

총설

# 유전체 역학연구의 동향

강대회, 이경무

서울대학교 의과대학 예방의학교실

## Current Status of Genomic Epidemiology Research

Daehee Kang, Kyoung-Mu Lee

Department of Preventive Medicine, Seoul National University College of Medicine

Genomic epidemiology is defined as "an evolving field of inquiring that uses the systematic application of epidemiologic methods and approaches in population-based studies of the impact of human genetic variation on health and disease (Khoury, 1998)". Most human diseases are caused by the intricate interaction among environmental exposures and genetic susceptibility factors. Susceptibility genes involved in disease pathogenesis are categorized into two groups: high penetrance genes (i.e., *BRAC1*, *RB*, etc.) and low penetrance genes (i.e., *GSTs*, *Cyps*, *XRCC1*, etc.), and low penetrance susceptibility genes has the higher priority for

epidemiological research due to high population attributable risk.

In this paper, the summarized results of the association study between single nucleotide polymorphisms (SNPs) and breast cancer in Korea were introduced and the international trends of genomic epidemiology research were reviewed with an emphasis on internet-based case-control and cohort consortium.

*Korean J Prev Med 2003;36(3):213-222*

**Key Words:** Genome, Epidemiology, SNP, Case-control study, Cohort study

## 서론

전통적인 역학연구에서 연령, 성별 등의 보정은 가능하지만 유전적인 요인이나 환경유해물질에 대한 개개인의 대사능력, 즉, 숙주요인 또는 개개인의 감수성에 대한 평가가 어려운 단점을 보완하기 위하여 분자 생물학 또는 유전학적인 방법으로 DNA, 염색체 등의 생화학적인 혹은 생물학적인 지표들을 이용하여 폭로 및 질병 추정의 유용성을 증대시키려는 노력이 역학 분야에서 진행되어 왔다. 그 결과 "역학적인 연구에서 생체지표를 이용하는 방법론"으로 정의되는 분자역학(molecular epidemiology) [1] 및 "유전적인 요인과 환경적인 요인의 상호작용에 관한 역학연구"로 정의되는 유전역학(genetic epidemiology) [2] 분야가 1980년대부터 새로이 생겨나게 되었다.

한편, 최근 인간게놈프로젝트(human genome project)의 완적으로 인해 방대한 양의 유전자 정보가 축적되고 있는데, 이러한 유전자 정보가 의학적 혹은 공중보

건의학적으로 이용될 수 있기 위해서는 인구집단을 대상으로 한 역학연구가 반드시 필요하다. 이런 맥락에서 Khoury 등 [3]은 기존의 분자역학 및 유전역학을 대신하여, 『인간 유전체 역학(human genome epidemiology: HuGE)』이라는 새로운 용어를 제안하면서, "인간의 유전적 변이가 건강 또는 질병에 미치는 영향을 규명하기 위해 인구집단을 대상으로 하는 연구에 역학적인 방법 및 접근법을 체계적으로 적용하는 분야(an evolving field of inquiry that uses systematic applications of epidemiologic methods and approaches in population-based studies of the impact of human genetic variation on health and disease)"로 정의하고 있다. 또한 Khoury 등 [3]은 이러한 유전체 역학은 분자역학과 유전역학이 교차되는 지점에 위치한다고 설명하고 있다.

본 논문에서는 유전체 역학연구의 동향을 단일염기서열 다형성(single nucleotide polymorphisms: SNPs)과 질병에 관한 환자-대조군 연구(association

study) 및 코호트 연구를 중심으로 살펴보고, 앞으로의 연구방향을 모색해 보고자 한다. 또한 유전적 다형성의 선정과 SNP 분석기법에 대해서도 간략히 고찰하고자 한다.

## 유전자 다형성과 질병에 관한 연관성 연구

질병 발생 특히 발암과정은 환경적인 인자와 유전적인 인자의 상호작용에 의해서 이루어진다고 판단되는데, 인체 감수성에 관여하는 유전자를 penetrance의 정도에 따라 두 가지로 분류할 수 있다 (Table 1). High penetrance를 갖는 유전자는 빈도는 드물게 나타나지만 돌연변이(<1%)를 갖는 사람들이 그렇지 않은 사람에 비해서 질병발생의 위험도(attributable risk or relative risk)는 상당히 높다. 하지만 빈도가 아주 낮기 때문에 일반 인구집단에서의 위험도 (population attributable risk, PAR)는 low penetrance를 갖는 유전자에 비해서 상당히 낮으며, PAR이 높은 유전자일수록 공중보건학적인 연구대상으로서의 우선순위가 높다고

할 수 있다 [4,5].

전체 역학연구에서의 특정 유전자 및 특정 유전적 다형성 선택의 기준은, ① 생물학적인 중요성(biological relevance), ② variant allele의 빈도, ③ 이용할 수 있는 유전자 분석 방법의 유무를 들 수 있다. 또한, 인터넷 웹상에서 많은 SNP 데이터베이스들이 운영되고 있어 유전적 다형성의 선정에 많은 자료를 얻을 수 있는데, 대표적인 예는 Table 2에 정리되어 있다 [6-8].

이상적인 유전자 다형성(SNP) 분석방법의 조건으로는, ① 이미 분석법이 확립되었거나, 유전자의 서열로부터 분석법이 쉽게 개발가능한 방법, ② 소요되는 장비, 재료 및 시간의 측면에서 상대적으로 효율적인 방법, ③ DNA sample이 최적의

상태가 아닐 경우에도 정확성과 신뢰성 있는 결과를 얻을 수 있는 방법, ④ 간편하면서도 자동화가 가능한 방법, ⑤ 분석법 및 분석 시료수의 조절이 유연한 방법과 같은 기준을 들 수 있다. 몇 가지 대표적인 SNP 분석 방법의 특징, 속도, 비용 및 정확도를 비교하면 Table 3과 같다 [9-16].

### 유전자 다형성과 유방암에 관한 연관성 연구

유전자 다형성과 질병에 관한 연관성 연구의 예로서 현재까지 서울의대 예방 의학교실에서 한국인 여성을 대상으로 수행한 유방암과 SNP의 관련성 연구를 소개하고자 한다.

유방암의 주요한 역학적 위험요인은

이른 초경, 늦은 폐경, 만삭 임신 경험이 없거나 늦은 경우, 폐경기 에스트로겐 요법 등으로서, 일생동안의 에스트로겐 폭으로 설명될 수 있다 [17]. 그러나, 아직도 많은 부분은 이와 같은 요인으로 설명되지 않으며 [18], 따라서, 개인 감수성에 영향을 미치는 유전자 다형성을 탐구하는 것은 유방암 병인의 이해에 큰 기여를 할 수 있다 [19].

#### 1. 연구대상 및 분석방법

1995년 이후 서울대학교병원, 아산재단 중앙병원, 보라매 병원 일반외과를 방문하여 유방암으로 확진 받은 여성을 환자군으로 하였으며, 대조군은 환자군과 동 기간동안 같은 병원 일반외과에 입원 하거나 외래를 방문한 여성 중 과거 다른 암종이나 만성질환을 앓은 병력이 없고, 난소 및 자궁 절제력, 또는 유방절제력이 없으며, 내분비 계통의 질환이 없는 사람으로 하였다. 모든 연구대상자에 대하여 설문조사를 통해 인적 사항, 질병 과거력, 가족력, 월경요인, 출산요인, 비만요인, 흡연, 음주습관, 식이습관 등을 조사하였다 [20,21].

