

*How do We Assess the Contribution of Complex Genotypes to the Population Burden of Common Diseases?*



**SAFER • HEALTHIER • PEOPLE™**



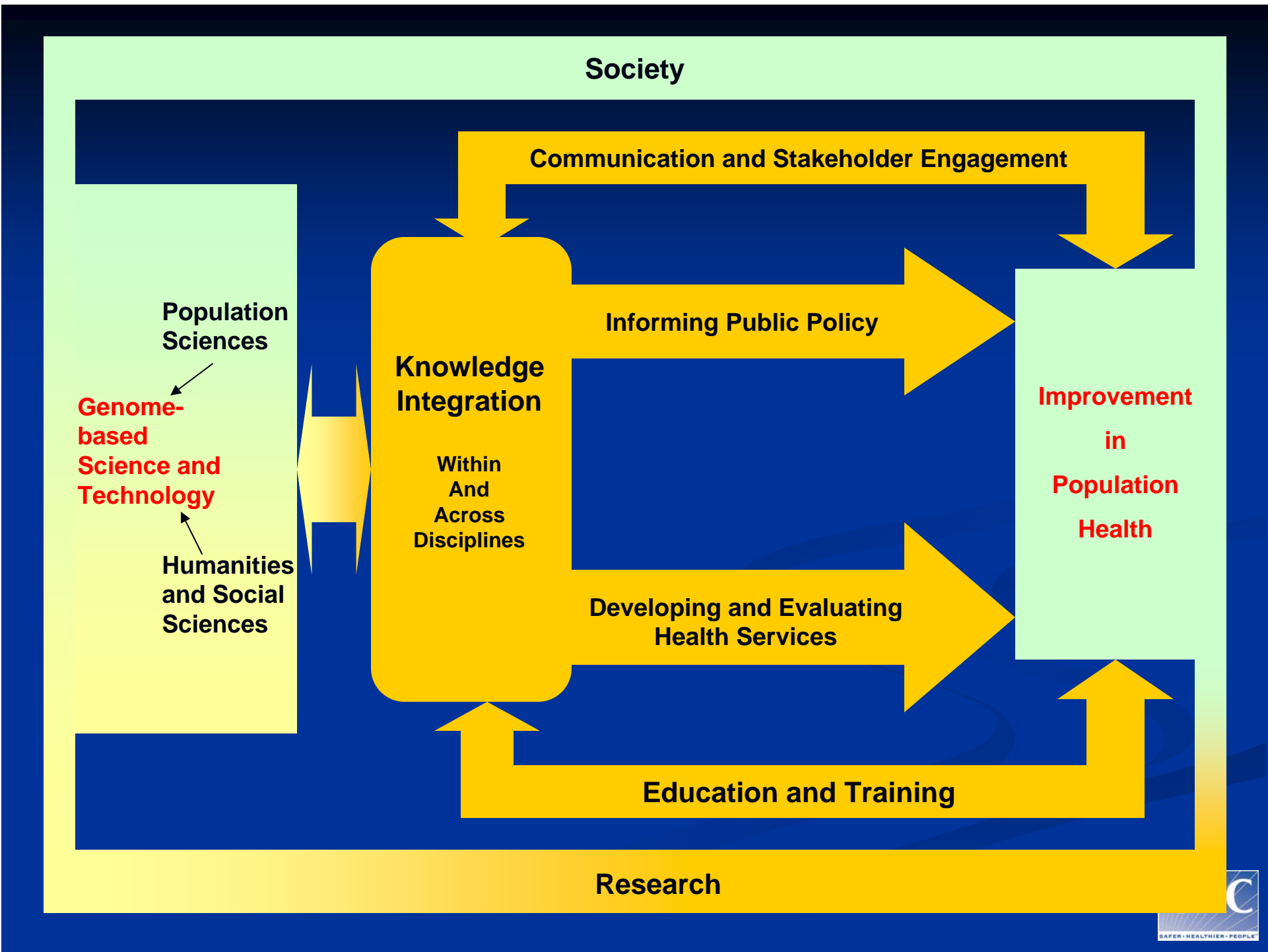
# *The Public Health Genomics Enterprise*

**Genome-  
based  
Science and  
Technology**

**Improvement  
in  
Population  
Health**

**Closing the Gap Between Gene Human  
Genome Discoveries and Population Health**





# Society

Communication and Stakeholder Engagement

Population Sciences

Genome-based Science and Technology

Humanities and Social Sciences

Last month, Dr Stephen Channock gave us a primer on Genome-based science and Technology

ment  
tion  
th

Education and Training

Research



*What do we do with a “gene”  
when we find one?*

*What do we do with a “gene”  
when we find one?*

The Emergence of Epidemiology  
in the Genomics Age!

