

NCBI

WWW  
110101

# GAIN in dbGaP

## Updating and Versioning



# GAIN in dbGaP

## Updating and Versioning

- Updating existing studies
- Adding new phenotype data
- Recalculating genotype data
- Submitting Associations
- Follow up Studies

NCBI

WWW  
110101

# Finding dbGaP

<http://view.ncbi.nlm.nih.gov/dbGaP>

Entrez PubMed - Windows Internet Explorer

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed

File Edit View Favorites Tools Help

Entrez PubMed

NCBI PubMed A service of the National Library of Medicine and the National Institutes of Health

My NCBI [Sign In] [Register]

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals Books

Search dbGaP for [ ] Go Clear

Limits Preview/Index History Clipboard Details

- To get started, enter one or more search terms.
- Search terms may be [topics](#), [authors](#) or [journals](#).

**My NCBI** Set up an automated PubMed update in less than 5 minutes.

(1) Get a [My NCBI account](#). (2) Save your search.  
(3) Your PubMed updates can be e-mailed directly to you.

Read the [My NCBI Help](#) material to explore other options, such as automated updates of other databases, setting search filters, and highlighting search terms.

PubMed is a service of the [U.S. National Library of Medicine](#) that includes over 17 million citations from MEDLINE and other life science journals for biomedical articles back to the 1950s. PubMed includes links to full text articles and other related resources.

Write to the Help Desk

Internet 100%

# What's in dbGaP

dbGaP Home - Windows Internet Explorer

http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap&cmd=search&term=

dbGaP Home

NCBI

dbGaP GENOTYPE and PHENOTYPE

My NCBI [Sign In] [Register]

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals Books

Search dbGaP for [ ] Go Clear

Limits Preview/Index History Clipboard Details

Browse dbGaP TUTORIAL ABOUT dbGaP

By Studies By Diseases Advanced Search

Study	Embargo Release	Variables	Documents	Participants	Type of Study
<a href="#">Collaborative Association Study of Psoriasis</a>	July 10, 2008	-	-	2902	Case-control
<a href="#">Framingham SHaRE</a>	October 1, 2008	<a href="#">13183</a>	<a href="#">58</a>	15876	Community-based, longitudinal, family-based cohort
<a href="#">Genotyping the 270 HapMap samples for GAIN by Broad</a>		-	-	-	Parent-offspring trios
<a href="#">Genotyping the 270 HapMap samples for GAIN by Perlegen</a>		-	-	-	Parent-offspring trios
<a href="#">International Multi-Center ADHD Genetics Project</a>	March 26, 2008	<a href="#">438</a>	<a href="#">12</a>	2835	Parent-offspring trios
<a href="#">LEAPS</a>		-	-	886	Case-control
<a href="#">Linking Genome-Wide Association Study of Schizophrenia</a>	August 1, 2008	-	-	2400	Case-control
<a href="#">Major Depression: Stage 1 Genomewide Association in Population-Based Samples</a>	July 16, 2008	-	-	3786	Case-control
<a href="#">NEI Age-Related Eye Disease Study (AREDS)</a>	June 11, 2007	<a href="#">174</a>	<a href="#">37</a>	600	Case-control
<a href="#">NINDS Parkinsonism Study</a>		<a href="#">43</a>	<a href="#">4</a>	1283	Case-set
<a href="#">NINDS Repository Neurologically Normal Control Collection</a>		<a href="#">66</a>	<a href="#">2</a>	2723	Control-set
<a href="#">Search for Susceptibility Genes for Diabetic Nephropathy in Type 1 Diabetes</a>	July 16, 2008	-	-	1835	Case-control
<a href="#">Whole Genome Association Study of Bipolar Disorder</a>	August 1, 2008	-	-	-	Case-control

Done Internet 100%

# Studies

Study: International Multi-Center ADHD Genetics Project - Windows Internet Explorer

http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?id=phs000016

File Edit View Favorites Tools Help

Study: International Multi-Center ADHD Genetics Project





## International Multi-Center ADHD Genetics Project

Accession: phs000016.v1.p1

**Description**

The goal of the project is to complete a 600,000 tag SNP genome-wide association scan of 958 parent-child trios from the International Multisite ADHD Genetics (IMAGE) project, in order to assess the association of SNP markers with ADHD, analyze quantitative ADHD phenotypes, complete copy number analyses, assess parent of origin effects and season of birth effects, and test for epistasis among apparently uncorrelated genes.

All consent forms stipulate that the samples can only be used by researchers who have been approved by the National Institute of Mental Health (NIMH), National Institute of Health. All consent forms, except those used at the Zürich site (N=141 subjects), explicitly indicate that the samples may be used by researchers from commercial enterprises seeking to benefit financially from the analysis of the samples. The Zürich consent does not prohibit such use. The Zürich consent form also included an "opt out" that allowed the subjects to indicate that they did not want their samples stored at the NIMH repository or used by researchers external to the project. No subjects enrolled in the project opted out.