설문자료 및 혈액시료를 모두 수집한 연구대상자는 2003년 4월 현재 환자군 1,147명, 대조군 1,016명이며, Table 4에 제시된 각각의 유전자 다형성은 PCR, PCR-RFLP, TaqMan Assay, Real-time PCR, PCR-CTPP, SBE, MALDI-TOF MS 방법으로 확인하였다.

#### 2. 연구결과

##### 1) variant allele 빈도

Table 4는 각 유전자 다형성 별로 분석

**Table 1. Characteristics of two categories of susceptibility genes**

	High Penetrance	Low Penetrance
Role	necessary and sufficient for disease	neither necessary nor sufficient for disease
Frequency in population	low to rare	often common
Familial hereditary pattern	clear	equivocal
Strength of association	high	low to moderate
Absolute risk of disease	high	low
Population attributable risk	low	high
Role of environmental exposure	secondary and variable	critical
Gene-environmental interaction	secondary and variable	primary and implicit
Study type	family	population
Examples	<i>BRCA1/2, RB, etc.</i>	<i>GSTs, CYPs, XRCC1, etc.</i>

modified from Caporaso [4] and Zheng [5]

**Table 2. Useful websites for SNP databases**

Database [Ref]	Website
NCBI SNP database [6]	www.ncbi.nlm.nih.gov/SNP
SNP Consortium (TSC) [7]	snpcshl.org
HGVbase [8]	hgvbase.cgb.ki.se
SNP500Cancer database (CGAP, Cancer Genome Anatomy Project)	snp500cancer.nci.nih.gov/home.cfm

**Table 3. Comparison of SNP genotyping methods**

Methods [Ref]	Characteristics	Throughput	Cost	Accuracy
Sequencing	identification of novel SNP locus	+	++++	++++
RFLP [9]	restriction fragment length polymorphism	+	++	+++
SSCP [10]	single-strand conformation polymorphism	+	+	+
DASH [11]	dynamic allele specific hybridization	++	+	++
TaqMan [12]	difference in hybridization (5'-exonuclease)	++	+++	++
PCR-CTPP [13]	allele-specific PCR with confronting two-pair primers	++	+	++
SBE [14]	single base extension (blocking nucleotide synthesis by ddNTP)	+++	++	+++
MALDI-TOF MS [15]	matrix-assisted laser desorption/ionization time of flight mass spectrometry	++++	++	+++
Oligonucleotide chip [16]	simultaneous genotyping for thousands of SNPs	++++	+++	++

**Table 4.** List of genes and polymorphisms analyzed in Korean breast cancer cases and controls

Gene	Polymorphism	Case	Control	Frequency in Controls		Genotyping Method
				Type	%	
<b>Xenobiotics metabolism</b>						
GSTM1	deletion	188	181	null	52	PCR
GSTT1	deletion	189	189	null	41	PCR
GSTP1	Ile105Val (A>G)	189	189	G	18	RFLP
NAT1	11 variants	284	366	*10	46	TaqMan
NAT2	24 variants	292	381	fast	15	TaqMan
CYP2E1	5' UTR 1259G>C	306	358	C	18	RFLP
EPHX	Thr113His (T>C)	169	175	C	57	DASH
NQO1	Pro187Ser (C>T)	307	390	T	44	RT-PCR
ALDH2	Glu487Lys (G>A)	592	523	A	19	PCR-CTPP
MTHFR	667C>T	314	364	T	41	RFLP
<b>Estrogen synthesis &amp; metabolism</b>						
ER- $\alpha$	PuvII (P/p)	310	376	p	60	RFLP
	XbaI (X/x)	310	376	x	75	RFLP
	Pro325Pro (C>G)	792	760	G	52	MALDI-TOF
CYP11A1	Thr594Thr (G>A)	820	785	A	18	MALDI-TOF
	NcoI Ile462Val (G>C)	537	445	G	27	SBE
	MspI 6235T>C	423	289	T	47	SBE
CYP17	Val432Leu (G>C)	308	399	G	11	RT-PCR
CYP19	5' UTR 1931C>T	500	426	T	42	SBE
CYP19	Arg264Cys (C>T)	389	345	T	21	DASH
COMT	Val158Met (G>A)	188	330	L	25	RFLP
<b>DNA repair</b>						
hOGG1	Ser326Cys (C>T)	307	350	T	55	RFLP
XRCC1	Arg194Trp (C>T)	306	315	T	33	RFLP
	Arg399Gln (G>A)	307	315	A	30	RFLP
	Arg188His (G>A)	864	786	A	0	MALDI-TOF
XRCC2	Thr241Met (C>T)	517	376	T	4	SBE
XRCC3	921G>T	844	812	T	28	MALDI-TOF
XRCC4	Gly593Gly (G>T)	477	464	T	0	MALDI-TOF
XRCC6	Asp312Asn (G>A)	557	445	A	5	SBE
ERCC2	2505T>C	386	337	C	23	DASH
ERCC4	Ser707Pro (C>T)	547	439	T	0	SBE
ATM	1066-6T>G	500	410	G	0	SBE
	Ile143Val (A>G)	442	424	G	0.1	SBE
AGT	Gly160Arg (G>A)	493	437	A	0	SBE
LIG4	Asp501Asp (T>C)	872	808	C	0	MALDI-TOF
RAD51	5' UTR 135G>C	834	734	C	16	MALDI-TOF
RAD51	5' UTR 172G>T	837	740	T	6	MALDI-TOF
	3' UTR 2259C>T	875	826	T	56	MALDI-TOF
<b>Cytokine &amp; growth factor</b>						
TGF- $\beta$ 1	Leu10Pro (T>C)	594	521	C	47	PCR-CTPP
TNF- $\beta$	252A>G	588	516	G	46	PCR-CTPP
IGF-1	3' UTR 2505T>G	595	518	G	24	PCR-CTPP
IL-1 $\beta$	5' UTR 31C>T	595	525	C	49	PCR-CTPP
IL1RN	86bp VNTR	595	522	*2	15	PCR-CTPP
HER2	Ile655Val (A>G)	586	519	G	12	PCR-CTPP
BAR2	Gln27Glu (C>G)	589	523	G	10	PCR-CTPP
<b>Cell cycle control</b>						
CCND1	870G>A	816	805	A	53	MALDI-TOF
CDK7	99T>C	730	653	T	56	MALDI-TOF
BCL-6	397G>C	847	785	C	21	MALDI-TOF

한국의 경우에는 다형성이 존재하지 않는 site로 나타나는 등, 백인 및 중국인과 비교하여 현저한 차이를 보이는 경우가 많다. 예를 들어, CYP1B1 432Val allele 을 가지는 대조군은 21%로서 일본인의 경우와 비슷하였으나 [22], 중국인(85%) [23], 백인(70-81%) 및 African-American(95%) [24,25]에 비하여 현저하게 낮게 나타났다.

2) 유전적 다형성과 유방암의 연관성 *GSTM1/T1/P1, COMT, CYP2E1, CYP19, CYP17, ER- $\alpha$ , XRCC1, XRCC3, RAD52, TGF- $\beta$ 1, TNF- $\beta$ , IL-1 $\beta$ , IL-1RN, CDK7* 등의 유전적 다형성에 따라 유방 암에 대한 감수성에 유의한 차이가 있었 으며, 유전자-환경간 또는 유전자-유전자 간에 상호작용을 관찰할 수 있었다 [20,21,26-31]. 각 유전적 다형성과 유방 암의 연관성에 대한 결과를 Figures 1-4 에 제시하였다.