# Society

Communication and Stakeholder Engagement

Population Sciences

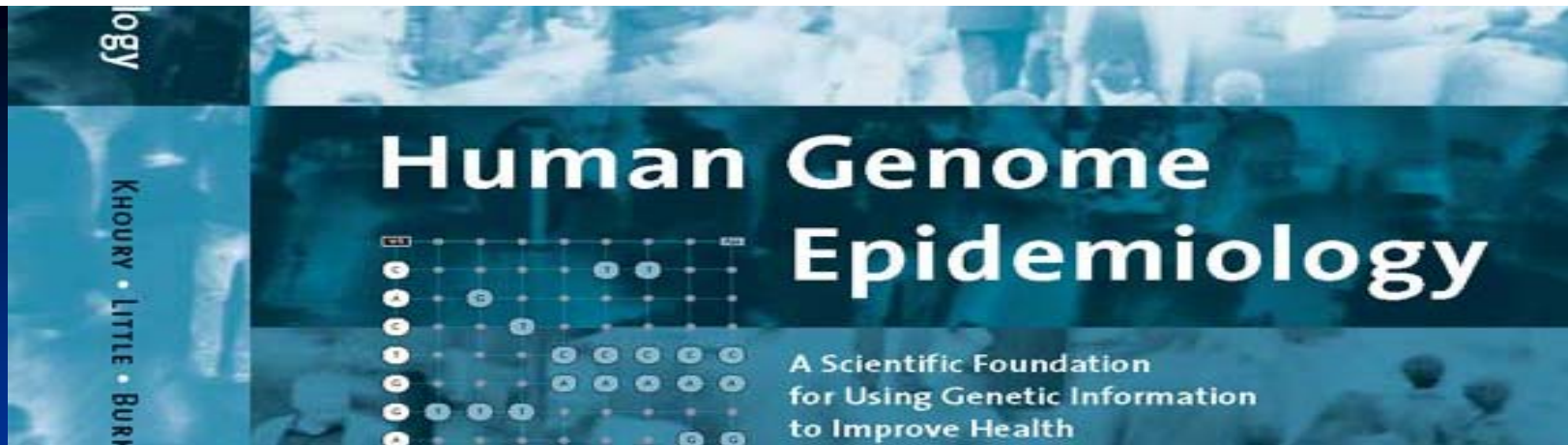
Genome-based Science and Technology

Behavioral and Social Sciences

Today, we will explore with our two speakers the role of population sciences with a focus on epidemiology

Education and Training

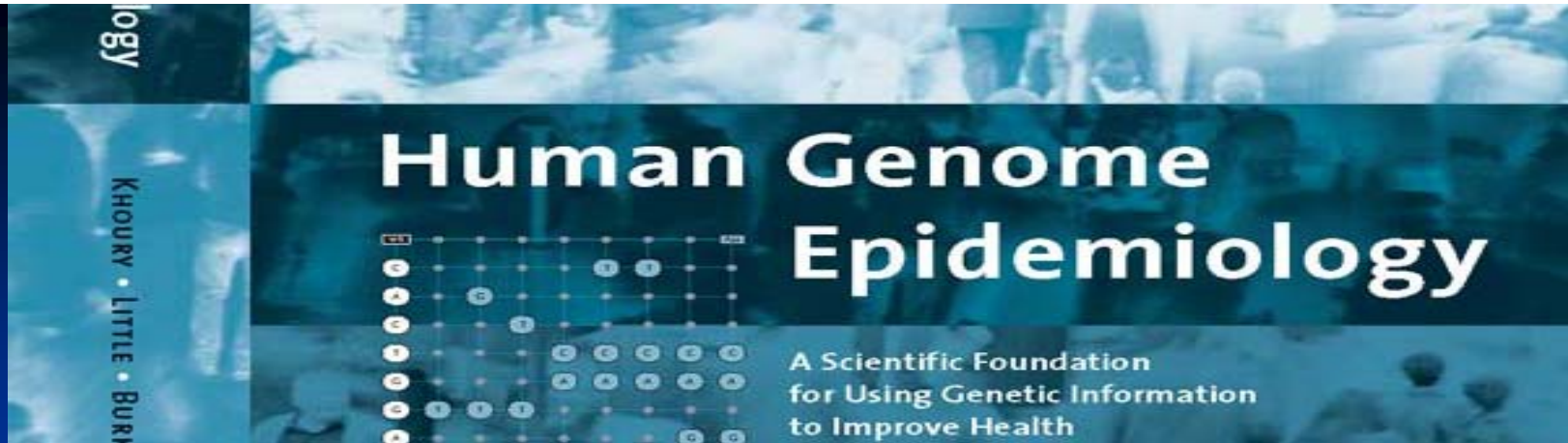
Research



“Systematic application of epidemiologic methods and approaches to assess the impact of human genetic variation on health and disease”