[GAIN The Genetic Association Information Network](#)

[Upstate Medical University - Medical Genetics Research Center](#)

- Participants: 2835
- Type: Parent-offspring trios

**Individual-Level Data**

- **Use restrictions**
  - **Consent Group**

ADHD

    - Consent limited to genetic studies of the pathophysiology or etiology of ADHD or its complications.
    - This consent group does not require IRB approval.
    - Participant set: 2835
- Embargo Release Date: March 26, 2008

**Search Within This Study**

Search for:  Go

**Associated Variables**

- International Multi-Center ADHD Genetics Project
- Sociodemography and Administration
- Affection Status
- Psychological and Psychiatric Observations
- Treatment

**Associated Documents**

- International Multi-Center ADHD Genetics Project
- Questionnaires
- Analysis Support Documents

Done Internet 100%

# Documents

Windows Internet Explorer - Description of Scoring Algorithms (dbGaP ID: phd000088) - Windows Internet Explorer  
 http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/GetDocument.cgi?id=phd000088

NCBI logo  
 dbGaP GENOTYPE logo



## Description of scoring algorithms

Accession: phd000088.1

>> [International Multi-Center ADHD Genetics Project](#)

[Download PDF version](#)



### International Multi-Site

### Description of Scoring

Desmond Campbell

15 November 2006

### Table of Contents

- [Purpose](#)
- + [Conners](#)
- + [Strengths and Difficulties Questionnaire](#)

### Purpose

This document describes the algorithm

- Conners N Subscale T score
- SDQ Hyperactivity subscale T score

used in the IMAGE project.

### Conners

Windows Internet Explorer - CONNERS' PARENTS RATING SCALE - Revised (L) (dbGaP ID: phd000095) - Windows Internet Explorer  
 http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/GetDocument.cgi?id=phd000095



## Conners' parents rating scale - revised (L)

Accession: phd000095.1

>> [International Multi-Center ADHD Genetics Project](#) >> [Conners' parents rating scale - revised \(L\)](#)

[Download PDF version](#)

### International Multi-Site ADHD Genetics Project

### CONNERS' PARENTS RATING SCALE - Revised (L)

C. Keith Conners Ph.D.

Identifier:

Center number

Family number

Individual number

Subject's date of birth

Gender  Male  
 Female

Age

In what situation was the child rated?  
 what are the values?




**Instructions:** Below are a number of common problems that children have. Please rate each item according to your child's behaviour in the last month. For each item, ask yourself 'How much of a problem has this been in the last month?', and check the best answer for each one. If none, not at all, seldom or very infrequently, you would check 0. If very much true or it occurs very often or frequently, you would check 3. You would check 1 or 3 for ratings in between. Please respond to all the items.

# Tables and Variables

Variable: spq2 - Windows Internet Explorer

http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/variable.cgi?id=phv00011497

Variable: spq2

**spq2**

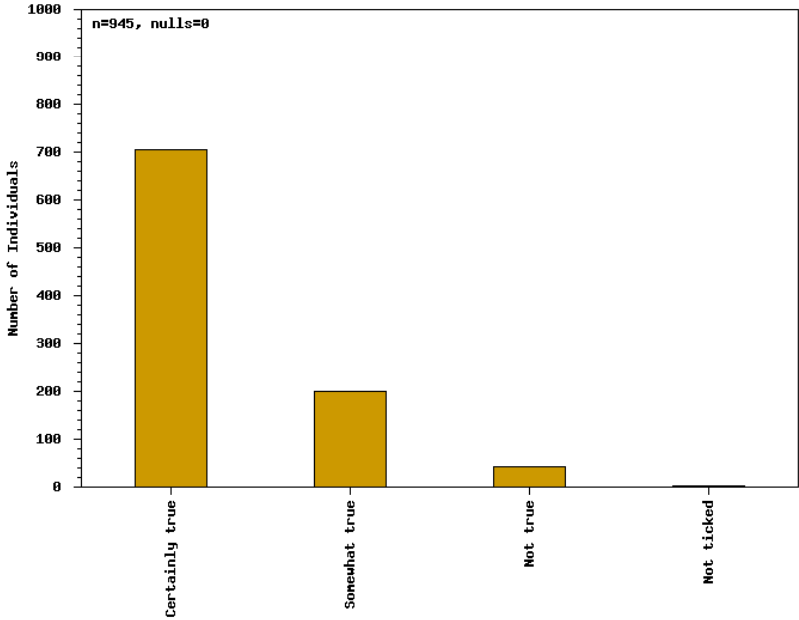
Accession: phv00011497.v1.p1

>> [International Multi-Center ADHD Genetics Project](#) >> [spq2](#)

**Description**

Is restless, overactive, cannot stay still for long

**Summary**



Response Category	Number of Individuals
Certainly true	~700
Somewhat true	~200
Not true	~40
Not ticked	~0

