

부분적으로 절편된 조직의 RNA를 이  
용한 위장관 암의 유전자  
발현 양상

연세대학교 대학원  
의과학과  
양 연 주

부분적으로 절편된 조직의 RNA를 이  
용한 위장관 암의 유전자  
발현 양상

지도교수 라 선 영

이 논문을 석사 학위논문으로 제출함

2005년 월 일

연세대학교 대학원

의과학과

양 연 주

양연주의 석사 학위논문을 인준함

심사의원\_\_\_\_\_인

심사의원\_\_\_\_\_인

심사의원\_\_\_\_\_인

연세대학교 대학원

2005년    월    일

## 감사의 글

대학원 생활을 시작한게 얼마전인 것 같은데 벌써 졸업을 눈앞에 앞두고 그동안 저에게 큰 힘이 되어주셨던 분들께 감사한 마음을 전하고자 합니다.

먼저 제가 자신감을 잃어버리고 앞이 막막할 때마다 북돋아 주시고 넓은 마음으로 감싸주신 라선영 교수님, 항상 인자하신 미소와 언제나 열심히 하시는 모습으로 모범이 되어주신 정현철 교수님, 바쁘신 와중에도 많은 관심을 보내주시고 격려해 주신 허만옥 교수님께 정말 깊은 감사를 드립니다.

그리고 대학원 공부가 어떤 것인지를 가르쳐 주시고 이끌어주신 양상화 선생님, 다방면에 박식하시고 실험실 모든 사람들을 편하게 대해주신 정희철 선생님, 언제나 뒤에서 학생들에게 용기를 주시고 응원해 주시는 박규현 선생님, 하나라도 더 챙겨주시려고 바쁘신 김태수 선생님, 실험실의 모든 머리 아픈 일을 도맡아 하시면서도 언제나 즐거운 표정의 강진석 선생님, 암전이 연구센터의 살림꾼이자 부탁하면 뭐든지 들어주시는 강언니와 보라...

처음 면접 보러 온 날부터 지금까지 정말 친 오빠처럼 챙겨준 우렁오빠, 항상 남에게 많은 것을 베풀어 주는 유근오빠, 똘똘이 스머프 같은 영이오빠, 때론 터프하고 때론 유머러스하며 패션리더인 정범오빠, 실험실에 들어왔을 때부터 짝궁이었고 실험의 기본을 가르쳐 주었던 옥이언니, 나와같이 석사를 시작했음에도 불구하고 배울것이 많았던 연호언니, 실험실 들어와서 더 많이 건강해지고 지금은 후배들까지 꼼꼼히 챙기는 우선이, 나이에 맞지 않게 의젓하고 항상 남을 배려하는 원주, 언제나 열심히 하려고 하고 귀여운 센터 막내 은정이, 언제나 차분하고 얌전하지만 가끔은 엉뚱한 큰 눈이 너무나도 부러운 명은이 정말 너무나 저에게 큰 힘이 되어 주었습니다.

연구동의 영원한 정신적 지주 태문오빠, 내가 힘들 때마다 격려해주고 더욱 열심히 할 수 있게 도와주고 나의 모든 것을 돌봐주느라 너무나도 힘들었을 지혜언니, 늘 열심히 공부하고 장난끼로 실험실 사람들에게 웃음을 주었던 민영언니, 후배들에게 하나라도 더 가르쳐 주려고 노력 많이 하는 세원오빠, 얼마 같이 지내진 않았지만 미소가 따뜻한 재희오빠, 실험실에 있는 듯 없는 듯 조용히 지내시지만 우리에게 많은 지식(?)을 주시는 면희 선생님 (나가기 전에 꼭 면희 오빠라고 불러 보고 싶어요), 겉으로는 무뎠는데 보이지만 젤루 맘씨 따뜻한 하진오빠 (원언니랑 오래 오래 행복하세요~~), 울 실험실 반장으로 고생 정말 많이 하고 있는 찬희오빠 (오빠 정말 카트라이더 캐릭터하고 닮았어요^^), 정상에 서는 그 날까지 항상 열심히 공부할 상철오빠 (정말 울 오빠 친구 맞아요?? ㅋㅋ), 성시경의 열렬한 팬이자 항상 소녀 같은 주혜언니, 외모는 모범생이지만 이 시대의 자유인 승희오빠 (머리 모양 넘 이뻐요!), 들어올 때만 해도 새색시 였는데 벌써 한 아이의 엄마가 된 선정언니, 언제나 자신감에 넘치는 동혁오빠와 자기 자신을 너무나도 아끼고 사랑하는 귀연이, 언제나 의젓해 보이려 하지만 아직도 애 같은 소심주의자 쓰리 제이, 언니로써 정말 힘들 때 도움이 못되 늘 미안한 경나미, 연구동의 귀염둥이 지은이 (들어오기 전에 각오로 잘 해내리라 믿어~히힛).

학부생 때부터 많은 가르침과 실험실 생활을 할 수 있는 기회를 주신 김창한 교수님, 늘 뒤에서 지켜봐 주시는 양종범 교수님, 인자하신 미소로 큰 가르침을 주신 심상국 교수님, 학생들에게 편안함과 웃음을 주시는 고명수 교수님, 언제나 소녀 같으신 이경혜 교수님 (자주 찾아뵙지 못해서 너무 죄송해요)께 깊은 감사의 마음을 전합니다.

결혼하고 나서 더욱 더 아름다워진 혜란언니, 늘 환한 미소로 웃으며 긍정적으로 생활하는 영미언니(언니 웃는게 젤로 이뻐요~), 차분한 이미지의 소유자이자 진용언니, 11월에 새 신랑되는 멋쟁이 기욱오빠, 영국에서 열심히 공부하고 있는 기태오빠, 짐 너무나도 행복한 신혼생활을 하고 있을 현

진, 항상 새로운 것을 추구하고 가꾸는 만인의 연인 주희가 있어 대학생활이 너무나도 즐거웠습니다. (지금도 마찬가지예요)

그 누구보다도 소중한 저를 언제나 믿고 지켜봐 주시며 사랑을 주신 사랑하는 엄마, 아빠, 어린 나이에 힘들게 생활하면서 누나를 도와준 내 동생 용권이, 내가 주말에 학교 올 때면 늘 옆에 있어주고 힘들 때 가장 큰 힘이 되어준 울오빠씨 (빈씨), 나의 부족함을 채워주고 나를 믿어준 나의 평생 친구이자 정신적 지주인 young... 녀 녀 감사합니다.

앞으로 힘든 날도 많겠지만 지금까지 해왔던 것처럼 곳곳이 해나가겠습니다. 그래도 저는 이렇게 많이 감사드릴 분이 있어 행복한 아이인 것 같습니다.

2005년 늦가을

양 연주 드림

# 목 차

국문요약	1
I. 서론	5
II. 재료 및 방법	8
1. 세포주와 임상 시료 수집	8
2. RNA 추출	8
3. RNA amplification	10
4. cDNA microarray 수행	11
5. Data analysis	13
III. 결과	18
1. 세포주에서 intact와 degraded RNA를 이용한 유전자 분석	18
가. 절편된 RNA 준비와 세포주 RNA 증폭	18
나. 양질의 RNA와 절편된 RNA를 이용한 cDNA microarray 결과 비교	20
2. 부분적으로 절편된 조직 RNA를 이용한 cDNA microarray의 수행	22

가. 부분적으로 절편된 RNA의 quality 확인과 RNA amplification.....	22
나. 부분적으로 절편된 조직 RNA를 이용한 cDNA microarray의 수행.....	24
(1) cDNA microarray 결과 전처리.....	24
(2) Unsupervised hierarchical clustering 을 이용한 부분 절편된 조직 RNA의 발현 양상 관찰.....	28
(3) 양질의 RNA와 부분 절편된 조직 RNA의 유전자발현 양상 비교.....	30
(4) 부분 절편된 조직의 RNA에서 절편된 정도에 따른 grouping.....	35
(5) 대장암에서 절편된 세 그룹과 양질의 RNA에서 SAM에 의해 선정된 유전자들의 비교.....	37
(6) 절편된 대장암 조직 RNA에서 선정된 유전자군의 유용도.....	40
IV. 고찰.....	45
V. 결론.....	50
참고문헌.....	51
영문요약.....	56



## 그림 차례

**Figure 1. Study scheme**

**Figure 2. Electropherograms of distribution of intact and degraded RNA with YCC-16**

**Figure 3. Electropherograms of distribution of intact and partially degraded tissue RNA**

**Figure 4. Representative scanned image of cDNA Microarray**

**Figure 5. M-A plot before and after normalization**

**Figure 6. Box plot before and after normalization**

**Figure 7.**

**A. Two-way hierarchical clustering of all experiments using partially degraded tissue RNA**

**B. Two-way hierarchical clustering using selected significant genes between normal and tumor**

**Figure 8. Unsupervised hierarchical clustering using intact and partially degraded RNA microarray data**

**Figure 9. Grouping of partially degraded colon tissue RNA according to degraded states**

**Figure 10. Venn diagram with selected significant genes between normal and tumor in colon intact RNA and three groups of partially degraded colon tissue RNA**

**Figure 11.**

**A. Two-way hierarchical clustering with selected significant 112 genes using Intact RNA microarray data in partially degraded colon tissue RNA microarray data**

**B. Two-way hierarchical clustering with selected significant 142 genes using partially degraded RNA microarray data in intact colon tissue RNA microarray data**

**Figure 12.**

**A. Two-way hierarchical clustering with selected significant**

**1320 genes using Intact RNA microarray data in partially  
degraded colon tissue RNA microarray data**

**B. Two-way hierarchical clustering with selected significant**

**1320 genes using partially degraded RNA microarray  
data in intact colon tissue RNA microarray data**

## 표 차례

**Table 1. List of top 20 selected significant genes using intact RNA in colorectal cancer tissues**

**Table 2. List of top 20 selected significant genes using degraded RNA in colorectal cancer tissues**

**Table 3. List of commonly selected top 20 significant genes between intact and mildly degraded A-group RNA in colon cancer tissues**

**Table 4. List of commonly selected top 20 genes between intact and degraded RNA in colon cancer tissues**

국문요약

부분적으로 절편된 조직의 RNA를 이용한

위장관 암의 유전자 발현 양상

<지도교수 라 선 영>

연세대학교 대학원 의과학과

**양 연 주**

cDNA microarray 기법은 암 연구 분야에서 여러 종류의 종양을 진단하는데 있어서 생물학적 표지자를 동정하거나 항암치료의 효과를 예측하는데 널리 적용되고 있다. 하지만 유용하게 사용되고 있는 cDNA microarray의 경우 많은 양의 RNA나 양질의 RNA를 필요로 하는 제한점을 가지고 있다. RNA amplification은 RNA의 양의 제한점을 극복하기 위해 고안되었으며, 현 시점에서 실험에 사용될 수 없어 절편된 RNA를 이용할 수 있는 방법이 필요로 되고 있다. 따라서, 본 연구에서는 절편된 RNA를 이용한 cDNA microarray 실험의 가능성과 그 유용성을 알아보려고 하였다.

YCC-16 위암 세포주의 양질의 RNA에 RNase와 NaOH를 처리하여 절편된 RNA를 만들어 RNA 증폭이 가능함을 확인하였고, 7.5K cDNA microarray를 이용하여 RNA 증폭을 기반으로 한 cDNA microarray를

수행하여 분석 가능한 결과를 얻을 수 있었다. 양질의 RNA와 절편된 RNA를 비교하기 위하여 상관계수와 discordance를 구하여 비교한 결과, 양질의 RNA와 절편된 RNA를 이용한 실험의 결과가 유사함을 확인하였다.

실제 임상시료 중 절편된 대장암과 위암의 정상 및 종양조직을 각각 10 예씩을 가지고 17K cDNA microarray를 이용하여 RNA 증폭을 기반으로 한 cDNA microarray를 수행하였다. 그 결과 소량의 ribosomal peak이 남아있을 정도로 절편된 경우 microarray 실험을 통하여 유전자 발현 양상을 관찰할 수 있음을 확인하였고, unsupervised clustering을 통해서 절편된 RNA로도 위 조직과 대장 조직 및 정상과 종양조직이 구분되는 것을 관찰하였다.

