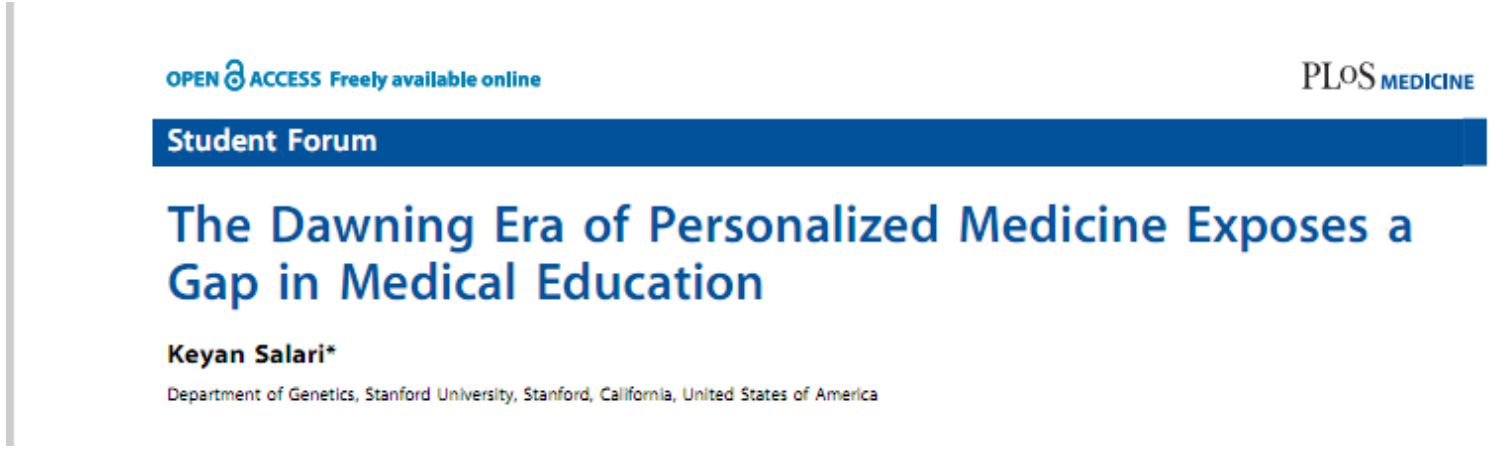
Please sign up for presentations and debates

1. Presentation to the class (>7 presentations needed)
Breast Cancer
Type 2 Diabetes
Alzheimer's/Parkinson's/ALS
other traits -

2. Debate (2 debates with two people each)
a. "Should DTC genetics be prescription only?".
b. "To know or not to know (Breast Cancer, Alzheimer's, ALS, Paternity)."

Readings for Wed. Mar. 30

http://stanford.edu/class/gene210/web/html/schedule.html



OPEN ACCESS Freely available online

PLoS MEDICINE

Student Forum

The Dawning Era of Personalized Medicine Exposes a Gap in Medical Education

Keyan Salari*

Department of Genetics, Stanford University, Stanford, California, United States of America

Perspective

JULY 22, 2010

The Path to Personalized Medicine

Margaret A. Hamburg, M.D., and Francis S. Collins, M.D., Ph.D.

Readings for Mon. Apr. 4

http://stanford.edu/class/gene210/web/html/schedule.html

How to Interpret a Genome-wide Association Study

Thomas A. Pearson, MD, MPH, PhD
Teri A. Manolio, MD, PhD

Genome-wide association (GWA) studies use high
nologies to assay hundreds of thousands of

Finding the missing heritability of complex diseases

Teri A. Manolio[1], Francis S. Collins[2], Nancy J. Cox[3], David B. Goldstein[4], Lucia A. Hindorff[5], David J. Hunter[6], Mark I. McCarthy[7], Erin M. Ramos[5], Lon R. Cardon[8], Aravinda Chakravarti[9], Judy H. Cho[10], Alan E. Guttmacher[1], Augustine Kong[11], Leonid Kruglyak[12], Elaine Mardis[13], Charles N. Rotimi[14], Montgomery Slatkin[15], David Valle[9], Alice S. Whittemore[16], Michael Boehnke[17], Andrew G. Clark[18], Evan E. Eichler[19], Greg Gibson[20], Jonathan L. Haines[21], Trudy F. C. Mackay[22], Steven A. McCarroll[23] & Peter M. Visscher[24]

