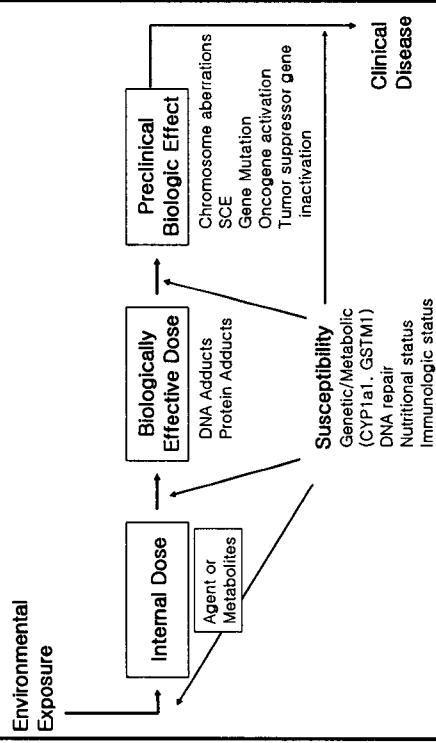


Application of Molecular Biomarkers for Epidemiology

Yun-Chul Hong
Department of Occupational and
Environmental Medicine
Inha University College of Medicine

Molecular Epidemiology

- Molecular epidemiology bridges basic research in molecular biology and studies of human disease causation by combining laboratory measurement of internal dose, biologically effective dose, biologic effects, and the influence of individual susceptibility with epidemiologic methodologies.



Examples of Markers of Internal Dose

Biomarker	Source of Exposure	Biologic Sample	Measurement
Cotinine	Cigarette smoke	Serum, urine, saliva	Higher levels in smokers compared to nonsmokers
Vitamine levels	Diet	Serum	Serum levels of Vitamines A,E,C, and D linked with risk for various cancers
Selenium	Diet	Hair, toe nails	Levels of selenium linked with risk for lung cancer
Levels of benzene And benzene Metabolites	Cigarette smoke	Urine, breath concentrations	Higher levels of benzene and metabolites in smokers compared to nonsmokers
Aflatoxin	Contaminated food	Urine	Higher levels of aflatoxin in urine of exposed compared to nonexposed

Examples of Markers of Biologically Effective Dose

Biomarker	Source of Exposure	Biologic Sample	Measurement
DNA and protein adducts	Cigarette smoke	Blood	More common in smokers than nonsmokers
8-OH-dG	Cigarette smoke	Blood, urine	Higher levels in smokers than nonsmokers

Early Biologic Effect or Response

Compound analyzed	Exposure source	Biologic Sample	Population
Single strand breaks	Styrene	WBC	Workers
Sister chromatid exchange	Various industrial exposures, radiation, air pollution	WBC	Workers, residents
Micronuclei	Organic solvents, heavy metals, cigarette smoke, beetel quid	WBC, oral mucosa	Workers
Chromosomal aberrations	Various industrial exposures, radiation, air pollution	WBC	Workers, residents
HPRT mutation	Chemotherapeutic agents, radiation	WBC	Patients, workers
GPA mutation	Chemotherapeutic agents, radiation	RBC	Patients, Japanese atom bomb survivors
Mutation in tumor suppressor genes	AFB1	Tumor tissue	Patients
Oncogene activation	PAH, cigarette smoke	Serum	Workers, cancer patients

Biomarker study - Example

Exposure Biomarkers

PAH Exposure biomarker (1-OHP, 2-Naphthol)
 VOC exposure biomarker (Hippuric acid, Muconic acid)
 Heavy metal exposure biomarker (Cadmium)
 Smoking (Cotinine)

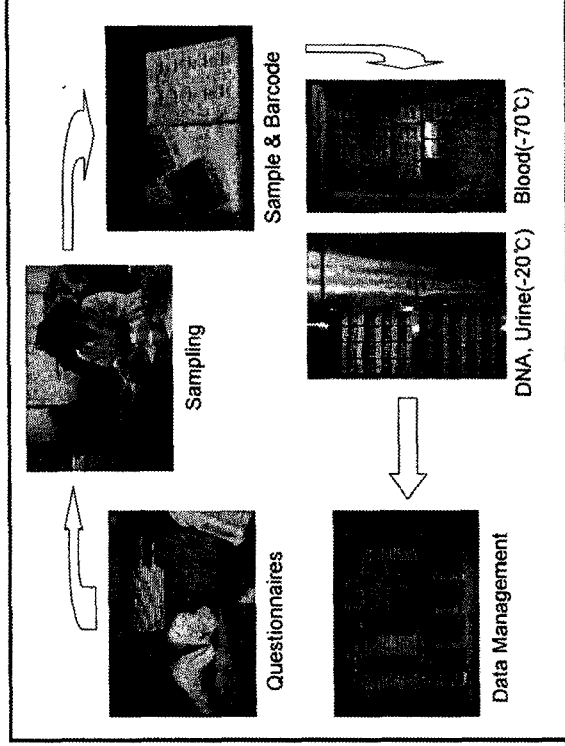
Early Biological Effect Biomarker

Oxidative injury biomarker (8-OH-dG, MDA)

Susceptibility biomarkers

- Xenobiotic metabolism enzymes
 - GSTs (glutathione S-transferases) M1/Π1/P1
 - CYP (cytochrome P-450)1A1, CYP2E1, CYP2A6
 - NAT2, UGT1A6
- Oxidative stress-related enzymes
 - MPO (myeloperoxidase)
 - MnSOD (manganese superoxide dismutase)
 - GPX1 (glutathione peroxidase)
- DNA repair genes
 - XRCC1 (X-ray repair cross-complementing 1)
 - hOGG1 (8-hydroxydeoxyguanosine DNA glycosylase/apurinic lyase)