조직을 이용한 microarray 실험에서 절편된 RNA와 양질의 RNA를 비교하고자 대장암과 위암에서 정상과 종양조직, 각각 절편된 RNA 10 예, 양질의 RNA 10 예씩의 microarray 결과를 비교하였다. Unsupervised clustering 결과, 양질의 RNA와 절편된 RNA의 유전자 발현 양상이 다르다는 것을 알 수 있었다. 또한 양질의 RNA에서 정상과 종양조직간의 유의한 유전자로 선정된 169 개 (FDR 0.39%)의 유전자와 절편된 RNA에서 선정된 114 개 (FDR 0.36%)의 유전자 중 12 개의 유전자 만이 겹침으로서, 양질의 RNA와 절편된 RNA의 발현 양상이 차이가 있다는 것을 확인할 수 있었다.

RNA의 절편된 정도가 유전자 발현 양상이 달라지는 데에 영향을 미치는지 확인하기 위하여 대장의 절편된 RNA를 깨진 정도에 따라 세 그룹으로 나누어 각 그룹에서 정상과 종양조직간에 유의한 유전자를 선정하였다. 선정된 유전자 군을 비교하였을 때 (A, B, C 각 84, 98, 98 개), 절편 정도가 심하지 않은 A 그룹에서 선별된 유전자 군과 절편 정도가 심한 B나 C그룹에서 선별된 유전자 군은 전혀 일치하지 않았다. 양질의 RNA에서 정상과 종양조직 간에 유의한 것으로 선정된 유전자 (FDR 1 % 142 개)와 비교한 결과, A그룹에서 선정된 유전자 중 38개가 일치하였고, B나 C그룹에서 선정된 유전자와는 일치하지 않음을 관찰하여, RNA의 절편된 정도가 결과에 영향을 미치는 것으로 사료된다.

한편 대장암에서 절편된 RNA에서 정상과 종양조직 간에 유의한 유전자로 선정된 유전자 군을 이용하여 양질의 RNA microarray data로 clustering을 시행한 결과 정상과 종양조직이 구분되는 것을 확인하였다. 또한 이는 양질의 RNA에서 선정된 유전자를 가지고 절편된 RNA 결과를 이용하였을 때도 같은 결과를 얻을 수 있었다. 즉, 절편된 RNA를 이용하였을 경우 양질의 RNA를 이용한 실험에서와는 다른 종양 관련 유전자 군이 선정되나, 그 유전자 군도 생물학적으로 유용함을 제시하고 있다.

본 연구의 결과를 통하여 부분적으로 절편된 RNA를 이용하여 유

전자 발현 양상의 관찰이 가능하다는 것은 보여주었고, 적절한 생물학적 의미를 위해서는 좀 더 주의 깊은 분석과 검증이 필요할 것으로 생각된다.

---

핵심되는 말: cDNA 마이크로어레이, RNA 증폭, 부분적으로 절편된 조직 RNA



## 부분적으로 절편된 조직의 RNA를 이용한

### 위장관 암의 유전자 발현 양상

<지도교수 라선영>

연세대학교 대학원 의과학과

양 연 주

#### I. 서론

최근 cDNA microarray 기법은 암 연구 분야에서 여러 종류의 종양을 진단하는데 있어서 생물학적 표지자를 동정하거나 항암치료의 효과를 예측하는데 널리 적용되고 있다<sup>1,2</sup>. cDNA microarray 기법의 주요한 제한점의 하나는 표준 방법상 최소한 1~2  $\mu\text{g}$ 의 mRNA 혹은 40~50  $\mu\text{g}$ 의 total RNA를 필요로 하는 것이다. 하지만 needle biopsy나 body fluid와 같은 대부분의 임상시료는 세포의 수가 적어 얻는 RNA의 양도 극히 적다. 1990년대 이후 이러한 양의 제한점을 극복하기 위해 고안된 T7 linear RNA amplification은 *in vitro*에서 RNA로부터 cDNA 합성 과정인 reverse transcription시 사용되

는 oligo-dT primer 대신 T7 promotor sequence를 포함하는 T7 oligo-dT primer를 사용하는 기법이다<sup>3,4</sup>. T7을 기반으로 한 RNA amplification 기법이 도입된 이후 여러 가지로 수정된 방법들이 확립되었으며, 여러 연구들을 통해 amplified mRNA를 이용해 얻은 cDNA microarray의 결과가 믿을 수 있고, 재현성 있음을 보여주었다<sup>5-14</sup>.

cDNA microarray 기법을 이용하는데 있어서 RNA 양의 제한점 외에도 또 다른 제한점은 좋은 양질의 RNA를 필요로 한다는 것이다. 임상시료는 환자의 질병상태와 특성을 나타내는 중요한 정보를 제공하므로 암 연구에서는 많은 실험들이 임상시료를 이용하여 수행되고 있다. 하지만 임상 시료는 세포주에 비교하여 적은 양 뿐만 아니라 불안정한 RNA quality로 microarray를 수행하는데 더 많은 제한점을 가지고 있다. 임상시료 사용시 불안정한 RNA 발생이 빈번한 이유로 다음과 같은 4 가지 이유를 들 수 있다. 첫째로, 조직 시료는 외과적 절개로부터 적절한 보관이 진행되는 동안에 hypoxia로부터 세포가 손상될 위험 요소를 가지고 있다. 둘째, 췌장 조직과 같은 몇몇 기관의 조직은 그 조직의 특성상 다른 조직에 비해 RNA가 쉽게 깨지는 경향이 있다. 셋째, RNA 손상은 오랜 기간 동안 보관하거나 적절하지 못한 조건하에서 보관하는 것으로 인하여 발생할 수 있다. 넷째, RNA 손상의 위험은 laser microdissection과

같이 RNA 추출까지 진행되는 단계가 많은 조직의 시료와 관련이 높아진다. 그러므로, 유전자 발현을 분석하는데 있어서 가능한 RNA가 깨지는 것을 감소시키려고 많은 노력이 시도되고 있지만 여전히 많은 정보의 손실이 있으므로 부분적으로 깨진 RNA도 이용할 수 있는 방법이 요구되고 있다<sup>15-17</sup>.

본 연구에서는 절편된 RNA를 이용하여 RNA증폭을 통한 cDNA microarray를 이용하여 유전자 발현 관찰이 가능한지를 확인하고자 하였다. 그리하여 조직의 절편된 RNA로부터 얻은 microarray data에서 정상과 종양조직 간에 유의한 유전자를 선택하여 양질의 RNA를 이용하여 얻은 microarray data와 비교함으로써 유용도를 평가하고자 하였다.

## II. 재료 및 방법

### 1. 세포주와 임상 시료 수집

2000년부터 2003년까지 연세대학교 의과대학 신촌 세브란스 병원, 연세 암센터에 내원하여 진단 당시 대장암 또는 위암으로 진단 받은 환자들을 대상으로 하였다. 이들로부터 수술 당시 대장의 종양 병소와 주위의 정상부위를 채취하였고, 실험 시까지  $-80^{\circ}\text{C}$ 에서 보관하였다. 모든 연구는 연세대학교 세브란스 병원 임상연구 심의 위원회 (Institutional Review Board)가 승인한 동의서에 자의로 동의한 환자를 대상으로 진행하였다.

세포주는 연세의대 암전이 연구 센터 (Cancer Metastasis Research Center, Yonsei University College of Medicine, Seoul, Korea)에서 한국인 진행성 위암 환자의 말초 혈액으로부터 확립된 YCC-16 위암 세포주를 사용하였다<sup>18</sup>. 세포주는 10 % fetal bovine serum, 100U/ml의 penicillin 그리고 0.1mg/ml의 streptomycin이 포함된 MEM 배지(GIBCO, Grand, NY, USA)에  $37^{\circ}\text{C}$ , 5%  $\text{CO}_2$ 에서 배양하였다.

### 2. RNA 추출

한 종류의 세포주로 부터 intact RNA와 degraded RNA를 얻기 위하

여 YCC-16 세포주에 RNase A ( $5 \times 10^{-5}$ U)를 상온에서 5 분간 반응시켰다. 또한 0.1 N NaOH 1  $\mu$ l를 65 °C에서 5 분간 처리하여 가수분해 시키고, 0.1 N HCl 1  $\mu$ l로 중화시켜 절편된 RNA를 확보하였다.

임상시료와 세포주로부터 total RNA추출은 TRIzol (Invitrogen, Carlsbad, CA, USA)을 이용하여 제조업자의 protocol에 따라 시행하였다. 조직을 액체질소에서 파쇄한 후, 1 ml의 TRIzol에 용해하였고 200  $\mu$ l의 chloroform을 넣고 4 °C에서 10 분간 반응 시킨 후, 4 °C에서 14,000 rpm으로 25 분간 원심분리를 하여 상층액에 RNA를 분리 시켰다. RNA를 Isopropanol과 -20 °C에서 30 분간 반응시킨 후, 4 °C에서 14,000 rpm으로 25 분간 원심분리를 하여 침전시켰다. 침전된 RNA는 70 % 에틸 알코올로 수세하고, 상온에서 남아있는 알코올을 건조시킨 후, ultra-pure water에 녹였다. 추출된 total RNA의 양과 질은 NanoDrop (NanoDrop Technologies, Rockland, USA), 1 % agarose gel electrophoresis와 Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA)을 이용하여 확인하였다.

대조군으로 사용할 Yonsei reference RNA (YRR)는 11 개의 human cancer cell line (YCC 3, YCC B1, HCT-116, SK-Hep-1, A549, HL-60, MOLT-4, HeLa, HT-1080, Caki-2, T98G)으로부터 추출한 같은 양의 total RNA를 섞어 준비하였다<sup>16</sup>.

### 3. RNA amplification

RNA의 amplification은 Eberwine과 coworker의 group에 의해 보고된 방법을 수정하여 만든 암전이 연구센터 protocol에 따라 수행하였다<sup>3,4</sup>.

먼저 일차 RNA 증폭 과정으로, 4  $\mu\text{g}$ 의 total RNA를 2  $\mu\text{g}$ 의 T7 oligo-dT primer (5' -GGC CAG TGA ATT GTA ATA CGA CTC ACT ATA GGG AGG CGG TTT TTT TTT TTT TTT TTT TTT TTT-3' )와 함께 혼합한 뒤, 9  $\mu\text{l}$ 의 RNase-free water에 녹였다. RNA/primer 혼합물은 65  $^{\circ}\text{C}$ 에서 10 분간 denature한 뒤, 4  $^{\circ}\text{C}$ 에서 5 분간 냉각하였다. First strand cDNA를 합성하기 위해서 4  $\mu\text{l}$ 의 5 X first strand buffer (Invitrogen, Carlsbad, CA, USA), 2  $\mu\text{l}$ 의 100 mM DTT, 2  $\mu\text{l}$ 의 10 mM dNTP mix (Invitrogen, Carlsbad, CA, USA), 1  $\mu\text{l}$ 의 RNasin (Promega, Madison, WI, USA) 그리고 2  $\mu\text{l}$ 의 SuperScript II (Invitrogen, Carlsbad, CA, USA)의 혼합물을 RNA/primer에 가하여 42  $^{\circ}\text{C}$ 에서 1 시간 동안 역 전사 반응을 시켰다. 그 후, second cDNA를 합성하기 위하여 91  $\mu\text{l}$ 의 RNase-free water, 30  $\mu\text{l}$ 의 5 X second strand buffer, 3  $\mu\text{l}$ 의 10 mM dNTP mix, 10 U의 DNA ligase, 4 U의 DNA polymerase I 그리고 2 U의 RNase H의 혼합물을, 생성된 first strand cDNA에 가하여 42  $^{\circ}\text{C}$ 에서 2 시간 동안 반응 시켰다. 반응은 10  $\mu\text{l}$ 의 1 N NaOH와 10  $\mu\text{l}$ 의 0.5 M EDTA를 가한

후 65 °C에서 10 분간 반응시킴으로써 중단 하였고, 25  $\mu$ l의 1 M Tris-HCl (pH 7.5)를 가하여 중화 시켰다. Double strand cDNA는 97  $\mu$ l의 phenol: chloroform: isoamyl alcohol (25:24:1)을 이용하여 생성하였고 Phase Lock Gel™ (Eppendorf, Westbury, USA)로 정제하였다. 정제된 double strand cDNA에 T7 MEGAscript™ (Ambion, Texas, USA)로 *in vitro* transcription을 수행하였다. 이는 8  $\mu$ l의 double strand cDNA에 2  $\mu$ l씩의 75 mM NTP, enzyme mix 그리고 10 X reaction buffer를 혼합하여 37 °C에서 5 시간 동안 반응시킴으로써 진행하였다. 생성된 mRNA는 RNeasy mini kit (Qiagen, Valencia, CA, USA)를 이용하여 정제하였으며, amplified mRNA의 양과 질은 NanoDrop (NanoDrop Technologies, Rockland, USA), 1 % agarose gel electrophoresis와 Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA)을 이용하여 확인하였다.

