

Genetics of extreme body size evolution in mice from Gough Island

Karl Broman

Biostatistics & Medical Informatics, UW–Madison

kbroman.org
github.com/kbroman
[@kwbroman](https://twitter.com/kwbroman)

bit.ly/BDS2017-07

This is a collaboration with Bret Payseur (Genetics, UW-Madison).

Gough Island is a small, isolated island in the South Atlantic. The mice on the island are considerably larger than mainland wild mice. We are seeking to dissect the underlying genetic mechanisms for this difference.

Gough Island



2

Gough Island is an isolated island in the South Atlantic, about half-way between South America and Africa.

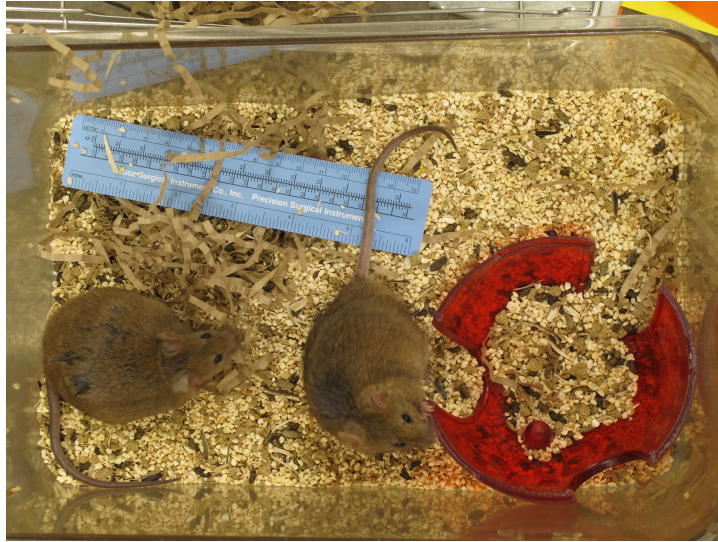
Big rodents



3

The mice on Gough Island are the largest wild mice on earth. And they've also developed rather nasty behaviors: much of the year, they eat young birds (while the birds are still alive).

Gough mice



4

This is a pair of Gough mice raised in the lab.

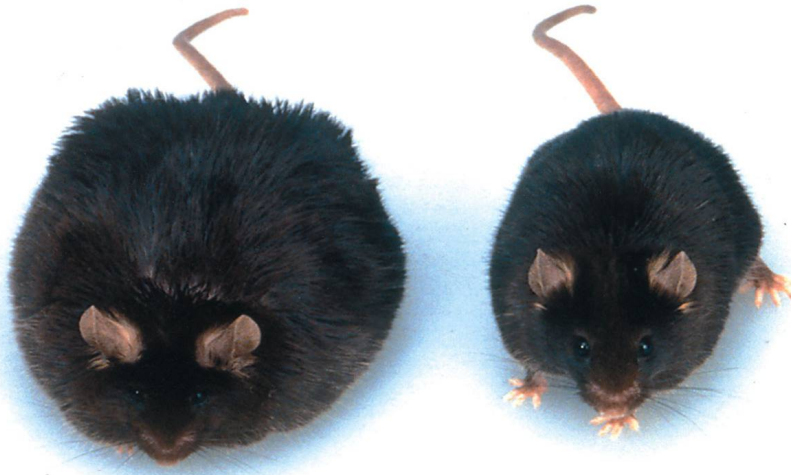
WSB and Gough mice



5

Here is a WSB mouse (left) next to a Gough mouse (right).

Leptin knockout



Biological Psychology 6e, Figure 13.25

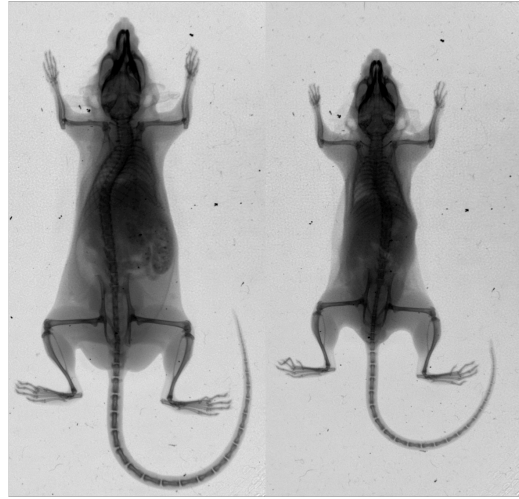
© 2010 Sinauer Associates, Inc.

The Gough mice didn't look all that large to me, because in other collaborations, I'd been studying mice whose leptin gene was knocked out.

On the left is a leptin knockout; on the right is a normal lab mouse.

The leptin gene is what makes you satisfied at the end of the meal. With the leptin gene knocked out (so that the gene doesn't function), the mice will just eat and eat and so become extremely obese.

WSB and Gough mice



7

But the Gough mice are indeed quite a bit larger than the most closely related wild strain.

