What Do We Know About Individual Variability and Its Contribution to Disease?

Nathaniel Rothman, MD, MPH, MHS
Occupational and Environmental Epidemiology Branch,
Division of Cancer Epidemiology and Genetics,
NCI, NIH, DHHS
Single Nucleotide Changes

10-12 Million Common SNPs
> 10% MAF

30-50 Million Uncommon SNPs
1-10% MAF

>100 Million Rare SNPs
< 1% MAF

All 3.1 Billion bp in Human Genome?

GWAS: Current chips
Interrogate ONLY
Common SNPs (> 10%)
Common Variation (MAF > 10%) Represents a Small Part of All Variation

SNPs by observations in CEU

Count (x10^6)

Minor Alleles

0 10 20 30 40 50 60

0.0 0.2 0.4 0.6 0.8 1.0

HapMap
1000 Genomes
Published Cancer GWAS Etiology Hits: April 1, 2012

~230 Disease Loci marked by SNPs
1 Locus marked by a CNV
Integrating Knowledge of Human Variability into Studies of Exposure and Disease

- Obtain mechanistic insight
- Clarify dose-response relationships, and more effectively evaluate low levels of risk
- Identify new environmental health hazards
- Develop more effective prevention, screening, and treatment strategies
Susceptibility for Bladder Cancer as a model for Gene-Environment Interactions
Spanish Bladder Cancer Study

- Hospital-based case-control study (1998-2001)
- DNA from 1,150 cases and 1,149 controls
## Data Collection

<table>
<thead>
<tr>
<th>Data Resources</th>
<th>Response Rate</th>
<th>Specific Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAPI</td>
<td>86%</td>
<td>Demographics</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Smoking</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Occupation/Environmental Family history</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Medical/Drugs</td>
</tr>
<tr>
<td>Blood/Buccal Cell</td>
<td>95%</td>
<td>Genetic Susceptibility</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Functional Assays</td>
</tr>
<tr>
<td>Diet Qx.</td>
<td>72%</td>
<td>Fluid intake</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Food Frequency</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Food Carcinogens</td>
</tr>
<tr>
<td>Urination Diary</td>
<td>60%</td>
<td>Urine pH</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Urinary freq</td>
</tr>
<tr>
<td>Toenails</td>
<td>77%</td>
<td>Arsenic/Selenium</td>
</tr>
<tr>
<td>Hair dye Qx.</td>
<td>85%</td>
<td>Hair Dye</td>
</tr>
</tbody>
</table>
Disinfection Byproducts, \textit{GSTT1/GSTZ1}, and Bladder Cancer

- Ubiquitous in all water supplies disinfected with chlorine

- Precursor levels much higher in surface than ground water sources, creating a major differential in exposure

- A complex mixture of halogenated organics with mutagenic properties (one class called Trihalomethanes – THM)
Exposure Assessment for Disinfection Byproducts

- Lifetime residential / water source history gathered from each study subject
- Detailed information on water treatments & sources gathered from all major water utilities in study area – average annual THM level at each utility assigned by expert evaluation
- Personal & utility information merged to create a lifetime year-by-year estimate of THM exposure for each study subject
- Long-term average, cumulative, peak, etc. exposures were calculated
GSTT1, GSTZ1 & Disinfection Byproducts

**GSTT1**

Mutagenic activation of trihalomethanes (THM) GSTT1 null deletion eliminates activity

**GSTZ1**

Clears haloacetic and other dihalogenated acids: The CT/TT: M82T (exon7+29C>T, rs 1046428) variant has decreased activity
Trihalomethane Concentration in Drinking Water and Risk of Bladder Cancer

OR=1.8

Villanueva et al., Am J Epidemiol 2007
Interaction between THM and GSTT1, \( p \) (interaction) = 0.021

Cantor et al., EHP, 2010
Interaction between THM and GSTZ1*

\[ p \text{ (interaction)} = 0.018 \]

\[ \text{OR} = 2.9, \quad p \text{ (trend)} = 0.0043 \]

\[ \text{OR} = 2.5, \quad p \text{ (trend)} = 0.28 \]

Cantor et al., EHP 2010

*CT/TT: M82T (exon7+29C>T, rs 1046428)
Trihalomethane Concentration, Combined GSTT1 +/++ and GSTZ1 CT/TT Genotype, and Bladder Cancer; p (interaction) = 0.005

OR=5.9

Cantor et al., Environ Health Perspect 2010
Study Design of Multi-stage NCI GWAS of Bladder Cancer

Includes ~12,000 cases and ~53,000 controls from 20 studies

Rothman et al., Nat Genet 2010
Bladder Cancer Susceptibility Loci

**Previous GWAS**
- 8q24.21
- 3q28 (*TP63*)
- 5p11.3 (*TERT-CLPTM1L*)
- 8q24.23 (*PSCA*)
- 4p16.3 (*TMEM129 TACC3-FGFR3*)

**NCI GWAS**
- 1p13.3 (*GSTM1*)
- 8p22 (*NAT2*)
- 2q37.1 (*UGT1A*)
- 19q12 (*CCNE1*)
- 22q13.1 (*CBX6, APOBEC3A*)

Rothman et al., Nat Genet 2010

**NCI + MD Anderson GWAS**
- 18q12.3 (*SLC14A1*)
- Garcia-Closas et al., Hum Mol Genet, 2011
In the Bladder Cancer GWAS, 7 out of 12 SNPs showed significant additive interactions with tobacco.