#### 4. cDNA microarray 수행

cDNA microarray는 임상 시료를 이용한 실험에는 17 K human cDNA chip (CMRC-GT, Seoul, Korea)을 이용하였고, 세포주를 이용한 실험에는 7.5 K human cDNA chip (CMRC, Yonsei University College of Medicine, Seoul, Korea)을 이용하였다. cDNA microarray 실험은 reference design으로 수행되어 test sample

RNA로는 조직과 세포주 RNA를 Cy5로 표지하였고, 각각의 sample을 Cy3로 표지된 YRR (CMRC, Seoul, Korea)과 co-hybridization 하였다. cDNA microarray 실험은 본 연세의대 암전이 연구센터 (CMRC)에서 확립된 방법<sup>18</sup>을 이용하여 수행되었으며, 세포주를 이용한 실험은 세 번씩 반복 수행하였다.

2  $\mu$ g의 amplified mRNA는 reverse transcription하는 동안 Cy3-또는 Cy5-dUTP로 표지하였고, mRNA는 6  $\mu$ g의 random primer와 65 °C에서 10 분간 incubation 하였다. 8  $\mu$ l의 5 X first strand buffer, 4  $\mu$ l의 100 mM DTT, 2  $\mu$ l의 SuperScript II RT, 2  $\mu$ l의 20 X low-dT/dNTP mix 그리고 1  $\mu$ l의 RNAsin을 RNA/random primer 혼합물에 가하여 42 °C에서 1 시간 동안 반응시켰다. 잔존하는 RNA는 5  $\mu$ l의 0.5 M EDTA와 10  $\mu$ l의 1 N NaOH로 65 °C에서 15 분간 반응시켜 가수분해 한 뒤, 25  $\mu$ l의 1 M Tris-HCl (pH7.5)를 가하여 중화시켰다.

형광 물질 표지 과정과 동시에 3.5 X SSC, 0.1 % SDS, 10 mg/ml Bovine serum albumin (BSA; AMRESCO, Ohio, USA)와 물로 구성된 blocking 용액을 MFS-25 (MFS Inc, Pleasanton, CA, USA)로 여과시킨 후, microarray를 42 °C에서 1 시간 pre-hybridization 시켰다. Pre-hybridization된 슬라이드는 물과 isopropanol에 담금질 후 500 rpm 에서 5 분 동안 원심 분리하여 건조시켰다. Cy3-dUTP와



Cy5-dUTP로 표지된 probe는 QIAquick PCR Purification Kit (Qiagen, Valencia, CA, USA)를 이용하여 정제한 뒤, 20  $\mu$ g의 Human Cot-1 DNA (GIBCO-BRL, Gaithersburg), 20  $\mu$ g의 yeast tRNA (Invitrogen, Carlsbad, CA, USA) 그리고 20  $\mu$ g의 poly(A) RNA (Sigma, Saint Louis, MO, USA)와 혼합하였다. Hybridization 용액은 Microcon-30 filter (Amicon, Bedford, MA, USA)를 이용하여 50  $\mu$ l로 농축시키고, 20 X SSC 25  $\mu$ l, formamide 25  $\mu$ l 그리고 0.1 % SDS 1  $\mu$ l로 최종 100  $\mu$ l로 맞추었다. 그 후, hybridization 용액은 100  $^{\circ}$ C에서 2 분간 denature 시킨 후, 13000 rpm에서 2 분 동안 원심 분리하여 슬라이드에 올리고, 42  $^{\circ}$ C에서 16 시간 동안 반응시켰다. 0.5 X SSC-0.01 % SDS, 0.06 X SSC-0.01 % SDS 와 0.05 X SSC 용액에 순서대로 수세하고 1000 rpm에서 5 분간 원심 분리하면서 건조시켰다. Hybridization이 끝난 microarray는 GenePix 4000B 스캐너 (Axon Instruments, Foster City, CA, USA)로 이미지를 읽고, GenePix Pro4.1 software (Axon Instruments, Foster City, CA, USA)로 수치화된 결과를 얻었다.

## 5. Data analysis

본 연구의 study scheme을 Figure 1에 정리하였다.

먼저 세포주 RNA를 이용하여 얻은 microarray 결과를 가지고 양

양질의 RNA와 절편된 RNA를 이용한 결과의 유사성을 비교하기 위하여 상관계수와 discordance rate를 구하여 확인하였다. 절편된 조직의 RNA에 적용해 보고자 절편된 대장암과 위암의 microarray 실험 결과를 이용하여 유전자 발현 양상은 unsupervised clustering으로 관찰하고 정상과 종양조직간의 유의한 유전자를 선정하여 비교하였다. 절편된 RNA를 이용한 microarray 결과를 양질의 RNA와 비교하기 위하여 대장과 위암 각각에서 유전자 발현 양상과 유의한 유전자를 선정하였다. 절편된 정도가 양질의 RNA를 이용하여 얻은 결과와 왜 다른지를 알아보려고 대장암에서 얻은 절편된 RNA를 절편 정도에 따라 세 군으로 나누고, 각 군에서 선정된 유의한 유전자들을 비교하였다. 또한 절편된 RNA의 microarray결과에서 선정된 유의한 유전자 군을 가지고 양질의 RNA를 이용하여 얻은 microarray data set에서 이들 유전자가 정상과 종양조직을 구분해 줄 수 있는지를 관찰하였고, 유전자 군과 data set을 바꾸어 반대로도 시행하였다.

임상시료를 이용한 실험에서는 절편된 RNA를 이용한 microarray 결과와 비교하기 위해 연세의대 암전이 연구센터에서 다른 project에 이용되었던 양질의 RNA를 사용한 위암 및 대장암 각각 정상조직 10개, 종양 조직 10개의 microarray결과를 사용하였다. 양질의 RNA를 이용한 실험은 total RNA를 이용하였으며 사용된 실험

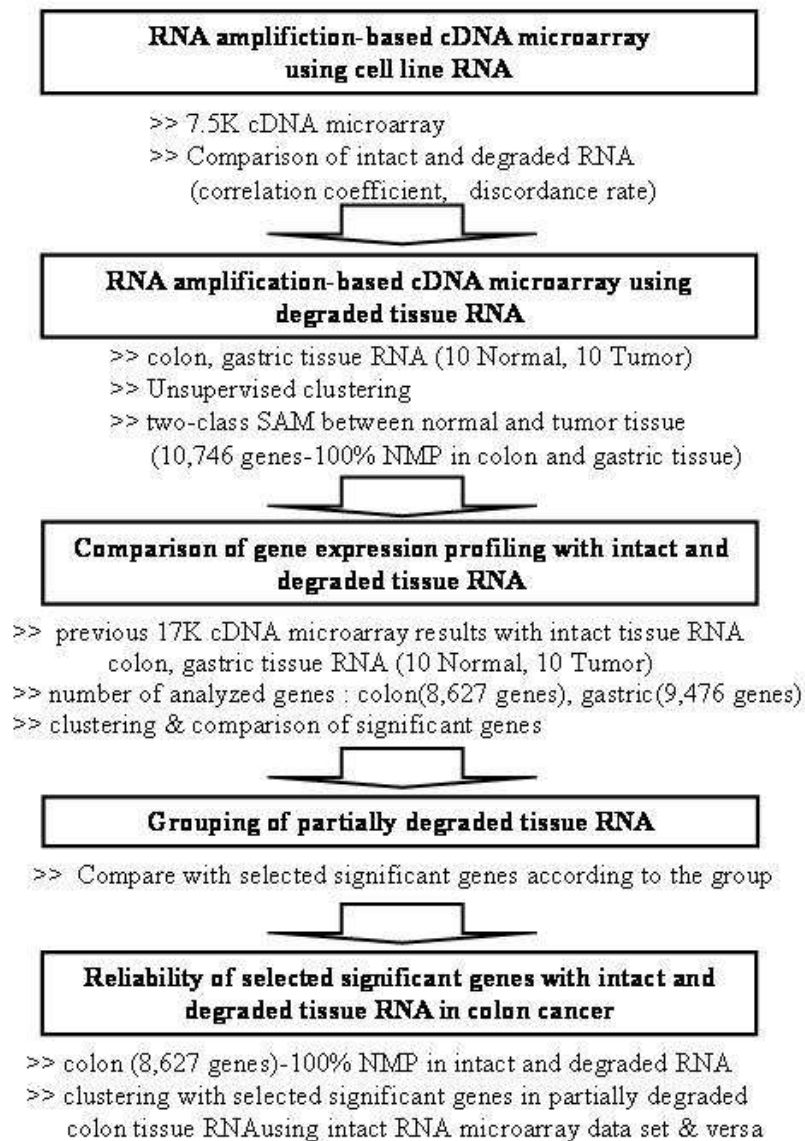
design과 cDNA chip는 동일한 조건으로 수행된 것이다.

절편된 RNA를 이용하여 얻은 microarray 결과는 실험간의 variation을 줄이기 위하여 intensity-dependent within pin-tip group Lowess normalization<sup>19,20</sup>방법을 이용하여 표준화 하였다. 표준화 작업 후, 본 연구에서는 결측치가 하나도 없는 유전자(Non-missing proportion 100 %)만을 선별하여 앞으로의 분석에 이용하였다. 유전자들의 발현 경향을 관찰하기 위하여 hierarchical clustering (Eisen software, Stanford Univ.)<sup>21</sup>을 시행하여 시각화하기 위해 TreeView software (Eisen software, Stanford Univ.)<sup>21</sup>를 사용하였다.

세포주를 이용한 세 번의 반복실험에는 재현성이 있는지를 확인하기 위하여 S-Plus 2000 software (Insightful, Seattle, WA, USA)를 이용하여 상관계수를 측정하였다. 양질의 RNA와 절편된 RNA간의 microarray data의 유사성을 비교하기 위하여 두 가지 종류의 discordance rate로 분석 하였다<sup>22</sup>. “quality discordance”는 비교하고자 하는 두 가지 실험에서 유전자의 signal이 present vs. absent로 변화를 나타내는 유전자로 정의되며, “quantity discordance”는 두 가지 실험에서 실험군의 유전자 발현이 대조군의 유전자 발현에 비하여  $\log_2$  ratio의 값이 positive vs. negative의 차이를 나타내는 유전자로 정의하였다. Discordance rate는 다

른 조건간의 차이를 비교하기에 앞서, 세 번 반복 실험간의 discordance rate를 구하여 실험상에서 유전자 발현의 변화가 생길 수 있는 유전자들을 제거한 뒤, 조건이 다른 실험상의 discordance rate를 구하였다.

비교하고자 하는 두 그룹간의 유의한 유전자를 선정하기 위하여, t-test 알고리즘을 바탕으로 분산의 영향을 최소화하고 두 그룹 이상을 비교할 수 있도록 개발된 Significant analysis of Microarray (SAM)<sup>23</sup> 이라는 통계학적인 program을 이용하였다. 정상 조직과 종양조직, intact RNA와 degraded RNA와 같은 두 그룹간의 유의한 유전자를 선정하기 위하여 2-class SAM을 시행하였다.