The Island Rule



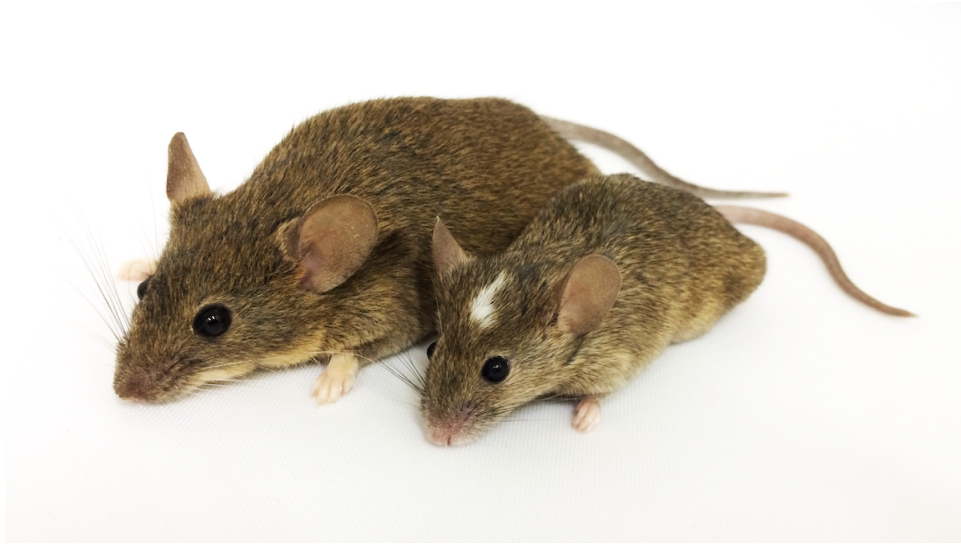
8

The “Island Rule” is that after colonizing an island, animals less than 1 kg tend to get bigger, and animals greater than 1 kg tend to get smaller.

On the left are a pair of fox skulls. The larger one is a mainland fox; the smaller one is an island fox.

On the right are a pair of scrub jays. The smaller one is a mainland jay; the smaller one is an island jay.

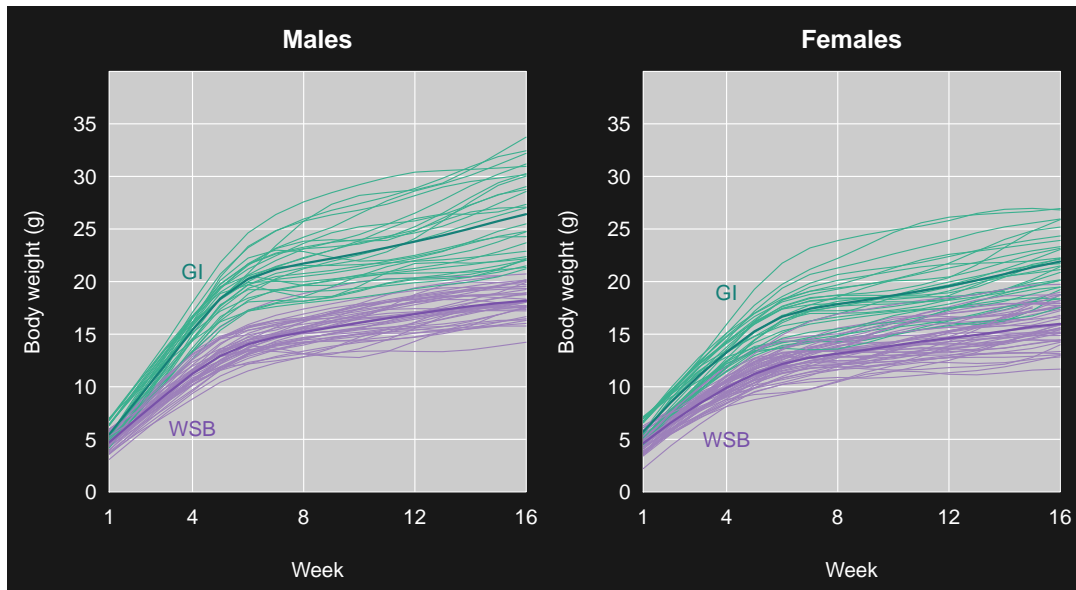
Gough and WSB mice



9

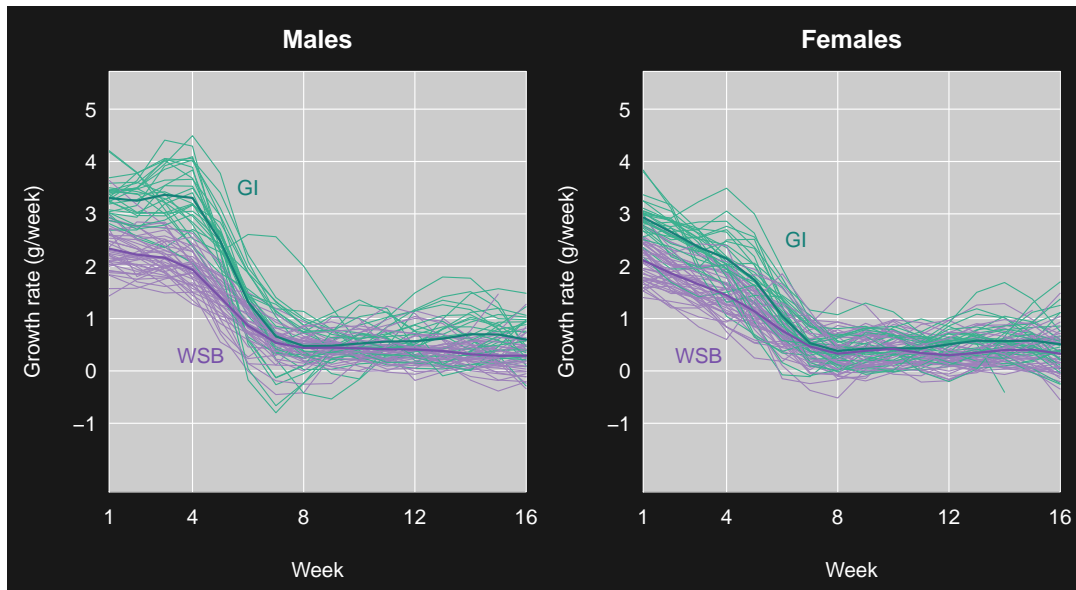
Why are the Gough mice larger?

