A paradigm shift in mouse genetics: Merging Reductionist and Holistic Approaches to Model Human Disease

Robert W Williams
rwilliams@uthsc.edu

Center for Integrative and Translational Genomics
1. **systems genetics**: a strategy of using built-in genetic perturbations to define relations and causality across scale and species

2. **experimental reference panels**: Permanent resources used to integrate data across scales

3. **GeneNetwork**: an open-source international data integrator/analyzer service

4. **high efficiency translation** from mouse to human and back via “synergistic technical bridging”
CEPH: 50 Mormon families

Framingham: n = 13,000

decode
BXD experimental reference panel

female
C57BL/6J

5 million
DNA variants

male
DBA/2J

isogenic
F1

hetero-
F2
gene-

The
BXD
family

BXD1
BXD43
BXD103

B vs D

5 million DNA variants
Current BXDs: very dense genotype data

- 540,000 SNPs (Affymetrix)
- SOLiD and Illuina sequencing of D parent ~100x
- 4.8 million SNPs, 500 k indels. All 80 strains are essentially “sequenced”

\[ B \neq DB \neq DB = D \]

\( B \) as in black  \( D \) as in dilute

\[ B = D \]

identical by descent  genetic polymorphisms

\[ B \neq D \]
Collaborative Cross: 2001–2012

Complex Trait Consortium
1st Workshop Report: Sept 2002
A Collaborative Cross for High-Precision Complex Trait Analysis

RW Williams, KW Broman, JM Cheverud, GA Churchill, RW Hitzemann, KW Hunter, J Mountz, D Pomp, RH Reeves, LC Schalkwyk, DW Threadgill

Scientists Dream of 1001 Complex Mice

Genetists want to create 1000 new lines of inbred mice to help them sort out the genes behind complex diseases such as diabetes and cancer.

Genetics, U.K.—As any lover of fine Scotch will tell you, a little complexity is a good thing. The trick for the distiller is to produce a complex, but consistent, product. Geneticists trying to sort out the genes that trigger some of medicine’s most common ailments face a similar challenge. They need to capture the natural variation that exists in human populations but in a way that is reproducible in the lab.

Commonly used mouse strains for four generations, producing litters with all 16,000 possible permutations of great-grandparents. They will then turn to inbreeding, performing brother-sister matings until the offspring of each strain are essentially clones. Not all strains will survive the inbreeding, but after many generations, scientists can use the Massachussetts Institute of Technology. “But there are different ways of going about [finding QTLs].

Controlled complexity. To help track down genes involved in complex traits, scientists want to create 1000 new lines of “recombinant inbred” mice. Each line would be a combination of eight existing strains; the scientists would outcross the mice for four generations and then breed brother-sister pairs for 20 generations to create inbred lines.
8 parental strains: 40 million SNPs

UK-Israel CC (Wellcome): ~ 250
US CC (NIH, DOE, Ellisson): ~ 500
UWA Perth Morahan CC: ~ 700

First inbred line complete May 2010
Consensus set of 200 lines by 2013
Consensus set of >500 lines by 2015

A = A/J
B = C57BL/6J
C = 129S1/SvImJ
D = NOD/LtJ
E = NZO/HILtJ
F = CAST/EiJ
G = PWK/PhJ
H = WSB/EiJ
BXD mice and HXB rats are analogous to Finnish or Icelandic populations.

4–5 million SNPs
~10,000 missense

CC more analogous to humanity—10x common alleles. 40 million SNPs and ~100,000 missense.

In both cases, very few rare alleles compared to humans.
deep experimental electronic medical records

High content

Rodent GRPs

$n = 30 - 1,000$

Genotypes and Sequence

environmental responsivity
disease resistance
structure
physiology
behavior
epigenetic variation
developmental variation
molecular expression variation

More is different.
high throughput: hepatocytes, ESC, MEFs

96 > 192 >
384 > 1536

can we predict *in vivo* from *in vitro*?
human expression data
search finds 11 items

2. Unpublished: RecordID/10685 – polyglucosan body number in hippocampus in 18-month-old females (unmodified values, unpublished) by Jucker M and colleagues (mjucker@uhbs.ch) unpublished data, et al
Polyglucosan bodies in hippocampus 18 months

Longevity (days) 24 strains (30% of available resource)

Only 24 of 80 strains have been characterized for longevity.
Genetic analysis of hematopoietic cell cycling in mice suggests its involvement in organismal life span.

De Haan G, Van Zant G.