turegenetics

Sizing up human height variation

Peter M Visscher

Reading for Wed. Apr. 6
Euan Ashley

nature genetics

Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease

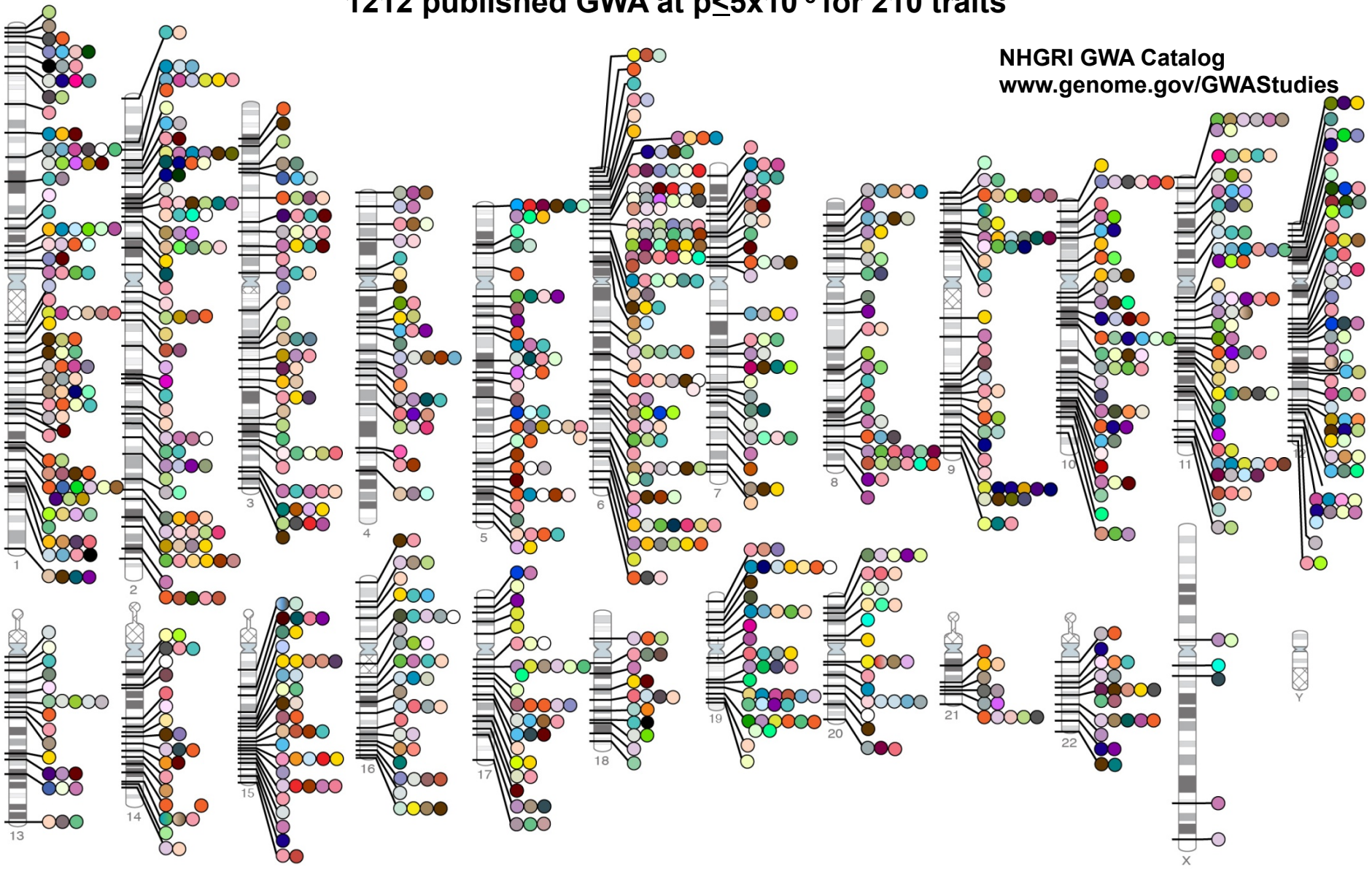
*Look up your genotype at 13 susceptibility loci at gene210.stanford.edu