Body weight



Body weight by age for the Gough and WSB mice.

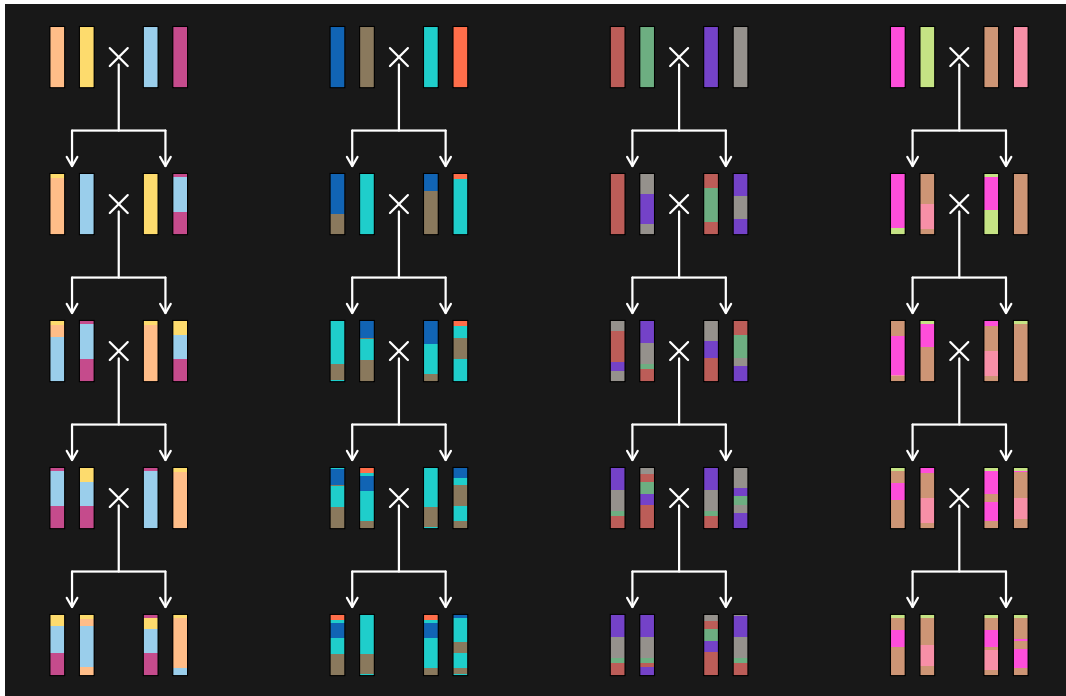
Growth rate



11

Growth rate (g/week) by age for the Gough and WSB mice.