Khoury, Little and Burke, HuGE 2004





“Systematic application of epidemiologic methods and approaches to assess the impact of human genetic variation on health and disease”

Khoury, Little and Burke, HuGE 2004

HuGE problem:  
25,000 genes, their  
combinations and  
interactions with risk  
factors

# *Human Genome Epidemiology: From Gene Discovery to Action*

*Genetic Epidemiology with a Capital E---D. Thomas*

- Gene Discovery (classical genetic epidemiology)

# ***Classical Genetic Epidemiology***

*(MC King, 1984)*

- Do diseases cluster in families?
- Causes of familial aggregation
- Search for genetic mechanisms
- Comparing disease risks among relatives of affected patients with relatives of unaffected persons
- Assessing risk factors; heritability analysis
- Segregation, linkage, mapping, TDTs...

# *Human Genome Epidemiology: From Gene Discovery to Action*

*Genetic Epidemiology with a Capital E---D. Thomas*

- Gene Discovery (classical genetic epidemiology)
- Assessing Genomic Effects on Population Health (molecular epidemiology)

# *Molecular Epidemiology: Assessing Genomic Effects on Population Health*

- Prevalence
- Assessing Disease burden
  - Relative risk
  - Absolute risk
  - Attributable fraction
- Gene-Gene and Gene-Environment interaction



# *The Importance of Population Data*

- **Issues in Prevalence**
- Non representative populations
- Too much focus on allele frequency
- Important information for research and practice

# NHANES III DNA BANK

## Prevalence of Genes of Public Health Significance

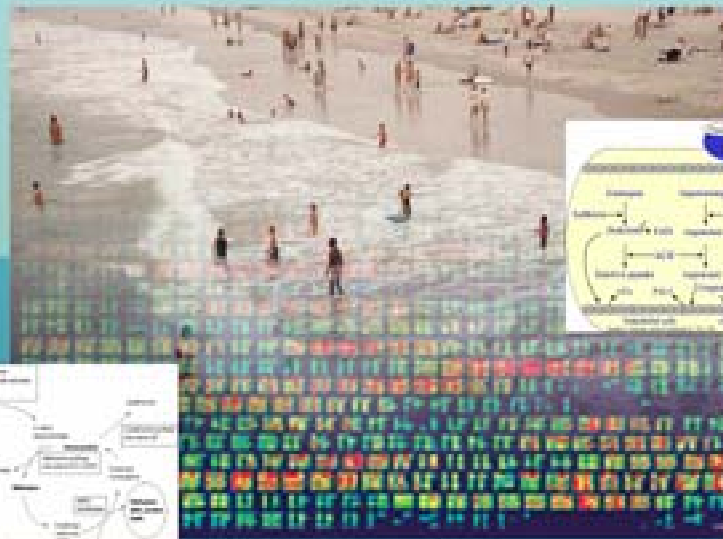
### Background

#### NHANES III DNA Bank

- > National Health and Nutrition Examination Survey (NHANES) is a nationally representative survey
- > Detailed interviews, clinical, laboratory and radiologic examinations are conducted
- > Phenotypic data, such as serostatus for many infectious exposures, blood count, chemistries, etc. were collected
- > During second phase NHANES III (1991-1994), white blood cells were frozen and cell lines were immortalized with EBV
- > NHANES III DNA bank is located at NCEH/CDC, with specimens available from over 7000 participants
- > In 2002, NCHS announced a call for proposals to use these specimens in the Federal Register

### Collaborative CDC-wide Proposal Objective

- > Determine the prevalence of genotypes of public health importance



### Criteria for Genetic Variants

#### Public Health Importance

- > Known or hypothesized association with diseases of public health importance
- > Role in pathways affecting multiple diseases
- > Identified functional variants
- > Relatively common (i.e., >2.0%)
- > Previously described gene-environment or gene-gene interactions
- > Relevant phenotypic data available in NHANES dataset
- > No current use for clinical risk assessment or intervention

### Challenges to Identifying Genes of Public Health Importance

- > Gaps in information in the literature
- > Methodological issues of many available studies
  - Selection bias, power, interaction
- > Non-replication of gene-disease associations



