Genetic susceptibility to the effects of environmental exposure to arsenic

BRANDON L. PIERCE
Assistant Professor
Department of Public Health Sciences
Department of Human Genetics
University of Chicago
Take-home messages

We all inherit slightly different versions of the human genome

These differences can affect health, including susceptibility to the effects of environmental exposures

For arsenic, we have identified genetic features associated with arsenic susceptibility through their effects on arsenic metabolism efficiency

Understanding the underlying biology is challenging
Overview

• Gene-environment interaction (GxE)
• Human genetic variation
• Genetics of arsenic metabolism and toxicity
• Moving from correlations to biology
• A few new projects
Susceptibility (or Resiliency) to harmful exposures
Somebody said to me this morning, 'To what do you attribute your longevity?' I don't know. I mean, I couldn't have planned my life out better. By all accounts I should be dead! The abuse I put my body through: the drugs, the alcohol, the lifestyle I've lived the last 30 years!

(Ozzy Osbourne)
The Gene x Environment Interaction (GxE)

Interpretation: The effect of an exposure on disease risk is modified by “genotype” status (i.e., “effect modification).
The Human Genome

- Cytosine
- Guanine
- Adenine
- Thymine
The human genome sequence consist of ~3 billion nucleotide pairs

...AGCTTCGCTCTCGGGAGAGAATTACAGGATCATATAATTTCG
AAGGGGTCCCGCTTTATTTCCCAGATGCATGCAAAGATTTCGATTCTC
GATATGCTACTTTAGGCTAAAGCTTCTTCGGCTCTCGGAGAGAATT
ACAGGATCATATAATTTCGAAAGGGGTCTATTATTCCCGATGCATG
GCAAAGATTTCGATTCTTTCTTCGGATATGCTTAGGCTAAGCTTTCTCG
GCTCTCGGGAGAGAATTACAGGATCATATAATTTCGAAAGGTC
TATTATTCCCGATGCAATGAAAGATTTCGATTCTTTCTTCGGATATGCT
AGGCTAGTTTTTTCGGGGGGCTTAAAAATGGGGGCTATGCATCGA
TTTCGATTTCGATGCAGCAGCATGCAGTCATGCAGCAGCTGGGCGGCTCGTGCATGATAAACGACAGGCCAAAAACACTAAAACGCTATGGATGCATGCATGCAGGGCTAGCATGCATGATATATCGCGCCTATTTATTGCTAGGCTAAGCTTCTCAGG...
There are ~10 million sites that vary among individuals (common variants)

SNP: Single Nucleotide Polymorphism
Two “alleles”: A and T
Most SNPs reside outside of protein-coding regions (i.e., genes)
How do we study genetic susceptibility to disease?

- Mid-2000s: technologies for measuring ~1 million of SNPs simultaneously became widely available
- This created the **Genome-Wide Association Study (GWAS)**
- GWAS is very simple:
  1. measure one million SNPs and one disease trait
  2. Test each SNP for association with disease, and correct for multiple testing issue
- Ideal for studying common genetic variation and common diseases
- GWAS has identified thousands of trait-association SNPs
GxEs have been hard to find!

- GWAS-identified variants show little evidence of interaction with risk factors
- Very few well-established examples
- Why so hard to find?
  - Poorly measured exposures?
  - Poor power for GxE tests?
  - Lots of options statistical modelling; which is optimal?
  - Perhaps GxE is an uncommon phenomenon

- Despite these challenges, there is still strong interest in GxE
  - Understanding mechanisms and pathways
  - Identify new genes involved in disease
  - Identify susceptible subgroups and establishing standards that protect them
Arsenic Epidemiology: a unique setting for GxE research

- **HEALS cohort** (Health Effects of Arsenic Longitudinal Study): Araihazar, Bangladesh (2001-present)

- Wide range of arsenic exposure through drinking water (>20,000 participants)

- Exposure is well-measured (drinking water and urine)

- Arsenic has common signs of toxicity (skin lesions) measured every 2 years

- Arsenic metabolism efficiency can be quantified

- Inter-individual variation in metabolism efficiency should...
  - Reflect differences in internal dose
  - Effect risk for arsenic-related disease

- **Genetic factors that influence arsenic metabolism should show GxE**

van Geen, et al
Inorganic Arsenic (iAs)

iAs absorbed into blood from GI tract as $\text{As}^\text{V}$

"Di-Methylated" (DMA)

Metabolism primarily occurs in liver

"Mono-Methylated" (MMA)

Gamble, et al. EHP 2005
Arsenic species measured in urine

In humans, DMA is typically the most abundant metabolite.