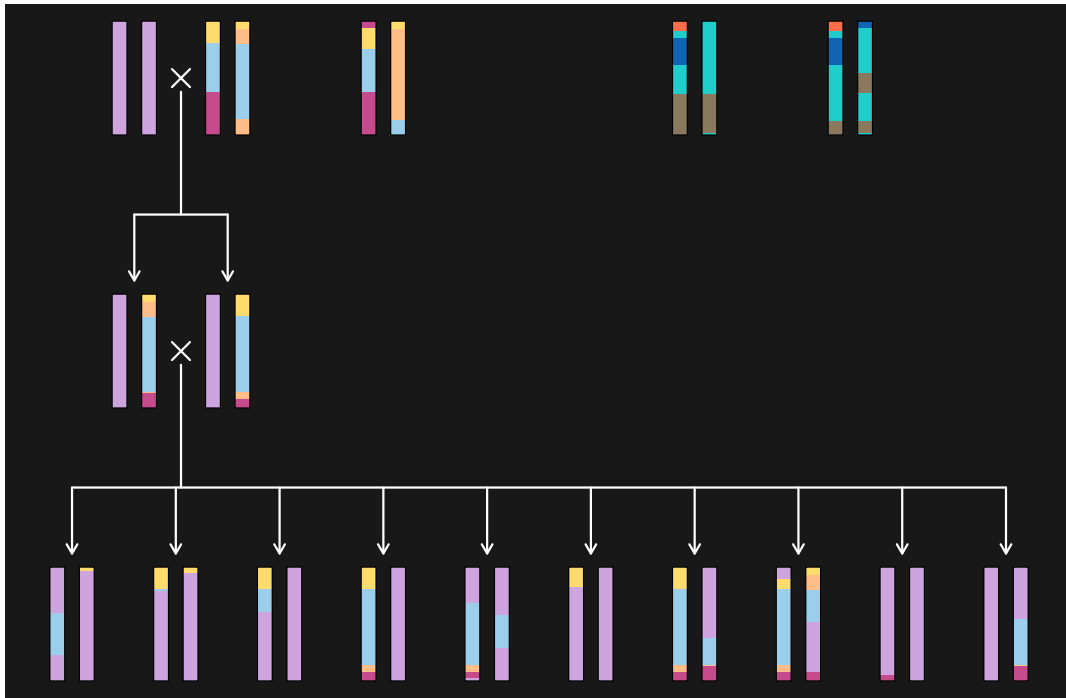
Partially-inbred Gough lines



12

Bret Payseur's lab is in the process of developing a set of inbred strains derived from Gough mice.

Gough × WSB crosses

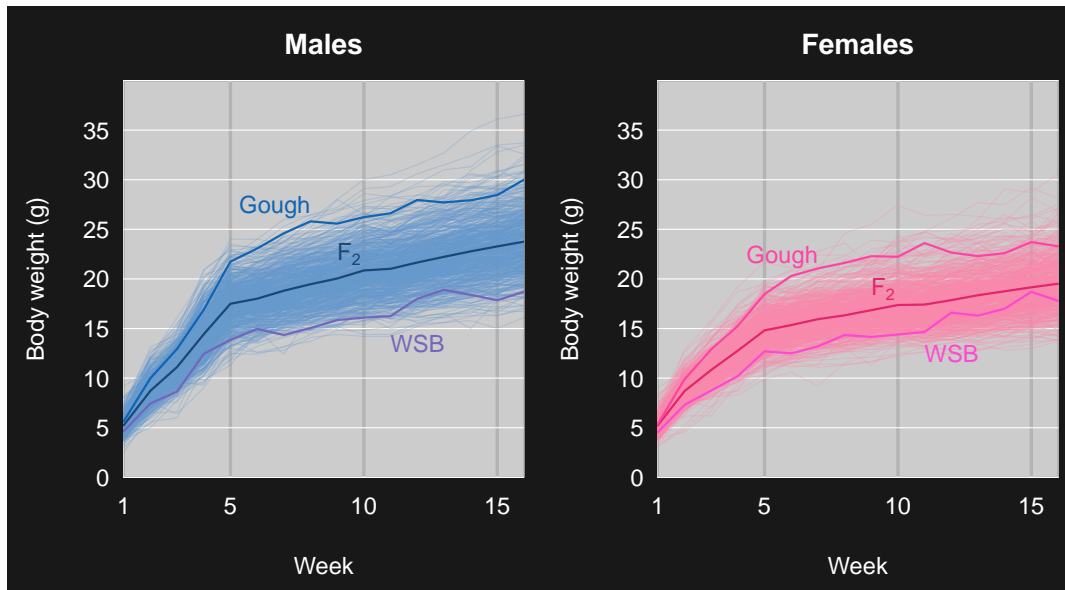


13

In order to identify the genes that contribute to the larger size of Gough mice, we're considering data from four large intercrosses, each between a Gough mouse (after three generations of inbreeding) and an inbred WSB mouse. There are a total of 1212 mice from four crosses.

The mice were genotyped with the MegaMUGA SNP array. For now, we're focusing on 11833 SNPs that are fixed in Gough, and so segregate like a standard F_2 intercross between inbred strains.

Growth curves

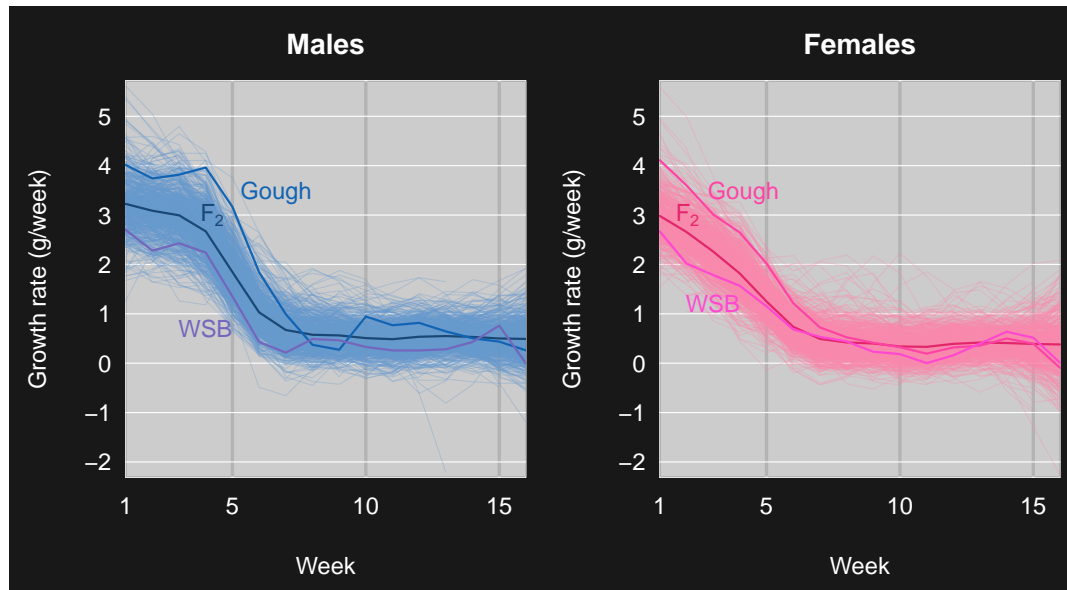


14

We don't have too much data on the body weights of the Gough and WSB parents, but they do show large differences.