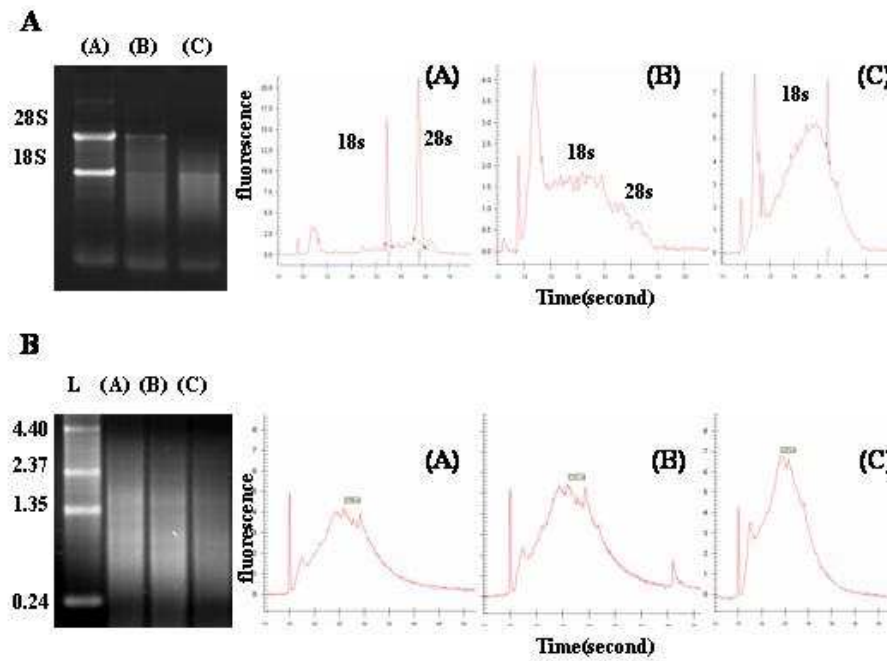
**Figure 1. Study scheme**

### III. 결과

#### 1. 세포주에서 intact와 degraded RNA를 이용한 유전자 발현 분석

##### 가. 절편된 RNA 준비와 세포주 RNA 증폭

YCC-16 세포주의 양질의 RNA로 부터 RNase와 NaOH를 이용하여 *in vitro* 상에서 인위적으로 절편된 total RNA의 quality를 electrophoresis와 Bioanalyzer를 이용하여 확인하였다(Figure 2A). 양질의 RNA에 비하여 RNase에 의해 절편된 RNA (Rd-RNA)는 소량의 28s와 18s만 남아있었고, NaOH에 절편된 RNA (Nd-RNA)는 18s만 남아있는 것을 확인하였다. 그 후, 세가지 조건의 total RNA를 이용하여 RNA amplification을 수행하였고, amplified mRNA의 quality는 electrophoresis와 Bioanalyzer를 이용하여 확인하였다 (Figure 2B). Amplified mRNA의 패턴은 양질의 RNA와 절편된 RNA간에 유사한 것으로 나타났으며, 특히 Rd-RNA가 electrophoresis에서 1.7 Kb 부근에 peak가 있어 양질의 RNA와 좀 더 유사하였다. 세 가지 조건의 RNA를 이용하여 얻은 amplification yield는 양질의 RNA의 경우 평균  $31.63 \pm 2.69$  ug, Rd-RNA의 경우 평균  $29.41 \pm 6.33$  ug, Nd-RNA의 평균이  $38.48 \pm 3.38$  ug으로 양질의 RNA와 절편된 RNA간에 큰 차이를 나타내지 않았다.



**Figure 2. Electropherograms of distribution of intact and degraded RNA with YCC-16.** Analysis was performed using agarose gel electrophoresis and the RNA 6000 Nano LabChip Kit on a bioanalyzer. A displays a gel image and Bioanalyzer images in total RNA with intact (A) and degraded (B, C) RNA. Also B displays a gel image and Bioanalyzer images in amplified mRNA with intact (A) and degraded (B, C) RNA.

나. 양질의 RNA와 절편된 RNA를 이용한 cDNA microarray 결과 비교

세포주를 이용한 cDNA microarray에서는 RNA amplification 과정의 재현성을 평가하기 위하여 세 번의 cDNA microarray는 독립적으로 수행된 amplification product로 수행되었다. 세 가지 조건으로부터 얻은 microarray data sets의 피어슨 상관계수를 array상의 모든 gene의 expression ratio를 이용하여 측정하였다. 양질의 RNA를 이용하여 amplification한 후 triplicate간의 상관계수는  $0.917 \pm 0.02$  였고, Rd-RNA를 이용하여 amplification한 후 triplicate간의 상관계수는  $0.917 \pm 0.02$ 이었으며, Nd-RNA를 이용하여 amplification한 것의 triplicate간의 상관계수는  $0.9 \pm 0.04$  이었다. 그리고 양질의 RNA그룹과 Rd-RNA그룹을 비교하였을 때  $0.927 \pm 0.03$ , intact한 RNA그룹과 Nd-RNA그룹을 비교하였을 때  $0.857 \pm 0.03$ , Rd-RNA그룹과 Nd-RNA그룹을 비교하였을 때를 비교하였을 때  $0.862 \pm 0.03$  이었다.

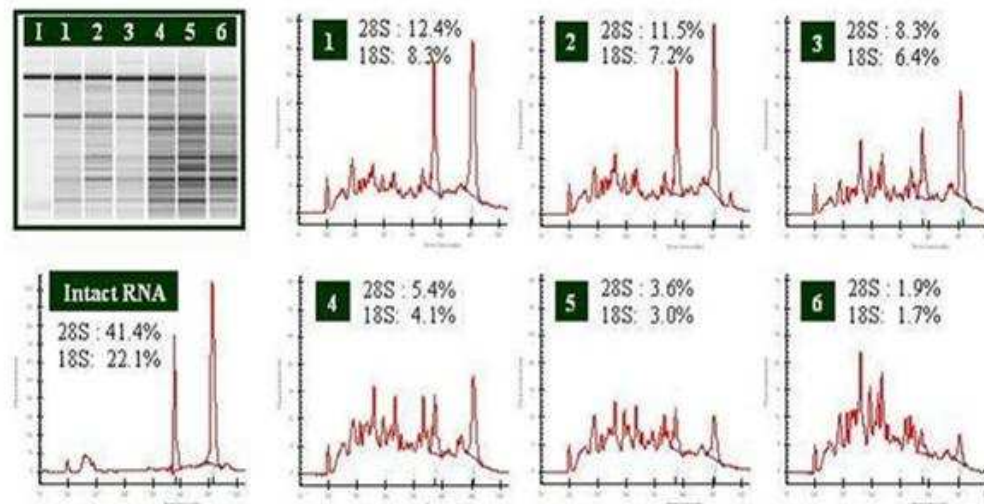
양질의 RNA와 절편된 RNA를 이용하여 amplification했을 때, 유사한 microarray data를 제공할 수 있는지를 평가하기 위해 세 가지 조건으로부터 discordance를 분석하였다. 7,680 개의 모든 spot으로 양질의 RNA, Rd-RNA 그리고 Nd-RNA를 이용한 triplicate간에 각각 4.4 %, 8.5 %, 12.6 %가 present 와 absent간의 변화가 있는



것으로 나타났으며, 양질의 RNA 그룹과 Rd-RNA 그룹간에는 5.6 %, 양질의 RNA 그룹과 Nd-RNA 그룹간에는 7.9 %의 변화가 있는 것으로 나타났다. Quantity discordance rate는 9 번의 실험에서 100 % NMP processing 후 5,002 개의 spot을 가지고 log ratio의 변화가 positive와 negative간의 변화를 나타내는 것을 조사하였다. 3 번의 반복 실험에 따른 intact한 조건간의 변화는 17.4 %, RNase에 의한 조건에서는 16.5 %, NaOH에 의한 조건에서는 16.4 %로 나타났다. 한가지 조건에서 나타날 수 있는 technical variation을 감안한 뒤 다른 두 조건끼리 비교해 봤을 때, 양질의 RNA와 Rd-RNA간의 discordance rate는 15.4 %였고, 양질의 RNA와 Nd-RNA간에는 21.6 %의 discordance rate를 나타내었다. 또한 양질의 RNA와 절편된 RNA 간의 quality discordance rate는 fold change가 증가함에 따라 거의 0 %로 감소하면서, 유의한 차이를 나타내는 유전자를 선택함에 있어서 양질의 RNA와 절편된 RNA간의 차이가 거의 없음을 보여 주었다.

## 2. 부분적으로 절편된 조직 RNA를 이용한 cDNA microarray의 수행 가. 부분적으로 절편된 RNA의 quality 확인과 RNA amplification

부분 절편 되어 표준 방법으로 이용할 수 없는 위암과 대장암의 환자로부터 얻은 각각 10 개의 정상 및 종양조직의 total RNA의 quality를 electrophoresis와 Bioanalyzer를 이용하여 확인하였다 (Figure 3). Ribosomal peak의 전체 비율 중 28s와 18s의 proportion을 Bioanalyzer를 통하여 조사한 결과 intact RNA의 28s와 18s의 비율은 각각 40 %와 18 %로 나타났으며, partially degraded RNA는 28s와 18s가 많게는 17 % 정도에서 적게는 Bioanalyzer에 detection되지 않을 정도로 존재하는 것까지 다양한 양상을 나타내었다. 이들 40 개의 부분 절편된 RNA를 이용하여 RNA amplification을 수행한 결과, 40 개의 모든 샘플은 다양한 size로 절편 되어 있었으나 증폭된 mRNA의 size는 0.2~4 Kb부근에 분포하며, amplification fold는  $10^3$  이상을 나타내었다.

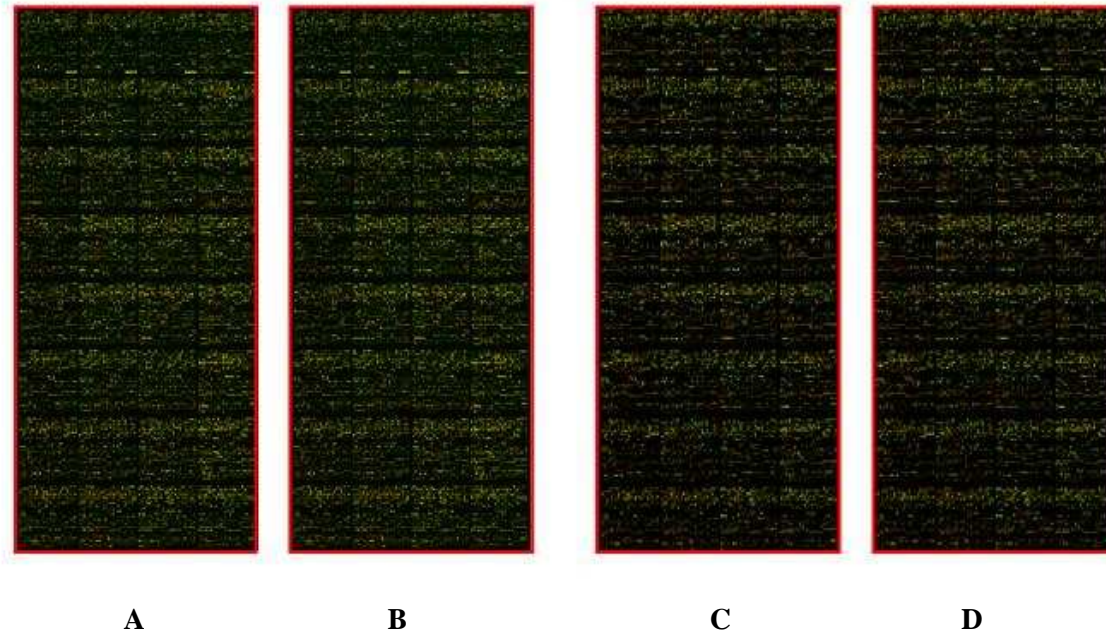


**Figure 3. Electropherograms of distribution of intact and partially degraded tissue RNA.** Analysis was performed using the RNA 6000 Nano LabChip Kit on a Bioanalyzer. 1~6 are electropherograms of partially degraded tissue RNAs from gastric or colon cancer patients.

나. 부분적으로 절편된 조직 RNA를 이용한 cDNA microarray의 수행

(1) cDNA microarray 결과 전처리

대장 또는 위암 조직으로부터 얻은 40 예의 부분 절편된 RNA를 이용하여 얻은 amplified mRNA를 이용하여 cDNA microarray를 수행하여 모든 예에서 flag가 20 % 내외로 분석이 가능한 데이터를 얻을 수 있었다 (Figure 4). cDNA microarray의 수치화된 결과는 실험간의 variation을 줄이기 위하여 표준화하였으며, 표준화 수행하기 전과 후의 변화 정도를 M-A plot과 Box plot을 이용하여 관찰하였다 (Figure 5, 6). M-A plot에서 검정색은 표준화 수행 전에 각 spot이 갖는 log ratio와 intensity를 나타내는 것이고, 주황색은 표준화 수행 후에 보정된 값을 의미한다. 표준화 전과 후에서 큰 변화가 나타나지 않아 실험에 이상이 없음과 실험의 재현성을 확인하였다. 표준화 수행 후, 부분 절편된 RNA를 이용한 40 예의 실험에서 NMP 100 %을 취하여 중복된 유전자는 평균을 구하여 보정한 후 10,736개의 유전자를 이후 분석에 사용하였다.