Basic Structure of Genomic Cohort Study



Specimens for DNA Banking for Epidemiologic Studies

Specimen Type	Blood spots
DNA Yield	12-42 ng/ μ l (adults) 43-78 ng/ μ l (neonates) 1/4 -inch punch from 75 μ l volumn yields about 12 μ l of blood
Advantages	Small sample size Ease of sample collection Ease of shipping (regular mail) Stability and low cost storage Offers a source for study of exogenous or endogenous compounds other than DNA (Genotyping generally requires 10 ng/genotype, and with current technology as little as 2.5 ng per SNP so that scores to hundreds of genotypes could be obtained from one blood spot)
Disadvantages	Low DNA yield : may not be suitable for whole-genome amplification Nonrenewable Smaller amplicons

Specimens for DNA Banking for Epidemiologic Studies

Specimen Type	Blood cells - Whole blood anticoagulated or blood clots - Buffy coat
DNA Yield	100-400 μ g/10 ml \approx 200 μ g/ml
Advantages	Relatively low-cost storage (-80°C) Yield large quantities of high-quality genomic DNA Offers a source for study of exogenous or endogenous compounds other than DNA
Disadvantages	Invasive sample collection Shipping (special requirements) Nonrenewable

Specimens for DNA Banking for Epidemiologic Studies

Specimen Type	Transformed lymphocytes
DNA Yield	10 ⁶ Cells = 6 µg 1-2 × 10 ⁸ Cells = 5-10 µg
Advantages	Renewable source of DNA Yields large quantities of High-quality, genomic DNA
Disadvantages	Labor-intensive preparation High-cost storage (liquid nitrogen and periodic reculture) Does not offer a source for study of exogenous or endogenous compounds other than DNA or RNA

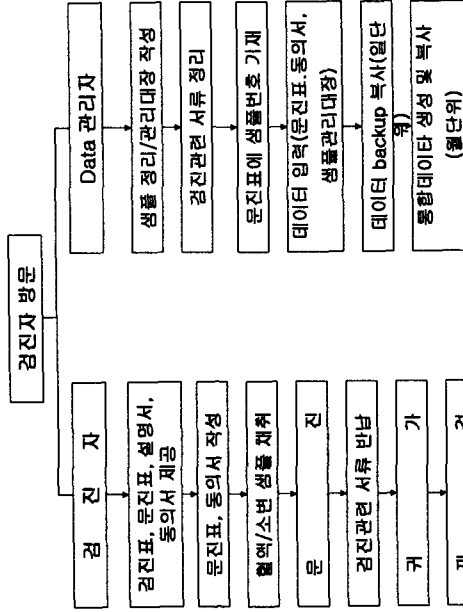
Specimens for DNA Banking for Epidemiologic Studies

Specimen Type	Buccal cells
DNA Yield	49.7 µg mean ; 0.2-134 µg range (mouthwash , total DNA) 12-60 µg range (mouthwash , total DNA) ≒ 16-30 µg median ; 1-290 µg range (mouthwash , hDNA) 32 µg median ; 4-196 µg range (mouthwash , hDNA) 1-1.5 µg/2 cytobrushes median ; 6 ng-13 µg range (hDNA) 1-2 µg/2 cytobrush (total DNA) 1-2 µg/2 swab (total DNA)
Advantages	Noninvasive collection Ease of sample collection (allows participant to collect and mail specimen) Genotyping generally requires 10 ng/genotype, and with current technology as little as 2.5 genotypes could theoretically be obtained from a buccal cell specimen
Disadvantages	Low DNA yield : not in general use for whole genome amplification Highly variable yield Does not offer a source for study of exogenous or endogenous compounds other than DNA or RNA Bacterial contamination must be addressed

Storage of Other samples

- Serum
- Plasma
- Urine
- Others (Mouthwash, Sputum, Hair, ...)

Data Flow



Data Backup & Storage

- Electronic data
 - Daily backup copy
 - Monthly 통합file 생성
 - Monthly Data 관리 보고서
- Data sheets
 - Informed consent forms
 - 문진표
 - 검진표
- Sample
 - Whole blood
 - Urine

Data Quality Control

- Problem data 확인
 - Missing values
 - Out-of-range values
 - Illogical data(sample number, ID, etc)
 - Inconsistent data
- Data base history table
 - Data field, date of update, user, previous value, updated value
 - electronic audit trail

Data Security

- Stand-alone PC
- 사용자 접근 권한 설정
 - 사용자별 접근가능한 Data 제한 (단순조회, 편집기능 등)
 - 사용자별 ID, Password로 접근 제한
- 암호화
- Data 갱신기록 유지

Basic Structure of Cohort

- Steering Committee (PI, Epidemiology, Biostatistics, Laboratory technician, Data manager)
- Questionnaire, Sample Collection and Storage
- Sample Analysis (Clinical Data, Genetic Data)
- Data Management (Data Entry and Merge, Data Storage, Security)

Setup of Cohort

- Approval of the Cohort Study by Hospital IRB
- Setup of Information System (Barcode, Clinical Information of Health Examinees, Database Entry)
- Setup of Sampling and Storage
- Establishing Database Management
- Initial Evaluation