We've measured body weight for each F₂ mouse, weekly, at ages 1–16 weeks. There is a large sex difference and considerable individual variation.

Growth rate

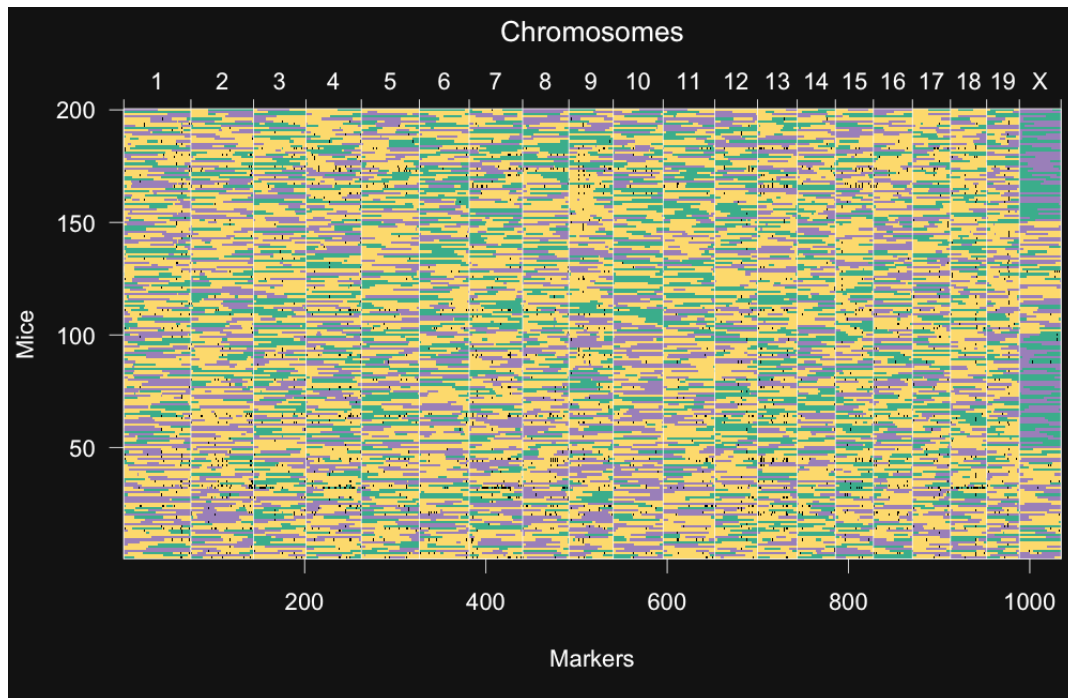


15

The estimated growth rate (the first derivatives of the body weight curves) are particularly interesting: high initial rate that slows during weeks 4–7 and then is flat thereafter.

The estimated average growth rates for the Gough and WSB parents are quite noisy, but it does seem that the biggest differences are in the initial growth, in weeks 1–5.

Genotypes

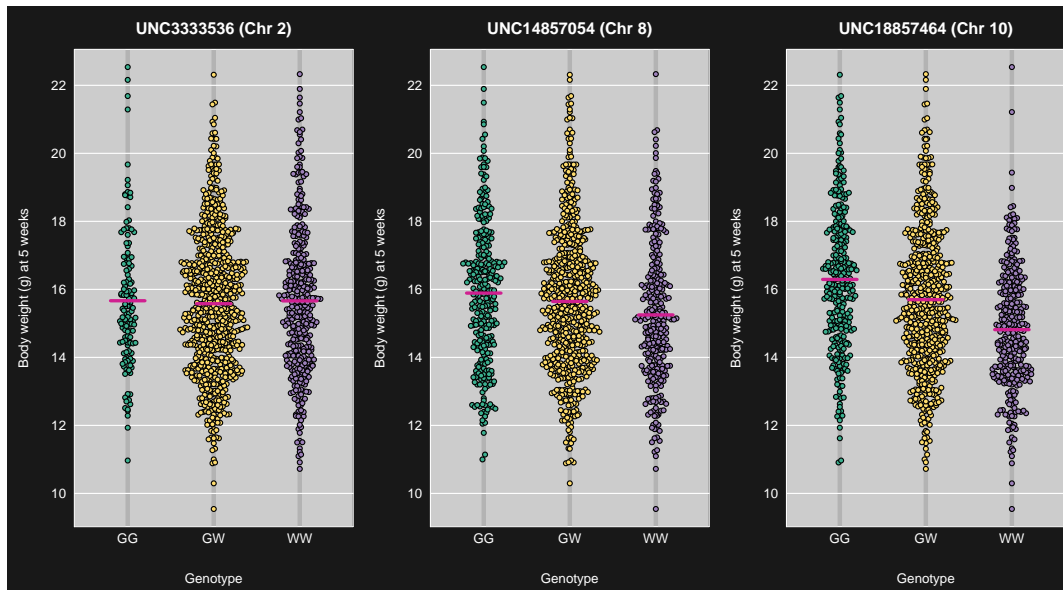


16

An image with a portion of the genotype data (green=Gough homozygote, Yellow=heterozygote, purple=WSB homozygote).