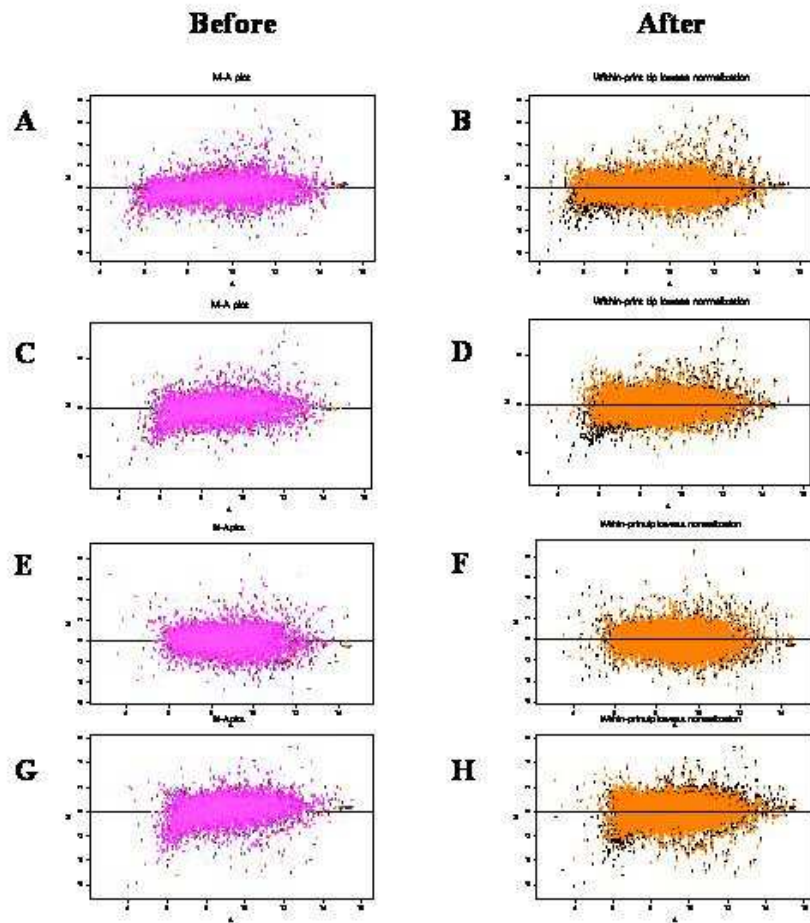
**Figure 4. Representative scanned image of cDNA microarray.**

A : Normal colon mucosa

B : Colon cancer mucosa

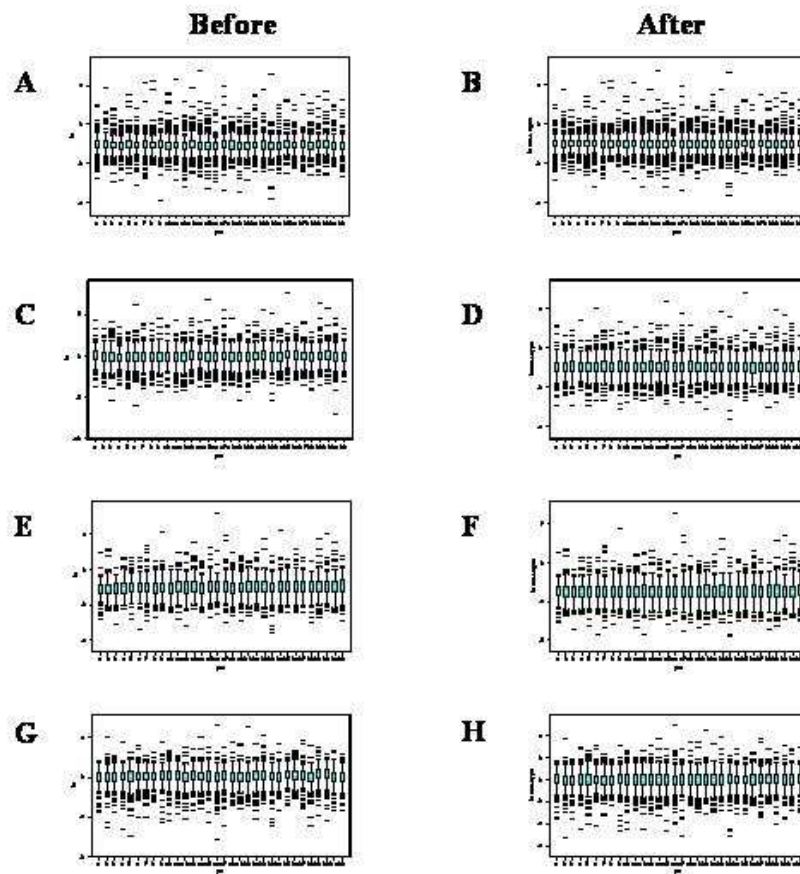
C : Normal gastric mucosa

D : Gastric cancer tissue



**Figure 5. M-A plot before and after normalization.**

A, B : Normal colon mucosa; C, D : Colon cancer tissue; E, F : Normal gastric mucosa; G, H : Gastric cancer tissue. The log ratio ( $M = \log_2 R/G$ ) are plotted on the y-axis against the log of the geometric mean of the signal intensities ( $A = \log_2 \sqrt{R/G}$ ) for each spot on the slides.



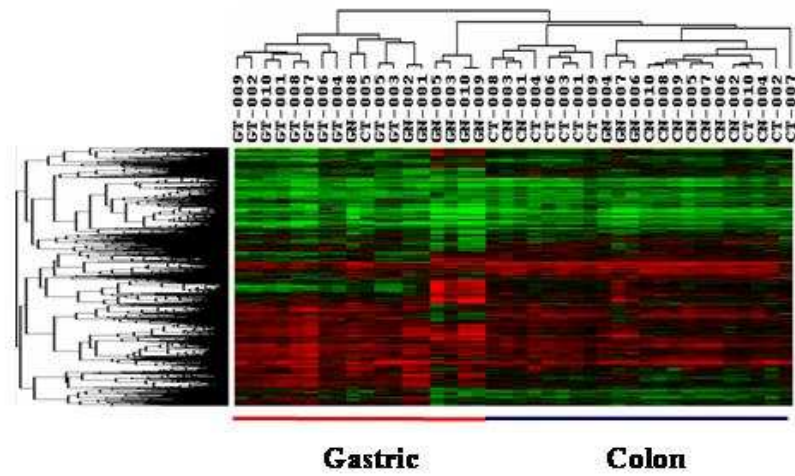
**Figure 6. Box plot before and after normalization.**

A, B : Normal colon mucosa; C, D : Colon cancer tissue; E, F : Normal gastric mucosa; G, H : Gastric cancer tissue. This plot displays a statistical summary consisting of the median, upper, and lower quartiles as well as the range. The central box in a plot represents the Inter Quartile Range, which is defined as the difference between 75<sup>th</sup> and 25<sup>th</sup> percentile or the upper and lower quartiles.

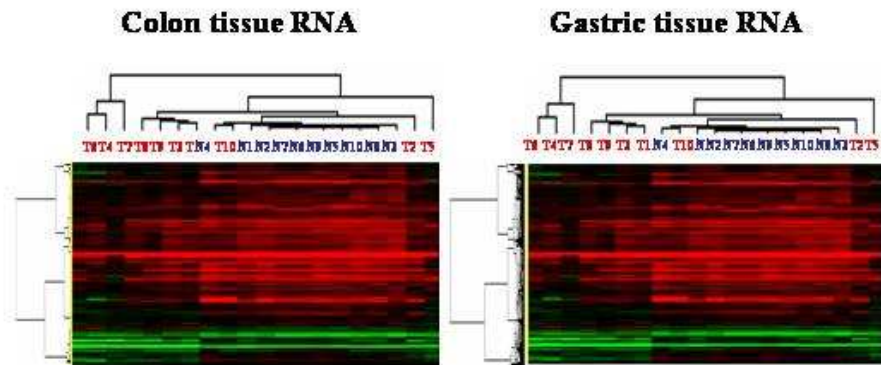
(2) Unsupervised hierarchical clustering 을 이용한 부분 절편된 조직 RNA의 발현 양상 관찰

40 예의 부분 절편된 조직의 RNA를 이용하여 cDNA microarray를 수행하여 얻은 결과로 유전자 발현 양상을 2 차원 계층 군집 분석을 수행하여 관찰한 결과, 장기와 조직 유형간에 구분됨을 알 수 있었다 (Figure 7-A). 또한 절편된 대장암과 위암의 정상 및 종양 조직을 이용하여 얻은 40 예의 microarray 결과에서 100 % NMP 처리를 하여 얻은 10,736 개의 유전자를 가지고 대장암과 위암에서 정상과 종양조직간에 유의한 유전자를 FDR 1.07 % 에서 251 개와 FDR 0.02 %에서 2,012 개의 유전자를 선별하였으며, 선별된 유전자들로 clustering을 시행한 결과 절편된 RNA를 이용하여 얻은 유의한 유전자로도 정상과 종양조직이 구분되는 것을 관찰 할 수 있었다 (Figure 7-B).





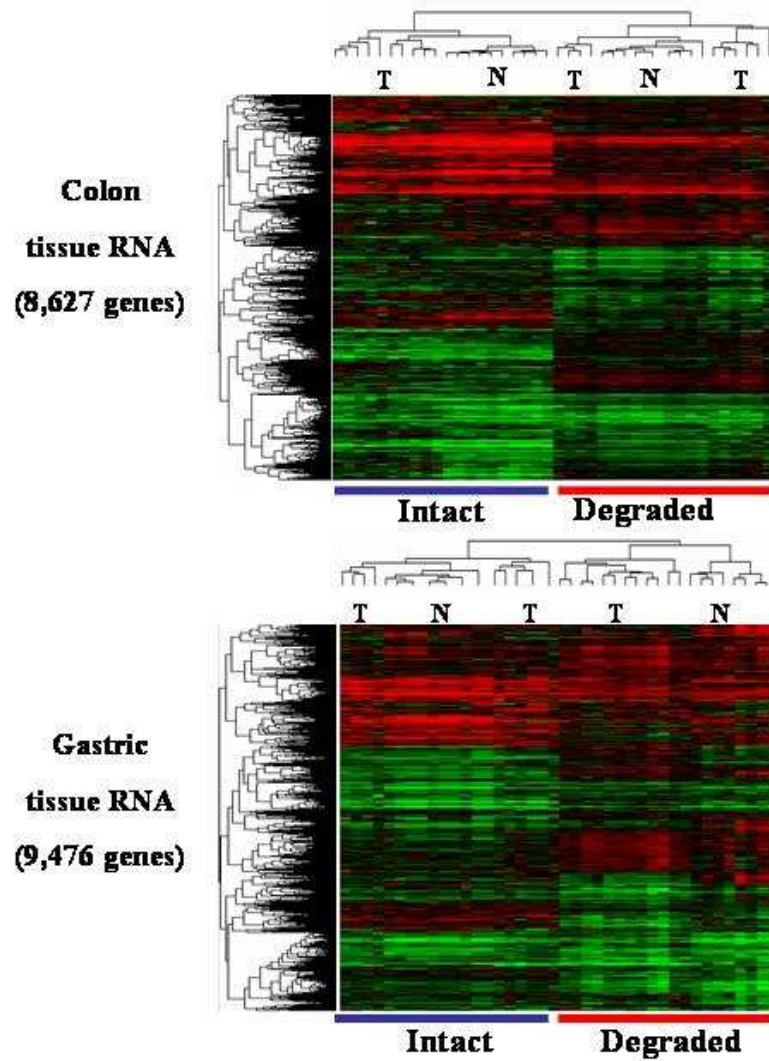
**Figure 7-A. Two-way hierarchical clustering of all experiments using partially degraded tissue RNA.** The CN, CT, GN and GT represent colon normal, colon tumor, gastric normal and gastric tumor, respectively. A number represents each patient's sample.



**Figure 7-B. Two-way hierarchical clustering using selected significant genes between normal and tumor.** The N, T represent normal tissue and tumor tissue, respectively.