### Public Health Significance of Proposal

- > Prevalence of gene variants
  - Basis for estimating population attributable fraction in combination with measure of gene-disease association
  - Enable assessment of potential for screening population subgroups for susceptibility genes
  - Prevalence of combinations of variants in pathways and at different loci
- > Examine gene-disease association, gene-environment and gene-gene interactions

### Selected Pathways of Gene Variants (87 variants of 57 genes)

- Nutrient Metabolism (e.g., folate and homocysteine, lipids, glucose, alcohol, vitamin D)
- Immune and Inflammatory responses (e.g., cytokines, receptors)
- Activation and detoxification pathways (e.g., drugs, carcinogens, environmental contaminants)
- DNA repair pathways (e.g., ionizing radiation, environmental toxins)
- Hemostasis pathway and renin-angiotensin (e.g., vasomotor) pathway
- Developmental (e.g., hearing loss)

### Laboratory Methods

- Genotyping
  - Assessing Capability of External Laboratories to conduct high throughput, accurate, low-cost, genotyping for >800,000 SNPs (~7300 specimens X 87 variants\*)

\*Table of gene variants available upon request



### Next Steps

- Pending approval from NCHS
- Laboratory Selected
- Genotype-Phenotype analysis

### CDC Working Group

ATSDR: Olivia Harris NCEH: Karen Abe, Cynthia Moore, Lorenzo Botto, Quanhe Yang NCHSTP: Mary Raichler NCI: Tom Hodges, Craig Hooper, Jai Lingappa, Janet McNeill, Anne Diley  
 NCEH: Amanda Brown, Peg Gallagher, Maria Gwinn, Omar Henderson, Bruce Lin, Mary Lou Lindgren, Julian Little, Karen Steinberg  
 NCCDPHP: Heidi Blain, Wayne Giles, Ingrid Hall, Giuseppina Imperatore, Ann Malarcher NIOSH: MaryAnn Butler, Ainsley Weston PHPPO: Bin Chen, NIP: Scott Campbell NCHS: Gery McQuillan



# *The Importance of Population Data*

## ■ Issues in Prevalence

- Non representative populations
- Too much focus on allele frequency
- Important information for research and practice

## ■ Issues in Associations

- Family studies overestimate risk
- Convenience studies may lead to spurious findings
- Important information for research and practice



# ***Association Studies***

- Used to be a dirty word in genetics-Linkage vs association
- Modern view of association studies as means to “discover” and “characterize” genes in populations
- Strong epidemiological basis for association studies
- Concept of “Mendelian Randomization”