Blood and Marrow Transplant Program, Division of Hematology/Oncology, University of Kentucky Medical Center, Lexington, Kentucky 40536-0093, USA.

Normal somatic cells undergo replicative senescence in vitro but the significance of this process in organismic aging remains controversial. We have shown previously that hematopoietic stem cells of common inbred strains of mice vary widely in cycling activity and that this parameter is inversely correlated with strain-dependent mean life span. To assess whether cell cycling and life span are causally related, we searched for quantitative trait loci (QTLs) that contributed to variation of these traits in BXH and BXD recombinant inbred mice. Two QTLs, mapping to exactly the same intervals on chromosomes 7 and 11, were identified that were associated with variation of both cell cycling and life span. The locus on chromosome 11 mapped to the cytokine cluster, a segment that shows synteny with human chromosome 5q, in which deletions are strongly associated with myelodysplastic syndrome. These data indicate that steady-state cell turn-over, here measured in hematopoietic progenitor cells, may have a significant effect on the mean life span of mammals.
longevity (QTL)

high gene density

high SNP-indel density
classical trait vs expression micro trait
SIRT1 sumoylation regulates its deacetylase activity and cellular response to genotoxic stress.

Yang Y, Fu W, Chen J, Olashaw N, Zhang X, Nicosia SV, Bhalla K, Bai W.

Departments of Pathology and Cell Biology, University of South Florida College of Medicine, 12901 Bruce B. Downs Blvd., Tampa, Florida 33612-4799, USA.

SIRT1 is the closest mammalian homologue of yeast SIR2, an important ageing regulator that prolongs lifespan in response to caloric restriction. Despite its importance, the mechanisms that regulate SIRT1 activity are unclear. Our study identifies a novel post-translational modification of SIRT1, namely sumoylation at Lys 734. In vitro sumoylation of SIRT1 increased its deacetylase activity. Conversely, mutation of SIRT1 at Lys 734 or desumoylation by SENP1, a nuclear desumoylase, reduced its deacetylase activity. Stress-inducing agents promoted the association of SIRT1 with SENP1 and cells depleted of SENP1 (but not of SENP1 and SIRT1) were more resistant to stress-induced apoptosis than control cells. We suggest that stress-inducing agents counteract the anti-apoptotic activity of SIRT1 by recruiting SENP1 to SIRT1, which results in the desumoylation and inactivation of SIRT1 and the consequent acetylation and activation of apoptotic proteins.

PMID: 17934453 [PubMed - indexed for MEDLINE]
millions of eQTL: mRNA as the first microphenotype

Prkce (2.4-fold)

Mpaz (4-fold)
thousands of eQTL: mRNA as microphenotype

Prkce (2.4-fold)

Mpdz (4-fold)
292 neurodevelopmental and autism-related expression traits from a neocortex database
rapid assay of translational relevance

Blood pressure QTL with Johan Auwerx using EMPReSS
Identification of the **UBP1** Locus as a Critical Blood Pressure Determinant Using a Combination of Mouse and Human Genetics

Hana Koutnikova, Markku Laakso, Lu Lu, Roy Combe, Jussi Paananen, Teemu Kuulasmaa, Johanna Kuusisto, Hans-Ulrich Häring, Torben Hansen, Oluf Pedersen, Ulf Smith, Markolf Hanefeld, Robert W. Williams, Johan Auwerx

**Table 1.** The effects of two SNPs of the **UBP1** gene locus on systolic and diastolic blood pressure (BP, mean±SD or SE) in three different Caucasian cohorts under the additive model.