### SNP Association Study를 위한 국제 공동연구

유전자와 질병의 관계를 규명하는 연 관성 연구(association study)에서 유전자- 환경 간 또는 유전자-유전자 간의 상호작용에 대한 검정력은 주효과에 대한 검정 력에 비하여 현저히 떨어진다 [32]. 따라서, 이러한 상호작용을 규명하기 위해서 는 훨씬 더 큰 표본수(sample size)가 필 요하다. 이러한 문제점을 극복하기 위해 서, 국제적으로 협동연구의 형태로 기존 연구성과를 종합하는 한편, 큰 표본수를 확보하여 유전자-환경간 또는 유전자-유전자간 상호작용 규명하고자 하는 시도 가 이루어지고 있다. 특이할만한 점은 이 러한 노력들이 인터넷을 주요 수단으로 하여 전개되고 있다는 사실이다.

#### 1. HuGE Net (The Human Genome Epidemiology Network)

HuGE는 인간 유전체에 관한 역학적 정보를 수집, 축적하며 정보를 교환하고 자 하는 목적으로 전 세계 연구자들 또는

된 환자군 및 대조군 시료의 수와 대조군 에서의 variant allele의 빈도를 제시하고

있다. 선행 연구에서의 variant allele 빈도 를 고려하여 선정하였음에도 불구하고,

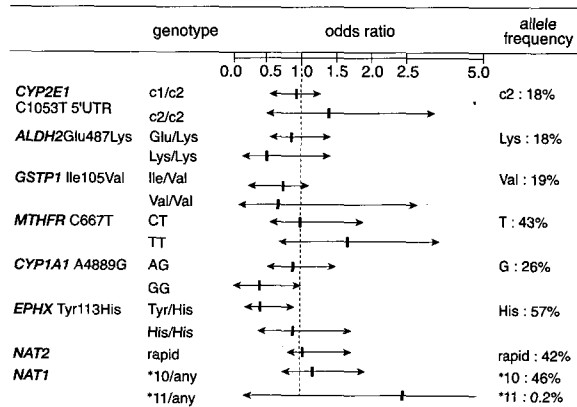


Figure 1. Carcinogen metabolizing enzyme genes and breast cancer.

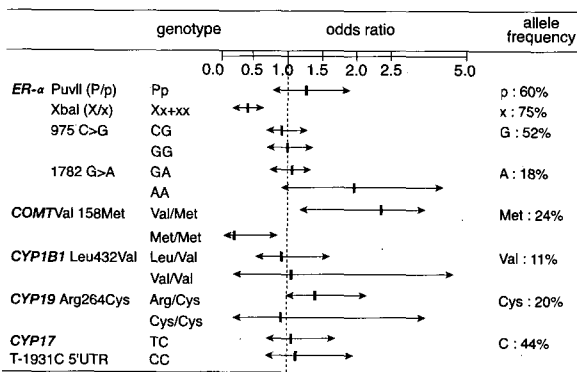


Figure 2. Estrogen synthesis and metabolism related genes and breast cancer.

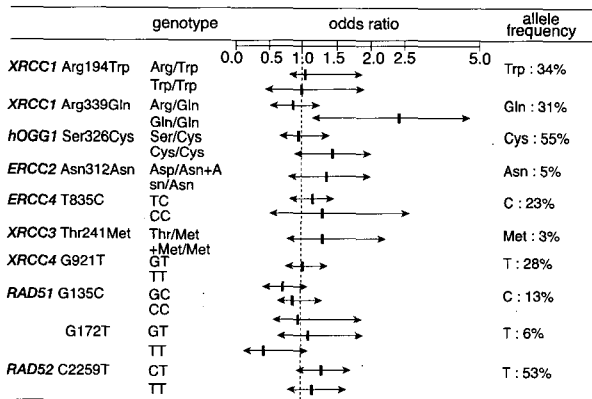


Figure 3. DNA repair genes and breast cancer.

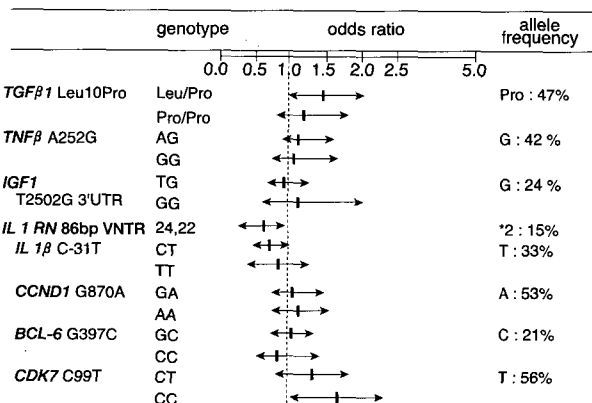


Figure 4. Cytokine, growth factor, and cell cycle control genes and breast cancer.

연구기관의 인터넷 네트워크 형태로 수립되었으며, 유전체 역학(genome epidemiology)이라는 용어를 제안하기도 하였다(<http://www.cdc.gov/genomics/hugenet>) [3]. 이 네트워크에 참여하는 연구자들은 인구집단별로 인간의 유전적 변이에 관한 데이터를 수집하며, 이러한 유전적 변이와 질병 간의 연관성을 다양한 인종집단에 대하여 규명하고, 인구집단을 기반으로 얻어진 유전자-환경간 상호작용에 대한 양적인 데이터를 축적하는 한편, 유전자 검사(genetic test) 및 서비스가 질병 예방과 건강증진에 미치는 영향을 규명하는 연구를 수행하고 있다.

HuGE의 연구결과들은 인터넷상에서 이용될 수 있는데, HuGE Reviews, Case Studies, Fact Sheets, HuGE Reports로 나누어 제공하고 있으며, E-Journal Club도 운영하고 있다. 이 중, HuGE Reviews는 주요 연구 성과로서, 2000년 이후 2003년 7월 현재까지 21편의 논문이 American Journal of Epidemiology와 Genetics in Medicine에 리뷰 형태로 계속적으로 게재되어 왔으며, 37편의 논문이 집필 중에 있다 [33,34]. 이러한 연구 결과들은 방대한 유전적 정보를 통해 질병예방과 건강증진이라는 목적을 달성하는데 기여할 것으로 기대된다.

## 2. GSEC (International Collaborative Study on Genetic Susceptibility to Environmental Carcinogen: International Pooled Data Analysis Project)

GSEC는 환경성 발암물질의 대사에 관련된 유전자와 질병의 연관성 및 유전자-환경간 상호작용을 충분한 검정력을 가지고 규명하고자 구축되었다(<http://www.gsec.net>) [35]. 연구대상이 되는 환경성 발암물질의 대사에 관련된 유전자는 *CYP1A1*, *CYP2D6*, *CYP2E1*, *GSTM1/T1*, *NAT2*, *EH*, *NAT1*, *NQO1*, *GSTP1* 등이며, 2003년 1월 현재, 78명의 연구자, 136개 기관에서 데이터를 공유하여 총 44,966명의 연구대상자(환자군 21,092

명, 대조군 23,874명)를 확보한 상황이다. 수집된 환자군을 암 종류별로 구분하면, 폐암 29%, 방광암 15%, 유방암 14%, 두경부암 10%, 피부암 8%, 대장암 7%, 백혈병 5%, 난소-자궁암 3%, 전립선암 3%, 기타 7%와 같다.

수집된 유전자 데이터의 현황은 Table 5에 제시되어 있으며, 환자군과 대조군 모두 약 80% 정도가 두 가지 이상의 유전자에 대한 정보를 가지고 있어 유전자 간 상호작용이 암발생에 미치는 영향도 밝혀낼 수 있을 것으로 기대되고 있다.