Workshop 2
Gene210 Personalized Medicine and Genomics
March 30, 2011

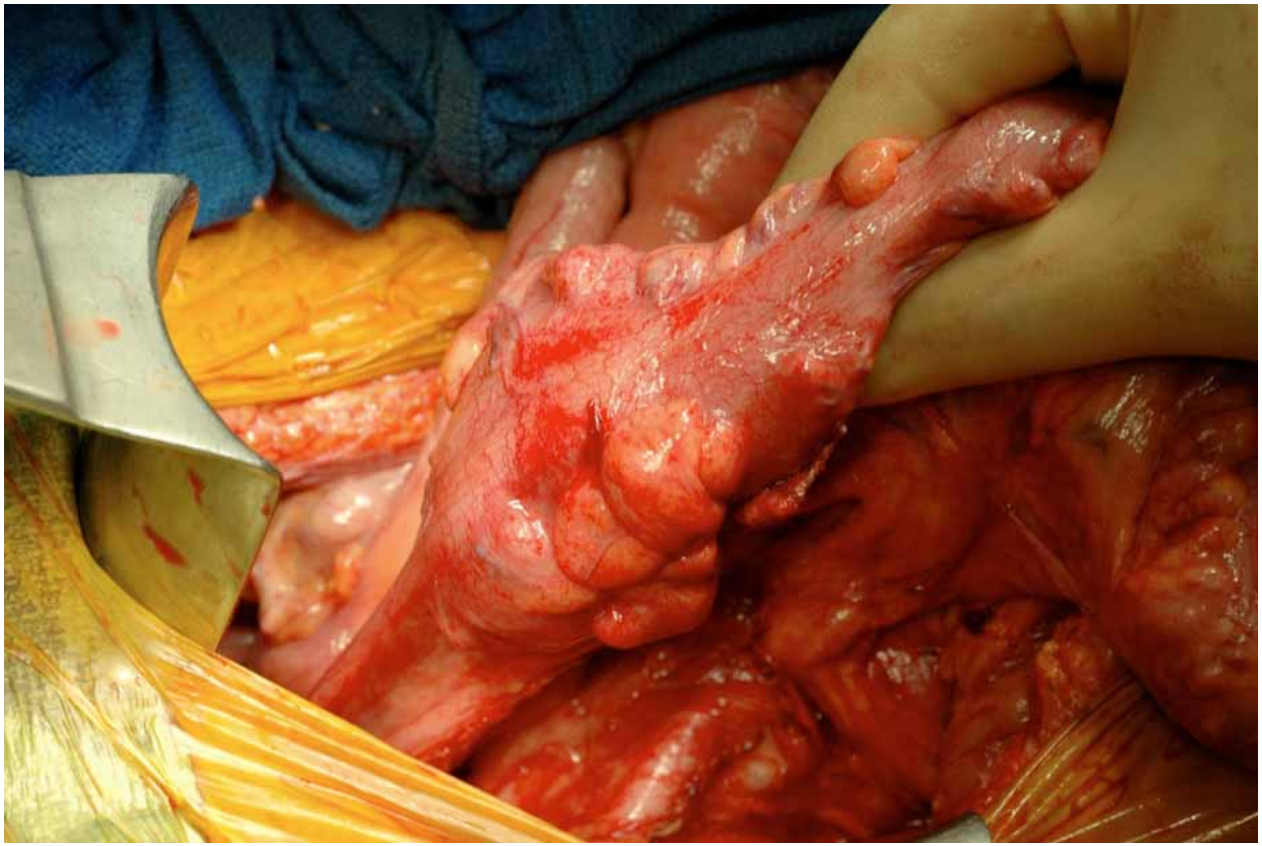
- Tour of dbSNP and hapmap (Nick)
- SNP imputer (Konrad)
- Example of colon cancer GWAS (Stuart)
- Gene210-wide association study (Konrad)

Published Genome-Wide Associations through 12/2010, 1212 published GWA at $p \leq 5 \times 10^{-8}$ for 210 traits

NHGRI GWA Catalog
www.genome.gov/GWAStudies



Colorectal cancer



1057 cases
960 controls

550K SNPs

Table 1 Risk of colorectal neoplasia associated with the 8q24 SNP rs6983267

Panel	Group	Total	Genotype			Frequency		Allele χ^2	*P* value	Allelic OR (for G against T)	95% c.i.
			GG	GT	TT	G	T				
A	All affected individuals	1,027[a]	352	486	189	0.579	0.421	31.79	1.72×10^{-7}	1.43	1.26–1.63
	Cancers only	620	202	302	116	0.569	0.431	18.96	1.34×10^{-5}	1.38	1.19–1.59
	Adenomas only	407[a]	150	184	73	0.595	0.405	25.01	5.70×10^{-7}	1.53	1.29–1.81
	Controls	960	235	471	254	0.490	0.510				
B	Colorectal cancers	4,261	1,334	2,216	811	0.560	0.440	29.71	5.02×10^{-8}	1.19	1.12–1.26

