작성하였다. 계수 값은 이들 4개의 마커에 대해 양이었으며, 이는 타액에서 이들 농도의 동시적인 상승이 OSCC 대상으로부터 샘플이 얻어질 확률을 증가시켰음을 나타내는 것이다.

제 2 모델인 "분류 및 회귀 트리(CART) 모델"이 생성되었다(도 4). 적합한 CART 모델은 OSCC에 대한 예측 변수로 IL8, H3F3A 및 SAT의 타액 mRNA 농도를 이용하였다. 초기 스플릿으로 선택되었으며 3.14E-18 mol/L의 역치를 가진 IL8은 총 64개 샘플을 함유한 모 그룹으로부터 두 개의 자식 그룹을 생산하였다. 3.14E-18 mol/L 미만의 IL8 농도를 갖는 30개 샘플은 "정상-1" 그룹으로 할당된 반면, 3.14E-18 mol/L 이상의 IL8 농도를 갖는 34개는 "암-1" 그룹으로 할당되었다. "정상-1" 그룹은 추가로 1.13E-14 mol/L의 역치를 갖는 SAT에 의해 분할되었다. 얻어진 하위 그룹인 "정상-2"는 1.13E-14 mol/L 미만의 SAT 농도를 갖는 25개 샘플을 함유하였으며, "암-2"는 1.13E-14 mol/L 이상의 SAT 농도를 갖는 5개

샘플을 함유하였다. 유사하게, "암-1" 그룹은  $2.07E-16$  mol/L의 역치를 갖는 H3F3A에 의해 추가로 분할되었다. 얻어진 하위 그룹인 "암-3"은  $2.07E-16$  mol/L 이상의 H3F3A 농도를 갖는 27개 샘플을 함유하였으며, "정상-3" 그룹은  $2.07E-16$  mol/L 미만의 H3F3A 농도를 갖는 7개 샘플을 함유하였다. 결과적으로, 본 연구에 포함된 64개 타액 샘플은 CART 분석에 의해 "암" 그룹과 "정상" 그룹으로 분류되었다. "정상" 그룹은 "정상-2" 그룹으로부터의 샘플 및 "정상-3" 그룹으로부터의 샘플로 구성되었다. "정상" 그룹에 할당된 샘플은 총 32개이며, 29개는 정상 대상으로부터, 3개는 암 환자로부터 왔다. 따라서, OSCC 예측을 위하여 IL8, SAT 및 H3F3A의 조합을 이용한 경우, 총 민감성은 90.6%이다(29/32). "암" 그룹은 "암-2" 그룹 및 "암-3" 그룹으로부터의 샘플로 구성되었다. "암" 그룹에 할당된 샘플은 총 32개이며, 29개는 암 환자로부터, 3개는 정상 대상으로부터 왔다. 따라서, OSCC 예측을 위해 이들 세 개의 타액 mRNA 바이오마커의 조합을 이용한 경우, 총 특이성은 90.6%이다(29/32).

암-스크리닝 프로그램의 목적은 치료가 성공적일 만큼 초기 단계에 종양을 검출하는 데에 있다. 고도의 민감성 및 고도의 특이성의 조합 특성을 나타내는 스크리닝 수단이 요구된다. 또한, 스크리닝 수단은 광범한 응용성을 위해 충분히 비침습성이며 저렴하여야 한다. 암 개시 및 진행에 대한 우리의 기본적인 지식의 진보 및 생명공학의 중대한 발전으로 종양이 있는 기관으로부터 추출된 체액에서 종양 유전자 및 종양-억제 유전자 변형과 같은 종양 징후를 확인하는 것이 가능하게 되었다(Sidransky, D. 1997 *Science* 278:1054-1059). 본 실시예에 제시된 결과는 타액 트랜스크립톰 진단법이 암의 비침습성 진단, 예후 및 추적 시험의 개발에 적합한 수단임을 보여준다.