<table>
<thead>
<tr>
<th>SNP ID</th>
<th>Cohort</th>
<th>N</th>
<th>Minor allele frequency</th>
<th>BP, (mmHg)*</th>
<th>Effect size, (mmHg) (SE)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Systolic BP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs 17030583</td>
<td>EUGENE2</td>
<td>867</td>
<td>0.24</td>
<td>123 (17)</td>
<td>+2.4 (0.9)</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>METSIM</td>
<td>2532</td>
<td>0.3</td>
<td>137 (18)</td>
<td>+0.8 (0.5)</td>
<td>0.076</td>
</tr>
<tr>
<td></td>
<td>RAD</td>
<td>237</td>
<td>0.26</td>
<td>131 (17)</td>
<td>+3.5 (1.8)</td>
<td>0.063</td>
</tr>
<tr>
<td></td>
<td>Pooled</td>
<td>3636</td>
<td>0.29</td>
<td></td>
<td>+1.3 (0.4)</td>
<td>0.00099</td>
</tr>
<tr>
<td>rs 2291897</td>
<td>EUGENE2</td>
<td>866</td>
<td>0.2</td>
<td>123 (17)</td>
<td>+2.8 (0.9)</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>METSIM</td>
<td>2537</td>
<td>0.27</td>
<td>137 (18)</td>
<td>+1.0 (0.5)</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>RAD</td>
<td>237</td>
<td>0.21</td>
<td>132 (18)</td>
<td>+2.2 (2.1)</td>
<td>0.297</td>
</tr>
<tr>
<td></td>
<td>Pooled</td>
<td>3640</td>
<td>0.25</td>
<td></td>
<td>+1.5 (0.4)</td>
<td>0.00036</td>
</tr>
<tr>
<td></td>
<td>Diastolic BP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs 17030583</td>
<td>EUGENE2</td>
<td>867</td>
<td>0.24</td>
<td>78 (12)</td>
<td>+2.1 (0.6)</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>METSIM</td>
<td>2532</td>
<td>0.3</td>
<td>90 (11)</td>
<td>+0.4 (0.3)</td>
<td>0.136</td>
</tr>
<tr>
<td></td>
<td>RAD</td>
<td>237</td>
<td>0.26</td>
<td>83 (11)</td>
<td>+0.9 (1.0)</td>
<td>0.452</td>
</tr>
<tr>
<td></td>
<td>Pooled</td>
<td>3636</td>
<td>0.29</td>
<td></td>
<td>+0.8 (0.3)</td>
<td>0.0063</td>
</tr>
<tr>
<td>rs 2291897</td>
<td>EUGENE2</td>
<td>866</td>
<td>0.2</td>
<td>78 (12)</td>
<td>+2.0 (0.7)</td>
<td>0.056</td>
</tr>
<tr>
<td></td>
<td>METSIM</td>
<td>2537</td>
<td>0.27</td>
<td>90 (11)</td>
<td>+0.6 (0.3)</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>RAD</td>
<td>237</td>
<td>0.21</td>
<td>83 (11)</td>
<td>+0.7 (1.2)</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>Pooled</td>
<td>3640</td>
<td>0.25</td>
<td></td>
<td>+0.8 (0.3)</td>
<td>0.003</td>
</tr>
</tbody>
</table>
Experimental reference panels combined with web-enabled analytic tools provide a powerful integrator for systems analysis of genome-phenome relations.

They provide efficient experimental methods that complement both human genetics and KO mouse models.

Phenotype revolution in progress: high throughput/high content methods. Gaps—proteome, metabolome, epigenetic, environmental,
new tactic: *reverse complex trait analysis*

- 4.8 million SNPs
- 0.5 million indels
- 55,000 CNVs
- 10,000 missense SNPs
- 2500 high impact

**SOLiD**

- 42X
- Stop Codon: 84
- Mis-sense: 10,012
- Coding SNPs: 30,553
- Non-Coding SNPs: 4,130,018
- All SNPs: 4,160,571

**GAIII**

- 12X
- Stop Codon: 74
- Mis-sense: 9,546
- Coding SNPs: 29,464
- Non-Coding SNPs: 3,117,536
- All SNPs: 3,147,000
~140 frameshift and stop mutations in BXDs