### (3) 양질의 RNA와 부분 절편된 조직 RNA의 유전자 발현 양상 비교

대장암과 위암에서 양질의 RNA와 절편된 RNA를 이용하여 수행한 cDNA microarray 결과를 가지고 유전자 발현 양상을 2 차원 계층 군집 분석을 수행하여 비교 관찰하였다 (Figure 8). 대장암에서는 양질의 RNA를 이용하여 얻은 microarray 결과 (정상 및 종양 조직 각각 10예)와 절편된 RNA를 이용하여 얻은 결과 (정상 및 종양 조직 각각 10예)에서 100 % NMP 하여 얻은 8,627 개의 유전자를 사용하였고, 위암에서는 동일하게 100 % NMP 하여 얻은 9,476 개의 유전자를 가지고 수행하였다. 그 결과 대장암과 위암 모두에서 양질의 RNA와 절편된 RNA간에 서로 다른 유전자 발현 양상을 나타내고 있음을 관찰 할 수 있었다. 또한 대장암에서 양질의 RNA를 이용하여 얻은 microarray 결과에서 정상과 종양조직간에 유의한 유전자로 선정된 114 개의 유전자 (FDR 0.36 %)와 절편된 RNA를 이용하여 얻은 결과에서 얻은 169 개의 유전자 (FDR 0.39 %)를 비교한 결과 12 개의 유전자만 동일한 것이 선별되었다. 이를 통하여 양질의 RNA와 절편된 RNA를 이용하여 얻은 결과가 다르다는 것을 알 수 있었다. 이때, 양질의 RNA를 이용한 실험에서만 선정된 유의한 유전자와 절편된 RNA를 이용한 실험에서만 선정된 유의한 유전자들은 Table 1과 Table 2에 나타내었다.



**Figure 8. Unsupervised hierarchical clustering using intact and partially degraded RNA microarray data.** The N, T represent normal tissue and tumor tissue, respectively.

**Table 1. List of top 20 selected significant genes using intact RNA in colorectal cancer tissues**

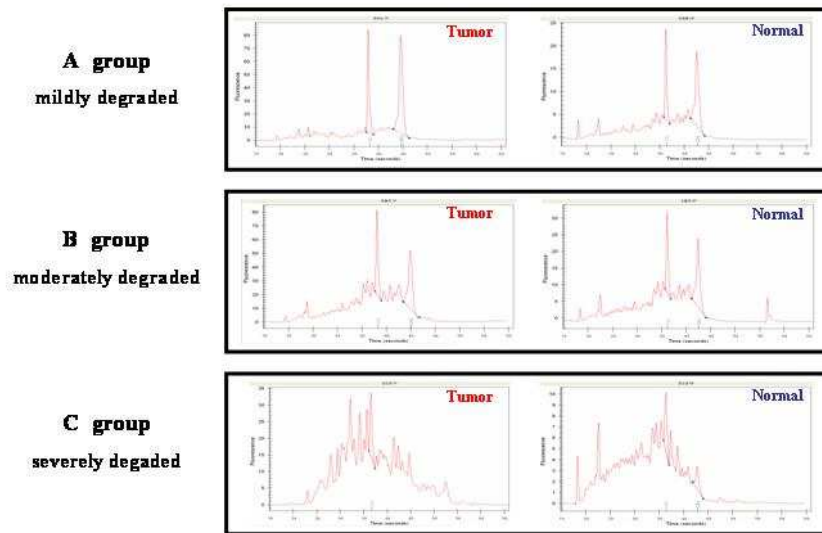
<b>Accession no.</b>	<b>Name</b>
W76339	nuclear factor (erythroid-derived 2)-like 3
AA629692	chaperonin containing TCP1, subunit 5 (epsilon)
AA485151	heat shock 105kD
AA398406	Homo sapiens EST00098 gene, last exon
AI669374	minichromosome maintenance deficient ( <i>S. cerevisiae</i> ) 3
AA052960	dyskeratosis congenita 1, dyskerin
AA912034	isoleucine-tRNA synthetase
AA446462	budding uninhibited by benzimidazoles 1 (yeast homolog)
AA402440	Homo sapiens EST00098 gene, last exon
AA504130	cytoskeleton associated protein 2
AW084638	carbonic anhydrase II
H98666	procollagen (type III) N-endopeptidase
W49672	wingless-type MMTV integration site family, member 5A
AA954935	matrix metalloproteinase 11 (stromelysin 3)
AA682719	chromobox homolog 3 ( <i>Drosophila</i> HP1 gamma)
W46577	endothelial cell-specific molecule 1
AA337434	ESTs, Moderately similar to A40493 DNA topoisomerase [ <i>H.sapiens</i> ]
AI350153	ATPase, Class VI, type 11A
AA459292	CDC28 protein kinase 1
AA400464	SRY (sex-determining region Y)-box 9 (campomelic dysplasia, autosomal sex-reversal)

**Table 2. List of top 20 selected significant genes using degraded RNA in colorectal cancer tissues**

<b>Accession no.</b>	<b>Name</b>
AA478724	insulin-like growth factor binding protein 6
R55490	phospholipase C, delta 1
AA815407	ryanodine receptor 1 (skeletal)
AA485893	ribonuclease, RNase A family, 1 (pancreatic)
AA609992	retinol dehydrogenase homolog
AA486324	proteasome (prosome, macropain) activator subunit 3 (PA28 gamma; Ki)
N24824	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog
AI271909	ESTs
AA975148	ESTs
W92603	cathepsin G
N93505	transmembrane 4 superfamily member 2
N59532	aminomethyltransferase (glycine cleavage system protein T)
H29475	pyruvate dehydrogenase kinase, isoenzyme 2
AA679278	nuclear factor of activated T-cells, cytoplasmic, calcineurin- dependent 1
W63749	B-cell CLL/lymphoma 2
AA281798	Homo sapiens mRNA; cDNA DKFZp564C1216 (from clone DKFZp564C1216)
AA699390	Novel human gene mapping to chromosome 13
AA634308	ATP-binding cassette, sub-family A (ABC1), member 8
AI017052	DNA methyltransferase 1-associated protein 1
AA608514	H3 histone, family 3B (H3.3B)

#### (4) 부분 절편된 조직의 RNA에서 절편된 정도에 따른 grouping

대장과 위암 모두에서 양질의 RNA와 절편된 RNA의 유전자 발현이 다른 것을 확인한 뒤, RNA의 절편된 정도가 영향을 미치는지 알아보기 위하여 대장의 절편된 RNA를 깨진 정도에 따라 세 그룹으로 나누어 보았다. A 그룹은 RNA의 ribosomal peak이 많이 남아있는 것들이고 B 그룹은 28s 보다 18s의 비율이 더 클 정도로 절편된 것, C 그룹은 ribosomal peak이 거의 남아있지 않을 정도로 절편된 RNA들로 나누었다 (Figure 9). 그룹을 나눈 결과 10 예씩의 대장의 정상과 종양 조직에서 A, B 그리고 C그룹을 각각 4, 4 그리고 2 예로 나눌 수 있었다. 본 연구에 이용된 위암 조직은 electrophoresis로 구분한 결과 깨진 정도가 편차가 심하여 그룹을 짓기에 적합하지 않아 이후 분석에서는 대장 조직의 결과만 사용하였고, 위암조직은 분석에서 제외하였다.



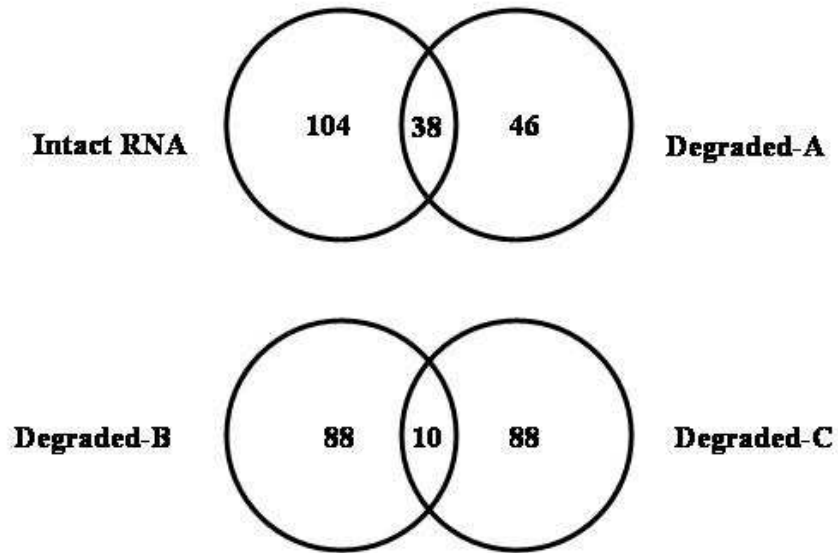
**Figure 9. Grouping of partially degraded colon tissue RNA according to degraded states. A group composed of RNA with remained ribosomal peak. B group composed of RNA with more 18s than 28s. C group composed of RNA with hardly remained ribosomal peak.**

(5) 대장암에서 절편된 세 그룹과 양질의 RNA에서 SAM에 의해 선정된 유전자들의 비교

대장암의 부분 절편된 RNA에서 그룹 지어진 세 그룹에서 정상과 종양 조직 간에 유의한 유전자를 FDR 1 % 수준에서 A, B, C그룹에서 각각 84, 98 그리고 98 개의 유전자를 선별할 수 있었다. 선별된 유전자들을 비교해 본 결과, 절편 정도가 심한 B와 C 그룹에서 10 개의 유전자가 공통적으로 선별되었고, 절편 정도가 심하지 않은 A 그룹에서 선별된 유전자는 B나 C그룹에서 선별된 유전자와 전혀 일치하지 않았다.

그리고 나서, 양질의 RNA를 이용한 microarray결과에서 정상과 종양조직 간에 유의한 유전자를 A, B 그리고 C그룹에서 선별된 유전자와 일치하는 지를 비교해 보았다. 비교한 결과, 양질의 RNA에서 선별된 유전자(FDR 1 % 142 개) 중 38 개의 유전자가 A 그룹에서 선별된 유전자와 공통인 유전자 였으며 B또는 C그룹에서 선별된 유전자와는 일치하지 않았다 (Figure 10, Table 3). 양질의 RNA에서 선별된 정상과 종양 조직간에 유의한 유전자가 B 또는 C그룹 보다는 A그룹 에서 선별된 유전자와 유사한 것을 통해, RNA의 degradation이 양질의 RNA와 다른 결과를 제공하는 데에 영향을 미친다는 것을 짐작할 수 있다.





**Figure 10.** Venn diagram with selected significant genes between normal and tumor in colon intact RNA and three groups of partially degraded colon tissue RNA.

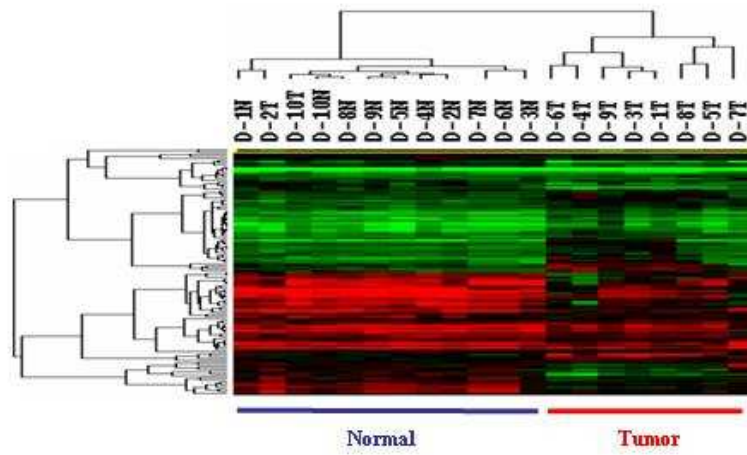
Table 3. List of commonly selected top 20 significant genes between intact and mildly degraded A-group RNA in colon cancer tissues

<b>Accession no.</b>	<b>Name</b>
W95082	hydroxysteroid (11-beta) dehydrogenase 2
AA447727	Homo sapiens cDNA FLJ12770 fis, clone NT2RP2001613, weakly similar to Mitochondrial import receptor subunit TOM40
AI017394	ESTs, Weakly similar to 2109260A B cell growth factor [H.sapiens]
AA609992	retinol dehydrogenase homolog
AI924753	aldo-keto reductase family 1, member B11 (aldose reductase-like)
AA454732	interleukin 16 (lymphocyte chemoattractant factor)
AI301329	aldo-keto reductase family 1, member B10 (aldose reductase)
AA455369	Na <sup>+</sup> /H <sup>+</sup> , amiloride solute carrier family 9 (sodium/hydrogen exchanger), isoform 1 (antiporter, sensitive)
AA293215	ubiquinol-cytochrome c reductase core protein I
W60057	keratin 13
AA461304	KIAA1605 protein
H03961	Homo sapiens CAC-1 mRNA, partial cds
R55490	phospholipase C, delta 1
T65736	selenium binding protein 1
H61193	Wiskott-Aldrich syndrome (eczema-thrombocytopenia)
N59532	aminomethyltransferase (glycine cleavage system protein T)
AI356451	CD19 antigen
AI369378	creatine kinase, mitochondrial 1 (ubiquitous)
AI990510	ATP-binding cassette, sub-family G (WHITE), member 2
AI343497	ESTs