기존의 연구는 인간 DNA 바이오마커가 타액에서 동정되어 구강암 검출에 이용될 수 있음을 보여주었다(El-Naggar, A.K et al. 2001 *J Mol Diagn* 3:164-170; Liao, P.H. et al. 2000 *Oral Oncol* 36:272-276). 타액내 인간 mRNA의 존재는 해독 및 임상적 응용에 대한 진단 분석물의 레퍼토리를 확대시킨다. 그러나, RNA는 DNA 보다 불안정하며, RNase에 따른 붕괴에 매우 민감할 것으로 여겨진다. 또한, 타액에서 RNase 활성은 암 환자에서 상승하는 것으로 보고되었다(Kharchenko, S.V. & Shpakov, A.A. 1989 *Izv Akad Nauk SSSR Biol* 58-63). 따라서, 인간 mRNA는 타액에서 세포외적으로 생존할 수 없다는 것이 통념이다. 놀랍게도, 본 발명자들은 RT-PCR을 사용하여 타액에서 인간 mRNA를 일관적으로 검출하였으며, 이에 따라 타액에 기초한 발현 프로파일링의 기회를 갖게 되었다. 개시된 수집 및 처리 프로토콜을 이용하여, 대조군 RNA의 존재가 RT-PCR/qPCR에 의해 모든 타액(환자 및 대조군)에서 확인되었다. RNA 품질은 PCR, qPCR 및 마이크로어레이 어세이의 요구를 만족한다. 본 실시예에서, 발명자들은 새로이 수집한 구강액에 즉시 RNase 저해제를 첨가한 후, 초저온 상태(-80 °C)로 저장하였다.

본 발명으로 보고된 결과는 암 및 질병 진단 분야에 상당한 관심을 불러 모을 것이다. 이와 같은 관심은 암의 타액에 기초한 진단 및 스크리닝 시험이 간단하고 관심을 끄는 구상이라는 사실과 통상적인 암 진단 시험이 불완전하다는 사실로부터 생겨난 것이다. 구강암을 예로 이용한 경우, 실망스러운 생존율은 아마도 대부분 진단 지연때문일 것이다(Wildt, J. et al. 1995 *Clin Otolaryngol* 20:21-25). 대부분의 구강암은 그의 초기 단계에 무증상의 작은 병소로 발생하기 때문에, 임상적 또는 환자가 이상 조직을 알아차리는 경우에만 형식적인 진단 절차를 시작하게 된다(Epstein, J.B. et al. 2002 *J Can Dent Assoc* 68:617-621). 진행 암에 대한 현미경적 수준이 성공적인 시술을 위해 너무 늦은 경향이 있다(Fong, K.M. et al. 1999 in: In: S.S. HD and G. AF(eds.), Molecular Pathology of Early Cancer, pp. 13-26: IOS Press). 시간을 요하고 비용이 많이 들기 때문에, 암 스크리닝을 위해 영상 기술을 실행할 수 없는 경우도 있다. 작은 병소에 대한 무감각 때문에, 상기 기술은 전형적으로 확인을 위해 이용된다(Myers, L.L. & Wax, M.K. 1998 *J Otolaryngol* 27:342-347). 연구에 따라 톨루이딘 블루로 구강암 조직을 염색하여 양호한 양성 예측값을 얻을 수 있음이 입증되었다(Mashberg, A. & Samit, A. 1995 *CA Cancer J Clin* 45:328-351). 그러나, 이러한 기술을 응용하고 그의 결과를 해석하는 데는 예의 경험이 따라야 한다. 탈락세포 진단학은 구강암 검출에 덜 침습적인 방법일 수 있다(Rosin, M.P. et al. 1997 *Cancer Res* 57:5258-5260). 그러나, 박리된 암 세포는 종양 적재량과 연관되는 경향이 있어서, 최소 또는 초기 질병 상태에서는 검출율이 낮다. 본 연구에서 확인된 타액 mRNA 바이오마커는 OSCC 검출에 새로운 장을 제공할 것이다. 타액 트랜스크립톰 진단법은 충분한 예측력으로 비침습성 진단 수단에 대한 요구를 만족시켜 준다.

정상 개체에서, 타액 RNA 공급원은 하기 세 공급원중 하나로부터 일 것이다: 타액선(귀밑샘, 턱밑샘, 혀밑샘 및 작은 샘), 잇몸 열구액 및 구강 점막 세포(내층 또는 박리). 구강암 환자의 경우, 검출된 암-관련 RNA 특징은 세 가지 각 주요 공급원(타액선, 잇몸 틈새액 및 구강 점막 세포)으로부터 유래한 전체 타액에서 스스로를 추가 반영하는 매치된 종양 및/또는 전신적 반응(국소 또는 원위)에서 비롯될 것으로 보인다. 질병-관련 RNA는 잇몸 틈새액을 통한 순환 또는 타액선을 통해 구강으로 유입될 것으로 생각된다. 좋은 예는 유방암 환자의 타액에서 HER-2 단백질의 존재 증가이다(Streckfus, C. et al. 2000 *Clin Cancer Res* 6:2363-2370). 구강암의 경우, 국소 종양은 고도의 타액 mRNA의 공급원이다. 본 발명자들은 최근에 가장 유의적으로 상승된 구강암 조직 전사물인 IL8을 선택하고 그의 단백질 수준(ELISA에 의해)이 구강암 환자의 타액에서도 크게 상승하였음을 확인하였다(St. John, M.A.R. et al. 2004 IL-6 and IL-8: Potential Biomarkers for Oral Cavity and Oropharyngeal SCCA. *Archives of Otolaryngology-Head & Neck Surgery*, in press). 첸(Chen) 등은 이