<table>
<thead>
<tr>
<th>Chr</th>
<th>Position</th>
<th>B</th>
<th>D</th>
<th>Symbol</th>
<th>GeneNetwork</th>
<th>SNP Loca</th>
<th>B</th>
<th>D</th>
<th>Functions For GeneWiki</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr10</td>
<td>76019442</td>
<td>T</td>
<td>A</td>
<td>Ftcd</td>
<td>done</td>
<td>exon 1</td>
<td>C</td>
<td>*</td>
<td>mis-ser Ftcd has a putative mutation</td>
</tr>
<tr>
<td>chr11</td>
<td>79434121</td>
<td>C</td>
<td>T</td>
<td>Hmha1</td>
<td>done, weak cis e</td>
<td>exon 22</td>
<td>R</td>
<td>*</td>
<td>mis-ser Hmha1 has a putative mutation</td>
</tr>
<tr>
<td>chr10</td>
<td>93782168</td>
<td>C</td>
<td>T</td>
<td>Gm4792</td>
<td>done, great supr</td>
<td>exon 1</td>
<td>W</td>
<td>*</td>
<td>mis-ser Gm4792 has a putative mutation</td>
</tr>
<tr>
<td>chr11</td>
<td>30366378</td>
<td>G</td>
<td>A</td>
<td>Sfi1</td>
<td>entered</td>
<td>possible exon 24</td>
<td>W</td>
<td>*</td>
<td>mis-ser Sfi1 has a putative mutation</td>
</tr>
<tr>
<td>chr11</td>
<td>30442555</td>
<td>G</td>
<td>A</td>
<td>Sfi1</td>
<td>entered</td>
<td>possible exon 18</td>
<td>Q</td>
<td>*</td>
<td>mis-ser Sfi1 has a putative mutation</td>
</tr>
<tr>
<td>chr11</td>
<td>3079453</td>
<td>G</td>
<td>A</td>
<td>Sfi1</td>
<td>entered</td>
<td>possible exon 5</td>
<td>W</td>
<td>*</td>
<td>mis-ser Sfi1 has a putative mutation</td>
</tr>
<tr>
<td>chr11</td>
<td>3084464</td>
<td>G</td>
<td>A</td>
<td>Sfi1</td>
<td>entered</td>
<td>possible exon 5</td>
<td>W</td>
<td>*</td>
<td>mis-ser Sfi1 has a putative mutation</td>
</tr>
<tr>
<td>chr11</td>
<td>5874307</td>
<td>G</td>
<td>A</td>
<td>Butr1</td>
<td>entered</td>
<td>highest exon 7</td>
<td>W</td>
<td>*</td>
<td>mis-ser Butr1 has a putative mutation</td>
</tr>
<tr>
<td>chr11</td>
<td>70467408</td>
<td>G</td>
<td>A</td>
<td>Rnf167</td>
<td>entered</td>
<td>strong c exon 9</td>
<td>*</td>
<td></td>
<td>synonyi Rnf167 has a putative mutation</td>
</tr>
<tr>
<td>chr11</td>
<td>89321323</td>
<td>G</td>
<td>A</td>
<td>Ankn1</td>
<td>entered</td>
<td>strongest exon 4</td>
<td>*</td>
<td></td>
<td>synonyi Ankn1 has a putative mutation</td>
</tr>
<tr>
<td>chr12</td>
<td>29239457</td>
<td>C</td>
<td>T</td>
<td>Ankn1</td>
<td>entered</td>
<td>exon 3</td>
<td>S</td>
<td>*</td>
<td>mis-ser Ankn1 has a putative mutation</td>
</tr>
<tr>
<td>chr12</td>
<td>40889100</td>
<td>C</td>
<td>T</td>
<td>Zfp277</td>
<td>entered</td>
<td>very strong exon 3</td>
<td>R</td>
<td>*</td>
<td>mis-ser Zfp277 has a putative mutation</td>
</tr>
<tr>
<td>chr12</td>
<td>55689655</td>
<td>C</td>
<td>T</td>
<td>Snx6</td>
<td>entered</td>
<td>strong c exon 5</td>
<td>Q</td>
<td>*</td>
<td>mis-ser Snx6 has a putative mutation</td>
</tr>
<tr>
<td>chr12</td>
<td>104618530</td>
<td>C</td>
<td>A</td>
<td>Serpina3f</td>
<td>entered</td>
<td>no human exon 3</td>
<td>Y</td>
<td>*</td>
<td>mis-ser Serpina3f has a putative mutation</td>
</tr>
<tr>
<td>chr13</td>
<td>41113731</td>
<td>C</td>
<td>T</td>
<td>Gcm2</td>
<td>entered, no cis e</td>
<td>exon 5</td>
<td>*</td>
<td></td>
<td>synonyi Gcm2 has a putative mutation</td>
</tr>
<tr>
<td>chr14</td>
<td>50639909</td>
<td>C</td>
<td>T</td>
<td>Gm5800</td>
<td>not entered</td>
<td>exon 5</td>
