An Atlas of Genetic Variation
Linking Pathogen-Induced Cellular Traits to Human Disease

Luyang Wang,1 Kelly J. Pittman,1 Jeffrey R. Barker,1 Raul E. Salinas,1 Ian B. Stanaway,2 Graham D. Williams,1 Robert J. Carroll,3 Tom Baimat,4 Andy Ingham,5 Anusha M. Gopalakrishnan,1 Kyle D. Gibbs,1 Alejandro L. Antonia,1 The eMERGE Network, Joseph Heitman,1,6 Soo Chan Lee,7 Gail P. Jarvik,8 Joshua C. Denny,3 Stacy M. Horner,1,6 Mark R. DeLong,5 Raphael H. Valdivia,1 David R. Crosslin,2 and Dennis C. Ko1,6,9,*

Data Expedition by
Rylee Hackley and Benjamin Schott
April 7, 2022
Chlamydia infection outcomes range from asymptomatic to severe.

- **Asymptomatic**: 80%
- **Symptomatic**: PID or subclinical PID
- **Cleared**

**Antibiotics**

- Ectopic pregnancy
- Infertility
- Preterm delivery
Chlamydia trachomatis is an obligate intracellular pathogen

Elwell et al., 2016

The Chlamydia trachomatis Ctd1 invasin exploits the human integrin β1 receptor for host cell entry

Sonja Stallmann 1, Johannes H Hegemann 1


Chlamydia pneumoniae exploits adipocyte lipid chaperone FABP4 to facilitate fat mobilization and intracellular growth in murine adipocytes

Nirwana Fitriani Walenna 1, Yusuke Kurihara 2, Bin Chou 2, Kazunari Ishii 2, Toshinori Soejima 2, Ryota Itoh 2, Akinori Shimitzu 2, Takeshi Ichinohe 2, Kenji Hiratsuka 4


Innate immunity in defense against Chlamydia trachomatis infections

J A Severin 1, J M Ossewaarde
Cellular models and this screen

**Advantages:**
Quantitative phenotypes instead of Case-Control Experimental platform

**Disadvantages:**
Possible detection of cell-type-specific phenomena
Bad assumptions about molecular correlates of disease
Today’s Question:

What, if any, host genetic variants affect CXCL10 protein levels?

Today’s Data:

The prevalence of variants within 2Mb of CXCL10 in 527 cells lines from 4 human populations, and their association with CXCL10/IP10 protein levels after Chlamydia infection.
Genome-wide association testing seeks to link DNA base-pair differences with differences in a trait.

**Case-control study for genetic association**

<table>
<thead>
<tr>
<th></th>
<th>Cases (n=1,000)</th>
<th>Controls (n=1,000)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(express the trait)</td>
<td>(do not express the trait)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases</td>
<td>62%</td>
<td>38%</td>
</tr>
<tr>
<td>Controls</td>
<td>49%</td>
<td>51%</td>
</tr>
</tbody>
</table>

**Associating One SNP with Quantitative Trait**

Linear Regression and Additive Risk Model

\[ y = \alpha + \beta x + \varepsilon \]

- \( T \) = risk allele
- \( x_{CC} = 0 \) if individual is CC
- \( x_{CT} = 1 \) if individual is CT
- \( x_{TT} = 2 \) if individual is TT

\[ \text{height} = \alpha + \beta x \]

\( \beta \): change in height for 1 risk allele
Student’s T-test

1-Sample t-test

Null Hypothesis
Mean = 5

2-Sample t-test

mean(EP) = 9.34
mean(No EP) = 9.61
Analysis of Variance (ANOVA)

Is the between group variation greater than the within group variation?
Why is the threshold for genome-wide significance so high?
Is high CXCL10 protective? And if so, why is the polymorphism maintained?

High CXCL10 -> protection from infection, predisposed to inflammatory disorders (IBS)

The pathogen fights back:
Chlamydia has evolved a protease, CPAF, which actively degrades host proteins including CXCL10

Balancing selection

\[
\text{Pearson } r = -0.50 \\
p < 0.0001
\]
You have to look at your data before you can analyze it.