I subsampled the genotypes to 1034 of the 11,833 total markers, and showing just 200 of the 1,212 mice (the top and bottom 50, by weight at 16 weeks, of each of the males and females).

Weight at 5 weeks vs genotype

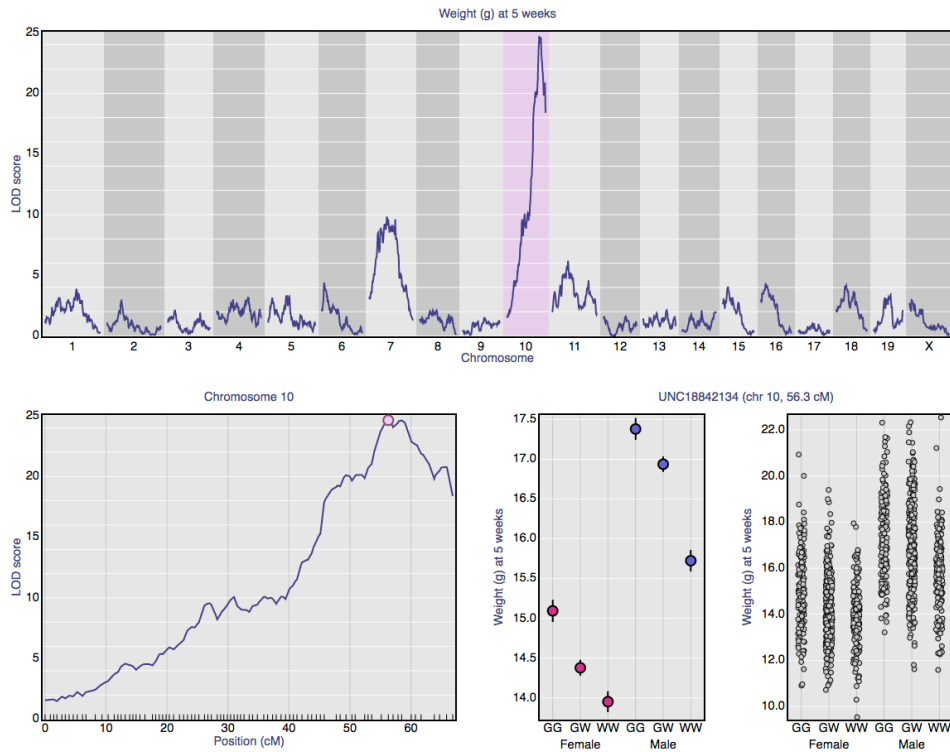


17

These dotplots show phenotype versus genotype for three selected markers, for body weight at 5 weeks.

Are the associations real?

Genome scan



18

Here is a plot of the scan across the genome.

Think of the test statistic, “LOD score”, as sort of like the $-\log_{10}$ p-value, though really it’s a \log_{10} likelihood ratio.

This links to an interactive plot, where you can explore the genotype-phenotype associations across the genome.

Permutation test



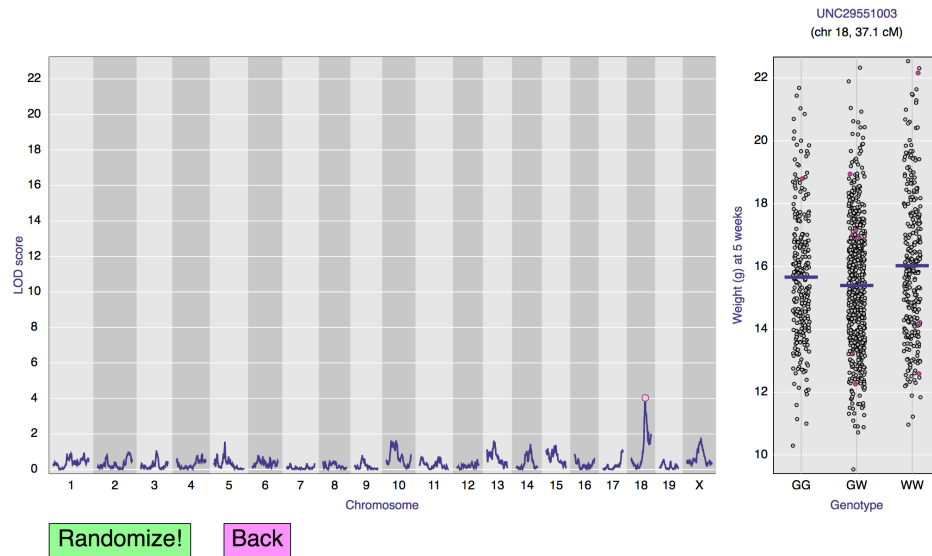
19

A key issue in this business is the need to adjust for the multiple statistical tests performed: that we did a “scan” across the genome, testing for the genotype/phenotype association at each marker.