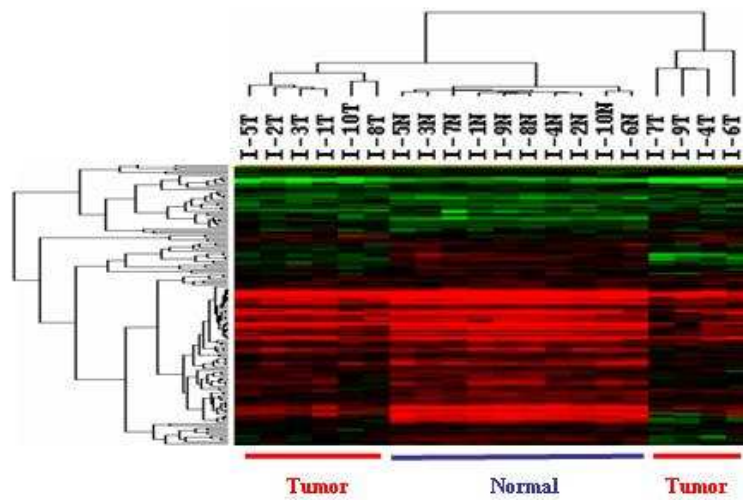
#### (6) 절편된 대장암 조직 RNA에서 선정된 유전자군의 유용도

양질의 RNA와 절편된 RNA를 이용한 실험에서 두 조건에서 선별된 유전자 군이 차이가 있음을 확인한 후, 양질의 RNA와 절편된 RNA에서 선별된 정상과 종양조직 간에 유의한 유전자군이 각각의 조건에서도 이를 정상과 종양조직으로 구분해 줄 수 있는지를 보았다. 양질의 RNA와 절편된 RNA를 이용하여 얻은 정상 및 종양 조직의 총 40 예의 실험에서 100 % NMP 처리하고 중복 유전자는 평균을 구하여 보정한 후 8,627 개의 유전자를 이후 분석에 사용하였다. 먼저 최소 FDR로 선별된 정상과 종양조직간의 유의한 유전자들을 가지고 비교하였다. 양질의 RNA에서 FDR 0.24 %로 선정된 112 개의 유전자를 가지고 절편된 RNA에서 얻은 microarray data를 가지고 clustering을 수행 하였고(Figure 11-A), 반대로 절편된 RNA에서 FDR 0.35 %로 선정된 142 개의 유전자들을 가지고 양질의 RNA로부터 얻은 microarray data를 가지고 clustering을 수행 하였다(Figure 11-B). 그 결과 절편된 RNA를 이용한 microarray data에서 선정된 유의한 유전자가 양질의 RNA를 이용한 microarray data에서도 정상과 종양 조직을 구분해 줄 수 있음을 확인하였고, 그 반대의 경우에도 같은 결과를 얻었다. 이때, 양질의 RNA에서 선정된 112개의 유전자와 절편된 RNA에서 선정된 142개의 유전자 중 공통

적으로 선정된 유전자들 중에서 상위 rank의 20 개의 유전자를 Table 4에 나타내었다. 이와 같은 결과는 양질의 RNA에서 더 많은 유전자인 1,320 개의 유전자 (FDR 0.03 %)를 가지고 절편된 RNA에서 얻은 microarray data를 가지고 clustering을 수행한 결과와 (Figure 12-A) 절편된 RNA에서 FDR 10.8 %로 선정된 1,320 개의 유전자들을 가지고 clustering을 수행한 것에서도 (Figure 12-B) 같은 결과를 확인할 수 있었다. 즉, 절편된 RNA를 이용하였을 경우 양질의 RNA를 이용한 실험에서와는 다른 종양관련 유전자가 선정되었으나, 이러한 유전자군도 정상과 종양조직의 특성을 구분해 주는 데 유용하다는 것을 제시할 수 있다.



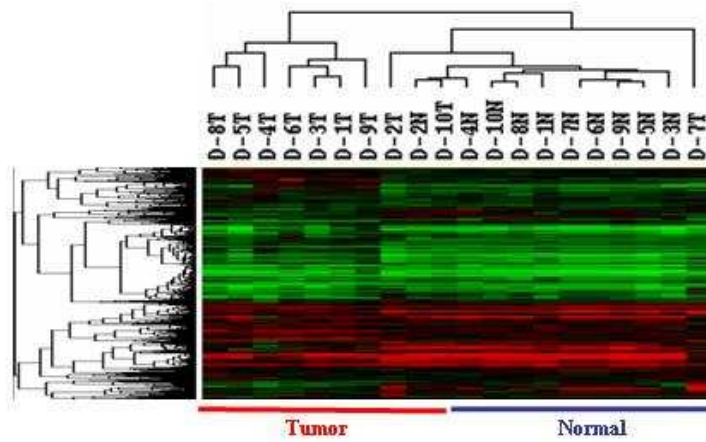
**Figure 11-A.** Two-way hierarchical clustering with selected significant 112 genes using Intact RNA microarray data in partially degraded colon tissue RNA microarray data.



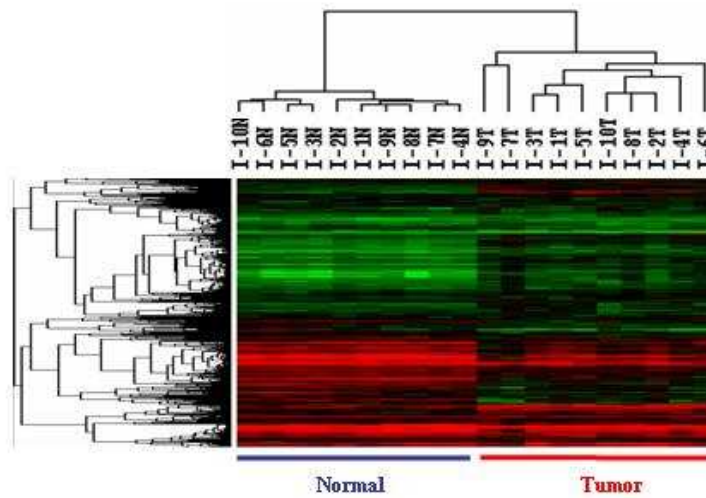
**Figure 11-B.** Two-way hierarchical clustering with selected significant 142 genes using partially degraded RNA microarray data in intact colon tissue RNA microarray data.

Table 4. List of commonly selected top 20 genes between intact and degraded RNA in colon cancer tissues

Accession no.	Name
AA046527	tetraspan transmembrane 4 super family
AA634308	ATP-binding cassette, sub-family A (ABC1), member 8
AI373245	hypothetical protein FLJ21511
AA171613	carbonic anhydrase XII
H45668	Kruppel-like factor 4 (gut)
N93505	transmembrane 4 superfamily member 2
H23187	carbonic anhydrase II
AI986336	adaptor-related protein complex 1, gamma 1 subunit
AA452278	solute carrier family 4, sodium bicarbonate cotransporter, member 4
R43532	neurexin 1
AI380522	integrin, beta 7
H72723	ESTs, Highly similar to SMHU1B metallothionein 1B [H.sapiens]
W95082	hydroxysteroid (11-beta) dehydrogenase 2
AA872383	metallothionein 1E (functional)
AA485893	ribonuclease, RNase A family, 1 (pancreatic)
W63749	B-cell CLL/lymphoma 2
AA775454	CMP-NeuAC:(beta)-N-acetylgalactosaminide (alpha)2,6-sialyltransferase member VI
AA034213	integral membrane protein 3
AA932983	UDP glycosyltransferase 1 family, polypeptide A9
AI924753	aldo-keto reductase family 1, member B11 (aldose reductase-like)



**Figure 12-A.** Two-way hierarchical clustering with selected significant 1320 genes using Intact RNA microarray data in partially degraded colon tissue RNA microarray data.



**Figure 12-B.** Two-way hierarchical clustering with selected significant 1320 genes using partially degraded RNA microarray data in intact colon tissue RNA microarray data.

#### IV. 고찰

최근 cDNA microarray에 의한 gene expression profiling은 다양한 암의 조기 진단을 위한 biomarker의 발견이나 환자의 anti-cancer drug에 대한 효과를 예측하는데 이용되어 왔다<sup>24,25</sup>. 이러한 gene expression profiling에 의해 제공된 정보는 암에 대한 새로운 기전을 밝혀내어 암 치료의 발전에 기여하고 있다. 이와 같이 유용한 tool을 사용하여 gene expression을 관찰하는데 있어서 환자의 암 조직뿐만 아니라, 소량의 blood나 biopsy와 같은 시료들도 이용에 필수적이다<sup>3,4</sup>. 그리하여 이러한 양의 제한을 극복하기 위하여 RNA amplification 방법이 고안되었으며, 지금까지 많은 논문을 통하여 non-amplification 방법과 amplification 방법과의 유사성이 보고 되었다<sup>5-14</sup>.

또한 cDNA microarray를 이용하는데 있어서 또 하나의 제한점이 양질의 RNA를 필요로 한다는 것이다. 환자의 상태를 나타내는 시점의 임상시료는 그 시기를 놓치면 다시 얻기 힘든데, 보관을 잘못하거나 endogenous enzyme에 의한 RNA degradation은 절편된 RNA가 실험에 중요한데도 불구하고 실험에 이용될 수 없게 한다. 따라서 본 연구에서는 절편된 RNA를 가지고 RNA amplification을 기반으로 한 cDNA microarray 실험의 가능성과 그 유용도를 확인하였다.



본 연구에서는 YCC-16 위암 세포주의 양질의 RNA로부터 RNase와 NaOH를 이용하여 절편된 RNA를 만들어 amplification-based cDNA microarray후, 양질의 RNA를 이용하여 얻은 cDNA microarray결과와 비교해보았다. 세포주를 이용한 실험에서는 microarray결과에서 양질의 RNA와 절편된 RNA간에 correlation coefficient와 discordance rate를 비교한 결과 intact RNA가 Rd-RNA와 더 높은 유사성을 나타내었으며 이는  $\beta$ -actin을 가지고 quantitative Real-time RT PCR을 한 결과에서도 알 수 있었다(data not shown). RNase와 NaOH에 의해 RNA가 degradation되는 기전은 높은 pH나 ribonuclease에 의해서 -OH 그룹으로부터 proton(H)을 제거하는 것에 의해 RNA chain cleavage가 일어남에 의한 것으로, 똑같은 기전에도 불구하고 이와 같은 결과가 나온 것은 intact한 total RNA가 이러한 것을 처리하는 조건에 있어서 차이가 있어 절편된 정도가 달랐기 때문으로 생각되며, 이는 후에 절편된 조직의 RNA를 이용한 실험의 결과에서도 알 수 있다. 실제 임상시료의 RNA degradation은 조직내 RNase에 의한 것이 주가 되므로, 본 연구 결과 RNase를 처리한 경우가 양질의 RNA를 사용하였을 때와 유사한 결과를 나타내는 것을 보아도 RNase처리가 좀더 세포 생리와 유사한 결과를 보여줄 수 있다.

임상시료를 이용하여 본 연구에 들어가기에 앞서, 무작위로 절편

된 RNA를 이용하여 RNA amplification을 통한 cDNA microarray를 수행한 결과 ribosomal peak이 남아있지 않을 정도로 절편된 RNA는 amplification되지 않아 cDNA microarray를 수행하여도 실험이 되지 않는다는 것을 알게 되었고, 실험이 되는 기준을 ribosomal peak이 있는 것으로 잡았다<sup>26</sup>. 그리하여 본 연구에서는 소량의 ribosomal peak이 남아있을 정도로 다양한 size로 절편된 RNA들을 실험에 사용하였다. 또한 양질의 RNA와 다양하게 절편된 RNA들을 이용하여 RNA amplification을 수행하였을 때, mRNA의 패턴에서 차이가 있는지를 알아보려고 0.2~4 Kb에 broad하게 분포하는 amplified mRNA를 크기 별로 6 구간으로 나누어 mRNA의 proportion을 quantitation한 결과, 정상과 종양조직에서 intact RNA와 degraded RNA간에 큰 차이가 없는 것을 알 수 있었다 (data not shown).