1027 Colorectal cancer

960 controls

Cancer: 0.57G 0.43T

controls: 0.49G 0.51T

Table 1 Risk of colorectal neoplasia associated with the 8q24 SNP rs6983267

			Genotype			Frequency			
Panel	Group	Total	GG	GT	TT	G	T	Allele χ^2	*P* value
A	All affected individuals	1,027[a]	352	486	189	0.579	0.421	31.79	1.72×10^{-7}
	Cancers only	620	202	302	116	0.569	0.431	18.96	1.34×10^{-5}
	Adenomas only	407[a]	150	184	73	0.595	0.405	25.01	5.70×10^{-7}
	Controls	960	235	471	254	0.490	0.510		
B	Colorectal cancers	4,361	1,334	2,216	811	0.560	0.440	29.71	5.02×10^{-8}

Are these different?

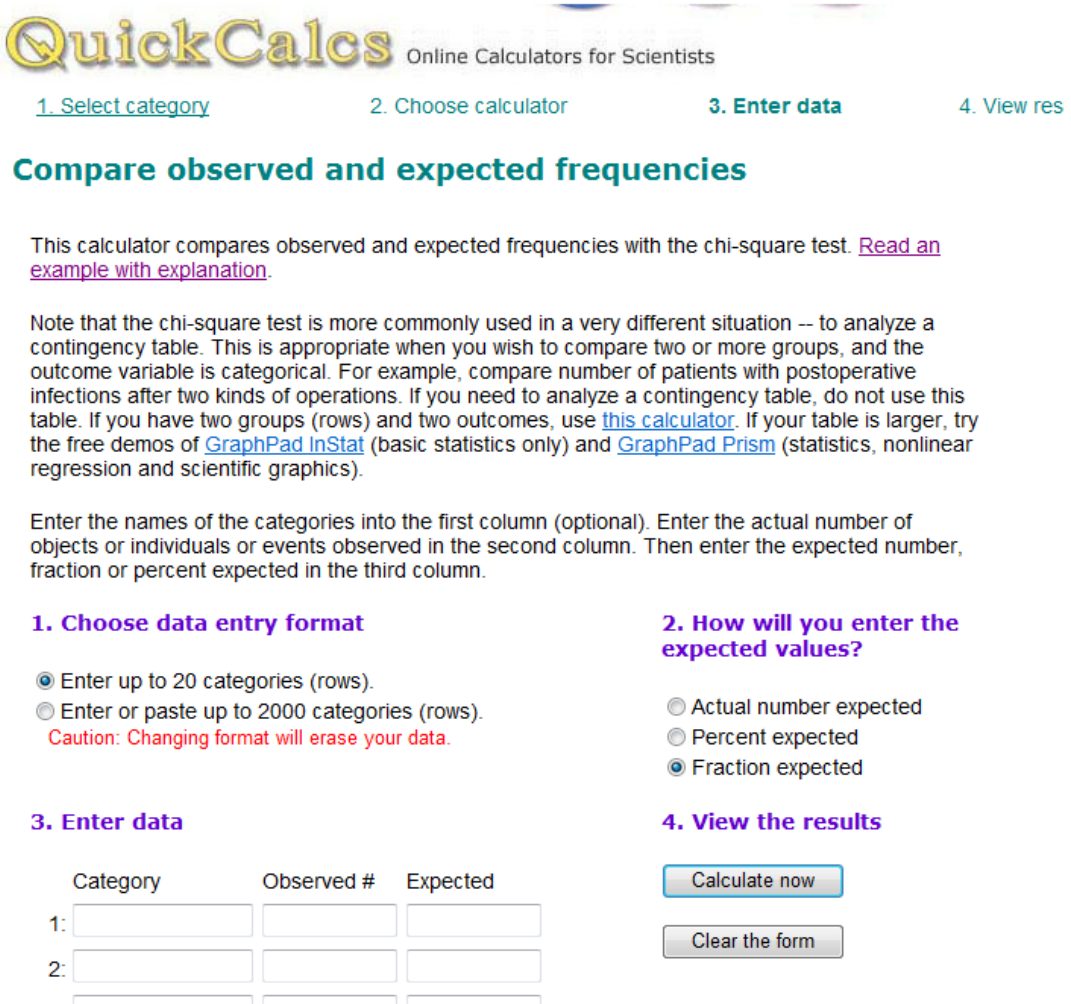
Cancer: 0.57G 0.43T

controls: 0.49G 0.51T

Chi squared

Chi squared

http://www.graphpad.com/quickcalcs/chisquared1.cfm



Chi squared

http://www.graphpad.com/quickcalcs/chisquared1.cfm

Table 1 Risk of colorectal neoplasia associated with the 8q24 SNP rs6983267

			Genotype			Frequency			
Panel	Group	Total	GG	GT	TT	G	T	Allele χ^2	*P* value
A	All affected individuals	1,027[a]	352	486	189	0.579	0.421	31.79	1.72×10^{-7}
	Cancers only	620	202	302	116	0.569	0.431	18.96	1.34×10^{-5}
	Adenomas only	407[a]	150	184	73	0.595	0.405	25.01	5.70×10^{-7}
	Controls	960	235	471	254	0.490	0.510		
B	Colorectal cancers	4,361	1,334	2,216	811	0.560	0.440	29.71	5.02×10^{-8}

Cancer: 352GG + 486GT = 1190 G alleles Controls: 0.49G

486GT + 189 TT = 864 T alleles 0.51T

Chi squared

http://www.graphpad.com/quickcalcs/chisquared1.cfm

1. Choose data entry format

- (•) Enter up to 20 categories (rows).
- ( ) Enter or paste up to 2000 categories (rows).

Caution: Changing format will erase your data.

2. How will you enter the expected values?

- ( ) Actual number expected
- ( ) Percent expected
- (•) Fraction expected

3. Enter data

	Category	Observed #	Expected
1:	G alleles	1190	.49
2:	T alleles	864	.51
3:			