미 두경부암 조직에서 IL8 단백질 발현이 증가하였음을 독자적으로 입증하였다(Chen, Z. et al. 1999 *Clin Cancer Res* 5:1369-1379). 이들 데이터를 조합하면 mRNA 및 단백질 수준에서 종양 조직 및 타액의 구강암 관련 발현 변화가 일치하여 변경된다는 것이 뒷받침된다.

IL8 외에, DUSP, H3F3A, OAZ1, SAT, S100P 및 IL-1B 등의 여섯개의 다른 암-관련 유전자가 구강암 환자로부터의 타액에서 상향조절되는 것으로 확인되었다. DUSP1 유전자는 이중 특이성 포스파타제를 코딩하며 종양 억제 PTEN 시그널링 경로의 매개자로서 연루된다(Unoki, M. & Nakamura, Y. 2001 *Oncogene* 20:4457-4465). DUSP1 발현은 난소 종양을 감소시키는 것으로 나타났으며, DUSP1 유전자내 새로운 단일-뉴클레오티드 다형체(SNP)가 확인되었다(Suzuki, C. et al. 2001 *J Hum Genet* 46:155-157). H3F3A mRNA는 통상 증식 마커로 이용되고 있으며, 그의 수준은 전립선암 및 결장암에서 상향조절되는 것으로 나타났다(Bettuzzi, S. et al. 2000 *Cancer Res* 60:28-34; Torelli, G. et al. 1987 *Cancer Res* 47:5266-5269). OAZ1은 오르니틴 데카복실라제(ODC)의 공지된 억제 작용에 기초해 종양 억제자로 추정된다(Tsuji, T. et al. 2001 *Oncogene* 20:24-33). 그러나, OAZ1 mRNA는 전립선암에서 상향조절되는 것으로 보고되었다(Bettuzzi, S. et al. 2000 *Cancer Res* 60:28-34). 흥미롭게도, 폴리아민 대사에 관여하기도 하는 SAT의 발현은 전립선암에서 상당히 높아지는 것으로 밝혀졌다(Bettuzzi, S et al. 2000 *Cancer Res* 60:28-34). S100P는 전립선암 진행과 관련이 있는 것으로 공지되었으며, 그의 과발현은 시험관내에서 인간 유방 상피 세포의 면역화 및 생체내에서 초기 유방암 발생 단계에 연루된다(Gribenko, A. et al. 1998 *Protein Sci* 7:211-215; Guerreiro Da Silva, I.D. et al. 2000 *Int J Oncol* 16:231-240; Mousses, S. et al. 2002 *Cancer Res* 62:1256-1260; Mackay, A. et al. 2003 *Oncogene* 22:2680-2688). 최근 연구는 S100P의 차등 발현이 췌장 암종과 관련이 있음을 보여주었다(Logsdon, C.D. et al. 2003 *Cancer Res* 63:2649-2657; Crnogorac-Jurcevic, T. et al. 2003 *J Pathol* 201:63-74). IL-1B의 발현은 또한 암과 관련되기도 한다. IL-1B의 혈청 수준은 구강암의 편평세포암종을 가진 환자에서 고 수준인 것으로 밝혀졌다(Jablonska, E. et al. 1997 *Pathol Oncol Res* 3:126-129). 또한, IL-1B의 수준은 난소암에 걸린 여성의 복수액에서 상당히 증가하는 것으로 보고되었다(Chen, C.K. et al. 1999 *J Formos Med Assoc* 98:24-30). IL-1B의 유전학적 다형성은 위암 및 유방암과 같은 질병에 걸릴 위험 가능성을 가지는 것으로 보고되었다(Hamajima, N. & Yuasa, H. 2003 *Nippon Koshu Eisei Zasshi* 50:194-207; El-Omar, E.M. et al. 2003 *Gastroenterology* 124:1193-1201).