현재까지, *GSTM1*, *GSTT1*, *CYP1A1*, *CYP2E1*, *EPHX*의 유전적 다형성과 폐암의 연관성에 관한 연구, 폐암발생에 미치는 흡연, 석면과 유전적 다형성의 상호작용에 관한 연구, *GSTM1*, *NAT2*의 유전적 다형성과 방광암의 연관성에 관한 연구 등은 이미 그 성과가 국제학술지를 통해 10여편 발표된 바 있으며 [36-44], 두경부암, 유방암의 경우도 유전적 다형성과의 연관성 및 상호작용에 관한 연구가 진행되고 있다.

## 유전체 역학연구를 위한 코호트 연구 동향

현재 활발히 수행되고 있는 국가간 공동연구 형태의 SNP 연관성 연구를 통하여 많은 정보를 얻을 수 있을 것으로 전망되나, 후향적 연구형태인 환자-대조군 연구의 한계점과 인종 간의 유전적인 차이로 인하여 더욱 큰 수의 연구대상자가 필요한 점을 감안하여, 최근에는 유전정보를 얻을 수 있는 시료를 확보한 코호트들을 모아 ‘코호트 연합(cohort consortium)’을 구성한 뒤, 코호트내 환자-대조군 연구형태(nested case-control study)로써 유전자 다형성과 질병간의 연관성과 질병발생에서 유전자-환경 및 유전자-유전자간 상호작용을 규명하려는 노력이 시도되고 있다.

Table 5. Percentage of SNPs analysed for the subjects in GSEC database

Gene	Controls, n (%)	Cases, n (%)
<i>GSTM1</i>	17,843 (74.7)	16,937 (80.3)
<i>GSTT1</i>	10,672 (44.7)	11,753 (55.7)
<i>CYP1A1</i>	10,133 (42.4)	8,780 (41.6)
<i>CYP2E1</i>	4,178 (17.5)	3,701 (17.5)
<i>CYP2D6</i>	5,811 (24.3)	5,805 (27.5)
<i>NAT2</i>	7,121 (29.8)	6,108 (29.0)
<i>GSTM3</i>	1,820 ( 7.6)	3,098 (14.7)
<i>GSTP1</i>	5,166 (21.6)	5,744 (27.2)
<i>HPX</i>	1,144 ( 4.8)	1,003 ( 4.8)
<i>NAT1</i>	1,714 ( 7.2)	1,999 ( 9.5)
<i>CYP2C19</i>	1,615 ( 6.8)	1,051 ( 5.0)
Two genes	4,846 (20.3)	4,962 (23.5)
Three genes	4,270 (17.9)	3,879 (18.4)
Four genes	3,532 (14.8)	2,685 (12.7)
Five or More genes	4,159 (17.4)	4,773 (22.6)
Total	23,874 (100)	21,092 (100)

(from <http://www.gsec.net>)

### 1. US National Cancer Institute Cohort Consortium

미국 NCI는 유전체 역학 연구의 효과적인 수행을 위하여 70만명 이상의 연구대상자를 확보하는 코호트연합(cohort consortium)을 추진하고 있다 (<http://ospahome.nci.nih.gov/cohort/default.htm>). 이는 NCI의 지원을 받은 Cancer Genetics Working Group (1996년 설립)이 1999년 “Genes and the Environment”라는 전략에 초점을 맞추면서 기존에 존재하고 있는 연구자원을 충분히 활용할 것을 권고한 결과로서, 코호트연합에 참여할 수 있는 최소한의 자격요건을 다음과 같이 제시하고 있다. 1) DNA sample (예. lymphocytes, buffy coat, cheek cells 등)이 구축되어 있어야 하며, 2) 2010년까지 최소한 1,500명의 암환자가 발생할 것으로 추정되어야 하고, 3) 신규 암발생을 추적조사하기 위한 암등록제도가 체계적으로 갖추어져 있고, 4) 알려진 또는 가능한 암의 위험요인에 대하여 충분한 정보가 시작단계에서 수집되어 있으면서, 5) 연합(consortium)의 활동 및 연구에 참여할 의사가 있어야 한다. Table 6은 주요 코호트들의 특징을 정리하여 제시해 주고 있다.

SNP와 암의 연관성 연구는 “Proof of

Principle Study”로서 수행되는데, 먼저 database등을 이용하여 후보유전자의 유전적 변이를 조사한 다음, haplotype를 결정하여 코호트내 환자-대조군 연구로써 암의 관련성을 살펴보는 과정으로 진행될 예정이다. 현재, 유방암 및 전립선암을 대상으로 스테로이드 호르몬과 관련된 49개의 후보유전자와 한 유전자당 25가지의 변이(variant)를 선정하였으며, 일부 연구는 유전자형이 IGF (insulin growth factor) 및 IGF-binding protein의 수준에 미치는 영향에 관한 연구를 수행하고 있다 [45].

### 2. COGENE (Co-ordination of Genomes Research Across Europe)

COGENE은 유럽연합(EU) 제5차 프레임워크 프로그램(FP5)의 ‘삶의 질 프로그램’에 포함된 프로젝트로서 유럽 내에서 이루어지고 있는 유전체 연구 관련 정보의 교류 및 협력을 강화하기 위한 목적으로 구축되고 있다(<http://forum.europa.eu.int/irc/rtd/cogene/info/data/pub/home.htm>).

25개 유럽국가가 참여하고 핀란드 아카데미에서 주관하고 있는 COGENE은 그동안 인간 유전체 연구에 대한 유럽 네트워크를 구축하며, 유럽연합 내의 유전

**Table 6.** List of cohorts in US NCI cohort consortium

Cohorts	Year began	No. at entry	Female (%)	Ethnicity (%)	Biospecimen n(%)	Types of Biospecimen	No. Cancer (by 2002) n (%)
AARP Diet and Health Study (US)	1995	540,833	40	White (92)	270,000 (50) (planning)	buccal cell (planning)	30,000 (6)
Agricultural Health Study (US)	1994	89,190	35	White (97)		buccal cell	2000 (2)
The Black Women's Health Study (US)	1995	64,500	100	Af-Am (100)		buccal cell	700 (1)
California Teachers Study (US)	1995	133,479	100	White (87)	432 (0.3)	buccal cell cheek brushes	13,400 (10)
Physicians' Health Study I/II (US)	1981/1997	29,571	0	White (94)	I: 15,000 (68) II: 6,150 (82)	plasma & red cells/buffy	3,607 (12)
Clue Cohort (US)	1974/1989	58,000	55	White (98)	58,000 (100)	serum, plasma, RBC, buffy coat	3,263 (6)
Women's Health Study (WHS) (US)	1992	39,876	100	White (94)	28,133 (71)	plasma, buffy coat, RBC	1,699 (4)
PLCO (US)	1993	148,000	50	White (90)	60,000 (41)	plasma, buffy coat, RBC	3,000 (2)
Nutrition Intervention Trials in Linxian (China)	1985	32,902	55	Asian (100)	32,902 (100)	plasma, buffy coat, RBC, toenails	3,656 (11)
Multiethnic Cohort (US)	1993	215,251	55	White (23)	8,900 (4)	serum, plasma, RBC, buffy coat, urine	18,300 (9)
The Shanghai Male Cohort Study (China)	1986	18,244	0	Asian (100)	18,244 (6)	serum, urine	1,178 (6)
The Singapore Chinese Health Study (Singapore)	1993	63,257	56	Asian (100)	1,500 (2)	plasma, serum, RBC, buffy coat, urine	2,355 (4)
Vitamins & Lifestyle Study (VITAL) (US)	2000	75,000	53	White (93)	35,000 (47)	buccal cell	N/A
Shanghai Women's Health Study (China)	1992	75,000	100	Asian (100)	60,000 (80)	plasma, RBC, WBC	1,200 (2)
Prevention Study (CPS) II Nutrition Cohort (US)	1982	184,000	53	White (97)	40,000 (22)	serum, plasma, buffy coat, RBC	20,000 (11)
Health Professionals Follow-up Study (US)	1986	51,529	0	White (97)	33,000 (64)	plasma, WBC, RBC, buccal cell	4,000 (8)
Nurses' Health Study (US)	1976	121,700	100	White (85)	32,826 (30)	plasma, buffy coat, RBC	
Nurses' Health Study II (US)	1989	116,671	100	White (94)	29,500 (25)	plasma, buffy coat, RBC, urine	
EPIC (Europe 8 countries)	1993	480,000	69	White (100)	400,000 (83)	serum, plasma, buffy coat, RBC	15,200 (3)
ATBC Study (US)	1985	29,133	0	White (100)	29,133 (100)	serum, RBC, toenail	7,000 (24)
New York Womens' Health Study (US)	1985	14,275	100	White (85)	14,250 (100)	serum, blood clots	2,000 (14)