To deal with this, we derive the distribution of the genome-wide maximum test statistic under the “global” null hypothesis that the phenotype is totally unrelated to the genotypes.

To determine that null distribution, we shuffle the phenotype relative to the genotypes, calculate the test statistics across the genome, derive the maximum test statistic, and repeat many times.

Permutation test

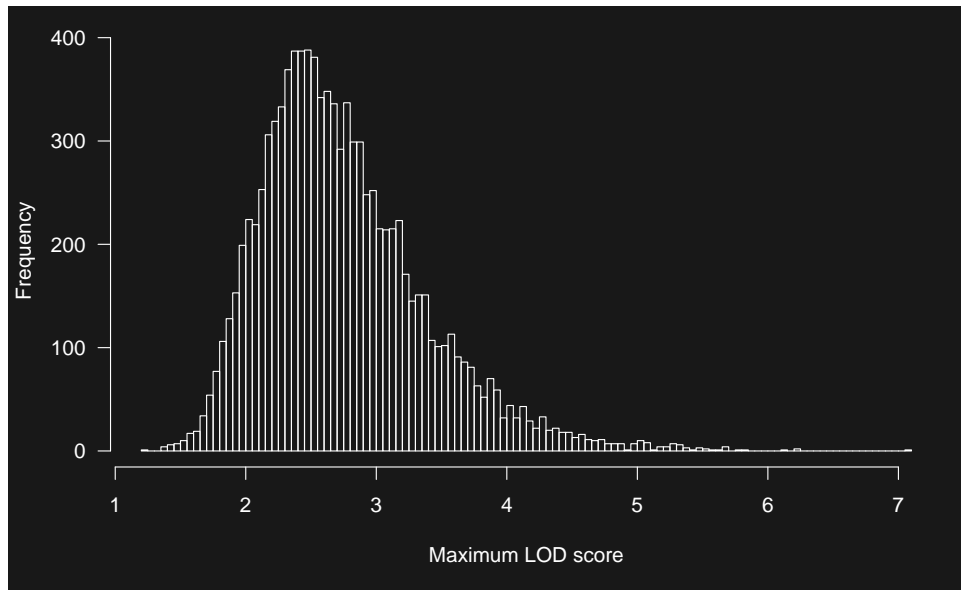


20

This is an interactive illustration of a permutation test.

Click the “Randomize” button to shuffle the phenotypes and re-draw the LOD curves; click the “back” button to go back.

Histogram of permutation results

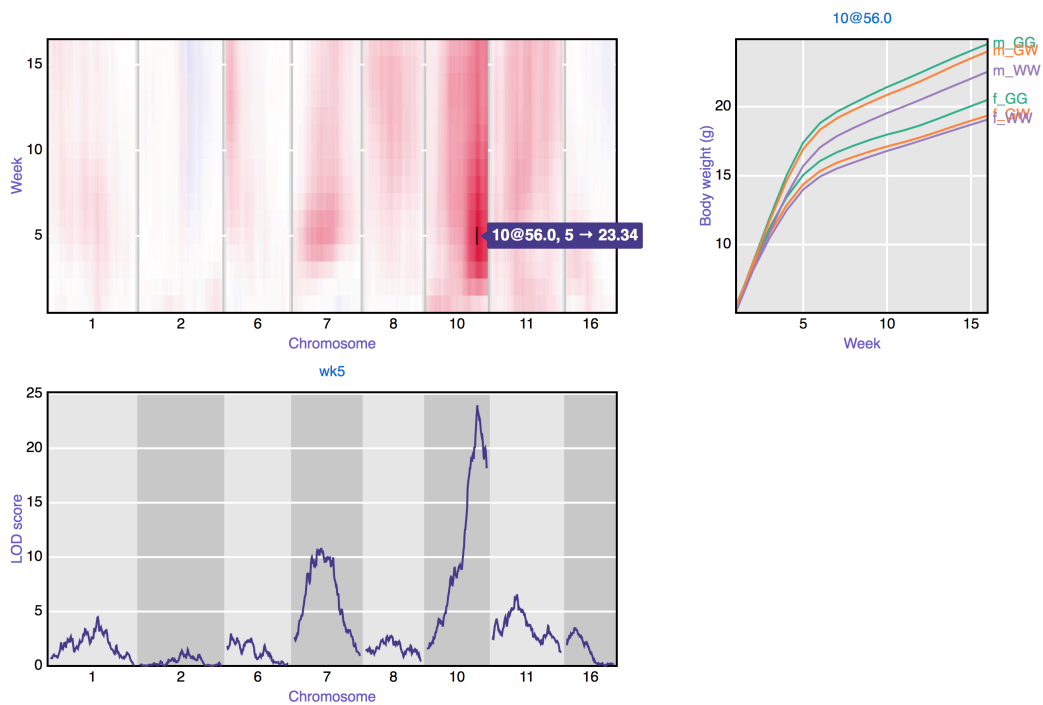


21

Here's a histogram of the results of a permutation test with 10,000 replicates.