타액은 HIV(Malamud, D. 1997 *Am J Med* 102:9-14), 당뇨병(Guven, Y. et al. 1996 *J Clin Periodontal* 23:879-881) 및 유방암(Streckfus, C. et al. 2000 *Clin Cancer Res* 6:2363-2370)과 같은 전신적 질환을 진단하는데 조사 보조제로서 그 사용이 증가하고 있다. 가장 주목할만한 것으로, 본 실시예에 적용된 다중 바이오마커의 개념, 기술 및 접근은 다른 질환들을 스크린 및 모니터링하기 위해 용이하게 변경이 가능하다. 구강암에 있어서, 타액 트랜스크립톰 진단법의 가장 중요한 응용중 하나는 구강 전암성 병소의 암 변환을 검출하는 것이다. 전체 악성 변환율은 11 내지 70.3% 범위이다(Lee, J.J. et al. 2000 *Clin Cancer Res* 6:1702-1710; Silverman, S., Jr. & Gorsky, M. 1997 *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 84:154-157). 구강 백색관증 세포의 DNA 함량 분석은 구강암의 위험을 예측하는데 유용한 것으로 입증되었다(Sudbo, J. et al. 2001 *N Engl J Med* 344:1270-1278). 그러나, 이는 여전히 생검후 방법이다. 본 발명자들은 "타액 트랜스크립톰 진단법"이 구강암 및 기타 인간 질병의 초기 진단에 새로운 장을 제공할 것이라 생각한다.

### 실시예 3

#### 타액 RNA에 대한 실용적인 실온 저장 프로토콜

진단용 타액 RNA를 최적으로 보존하기 위한 실용적이면서 사용하기에 친숙한 실온 프로토콜이 개발되었다. 본 발명의 구체에는 타액 트랜스크립톰 진단법에 가장 적합한 품질 및 양의 타액 RNA를 제공한다.

인간 mRNA의 검출 및 정량화를 RNeasy<sup>TM</sup>-처리된 타액에서 실시하였다. 타액을 1 또는 2 부피(들)의 RNeasy<sup>TM</sup>(라인 1 또는 2)과 혼합하였다. 140  $\mu$ l 타액 상등액으로부터 모든 RNA를 Qiagen 키트를 사용하여 분리하였다. 분리된 RNA 분취물을 DNase I(Ambion)로 처리하였다. RT-PCR을 이용하여 베타-액틴(ACTB), 글리세르알데히드-3-포스페이트 데하이드로게나제(GAPDH) 및 인터류킨 8(IL-8)의 세 유전자로부터의 전사물을 검출하였다(도 5A). Ribogreen<sup>®</sup> 키트(Molecular Peobes)를 이용한 RNA의 정량화는 Superase-In(Ambion)으로 처리된 샘플이 아닌 RNeasy<sup>TM</sup>으로 처리된 샘플에서 RNA 수율이 더 높다는 것을 보여준다(도 5B). 1 부피의 RNeasy<sup>TM</sup>(L1) 또는 2 부피의 RNeasy<sup>TM</sup>(L2)은 각각 Superase-In(S) 보다 10배까지 및 3.3배 이상까지 RNA를 제공하였다. 이들 데이터는 한명의 동일 개체로부터 상이한 시점에 수집한 샘플 및 다섯명의 상이한 개체에서 동일한 시점에 수집한 샘플에서 재현되었다.



타액 GAPDH 및 IL-8을 정량하기 위하여 정량적 PCR(qPCR)을 수행하였다. 타액 샘플을 RNAlater™(1:1 비) 또는 Superase-In으로 처리한 분취물로 분할하였다. 처리하지 않은(비처리) 타액을 대조군으로 사용하였다. 샘플을 실온에서 24 시간동안 유지한 후, 4 °C에서 저장하였다. 140 µl 타액 상등액으로부터 모든 RNA를 5일 연속 분리하였다. RT-qPCR을 1일에서 5일까지 행하여 GAPDH 및 IL-8에 의해 코딩된 cDNA/RNA를 정량하였다. 도 6에 제시된 데이터는 RNAlater™이 타액 RNA 보존성에 더 좋은 보호 효과를 가짐을 보여준다. 용어 "RNAlater™"은 엠비온사(Ambion, Inc.)의 등록상표이다(USP 6,528,641 및 USP 6,204,375).

인간 타액 mRNA를 2.0 어레이의 HG U133(Affymetrix)을 사용하여 프로파일링하였다. 표 7에 있는 수치는 MAS 5.0 및 Dchip 1.3에 의해 할당된 존재하는 mRNA 수를 나타낸다.