체 연구 현황에 대해 조사하고, 대중에 대해 홍보하는 한편, 집단유전학(population genetics)에 대한 워크샵(Co-ordination of European Population Databases for Investigating Genetic Susceptibility to Disease: Oct 17-19, 2002, Cyprus)과 약물유전체학(pharmacogenomics)에 대한 워크샵(The 2nd COGENE workshop on Pharmacogenomics: Sep 26-27, 2003, Frankfurt)등 다양한 워크샵을 계획하고 있다.

인터넷 웹사이트에 참여국가의 유전체 연구 현황을 공유하고 있으며, 또한, 향후 유전체 역학 연구를 위하여 참여 유럽국가 내의 1,000개 이상의 시료를 가지고 있는 코호트(population cohort) 연구기반

을 조사하고 이들의 연합을 시료은행으로 구성하고 있다 (Table 7).

### 3. UK Biobank Project

기존의 유전체 역학 연구는 작은 표본수로 인해 불안정한 결과를 얻거나, 환경적인 노출에 대한 정보가 충분하지 못한 경우가 많으며, 상호작용에 관한 일치된 검정방법이 없고, 대부분 후향적인 환자-대조군 연구설계를 가지는 등의 문제점을 가지고 있었으며, 이를 해결하기 위한 노력으로서 코호트연합이 추진되고 있으나, 연구방법 상의 사용도구 및 연구설계가 상이한 부분이 많다는 점은 여전히 문제점으로 남아있다.

이러한 배경 하에, 영국에서는 Medical

Research Council (MRC), Wellcome Trust Biomedical Charity 및 Department of Health의 연구비 지원 아래 유전자, 생활양식 및 환경적인 요인이 다병인 질병(multifactorial disease) 위험에 어떠한 영향을 미치는지를 규명하고자 하는 목적으로 UK Biobank(2002년-2009년)라는 대규모의 단일 코호트 구축 프로젝트를 진행하고 있으며, 2004년부터 본격적으로 수행될 예정이다(<http://www.ukbiobank.ac.uk/Welcome.htm>).

이 프로젝트는 전국에서 45~69세에 연령대에 해당하는 사람들을 최대 50만 명까지 대표성 있는 표본으로 선정할 계획이며, 이를 위해서 National Health Service (NHS) 참여 의원(General Prac-

tion) 중 500~600곳이 선정될 것으로 추정된다. 연구참여 대상자들로부터 자기기입식 설문을 통해 사회경제적 상태, 인구학적 변수, 생활양식,식이습관(식품빈도 조사지 이용), 임신관련 요인, 질병의 가족력을 조사하며, 개별 인터뷰에서 연구 대상자 자신의 질병력을 조사하고, 향후 추적조사에 대한 동의서를 받으며, 혈압, 체격(키, 몸무게, 허리둘레, 엉덩이둘레)

및 폐활량(FEV1) 검사를 실시한다. 또한, 7일간의 식품섭취를 식사기록법을 통해 조사한다. 한편 다량의 혈액(50ml)을 채취한 다음, 일부는 혈장(plasma)과 buffy coat로, 일부는 전혈로 액체 질소에 보관하고, 향후 유전자를 포함한 다양한 성분을 분석한다. 질병에 대한 추적조사는 암, 심장병, 당뇨, 치매 등을 대상으로 10년 동안 NHS

체계를 기반으로, 통계청(Office of National Statistics) 자료를 통해 질병원인별 사망률과 암 발생률을, 병·의원(General Practice)의 기록을 통하여 상병률을 산출한다. 한편, 노출상태의 변화 및 상병상태를 조사하기 위하여 2년마다 2,000명을 대상으로 재조사를 실시하며, 5년마다 전체 코호트를 대상으로 우편설문을 실시한다.

Table 7. Major European population cohorts/biobanks collected for genetic studies (>1000 samples)

Country	Cohorts/Registry	DNA Biobanks/Subjects
Austria	Austrian Disease and Health Bank Project	Frozen tissues: >20,000; Paraffin embedded tissues >2,5 M; Patient data >600,000
Cyprus	Histocompatibility Laboratory at the Paraskeviadion Transplant Center	Approx. 3,000 DNA samples
Czech	Masaryk Memorial Cancer Institute Cystic Fibrosis Registry Thrombophilia Registry	Breast ca. pts & healthy >1000 DNA samples >2,500 DNA samples >1,000 DNA samples
Estonia	Estonian Genome Project Cancer Register	10,000 participants (DNA & health records) 30,000 samples?
Finland	Finnish Twin Cohort Health 2000 Cohort Northern Finland Cohorts FINRISK 92 FINRISK 97 FINRISK 02 Diabetes Cohort Autism Cohort Psychosis Cohort	170,000 samples 11,500 samples 21,000 samples 8,000 samples (cardiovascular diseases) 10,000 samples (cardiovascular diseases) 11,000 sample (cardiovascular diseases) 6,000 samples 2,000 samples 3,000 samples
Germany	KORA/MONICA EPICS PopGen	>20,000 individuals (for cardiovascular disease & diabetes) >30,000 individuals 2.3 million individuals (for 12 “civilization diseases)
Israel	Ashkenazi Jewish population The Israel Ischaemic Heart Disease Jerusalem Lipid Research Clinic Jerusalem Perinatal Study	several thousands samples (IDgene Pharmaceuticals Ltd) 10,000 civil servant men 7,000 middle aged and 8600 offspring 92,000 children & 40,000 of their mothers.
Latvia	Latvian Diabetes Registry	31,000 patients
Netherlands	Cohorts of elderly subjects Twin Cohorts	from many studies Netherlands Twin Register (adult/young twins): depression, cardiovascular etc.
Norway	CONOR - Cohort of Norway MOBA - Norwegian mother and child cohort	175,000 participants 16,000 mothers
Portugal	Bleeding disorder register Cardiovascular disease register Cystic fibrosis register	2,800 samples 1,746 samples 4,100 samples
Sweden	The Vasterbotten Project Cohort The Northern Sweden Monica Cohort The Vasterbotten Mammary Screening Cohort	70,000 samples 7,000 samples 35,000 samples
UK	Biobank UK MRC - Genetic Materials Collection National Child Development Study EPIC-Norfolk: Prospective study on cancer ALSPAC Million Women Study ProtectT Study	500,000 participants (planning) 14 collections of genetic material (complex disease patients) 17,000 participants 30,000 aged 45-74 yrs 14,000 children born in Bristol area 1991-2 Breast cancer screening programme Prostate cancer study

결과 분석은 각 질병에 대한 코호트내 환자-대조군 연구에 대한 분석을 실시하는데, 이 프로젝트는 상호작용에 대한 분석방법을 multiplicative model에서 벗어나는지 여부를 검토하는 것으로 사전에 결정하고 있을 뿐 아니라, 연구계획 시점부터 주효과(main effect)뿐만 아니라, 상호작용 효과에 대해서도 통계학적 검정력을 충분히 고려하고 있는 것을 장점으로 들 수 있다 [46].