<table>
<thead>
<tr>
<th>Location</th>
<th>Gene</th>
<th>H^2</th>
<th>Cases</th>
<th>Controls</th>
<th>Var^2</th>
<th>Observed OR^2</th>
<th>Expected OR^2</th>
<th>P-value Interaction^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chr 6p21</td>
<td>SMARC</td>
<td>0.38</td>
<td>3.92</td>
<td>5.062</td>
<td>0.77</td>
<td>0.87</td>
<td>2.04</td>
<td>3.39</td>
</tr>
<tr>
<td>Chr 12q13</td>
<td>BUB1B</td>
<td>0.37</td>
<td>3.95</td>
<td>5.062</td>
<td>0.77</td>
<td>0.87</td>
<td>2.04</td>
<td>3.39</td>
</tr>
<tr>
<td>Chr 8q21</td>
<td>EGFR</td>
<td>0.37</td>
<td>3.95</td>
<td>5.062</td>
<td>0.77</td>
<td>0.87</td>
<td>2.04</td>
<td>3.39</td>
</tr>
<tr>
<td>Chr 10p13</td>
<td>ESR1</td>
<td>0.37</td>
<td>3.95</td>
<td>5.062</td>
<td>0.77</td>
<td>0.87</td>
<td>2.04</td>
<td>3.39</td>
</tr>
<tr>
<td>Chr 1q12</td>
<td>PCCL2</td>
<td>0.37</td>
<td>3.95</td>
<td>5.062</td>
<td>0.77</td>
<td>0.87</td>
<td>2.04</td>
<td>3.39</td>
</tr>
<tr>
<td>Chr 22q11</td>
<td>GRB10</td>
<td>0.37</td>
<td>3.95</td>
<td>5.062</td>
<td>0.77</td>
<td>0.87</td>
<td>2.04</td>
<td>3.39</td>
</tr>
<tr>
<td>Chr 1q12</td>
<td>PCCL2</td>
<td>0.37</td>
<td>3.95</td>
<td>5.062</td>
<td>0.77</td>
<td>0.87</td>
<td>2.04</td>
<td>3.39</td>
</tr>
<tr>
<td>Chr 1q12</td>
<td>PCCL2</td>
<td>0.37</td>
<td>3.95</td>
<td>5.062</td>
<td>0.77</td>
<td>0.87</td>
<td>2.04</td>
<td>3.39</td>
</tr>
</tbody>
</table>

Odds ratios for joint associations of smoking status (ever vs never smokers) and 12 susceptibility variants with bladder cancer risk.

Garcia-Closas et al., Submitted
The Level of Risk Discrimination Could be of Public Health Relevance

Cumulative 30-year absolute risk for bladder cancer in a 50 year old male in the USA, overall and by quartiles of a polygenetic genetic score.

P-additive = $1 \times 10^{-4}$

RD = 8.2%

RD = 2.0%

RD are risk differences for current vs never smokers

Garcia-Closas et al., Submitted
Impact of Eliminating Smoking in 100,000 Current Smokers in Highest vs. Lowest Genetically Susceptible Subgroup of the Population

- 8,000 vs. 2,000 cases of bladder cancer eliminated
Tobacco smoking

Occupational exposures:
e.g., aromatic amine dyes, cutting oils

Water contaminants:
arsenic disinfection by-products

Liver
NAT2
GSTM1
GSTT1
UGT1A6

Kidney
SLC14A1

Urinary Bladder?
CCNE1
CBX6, APOBEC3A
TMEM129 TACC3-FGFR3

8q24: PSCA, MYC
TP63

Urine pH<6
Voiding frequency

pre-carcinogens
carcinogenic metabolites

Urinary Bladder
UGT1A6

TP63
Integrating Knowledge of Human Variability into Studies of Exposure and Disease

• Obtain mechanistic insight

• Clarify dose-response relationships, and more effectively evaluate low levels of risk

• Identify new environmental health hazards

• Develop more effective prevention, screening, and treatment strategies
Advances will be accelerated by "Collective Intelligence"

"I not only use all of the brains I have, but all I can borrow"

Woodrow Wilson
Acknowledgements – Studies in Europe and US, Genomics, and Analysis

NCI
D. Silverman
M. Garcia-Closas
N. Chatterjee
K. Cantor
L. Prokunina
K. Jacobs
D. Baris
J. Figueroa
S. Chanock

New England
M. Schwenn
A. Johnson
M. Karagas

ACS
M. Thun
E. Jacobs
R. Diver

MD Anderson
Xifeng Wu

Harvard
I. De Vivo

Spain
N. Malats
F. Real
M. Kogevinas