# *The End of Black Box Epidemiology?*

## Risk Factors

Demographics

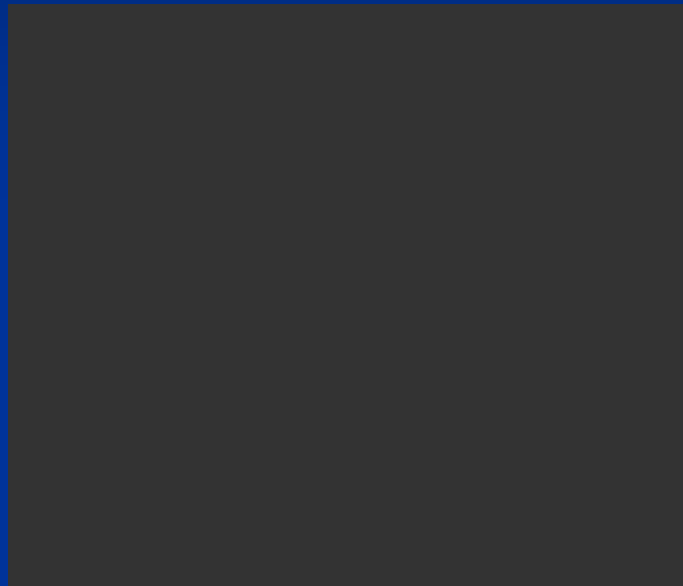
Diet

Occupation

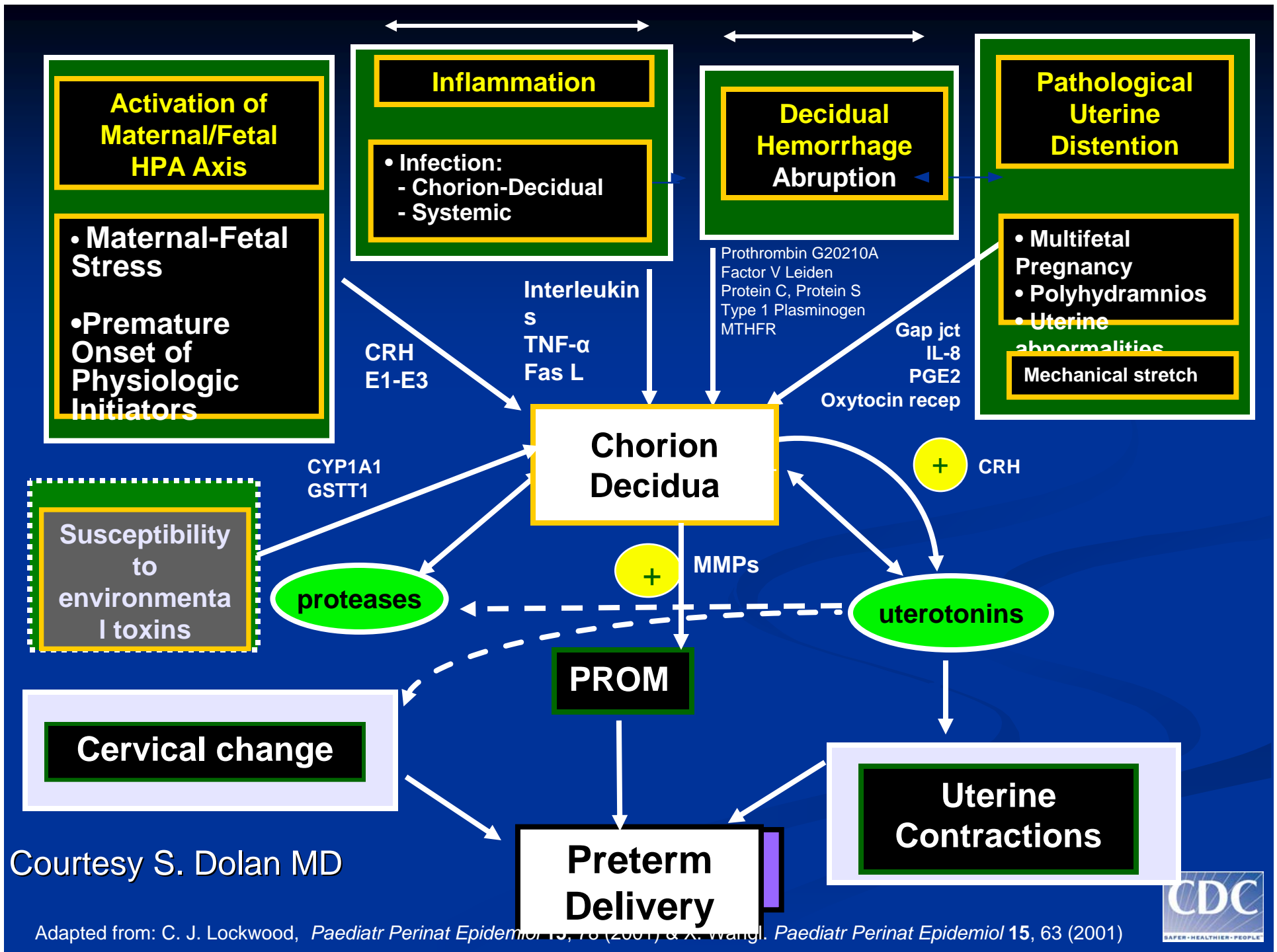
Smoking

Alcohol

Environment



**Adverse  
Health  
Outcomes**



## *Number of Published HuGE Papers\* 2001-2006*

■ Year	Prevalence	Associations	Interactions
■ 2001	308	2141	436
■ 2002	349	2799	569
■ 2003	328	3021	600
■ 2004	430	3772	664
■ 2005	404	4486	885
■ 2006	365	4761	905

\* Data from CDC HuGE Published Literature Database  
January 25, 2007



# *Human Genome Epidemiology: From Gene Discovery to Action*

*Genetic Epidemiology with a Capital E---D. Thomas*

- Gene Discovery (classical genetic epidemiology)
- Assessing Genomic Effects on Population Health (molecular epidemiology)
- Assessing Genetic Tests for Screening and Prevention (applied epidemiology and health services research)

Society

Communication and Stakeholder Engagement

Population Sciences

Genome-based Science and Technology

Humanities and Social Sciences

Today's Speakers  
Robert Hoover : Cancer genetic epidemiology  
Teri Manolio: Case-control and cohort studies  
in the era of genome wide associations

Education and Training

Research