4. View the results

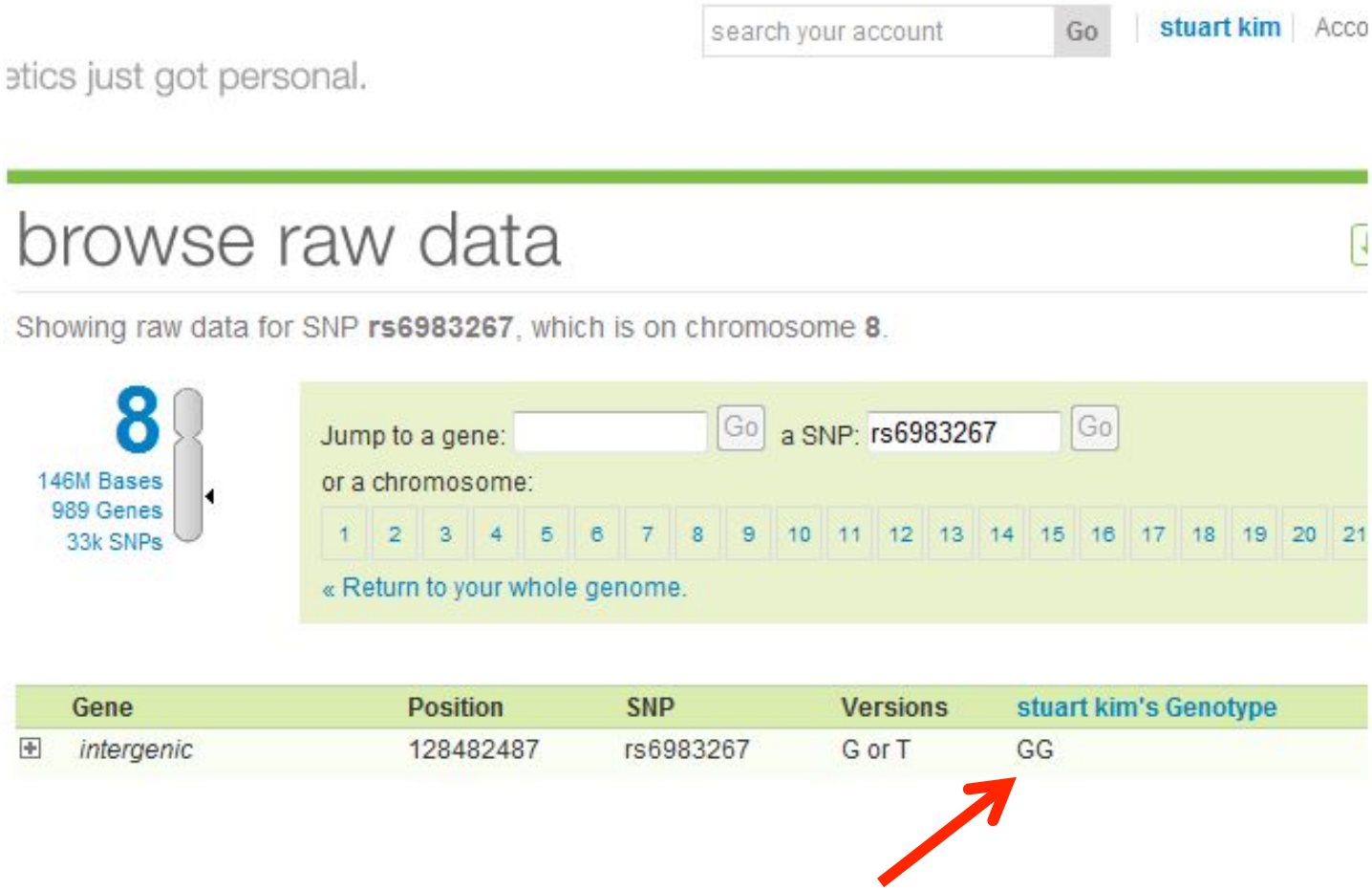
Calculate now

Clear the form

Chi squared = 31

P values = 10^{-7}

Stuart's genotype



Homozygous bad allele ☹

Allelic odds ratio

Cancer: 0.57G 0.43T

Cancer: G:T ratio = 0.57/.43 = 1.32

controls: 0.49G 0.51T

controls: G:T ratio = .49/.51 = .96

Allelic odds ratio = 1.32/.96 = 1.37

A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21

Ian Tomlinson[1,2], Emily Webb[3,13], Luis Carvajal-Carmona[1,13], Peter Broderick[3,13], Zoe Kemp[1,13], Sarah Spain[1,13], Steven Penegar[3], Ian Chandler[3], Maggie Gorman[1], Wendy Wood[3], Ella Barclay[1], Steven Lubbe[3], Lynn Martin[1], Gabrielle Sellick[3], Emma Jaeger[1], Richard Hubner[3], Ruth Wild[3], Andrew Rowan[1], Sarah Fielding[3], Kimberley Howarth[1], the CORGI Consortium, Andrew Silver[2], Wendy Atkin[4], Kenneth Muir[5], Richard Logan[5], David Kerr[6], Elaine Johnstone[6], Oliver Sieber[7], Richard Gray[8], Huw Thomas[9], Julian Peto[10,11], Jean-Baptiste Cazier[12] & Richard Houlston[3]

Much of the variation in inherited risk of colorectal cancer (CRC) is probably due to combinations of common low risk variants. We conducted a genome-wide association study of 550,000 tag SNPs in 930 familial colorectal tumor cases and 960 controls. The most strongly associated SNP ($P = 1.72 \times 10^{-7}$, allelic test) was rs6983267 at 8q24.21. To validate this

wide studies for disease associations. This approach is unbiased and does not depend upon prior knowledge of function or presumptive involvement of any gene in disease causation. Moreover, it minimizes the possibility of failing to identify important variants in hitherto-unstudied genes.