Methylated species are more readily excreted in urine.

Increasing methylation efficiency increases rate of excretion.

There is inter-individual variation methylation efficiency (expressed as DMA%, MMA%, and iAs%).

Influenced by age, sex, exposure level, genetic variation, BMI, etc.
Genome-wide association study (GWAS)

- Measure 300,000 genome-wide SNPs using bead array technology for >5,000 participants
- Lots of QC
- Impute to >1 million SNPs

- Urinary arsenic metabolites measured for 2,060 of these participants
- DMA% represents “arsenic methylation capacity”

- Test each SNP for association with lnAs%, MMA%, DMA% (mixed models to deal with relatedness)
GWAS results for DMA%

1,211,988 SNPs
2,056 participants

10q24.32

Similar results for MMA% and iAs% (opposite direction)

1,211,988 SNPs
2,056 participants
Pierce et al, 2012 PLoS Genetics
Pierce et al, 2013 IJE
DMA% and 10q24.32 SNPs

AS3MT
(arsenite methyltransferase)
The *AS3MT* gene

- Conserved across a wide range of species
- First characterized as an arsenic methyltransferase in rabbit and mouse (1990s-2000s).
- *as3mt* knock-out mice
  - have more much inorganic arsenic in their body and the rate of clearance is much slower (methylated arsenic clear fasters)
  - These phenotypic changes are associated susceptibility to damage to the uroepithelial cells following exposure
- Studies of various human populations show associations between genetic variation near AS3MT and arsenic metabolism efficiency (reviewed by Agusa et al 2011)
Are 10q24/AS3MT variants also associated with toxicity?

Arsenic is a class I human carcinogen and chronic exposure can increase risk for lung, bladder, liver, kidney, and skin cancer.

Accumulates in extremities, with skin lesions being the most common sign of arsenic toxicity.

Skin lesions may reflect susceptibility to arsenic-related disease including cancer (Hsu et al 2012)
10q24.32 SNPs and skin lesion risk

2,073 skin lesion cases & 2,857 controls

Is this a "gene-environment interaction" (GxE)?

i.e., does exposure have a larger impact on toxicity risk among genetically-susceptible individuals?
Association b/t arsenic exposure and skin lesion risk (by genotype)

1 0.1-16 ug/L
2 17-87 ug/L
3 >88 ug/L
Association b/t arsenic exposure and skin lesion risk (by genotype)

<table>
<thead>
<tr>
<th>Arsenic Exposure Tertiles</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 0.1-16 ug/L</td>
<td>1</td>
</tr>
<tr>
<td>2 17-87 ug/L</td>
<td>1.47</td>
</tr>
<tr>
<td>3 &gt;88 ug/L</td>
<td>6.66</td>
</tr>
</tbody>
</table>

Arsenic Methylation Capacity
- Low
- Intermediate
- High
Association b/t arsenic exposure and skin lesion risk (by genotype)

1 0.1-16 ug/L
2 17-87 ug/L
3 >88 ug/L

Global Test of Additive Interaction P=0.03
Now that we’re confident an association/interaction exists…

What is the **biological explanation**?

1. Can we identify the causal SNP/variant?

2. Which gene is affected by this variant?

3. What is the mechanism by which this variant influences gene function?

Genetic epidemiology often cannot provide definite answers
Which variant is causal?

- Lots of highly correlated SNPs
- Typically, there’s no “smoking gun” (e.g., protein-coding SNP)
- Causal SNPs are often “regulatory” in nature. Not easy to recognize.
- Larger studies of diverse populations can narrow the search
Identifying causal variants

We will sequence the 10q24.32 region in 3 arsenic exposed populations (n >4,500)

Correlations among SNPs differ across populations

Aim 1: prioritize variants

Aim 2: assess rare variants

Aim 3: skin outcomes
Which gene is affected by the variant? And by what mechanism?

Do our SNPs influence gene activity?

GTEx (Gene-Tissue Expression Project)
~900 multi-tissue donors
>40 different tissues collected
• How do they alter AS3MT function?

(eQTLs): GTEx

• AS3MT is expressed in many tissues
In most cases, to prove a specific variant is causal, some sort of experimental validation is needed
Are there arsenic metabolism variants outside of AS3MT region?