4. Japanese Cohorts for Genomic Epidemiology

일본에서는 현재 JPHC (Japan Public Health Center-based prospective study) (<http://www2.ttcn.ne.jp/~epidemiology/jphc>) [47] 및 JACC (Japan Collaborative Cohort Study) [48]와 같은 대규모의 일 반인구집단 코호트가 구축되어 있으며, 구축된 생체시료를 이용하여 유전체 역학 연구를 수행하고 있다.

JPHC는 식이요인이 암과 심혈관계 질환에 미치는 영향을 알아보기 위하여 구축된 코호트로서 1990년에 40~59세 62,000명(Cohort I), 1993년에 40~69세 79,000명(Cohort II)이 구축되어 총 141,000명의 연구대상자를 확보하고 있다. 기초데이터는 자기기입식 설문지(과거 질병력, 흡연 및 음주, 질병 가족력, 운동, 스트레스 및 사회적인지, 거주력, 직업력, 성격, 임신 관련요인 등)와 식이 섭취 빈도 설문지(46 항목, 4~5개의 빈도 범주, 음식의 분량은 조사하지 않음)를 통하여 조사하였다. 혈액 시료는 건강검진 서비스를 제공하면서 10ml의 혈액을 헤파린 튜브에 채취하였다. 채취한 혈액은 12시간 안에 원심분리한 뒤 혈장과 buffy coat를 -80℃에 보관하였다. 추적조사는 인구 등록자료, 사망진단서, 의무기록을 통하여 실시하되, 5년 또는 10년 마다 재조사를 실시하고 있다.

JACC는 암발생(lung, stomach, pancreas, liver, large intestinal, urinary tract and prostate, gallbladder/bile duct)에 미치는 위험요인 규명을 위한 코호트 연구로서 1988~1990년 사이에 29개 지역에

서 40~79세 110,792명의 주민들로 구축되었으며, 이중 40,000개의 혈청시료를 구축하고 있다. 추적조사(1999년 현재)를 통하여 얻어진 자료에 따르면 총 12,178명이 사망(암사망은 37%를 차지)하였고, 5,228명의 암환자가 발생한 상황이며, 생활요인 및 혈청내 생체지표와 폐암의 연관성에 관한 연구 등을 수행하고 있다.

한편, 2003년부터 다기관 분자역학 코호트(multi-institutional molecular epidemiology cohort) 구축을 추진하고 있다 [49]. 전국적인 코호트 구축에 앞서 히로시마 지역을 대상으로 예비연구의 형태로서 40세 이상 주민 7,000명을 대상으로 착수할 계획에 있으며, 지방 행정기관과 의료기관, 히로시마 대학 간에 긴밀한 협조 하에 수행될 예정이다. 이 연구는 주민들에 대하여 생활양식에 대한 교육을 실시하면서 정기적인 재조사를 병행하는 연구디자인(interactive cohort study design)을 채택하고 생활양식의 변화가 phenotypic 또는 genotypic 지표에 미치는 영향에 대한 연구를 진행한다.

5. Korean Multi-center Cancer Cohort (KMCC)

1993년부터 구축되기 시작한 KMCC 코호트(<http://www.kmcc.or.kr>)는 한국인에서 빈발하는 암의 원인을 역학적으로 규명하기 위한 지역사회 단위 대규모 전향적 코호트이며, 특정 암과 개인 생활습관, 환경적 요인 및 숙주 요인과의 관련성을 역학적으로 규명함은 물론, 암 발생에 관련된 여러 생체지표, 개인 숙주 감수성 및 종양 표지자와의 관련성을 규명함으로써 향후 암의 원인에 관한 새로운 가설을 증명할 수 있도록 생체시료은행을 연구 개시단계에서부터 구축하고 있다. KMCC 코호트 연구진은 국내 및 미국-일본 등에서 개최된 주요 해외학회에서 이 코호트 구축현황을 발표한 바 있으며, 미국 국립암연구소 암유전역학과(Division of Cancer Epidemiology and Genetics) 연구진으로부터 조직은행과 역학적 설문조사 데이터베이스를 갖춘 세계 주요 암 코호트의 하나로 채택된 바 있다 [50].

2002년 현재 15,000명 이상에 대한 설문조사/신체계측/임상검사 자료와 생체시료를 냉동-보관하고 있으며, 추적 연인원은 이미 7만 인-년(person-year)을 상회하고 있다. 추적조사를 위하여 각 지역사회 단위로 암 발생 감시체계를 구축하고 있는데, 의료이용자료와 암등록자료 그리고 사망자료와의 통합 전산검색을 통해 암환자를 색출하는 수동적 감시체계와 더불어 지역사회 단위에서 전화면접 및 직접 조사를 통해 추가로 암 발생환자를 색출하는 능동적 감시체계를 동시에 가동 중에 있다.

결 론

이상에서 세계적인 유전체 역학 연구 동향을 살펴보았는데, 인종적인 차이와 생활습관 등의 국가간 차이를 고려할 때, 국내에서도 대규모 유전체 역학 코호트를 구축할 필요성이 있으며, 이를 통해 충분한 표본수를 대상으로 질병 발생원인과 질병발생 후 생존 및 예후에 미치는 유전자-환경 및 유전자-유전자간 상호작용을 규명할 수 있을 것으로 전망된다. 또한, 기존의 코호트연합은 주로 백인과 흑인을 대상으로 하는 경우가 대부분이므로 아시아인종 간의 코호트 연합(Asian cohort consortium)의 접근이 필요할 것으로 생각된다. 그리고, 유전자형과 표현형 간의 관련성(genotype-phenotype relationship)에 관한 연구, 약물치료에 대한 반응에서 나타나는 개인적인 차이에 관한 연구도 활발히 이루어질 필요성이 있다.

이러한 유전체 역학연구를 통해 질병 발생에 있어서 서구와 다른 유전적 요인 및 환경적 요인을 규명하고, 유전자-환경 및 유전자-유전자 간 상호작용을 파악하여, 궁극적으로 질병 발생기전 중 기존의 위험요인으로 설명될 수 없었던 부분을 이해하는데 도움이 될 것으로 사료되며, 유전자 또는 분자생물학적 표지자를 이용하여 특정 질병발생의 고위험 집단을 찾아내고 효과적인 예방 대책을 수립하는데 기여할 수 있을 것으로 기대된다.