환자로부터 얻은 총 40 예의 대장암과 위암의 절편된 RNA를 이용한 본 연구에서는 세포주와 마찬가지로 절편된 RNA의 quality를 확인한 후 total RNA로 cDNA microarray에 사용할 수 없는 시료들을 amplification-based cDNA microarray를 수행하여, cDNA microarray가 가능하다는 것을 확인할 수 있었다. 이는 각 microarray의 실험에서 flag %를 구하였을 때, 20% 이하의 분석이 가능한 데이터를 얻었으며 표준화 전후의 데이터를 plot으로 비교

한 결과 차이가 거의 없으므로 양질의 실험임을 확인하였다.

하지만 같은 조직의 양질의 RNA를 이용하여 얻은 microarray data(대장암과 위암의 정상과 종양 각각 10 예씩)를 가지고 unsupervised hierarchical clustering을 이용하여 gene expression을 비교한 결과 절편된 RNA와 발현 패턴이 다르다는 것을 알게 되었다. 또한 양질의 RNA와 부분적으로 절편된 RNA를 이용하여 수행한 microarray의 결과의 유사성을 알아보기 위하여 SAM을 통한 정상과 종양 조직간의 유의한 유전자를 선별하여 비교한 결과, 양질의 RNA와 절편된 RNA에서 선별된 유전자가 일치하지 않아 본 연구에서는 양질의 RNA와 절편된 RNA가 동일하지 않음을 제시하고 있다.

RNA의 절편된 정도가 유전자 발현 양상이 양질의 RNA를 사용한 결과와 차이가나는 것에 영향을 미치는지 알아보기 위하여 절편된 정도에 따라 분류를 해 보았다. Oliver Schoor 등은 절편된 RNA를 28s/18s의 ratio로 분류하여 양질의 RNA를 이용하여 얻은 결과와 상관계수를 이용해서 비교한 후, A그룹의 정도까지는 0.9의 상관계수를 나타냄으로써 믿을 수 있는 결과를 얻을 수 있다고 보고한바 있으며<sup>16</sup>, 조직의 특성에 따라 정상조직이 종양조직에 비하여 좀 더 불안정적이라고 생각하여 정상과 종양조직을 다른 비율로 분류하였다. 그러나 본 연구에 이용된 절편된 시료들은 정상과 종양 조직에

따라 28s/18s의 ratio에 큰 차이를 나타내지 않았다. 비록 연구에 사용된 시료수가 많지 않아 RNA가 절편된 정도에 따라 어느 정도 절편된 RNA까지 사용해야 하는지 정확한 기준을 제시하지는 못하나, 유의한 유전자를 선정하는 방법만을 이용하여 28s가 많이 남아있는 A그룹만이 양질의 RNA와 더 유사하다는 것을 보여주었다. 하지만 이와 같이 그룹을 나누는 것은 육안에 의한 것이고 ribosomal peak의 양이 적을수록 Bioanalyzer와 같은 기계를 사용할 때에도 noise가 심하여 error가 있을 수 있으므로 세심한 주의가 필요하다.

마지막으로 본 연구에서 양질의 RNA를 이용하여 얻은 유전자 군은 절편된 RNA를 이용한 실험에서 얻은 정상과 종양조직간의 유의한 유전자 군과 다르지만 생물학적으로 유용하게 이용할 수 있는지를 알아보기 위해서, 절편된 RNA에서 선정된 유전자군을 가지고 양질의 RNA로부터 얻은 microarray data에서 유전자 발현 수치를 가지고 clustering을 수행한 것과 그 반대로 수행한 경우 모두 정상과 종양조직을 구분해 줌으로써 절편된 RNA를 이용하여 얻은 유의한 유전자도 유용하게 사용될 수 있음을 보여주었다.

## V. 결론

본 연구를 통하여 세포주와 임상시료를 이용해서 절편된 RNA로도 RNA 증폭 및 gene expression 관찰이 가능하다는 것을 확인하였다. 절편된 RNA를 이용한 실험 결과를 양질의 RNA를 이용한 실험과 비교한 결과 다른 유전자가 선정됨으로써 양질의 RNA만을 이용한 실험과는 다른 결과를 얻을 수도 있음을 알 수 있었으나, 절편된 RNA를 이용한 결과에서 얻은 유전자군도 정상과 종양조직을 구분할 수 있다는 것을 보여줌으로써 종양 관련 유전자를 선정하는 데에 유용하게 사용될 수 있을 것으로 사료된다.

## VI. 참고문헌

1. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999;286:531-537.
2. Andrew IS, John BW, Lisa MS, Suzanne GK, Petre D, Hilmar L, et al. Molecular classification of human carcinomas by use of gene expression signatures. *Cancer Res* 2001;61:7388-7393.
3. James E, Hermes Y, Kevin M, Yanxiang C, Suresh N, Richard F, et al. Analysis of gene expression in single live neurons. *Proc. Natl. Acad. Sci.* 1992;89:3010-3014.
4. Jennifer P, James HE. Antisense RNA amplification: A linear amplification method for analyzing the mRNA population from single cells. *Methods* 1996;10:283-288.
5. Ena W, Lance DM, Galen AO, Edison TL, Francesco MM. High-fidelity mRNA amplification for gene profiling. *Nature Biotechnology* 2000;18:457-459.
6. Feldman AL, Costouros NG, Wang E, Qian M, Marincola FM, Alexander HR et al. Advantages of mRNA amplification for

- microarray analysis. *BioTechniques* 2002;33:906-914.
7. Limei H, Jing W, Keith B, Hua W, Gregory NF, Stanley RH, et al. Obtaining reliable information from minute amounts of RNA using cDNA microarrays. *BMC Genomics* 2002;3:16-24.
  8. Laszlo GP, Agnes Z, Laszlo HJ, Paul VH. RNA Amplification Results in Reproducible Microarray Data with Slight Ratio Bias. *BioTechniques* 2002;32:1330-1340.
  9. Hastie T, Zhao H, Whitfield ML, Dale AL, Jeffrey SS. Optimization and evaluation of T7 based RNA linear amplification protocols for cDNA microarray analysis. *BMC Genomics* 2002;3:31-46.
  10. Jenson SD, Robetorys RS, Bohling SD, Schumacher JA, Morgan JW, Lim MS, et al. Validation of cDNA Microarray gene expression data obtained from linearly amplified RNA. *Molpath* 2003;56:307-312.
  11. Yiwei L, Shadan A, Philip AP, Fazlul HS. Direct comparison of Microarray gene expression profiles between non-amplification and a modified cDNA amplification procedure applicable for needle biopsy tissues. *Cancer Detection and Prevention* 2003;27:405-411.

12. Luciana IG, Ricardo LA, Beatriz SS, Elier BC, Roberto HJ, Fernando AS, et al. Comparative analysis of amplified and nonamplified RNA for hybridization in cDNA microarray. *Analytical Biochemistry* 2003;321:224-251.
13. Sandra K, Karen T, Mickey PW, Ulrich C. Evaluation of procedures for amplification of small-size samples for hybridization. *Genomics* 2004;83:508-517.
14. Zhu B, Xu F, Baba Y. An evaluation of linear RNA amplification in cDNA Microarray gene expression analysis. *Mol Genet and metab* 2005;
15. Charlie CX, Mei C, Li M, Quang NP, Jason MI, Olga AK, et al. A new strategy to amplify degraded RNA from small tissue samples for microarray studies. *Nucleic Acids Research* 2003;31:e53.
16. Schoor O, Weinschenk T, Hennenlotter J, Corvin S, Stenzl A, Rammensee HG, et al. Moderate degradation does not preclude microarray analysis of small amounts of RNA. *BioTechniques* 2003; 35: 1192-1201.
17. Katsuta H, Koyanagi KR, Shiiba M, Anzai K, Irie T, Aida T, et al. cDNA Microarray analysis after laser



- microdissection in proliferating islets of partially pancreatectomized mice. *Med Mol Morphol* 2005; 38:30-35.
18. Kim TM, Jeong HJ, Seo MY, Kim SC, Cho G, Park KH, et al. Determination of genes related to gastrointestinal tract origin cancer cells using a cDNA microarray. *Clin Cancer Res* 2005;11:79-86.
19. Kim BS, Lee S, Rha SY, Chung HC. cDNA microarray experiment: Design issues in early stage and the need of normalization. *Cancer Res Treat* 2004; 35:533-540.
20. Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, et al. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* 2002;30:e15.
21. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci* 1998;95:14863-14868.
22. Li Y, Li T, Liu S, Qiu M, Han Z, Jiang Z, et al. Systematic comparison of the fidelity of aRNA, mRNA and T-RNA on gene expression profiling using cDNA microarray. *J Biotechnol* 2004;107:19-28.

23. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to ionizing radiation response. *Proc Natl Acad Sci* 2001;98:5116-5121.
24. Alain B, Antoinette L, Pierre YB, Chantal T, Didier B, Franck C, et al. Colon cancer prognosis prediction by gene expression profiling. *Oncogene* 2005;24:6155-6164.
25. Borovecki F, Lovrecic L, Zhou J, Jeong H, Then F, Rosas HD, et al. Genome-wide expression profiling of human blood reveals biomarkers for Huntington's disease. *PNAS* 2005;102:11023-11028.
26. Dirk ML, Ulrike S, Ingo BA, Till TB. Development of a DNA Microarray for Detection and Identification of Fungal Pathogens Involved in Invasive Mycoses. *J Clin Microbiol* 2005;43:4943-4953.

ABSTRACT

Gene expression profiling of partially degraded tissue RNA  
using RNA amplification-based cDNA microarray

Yeon Ju Yang

*Department of Medical Science*

*The Graduate School, Yonsei University*

(Directed by professor Sun Young Rha)

Recently, cDNA microarrays are widely applied in cancer research to identify the candidate biomarkers for prognosis of various cancers or predicting the efficacy of cancer therapy. Besides the limitation of RNA amount in cDNA microarray, the other significant limitation is the requirement of good quality of RNA. Various RNA amplification methods have been devised to overcome limitation of amounts and there had been strong demands to use even the partially degraded RNAs for gene expression analysis.

The purposes of this study are to evaluate the possibility of using degraded tissue RNAs for gene expression analysis and to evaluate the reliability of the microarray data from the degraded RNA.

Intact RNA from YCC-16 gastric cancer cell and their degraded products were prepared by RNase A or NaOH treatment. T7 linear RNA amplification-based cDNA microarray with intact and degraded RNA using 7.5K human cDNA microarray in an indirect design was performed. The degraded RNA was successfully amplified and the quality of microarray data was acceptable for gene expression profiling. To evaluate the similarity between intact and degraded RNA a Pearson correlation coefficient ( $r^2$ ) and discordance rates were analyzed. The correlation coefficients between 3 conditions were 0.93 between I-RNA and Rd-RNA and 0.86 between I-RNA and Nd-RNA. The quantity discordance rate was decreased at the larger fold-changes.

Next, T7 linear RNA amplification-based cDNA microarray with partially degraded colon and gastric tissue RNAs (10 normal & 10 tumor tissues for each tumor type) using 17K human cDNA

microarray in an indirect design were performed. After the preprocessing, the gene expression profiling was compared with previous microarray data of intact tissue RNAs (10 normal & 10 tumor tissues for each tumor type) using unsupervised clustering and the differentially expressed genes between normal and cancer tissues using SAM.

The degraded tissue RNAs with at least one ribosomal peak could be successfully amplified and resulted in reliable gene expression profiling. However, the gene expression patterns of intact and degraded RNAs were different both in clustering and selected genes. Among the selected colon cancer specific genes, 114 (FDR 0.36%) in intact and 169 (FDR 0.39%) in degraded tissues, only 12 genes between them were overlapped, and the functional category of selected genes in each condition was not exactly identical.

In order to evaluate the influence of degraded states, a partially degraded colon tissue RNAs were divided into three groups. As the selected significant genes from intact RNA data tend to more similar to A group of mild degraded RNA than B, C group of moderate to severely degraded RNA, it is suggested

that the degree of RNA degradation might affect the expression profiling.

Although the selected genes differ with in intact RNA and degraded RNA, to observe the reliability of selected significant genes in degraded RNA microarray data, results of clustering using the selected significant genes in degraded RNA to intact microarray data set showed could classify normal and tumor, and vise versa.

In summary, the gene expression profiling is possible using partially degraded tissue RNAs, but the cautious data analysis for relevant biological meaning is needed.

---

Key Words: cDNA microarray, RNA amplification, partially degraded tissue RNA