<td>E</td>
<td>*</td>
<td>mis-ser Gm5800 has a putative mutation</td>
</tr>
<tr>
<td>chr14</td>
<td>34109477</td>
<td>C</td>
<td>T</td>
<td>Bnln</td>
<td>entered</td>
<td>exon 9</td>
<td>G</td>
<td>*</td>
<td>mis-ser Bnln has a putative mutation</td>
</tr>
<tr>
<td>chr13</td>
<td>34923242</td>
<td>C</td>
<td>A</td>
<td>Pel3</td>
<td>entered, covaexes exon 7</td>
<td>E</td>
<td>*</td>
<td>mis-ser Pel3 has a putative mutation</td>
<td></td>
</tr>
<tr>
<td>chr1</td>
<td>47949981</td>
<td>T</td>
<td>C</td>
<td>Itpr</td>
<td>entered</td>
<td>exon 2</td>
<td>*</td>
<td></td>
<td>synonyi Itpr has a putative mutation</td>
</tr>
<tr>
<td>chr1</td>
<td>17341104</td>
<td>C</td>
<td>T</td>
<td>Cd244</td>
<td>entered</td>
<td>exon 3</td>
<td>S</td>
<td>*</td>
<td>mis-ser Cd244 has a putative mutation</td>
</tr>
<tr>
<td>chr2</td>
<td>178342851</td>
<td>A</td>
<td>T</td>
<td>Pppr13d</td>
<td>entered</td>
<td>exon 1</td>
<td>Y</td>
<td>*</td>
<td>mis-ser Pppr13d has a putative mutation</td>
</tr>
<tr>
<td>chr2</td>
<td>43448391</td>
<td>A</td>
<td>T</td>
<td>Fam166b</td>
<td>entered</td>
<td>exon 2</td>
<td>E</td>
<td>*</td>
<td>mis-ser Fam166b has a putative mutation</td>
</tr>
<tr>
<td>chr2</td>
<td>111480193</td>
<td>A</td>
<td>T</td>
<td>Skint7</td>
<td>entered</td>
<td>exon 4</td>
<td>Q</td>
<td>*</td>
<td>mis-ser Skint7 has a putative mutation</td>
</tr>
<tr>
<td>chr2</td>
<td>117860143</td>
<td>A</td>
<td>T</td>
<td>Hya</td>
<td>entered</td>
<td>exon 6</td>
<td>S</td>
<td>*</td>
<td>mis-ser Hya has a putative mutation</td>
</tr>
<tr>
<td>chr2</td>
<td>143424566</td>
<td>A</td>
<td>T</td>
<td>Oog3</td>
<td>entered</td>
<td>exon 4</td>
<td>E</td>
<td>*</td>
<td>mis-ser Oog3 has a putative mutation</td>
</tr>
<tr>
<td>chr2</td>
<td>14797487</td>
<td>T</td>
<td>C</td>
<td>Apitd1</td>
<td>entered</td>
<td>strong c exon 5</td>
<td>Q</td>
<td>*</td>
<td>synonyi Apitd1 has a putative mutation</td>
</tr>
<tr>
<td>chr4</td>
<td>48519551</td>
<td>C</td>
<td>A</td>
<td>Gm5111</td>
<td>entered</td>
<td>strong exon 3</td>
<td>S</td>
<td>*</td>
<td>mis-ser Gm5111 has a putative mutation</td>
</tr>
<tr>
<td>chr4</td>
<td>14879479</td>
<td>G</td>
<td>T</td>
<td>Gnpmb</td>
<td>entered</td>
<td>possible exon 2</td>
<td>S</td>
<td>*</td>
<td>mis-ser Gnpmb has a putative mutation</td>
</tr>
<tr>
<td>chr6</td>
<td>122001620</td>
<td>C</td>
<td>T</td>
<td>Cpammd8</td>
<td>entered</td>
<td>exon 9</td>
<td>Q</td>
<td>*</td>
<td>mis-ser Cpammd8 has a putative mutation</td>
</tr>
</tbody>
</table>
GRP can map/define causes (QTL and QTGs)

- **cis eQTL for Ahr**

- **trans eQTL for Cyp1a1**

  **AHRR**: aryl hydrocarbon receptor repressor
Collaborators

- Lu Lu
- Xusheng Wang
- Megan Mulligan
- Khyobeni Mozhui
- Ash Panday
- David Li
- Jeremy Peirce
- Elissa Chesler
- Daniel Ciobanu

Xiaodong Zhou
Lei Yan
Arthur Centeno
Zachary Sloan
David Crowell
Evan Williams
Zhaohui Sun
Alex Williams

- Guy Mittleman
- Melloni Cook
- Ramin Homayouni
- Doug Matthews
- Tom Sutter
- Ken Manly & Jintao Wang (GN)
- Bill Taylor (SOLiD)
- Stan Nelson (Illumina)
- Mike Miles (NAc)
- Glenn Rosen (BXD)
- Johan Auwerx (BXD)
- Richard Nowakowski

www.GeneNetwork.org

NIAAA, NIDA, NCRR, NIMH, CITG