A 5% significance threshold can be taken to be the 95th percentile of the results, which is about 3.9 in this case.

Genome scan for body weight



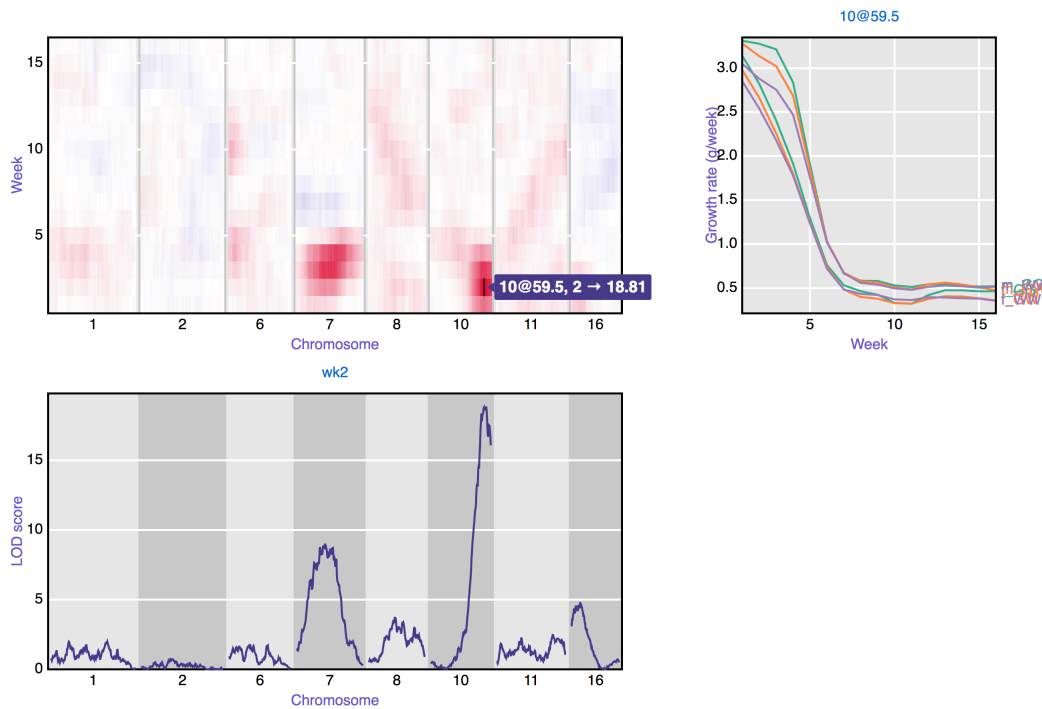
22

These are the results of genome scans for body weight, considering each time point individually. It is a snapshot of an interactive graph produced with R/qlcharts kbroman.org/qlcharts.

The top-left panel is a heat map of the LOD scores for selected chromosomes, with red indicating that the Gough allele causes increased body weight and blue indicating that the Gough allele causes decreased body weight. In the interactive version, if you hover over a point in the top-left panel, the LOD curves at that time point are shown below, and the estimated QTL effects as a function of time are shown on the right.

Strong QTL are seen on chr 7 and 10.

Genome scan for growth rate



23

These are analogous results to the previous slide, but for growth

The QTL on chr 7 and 10 are seen to affect growth rate in weeks 2–5.

Summary

- ▶ Multiple genetic loci contribute to body size in Gough × WSB.
- ▶ Gough alleles result in increased size.
- ▶ The major effects are on growth rate in first five weeks.

A summary of the main points.

Acknowledgments

Bret Payseur
Melissa Gray
Michelle Parmenter
Caley Hogan
Irene Ford

University of Wisconsin–Madison

Richard Cuthbert

Royal Society for the Protection of Birds

Peter Ryan

Percy FitzPatrick Institute of African
Ornithology, University of Cape Town

NIGMS/NIH

25

Life of a data science prof

- ▶ Collaborative data analysis
 - Gough Island mice
 - Obesity/diabetes in mice
 - Flu infection in mice
 - Meiotic recombination in mice and humans
 - Other projects in maize, rice, deer mice, ...
- ▶ Methods and software development
 - QTL mapping in multi-parent populations
 - Interactive graphics for statistical genomics
 - R/qlt2, R/qltcharts, GeneNetwork
- ▶ Teaching
 - Tools for Reproducible Research
 - Statistical Genomics
 - Data Analysis and Visualization
 - Software and Data Carpentry workshops
- ▶ Mentoring
 - One grad student + 2 summer students

26

Life of a data science prof (2)

- ▶ Program administration
 - Director of new PhD program in Biomedical Data Science
- ▶ Journal editing/reviewing
 - Senior editor for *Genetics*
 - Editorial boards at *PeerJ* and *BMC Biology*
 - Frequent reviewer for other journals + book publishers
- ▶ Committees
 - Faculty Senate
 - Steering committee, MD/PhD program
 - Genomics Advisory Committee, SMPH
 - Departmental committees
 - PhD thesis committees
- ▶ Travels
 - Scientific meetings
 - Departmental seminars
 - Various workshops