**[표 7]**

**마이크로어레이상에 존재하는 mRNA의 수**

	<b>RNAlater™</b>	<b>Superase In</b>
<b>MAS 5.0</b>	<b>5,566</b>	<b>2,868</b>
<b>Dchip 1.3</b>	<b>10,202</b>	<b>7,566</b>

데이터는 "Superase-In"-처리 샘플이 아닌 RNAlater™-처리 샘플에서 보다 많은 mRNA가 회수됨을 보여준다.

본 발명에 따른 본 구체에는 타액내 RNA의 보존을 원하는(예: 소아과 의사, 주치의, 치과 의사, 기타 다른 건강 관리 요원의 연구실, 공중 진료소, 가정 치료 키트) 임의 세팅에 사용될 것으로 판단된다. 이어서, 보존된 RNA는 실시예 1 및 2에 개시된 특이적 RNA에 기초한 스크리닝 또는 진단을 위해 진단 센터로 보내진다. 본 발명자들은 RNAlater™-타입 RNase 저해 조성물과 조합된, 예를 들어 미국 특허 제 6,652,481호; 6,022,326호; 5,393,496호; 5,910,122호; 5,376,337호; 4,019,255호; 및 4,768,238호에 개시된 타액 수집용 키트를 계획하였다.

이제까지 본 발명이 상세히 기술되었지만, 당업자들은 본 발명의 영역 또는 그의 임의적인 구체에 영향을 주지 않으면서 광범하고 등가적인 범위의 조건, 제제 및 기타 파라미터를 이용하여 본 발명이 수행될 수 있음을 알 것이다. 본원에 인용된 모든 특허 및 문헌은 그의 전체내용이 본원에 참고로 포함된다.

[서열목록]



1788821\_1.TXT

SEQUENCE LISTING

<110> Regents of the University of California, The  
Wong, David

<120> SALIVARY TRANSCRIPTOME DIAGNOSTICS

<130> UCLA020.001VPC

<150> US 60/589,627

<151> 2004-07-21

<160> 30

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic primer

<400> 1

tcaccagggc tgcttttaac tc 22

<210> 2

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic primer

<400> 2

atgacaagct tcccgttctc ag 22

<210> 3

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic primer

<400> 3

aggatgcaga aggagatcac tg 22

<210> 4

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic primer

<400> 4

atactcctgc ttgctgatcc ac 22

<210> 5

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

1788821\_1.TXT

<223> synthetic primer

<400> 5  
gacccttcga gaaatctcgt ctc 23

<210> 6  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic primer

<400> 6  
tctcatcaag cgtcagcagt tc 22

<210> 7  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic primer

<400> 7  
gtgctgaatg tggactcaat cc 22

<210> 8  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic primer

<400> 8  
accctaaggc aggcagttg 19

<210> 9  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic primer

<400> 9  
cctgcgaaga gcgaaacctg 20

<210> 10  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic primer

<400> 10  
tcaataactgg acagcaccct cc 22

<210> 11  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic primer

1788821\_1.TXT

<400> 11  
agcgtgcctt tgttcactg 19

<210> 12  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic primer

<400> 12  
cacaccaacc tcctcataat cc 22

<210> 13  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic primer

<400> 13  
cctaccagta ttattcccga cg 22

<210> 14  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic primer

<400> 14  
ttgtgaaggc agacacctac ac 22

<210> 15  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic primer

<400> 15  
aaagcaccca ggaagcaac 19

<210> 16  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic primer

<400> 16  
gcgaatcaga agttcagtgg ac 22

<210> 17  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic primer

<400> 17  
gtgctgaatg tggactcaat cc 22

1788821\_1.TXT

<210> 18  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> synthetic primer  
  
 <400> 18  
 accctaaggc aggcagttg 19  
  
 <210> 19  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> synthetic primer  
  
 <400> 19  
 gaggggttg gagaagttt tg 22  
  
 <210> 20  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> synthetic primer  
  
 <400> 20  
 ctggcatctt cactgattct tg 22  
  
 <210> 21  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> synthetic primer  
  
 <400> 21  
 agagagagtc ttcgggagag g 21  
  
 <210> 22  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> synthetic primer  
  
 <400> 22  
 agatgagcga gtctacggtt c 21  
  
 <210> 23  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> synthetic primer  
  
 <400> 23  
 gaggatcatcg tggtcgtggc tg 22  
  
 <210> 24

1788821\_1.TXT

```

<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic primer

<400> 24
ctccagggca tcatttgagt cc 22

<210> 25
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic primer

<400> 25
ccagtgaaga gggttggaga c 21

<210> 26
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic primer

<400> 26
tggaggttgt catctacagc ag 22

<210> 27
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic primer

<400> 27
tgatgaatgt ggaccagac 20

<210> 28
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic primer

<400> 28
gagcgtgaag tggatttc 19

<210> 29
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> synthetic primer

<400> 29
cctgccataa agactgacct tg 22

<210> 30
<211> 22
<212> DNA