We initially genotyped 550,163 tag SNPs in 940 individuals with

"Of the 547,647 polymorphic tag SNPs, 27,673 showed an association with disease at P < .05."

Tomlinson et al., Nat. Genet. 39, 984, 2007

Multiple hypothesis testing

"Of the 547,647 polymorphic tag SNPs, 27,673 showed an association with disease at $P < .05$."

- $P = .05$ means that there is a 5% chance for this to occur randomly.
- If you try 100 times, you will get about 5 hits.
- If you try 547,647 times, you should expect $547{,}647 \times .05 = 27{,}382$ hits.
- So 27,673 (observed) is about the same as one would randomly expect.

Multiple hypothesis testing

"Of the 547,647 polymorphic tag SNPs, 27,673 showed an association with disease at $P < .05$."

- Here, have 547,647 SNPs = # hypotheses
- False discover rate = q = p x # hypotheses. This is called the Bonferroni correction.
- Want q = .05. This means a positive SNP has a .05 likelihood of rising by chance.
- At q = .05, p = .05 / 547,647 = $.91 \times 10^{-7}$
- This is the p value cutoff used in the paper.

Multiple hypothesis testing

"Of the 547,647 polymorphic tag SNPs, 27,673 showed an association with disease at $P < .05$."

- The Bonferroni correction is too conservative. It assumes that all of the tests are independent.
- But the SNPs are linked in haplotype blocks, so there really are less independent hypotheses than SNPs.
- Another way to correct is to permute the data many times, and see how many times a SNP comes up in the permuted data at a particular threshold.