## 참고문헌

- Schulte PA, Perera FP. A conceptual and historical framework for molecular epidemiology. In: Schulte PA, Perera FP, editors. *Molecular epidemiology: principles and practices*, San Diego: Academic Press; 1993. p. 3-44
- Khoury MJ, Beaty TH, Cohen BH. *Fundamentals of genetic epidemiology*, New York, NY: Oxford University Press; 1993. p. 3-25
- Khoury MJ, Dorman JS. The human genome epidemiology network (HuGE Net). *Am J Epidemiol* 1998; 148: 1-3
- Caporaso N. Selection of candidate genes for population studies. In: Vineis P, Malats N, Lang M, d'Errico A, Caporaso N, Cuzick J, Boffetta P, editors. *Metabolic polymorphisms and susceptibility to Cancer*. Lyon: IARC Scientific Publishers; 1999. p. 23-36
- Zheng W. Epidemiological studies of genetic factors for cancer. In: Nasca PC, Pastides H, editors. *Fundamentals of cancer epidemiology*. Maryland: An Aspen Publications; 2001. p. 103-121
- Sherry ST, Ward M, Sirotkin K. Use of molecular variation in the NCBI dbSNP database. *Hum Mutat* 2000; 15(1): 68-75
- Thorisson GA, Stein LD. The SNP consortium website: past, present and future. *Nucleic Acids Res* 2003; 31(1): 124-127
- Fredman D, Siegfried M, Yuan YP, Bork P, Lehvaslaiho H, Brookes AJ. HGVbase: a human sequence variation database emphasizing data quality and a broad spectrum of data sources. *Nucleic Acids Res* 2002; 30(1): 387-391
- Inoko H. PCR-RFLP method holds great promise for complete HLA class II genotyping. *Tissue Antigens* 1990; 36(2): 88-92
- Hayashi K, Yandell DW. How sensitive is PCR-SSCP? *Hum Mutat* 1993; 2(5): 338-346
- Howell WM, Jobs M, Gyllensten U, Brooks AJ. Dynamic allele-specific hybridization. *Nature Biotechnology* 1999; 17: 87-88
- Tapp I, Malmberg L, Rennel E, Wik M, Syvanen AC. Homogeneous scoring of single-nucleotide polymorphisms: comparison of the 5'-nuclease TaqMan assay and Molecular Beacon probes. *Bio-techniques* 2000; 28(4): 732-728
- Hamajima N, Saito T, Matsuo K, Tajima K. Competitive amplification and unspecific amplification in polymerase chain reaction with confronting two-pair primers. *J Mol Diagn* 2002; 4: 103-107
- Nyren P, Karamohamed S, Ronaghi M. Detection of single-base changes using a bioluminometric primer extension assay. *Anal Biochem* 1997; 244(2): 367-373
- Ross P, Hall L, Smirnov I, Haff L. High level multiplex genotyping by MALDI-TOF mass spectrometry. *Nature Biotechnology* 1998; 16: 1347-1351
- Santacrose R, Ratti A, Caroli F, Foglieni B, Ferraris A, Cremonesi L, Margaglione M, Seri M, Ravazzolo R, Restagno G, Dallapiccola B, Rappaport E, Pollack ES, Surrey S, Ferrari M, Fortina P. Analysis of clinically relevant single-nucleotide polymorphisms by use of microelectronic array technology. *Clinical Chemistry* 2002; 48(12): 2124-2130
- Yager JD. Endogenous estrogens as carcinogens through metabolic activation. *J Natl Cancer Inst Monogr* 2000; 27: 67-73
- Madigan MP, Ziegler RG, Benichou J, Byrne C, Hoover RN. Proportion of breast cancer cases in the United States explained by well-established risk factors. *J Natl Cancer Inst* 1995; 87(22): 1681-1685
- Kang D. Genetic polymorphisms and cancer susceptibility of breast cancer in Korean women. *J Biochem Mol Biol* 2003; 36(1): 28-34
- Park SK, Yoo KY, Lee SJ, Kim SU, Ahn SH, Noh DY, Choe KJ, Strickland PT, Hirvonen A, Kang DH. Alcohol consumption, glutathione S-transferase M1 and T1 genetic polymorphisms and breast cancer risk. *Pharmacogenetics* 2000; 10: 301-309
- Yim DS, Park SK, Yoo KY, Yoon KS, Chung HH, Ahn SH, Noh DY, Choe KJ, Cho SH, Jang IJ, Shin SG, Strickland PT, Hirvonen A, Kang D. Relationship between the Val158Met polymorphism of catechol O-methyl transferase and breast cancer. *Pharmacogenetics* 2001; 11: 279-286
- Watanabe J, Shimada T, Gillam EMJ, Ikuta T, Suemasu K, Higashi Y, Gotoh O, Kawajiri K. Association of *CYP1B1* genetic polymorphism with incidence to breast and lung cancer. *Pharmacogenetics* 2000; 10: 25-33
- Zheng W, Xie DW, Jin F, Cheng JR, Dai Q, Wen WQ, Shu XO, Gao YT. Genetic polymorphism of cytochrome P450-1B1 and risk of breast cancer. *Cancer Epidemiol Biomarkers Prev* 2000; 9: 147-150
- Bailey LR, Roodi N, Dupont WD, Parl FF. Association of cytochrome P450 1B1 (*CYP1B1*) polymorphism with steroid receptor status in breast cancer. *Cancer Res* 1998; 58: 5038-5041
- Ko Y, Abel J, Harth V, Brde P, Antony C, Donat S, Fischer HP, Oritz-Pallardo ME, Thier R, Sachinidis A, Vetter H, Bolt HM, Herberhold C, Bruning T. Association of *CYP1B1* codon 432 mutant allele in head and neck squamous cell cancer is reflected by somatic mutations of p53 in tumor tissue. *Cancer Res* 2001; 61: 4398-4404
- Kim SU, Park SK, Yoo KY, Yoon KS, Choi JY, Noh DY, Ahn SH, Choe KJ, Strickland PT, Hirvonen A, Kang D: *XRCC1* genetic polymorphism and breast cancer risk. *Pharmacogenetics* 2002; 12: 335-338
- Park SK, Kang D, Noh DY, Lee KM, Kim SU, Choi JY, Choi IM, Shin AS, Kim JH, Ahn SH, Choe KJ, Hirvonen A, Strickland PT, Yoo KY: Reproductive factors, glutathione S-transferase M1 and T1 genetic polymorphism and breast cancer risk. *Breast Cancer Res Treat* 2003; 78: 89-96
- Lee KM, Abel J, Ko Y, Harth V, Park WY, Seo JS, Yoo KY, Choi JY, Shin A, Ahn SH, Noh DY, Hirvonen A, Kang D: Genetic polymorphisms of *CYP19* and *CYP1B1*, alcohol use, and breast cancer risk. *Br J Cancer* 2003; 88: 675-678
- Choi JY, Abel J, Neuhaus T, Ko Yon, Harth V, Hamajima N, Tajima K, Yoo KY, Park SK, Noh DY, Han W, Choe KJ, Ahn SH, Kim SU, Hirvonen A, Kang D: Role of alcohol and genetic polymorphisms of *CYP2E1* and *ALDH2* in breast cancer development. *Pharmacogenetics* 2003; 13(2): 67-72
- Lee KM, Park SK, Kim SU, Doll MA, Yoo KY, Ahn SH, Noh DY, Hirvonen A, Hein DW, Kang D: N-acetyltransferase (*NAT1*, *NAT2*) and glutathione S-transferase (*GSTM1*, *GSIT1*) polymorphisms in breast cancer. *Cancer Lett* 2003; 196(2): 179-186
- Choi JY, Hamajima N, Tajima K, Yoo KY, Yoon KS, Park SK, Kim Su, Lee KM, Noh DY, Ahn SH, Choe KJ, Han W, Hirvonen A, Kang D: *OGG1* Ser326Cys polymorphism and breast cancer risk among Asian women. *Breast Cancer Res and Treat* 2003; 79(1): 59-62
- Cuzick J. Interaction, subgroup analysis and sample size. In: Vineis P, Malats N, Lang M, d'Errico A, Caporaso, Cuzick J, Boffetta P, editors. *Metabolic polymorphisms and susceptibility to Cancer*, Lyon: IARC Scientific Publishers; 1999. p. 109-121
- Khoury MJ, Little J. Human genome epidemiologic reviews: The beginning of something HuGE. *Am J Epidemiol* 2000; 151(1): 2-3
- Engel LS, Taioli E, Pfeiffer R, Garcia-Closas M, Marcus PM, Lan Q, Boffetta P, Vineis P, Autrup H, Bell DA, Branch RA, Brockmoller J, Daly AK, Heckbert SR, Kalina I, Kang D, Katoh T, Lafuente A, Lin