```

1788821\_1.TXT

<213> Artificial Sequence

<220>  
<223> synthetic primer

<400> 30  
gcttcctgat tcactacca ac

22

### 도면의 간단한 설명

도 1은 RT-PCR을 사용하여 무세포 타액에서 유전자 특이적 RNA를 검출한 결과를 나타낸다. (A) 1개월, 3개월 및 6개월 (각각 레인 2, 3 및 4)간 저장 후, 액틴-β(ACTB)에 대한 RT-PCR 형결정(typing)으로 시험된 타액내 RNA 안정성. 레인 1, 분자량 마커(100 bp 래더(ladder)); 레인 5, 음성 대조군(템플레이트 생략). (B) 글리세르알데히드-3-포스페이트 데하

이드로게나제(GAPDH), 리보솜 단백질 S9(RPS9) 및 ACTB가 10개의 모든 케이스에서 일관적으로 검출되었다. 레인 1, 2 및 3은 각각 타액 RNA, 양성 대조군(인간 전체 RNA, BD Biosciences Clontech, Palo Alto, CA, USA) 및 음성 대조군(템플레이트 생략)을 나타낸다.

도 2는 마이크로어레이 연구를 위한 무세포 타액으로부터 분리된 RNA의 증폭 결과를 보여준다. (A) 아가로스 겔 전기영동에 의한 RNA 증폭 모니터링. 레인 1 내지 5는 각각 1kb DNA 래더, RNA 분리후 5  $\mu$ l 타액(미검출), 2 라운드로 증폭된 cRNA 1  $\mu$ l(200 bp 내지 ~ 4kb 범위), 단편화(약 100 bp) 후 cRNA 및 Ambion RNA Century 마커를 나타낸다. (B) ACTB는 타액 RNA 증폭동안 매 주요 단계마다 검출될 수 있다. 아가로스 겔은 PCR 산물의 예측되는 단일 밴드(153 bp)를 나타낸다. 레인 1 내지 8은 각각 100 bp DNA 래더, 무세포 타액에서 분리된 전체 RNA, 첫번째 라운드 cDNA, RT후 첫번째 라운드 cRNA, 두번째 라운드 cDNA, RT후 두번째 라운드 cRNA, 양성 대조군(인간 전체 RNA, BD Biosciences Clontech, Palo Alto, CA, USA) 및 음성 대조군(템플레이트 생략)을 나타낸다. (C) 마이크로어레이에서 하이브리드화 전 Agilent 2100 바이오어널라이저(bioanalyzer)로 분석한 표적 cRNA. 좁은 범위(50-200 bp)에서 하나의 단일 피크만이 검출되었으며, 이는 산물이 고순도임을 입증한다.

도 3은 조합된 타액 mRNA 바이오마커의 예측력에 대한 수용체 작동 특성(Receiver Operating Characteristic)(ROC) 곡선 분석을 나타낸다. 최종 로지스틱 모델(logistic model)은 다음과 같은 네 개의 타액 mRNA 바이오마커를 포함한다: 인터류킨 1 $\beta$ (IL1B), 오르니틴 데카복실라제 안티자임 1(OAZ1), 스퍼디민/스퍼민 N1-아세틸트랜스퍼라제(SAT) 및 인터류킨 8(IL-8). 50%의 컷-오프(cut-off) 확률을 이용하여 ROC에 의해 91%의 민감성 및 91%의 특이성을 얻었다. ROC 곡선 아래 계산된 면적은 0.95 이었다.