• Using genome-wide SNP data, we can estimate the “heritability” of a trait that is due to measured SNPs.

• **Heritability**: the proportion of variance in a trait that is due to genetic factors

• Our heritability estimate: 16%

• After removing AS3MT region: 5%

• **Bottom line**: Among common variants, the AS3MT region appears to be the major genetic determinant of arsenic metabolism efficiency (in our population)

• There may be additional heritability due to **rare** variants
Urinary Arsenic Species Percentages
What about GxEs that affect disease, but not metabolism?

- We will first search for SNPs that modify the effect of arsenic on molecular phenotypes. Then test SNP-arsenic interactions in relation to arsenic-related health conditions.
What about GxEs that affect disease, but not metabolism?

• We will search for SNPs that modify the effect of arsenic on molecular phenotypes (gene expression and DNA methylation)

• **Step 1:** genome-wide search for SNPs that modify the effect of arsenic exposure on molecular (“omic”) phenotypes

• **Step 2:** Using SNPs identified in Step 1, test SNP-arsenic interactions in relation to arsenic-related health conditions: skin lesion status and diabetes-related phenotypes

• **Innovation:** Most genome-wide GxE approaches only leverage information on how GxE influences disease risk. Our approach is leverages information on how GxE influences biological systems
Can we use genetic information to improve health?

- HEALS has an exposure-reduction program, but some individuals do not appear to respond.
- Could returning genetic information to highly-susceptible individuals increase the efficacy of this program?
- We assessed attitudes toward genetic testing to 200 HEALS participants.
Attitudes towards genetic testing (in HEALS)

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Are you Concerned about the health effects of arsenic?</td>
<td>173</td>
<td>27</td>
</tr>
<tr>
<td>2. Are you interested in receiving genetic information on susceptibility to arsenic toxicity?</td>
<td>200</td>
<td>0</td>
</tr>
<tr>
<td>3. Is it possible for you to take steps to further reduce you and your family’s exposure? (i.e., well-switching, drilling new well)</td>
<td>98</td>
<td>102</td>
</tr>
<tr>
<td>4. If your genes indicated you or your family were at elevated risk, would this motivate to you reduce your family’s exposure?</td>
<td>182</td>
<td>18</td>
</tr>
<tr>
<td>5. If genetic information indicated high risk, would you disclose this information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. To your family?</td>
<td>199</td>
<td>1</td>
</tr>
<tr>
<td>b. To non-family members</td>
<td>59</td>
<td>141</td>
</tr>
</tbody>
</table>
Final Thoughts

- Genetic variation at 10q24.32 is a compelling example of GxE interaction (i.e., SNPs modify the effect of arsenic)

- Larger studies needed to confirm that these metabolism-related SNPs modify the effect of arsenic exposure on other arsenic-related outcomes.

- What makes this work unique: 1. Studying GxE in a cohort designed to study a specific exposure and 2. using “metabolism” as an intermediate phenotype

- Understanding the biology underlying genetic associations/interactions is challenging.

- Additional research is needed to identify SNPs that modify the effect of arsenic on health through pathways other than arsenic metabolism efficiency

- Returning genetic information could potentially have public health benefits, but more research is needed
Acknowledgements

University of Chicago
Habibul Ahsan
Muhammad Kibriya
Chenan Zhang
Farzana Jasmine
Maria Argos (UIC)
Shantanu Roy
Rachelle Paul-Brutus
Ronald Rahaman
Lin Tong
Lin Chen
Jianjun Gao

Columbia University
Joseph Graziano
Faruque Parvez
Mary Gamble

New York University
Yu Chen

U Chicago and Columbia U
Research Office, Bangladesh
Muhammad Rakibuz-Zaman
Alauddin Ahmed
Iftekhar Quasem
Samar K Hore
Shafiul Alam
Tariqul Islam
Golam Sarwar
Mahfuzar Rahman
Mohammad Yunus (ICDDR,B)

Rutgers
Abraham Aviv

University of Utah
Steve Hunt

Dartmouth
Jen Doherty
Margaret Karagas

Johns Hopkins
Ana Navas-Acien

Funding
R01ES020506 (ONES)
U01HG007601
R21ES024834
R01ES023834
DoD W81XWH-14-1-0529

Collaborator Grants
P42ES010349
R01CA102484
R01CA107431
P30CA014599