- HJ, Romkes M, Taylor JA, Rothman N. Pooled analysis and meta-analysis of Glutathione S-transferase M1 and bladder cancer: a HuGE review. *Am J Epidemiol* 2002; 156(2): 95-109
35. Taioli E. International collaborative study on genetic susceptibility to environmental carcinogens. *Cancer Epidemiol Biomarkers Prev* 1999; 8:727-728.
36. Garte S, Gaspari L, Alexandrie AK, Ambrosone C, Autrup H, Autrup JL, Baranova H, Bathum L, Benhamou S, Boffetta P, Bouchardy C, Breskvar K, Brockmoller J, Cascorbi I, Clapper ML, Coutelle C, Daly A, Dell' Omo M, Dolzan V, Dresler CM, Fryer A, Haugen A, Hein DW, Hildesheim A, Hirvonen A, Hsieh LL, Ingelman-Sundberg M, Kalina I, Kang D, Kihara M, Kiyohara C, Kremers P, Lazarus P, Le Marchand L, Lechner MC, van Lieshout EM, London S, Manni JJ, Maugard CM, Morita S, Nazar-Stewart V, Noda K, Oda Y, Parl FF, Pastorelli R, Persson I, Peters WH, Rannug A, Rebbeck T, Risch A, Roelandt L, Romkes M, Ryberg D, Salagovic J, Schoket B, Seidegard J, Shields PG, Sim E, Sinnet D, Strange RC, Stucker I, Sugimura H, To-Figueras J, Vineis P, Yu MC, Taioli E. Metabolic gene polymorphism frequencies in control populations. *Cancer Epidemiol Biomarkers Prev* 2001; 10(12): 1239-1248
37. Vineis P., Marinelli D., Autrup H., Brockmoller J., Cascorbi I., Daly AK., Golka H., Okkels K., Risch A., Rothman N., Sim E., Taioli E. Current smoking, occupation, N-acetyltransferase-2 and bladder cancer: a pooled analysis of genotype-based studies. *Cancer Epidemiol Biomarkers Prev* 2001; 10: 1249-1252
38. Stucker I, Boffetta P., Anttila S., Benhamou S., Hirvonen A., London S., and Taioli E. Lack of interaction between Asbestos Exposure and Glutathione S-Transferase M1 and T1 genotype in lung carcinogenesis. *Cancer Epidemiol Biomarkers Prev* 2001; 10: 1253-1258
39. Lee WJ, Brennan P, Boffetta P, London SJ, Benhamou S, Rannug A, To-Figueras J, Ingelman-Sundberg M, Shields P, Gaspari L, Taioli E. Microsomal epoxide hydrolase polymorphisms and lung cancer risk: a quantitative review. *Biomarkers* 2002; 7(3): 230-241
40. Benhamou S, Lee WJ, Alexandrie AK, Boffetta P, Bouchardy C, Butkiewicz D, Brockmoller J, Clapper ML, Daly A, Dolzan V, Ford J, Gaspari L, Haugen A, Hirvonen A, Husgafvel-Pursiainen K, Ingelman-Sundberg M, Kalina I, Kihara M, Kremers P, Le Marchand L, London SJ, Nazar-Stewart V, Onon-Kihara M, Rannug A, Romkes M, Ryberg D, Seidegard J, Shields P, Strange RC, Stucker I, To-Figueras J, Brennan P, Taioli E. Meta- and pooled analyses of the effects of glutathione S-transferase M1 polymorphisms and smoking on lung cancer risk. *Carcinogenesis* 2002; 23(8): 1343-1350
41. Le Marchand L, Guo C, Benhamou S, Bouchardy C, Cascorbi I, Clapper ML, Garte S, Haugen A, Ingelman-Sundberg M, Kihara M, Rannug A, Ryberg D, Stucker I, Sugimura H, Taioli E. Pooled analysis of the *CYP1A1* exon 7 polymorphism and lung cancer (United States). *Cancer Causes Control* 2003; 14(4): 339-346
42. Hung RJ, Boffetta P, Brockmoller J, Butkiewicz D, Cascorbi I, Clapper ML, Garte S, Haugen A, Hirvonen A, Anttila S, Kalina I, Le Marchand L, London SJ, Rannug A, Romkes M, Salagovic J, Schoket B, Gaspari L, Taioli E. *CYP1A1* and *GSTM1* genetic polymorphisms and lung cancer risk in Caucasian non-smokers: a pooled analysis. *Carcinogenesis* 2003; 24(5): 875-882
43. Taioli E, Gaspari L, Benhamou S, Boffetta P, Brockmoller J, Butkiewicz D, Cascorbi I, Clapper ML, Dolzan V, Haugen A, Hirvonen A, Husgafvel-Pursiainen K, Kalina I, Kremers P, Le Marchand L, London S, Rannug A, Romkes M, Schoket B, Seidegard J, Strange RC, Stucker I, To-Figueras J, Garte S. Polymorphisms in *CYP1A1*, *GSTM1*, *GSTT1* and lung cancer below the age of 45 years. *Int J Epidemiol* 2003; 32(1): 60-63
44. Vineis P, Veglia F, Benhamou S, Butkiewicz D, Cascorbi I, Clapper ML, Dolzan V, Haugen A, Hirvonen A, Ingelman-Sundberg M, Kihara M, Kiyohara C, Kremers P, Le Marchand L, Ohshima S, Pastorelli R, Rannug A, Romkes M, Schoket B, Shields P, Strange RC, Stucker I, Sugimura H, Garte S, Gaspari L, Taioli E. *CYP1A1* T3801C polymorphism and lung cancer: a pooled analysis of 2451 cases and 3358 controls. *Int J Cancer* 2003; 104(5): 650-657
45. Meeting Summary. NATIONAL CANCER INSTITUTE Consortium of Cohorts Minutes, Maryland, November 15-16, 2001 (<http://ospahome.nci.nih.gov/cohort/summaries/nov01summary.html>)
46. Protocol for the UK biobank - a study of genes, environment and health - Feb. 14, 2002. ([http://www.ukbiobank.ac.uk/documents/draft\\_protocol.pdf](http://www.ukbiobank.ac.uk/documents/draft_protocol.pdf))
47. Watanabe S, Tsugane S, Sobue T, Konishi M, Baba S. Study design and organization of the JPHC study. Japan Public Health Center-based prospective study on cancer and cardiovascular diseases. *J Epidemiol* 2001; 11(6 Suppl): S3-7
48. Ohno Y, Tamakoshi A; JACC Study Group. Japan collaborative cohort study for evaluation of cancer risk sponsored by monbusho (JACC study). *J Epidemiol* 2001; 11(4): 144-150
49. Tajima K. Protocol of large-scaled multi-center genome cohort in Japan. Korea-Japan Joint Meeting on cohort study for cancer research. Korea National Cancer Center. Seoul, Korea. March 24, 2003
50. Yoo KY, Shin HR, Chang SH, Lee KS, Park SK, Kang D, Lee DH. Koan multi-center cancer cohort study including a biological material bank (KMCC-1). *Asian Pac J Cancer Prev* 2002; 3(1): 85-92