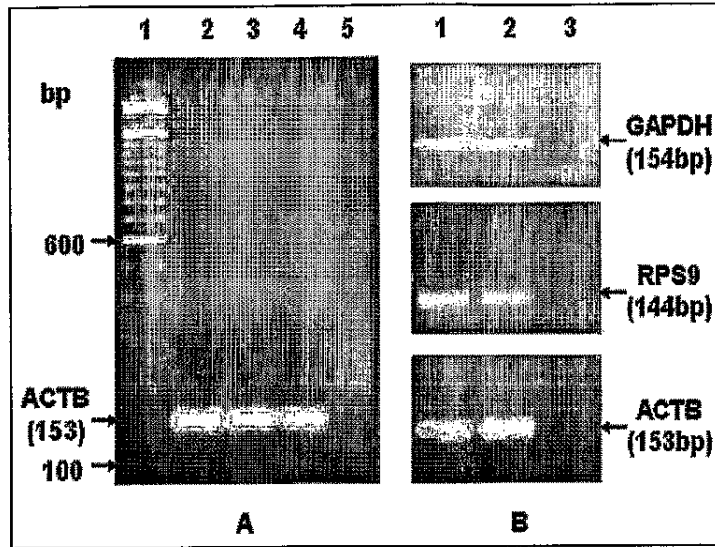
도 4는 구강 편평세포암종(OSCC)을 위한 타액 mRNA 예측자를 평가하기 위한 분류 및 회귀 트리(CART) 모델을 나타낸다. 초기 스플릿으로 선택된 IL-8(컷오프 값 = 3.14E-18)은 총 64개의 샘플을 함유하는 모 그룹으로부터 두개의 자식 그룹을 생산한다. 정상-1 그룹을 SAT(컷오프 값 = 1.13E-14)에 의해 추가로 분할하고, 암-1 그룹을 H3F3A(컷오프 값 = 2.07E-16)에 의해 추가로 분할하였다. 본 연구에 포함된 64개 샘플은 CART에 의해 최종 암 또는 정상 그룹으로 분류되었다. OSCC 분류에 대해 총 민감도는 90.6%(29/32, 정상 그룹)이고, 특이성은 90.6%(29/32, 암 그룹)이다.

도 5는 RNeasy<sup>TM</sup>-처리된 타액에서 인간 mRNA의 검출 및 정량 결과를 나타낸다. (A) RT-PCR을 이용하여 세 유전자, 즉 베타-액틴(ACTB), 글리세르알데히드-3-포스페이트 데하이드로게나제(GAPDH) 및 인터류킨 8(IL-8)로부터 전사체를 검출하였다. (B) Ribogreen<sup>®</sup> 키트(Molecular Probes)를 이용한 RNA 정량은 Superase-In(Ambion)으로 처리된 샘플이 아닌 RNeasy<sup>TM</sup>-처리된 샘플에서 RNA 수율이 더 높다는 것을 보여준다.

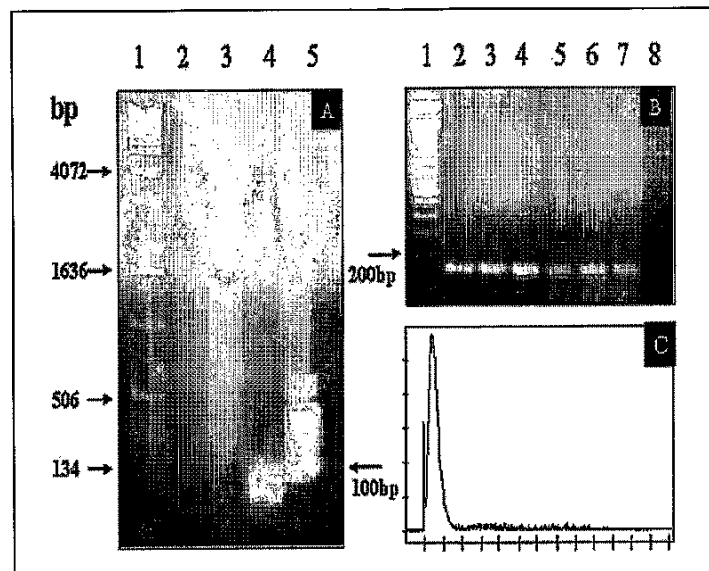
도 6은 타액 GAPDH 및 IL-8을 정량화하기 위한 정량적인 PCR(qPCR)을 나타낸다.

## 도면

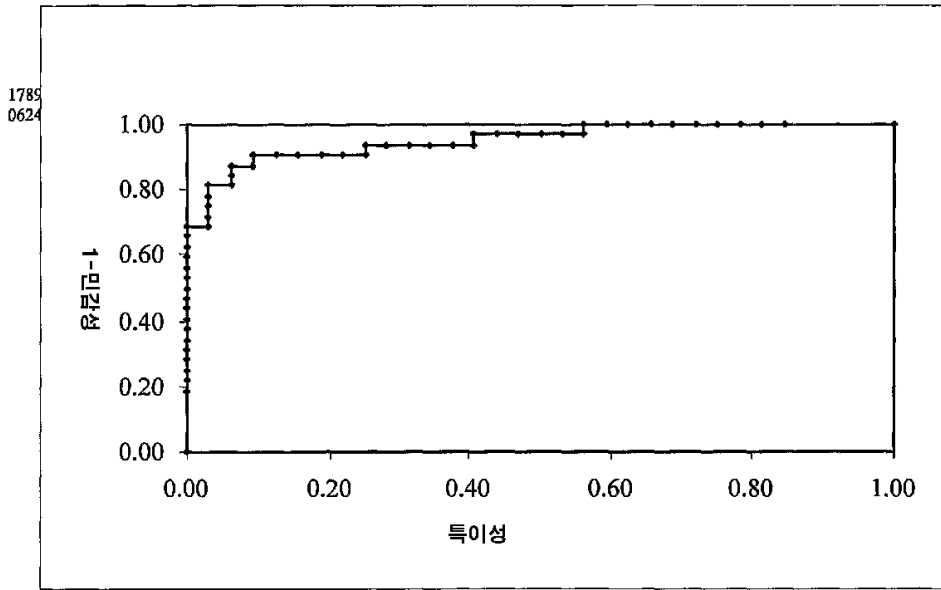
도면1



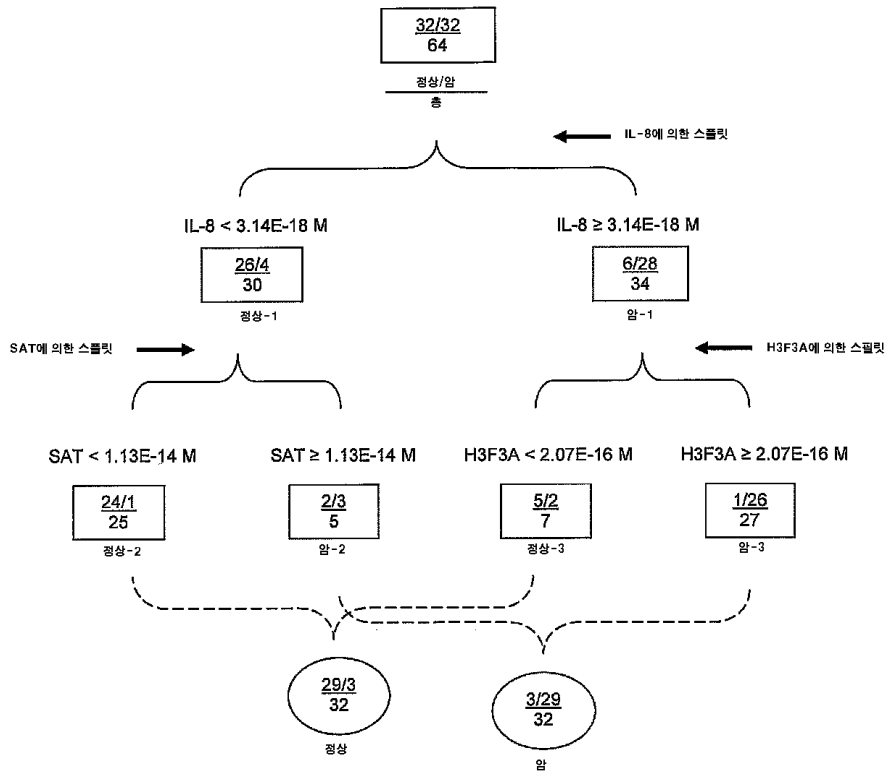
도면2



도면3

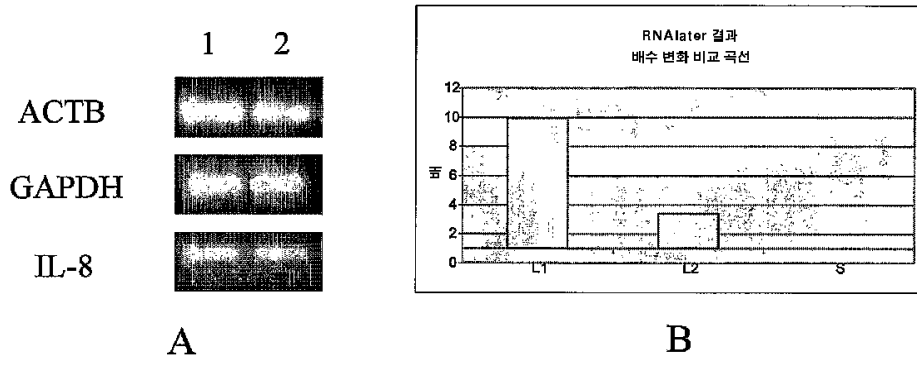


도면4





도면5



도면6