Done

Internet 100%

# Consent groups

Adobe Acrobat Professional - [Study\_Report phs000007 FHS v1 p1.pdf]

File Edit View Document Comments Tools Advanced Window Help

Note Tool Text Edits Stamp Tool Show

Select Object Data Tool 110%

Bookmarks Signatures Pages Model Tree Attachments Comments

**dbGaP Study Configuration Report**  
The Framingham Heart Study

**Distribution Set** [General\\_research\\_use](#)

This consent group is released with the filename pattern [phs000007.FHS.v1.p1.c1.GRU](#)

**Data Use Restrictions**

restriction label **General Research Use**

full restriction text  
The informed consent document signed by the Framingham Heart Study Participants allows use of these data by investigators employed by non-profit and for-profit organizations. These data may be used by private companies in the development of diagnostics and therapeutics under the current consent.

IRB approval required to use the data **Yes**

public link <http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?id=phs000007#restricted-access>

**Set Information**

distribution set id **1**

number of individuals **6907**

distribution set comments  
This data set phs000007.v1.p1.c1 (SHARe Framingham) is composed of 6907 individuals who gave consent for GENERAL RESEARCH USE. These individuals are specific to this distribution set within dbGaP and separate from permission set 2 – NOT-FOR-PROFIT USE ONLY. This distribution set includes the following components: 117 substudy files of phenotype traits; complete set of genotype calls (unfiltered set) in two formats (INDIVIDUAL, MATRIX) for 6775 participants; subsets of high quality genotype calls (filtered to exclude Mendelian transmission errors) in MATRIX format; genotype QC information; and genotype intensity files (Affymetrix CEL format).

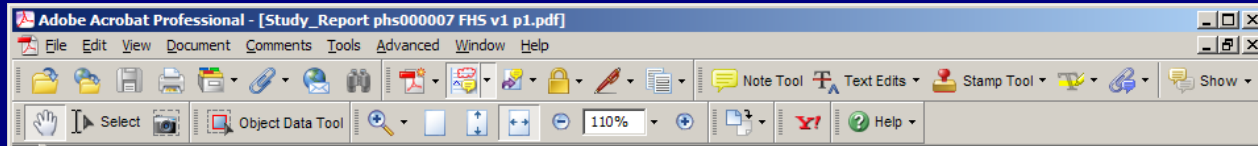
**Files distributed in this set**

[Refer to index number at the end of this report for a detailed description of each file.](#)

4 of 282

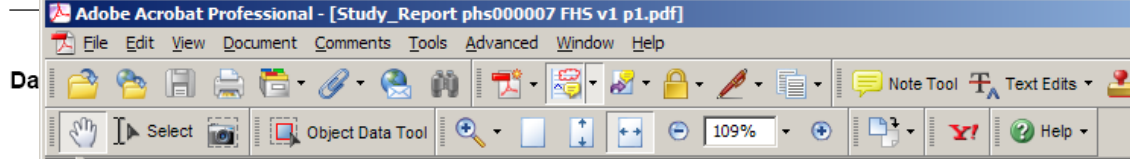


# Tables



**Distribution Set** General\_research\_use

This consent group is released with the filename pattern `_phs000008_FHS.v1.p1.c1.GRU`



204

content type	phenotype-individual-traits
description	hearing test exam 22 GRU participants
location	NHLBI/SHARE/Framingham/phs000068v1/p1/phs000068.pht000158.v1.p1.c1.heard0_22s.GRU.txt.gz
release date	2007-10-01
embargo release date	2008-10-01
filesize, kilobytes	9

General research use (GRU) participants. Table columns include IDTYPE, RELIAB\_A, OTOSCOPI, EXAM\_15, CONSENT, HHIE, EXAM\_18\_, EXAM\_181, Q\_COMPLE, QUAL\_PAR, PART\_B\_R, PTA\_R, PTA\_L, W22\_R\_PC, W22\_R\_HL, W22\_L\_PC, W22\_L\_HL, DPOAE\_CO, RELIAB\_B, CID\_R\_PC, CID\_R\_HL, CID\_L\_PC, CID\_L\_HL, OD\_R\_50, OD\_R\_70, OD\_R\_90, OD\_L\_50, OD\_L\_70, OD\_L\_90, PL\_R\_HL, PL\_L\_HL, MCR\_R\_10, MCR\_L\_10, MCR\_R\_0, MCR\_L\_0, MCR\_RTST, MCR\_RTS1, FIRST\_EA, DSI\_R\_PL, DSI\_L\_SC, DSI\_L\_PL, DSI\_R\_SC, SSW\_R\_PL, SSW\_L\_PL, SSW\_RNC, SSW\_RC, SSW\_LC, SSW\_LNC, SPIN\_R\_B, SPIN\_L\_B, SPIN\_R\_P, SPIN\_L\_P, SPIN\_R\_F, SPIN\_L\_F, SPIN\_R\_H, SPIN\_L\_H, SPIN\_R\_L, SPIN\_L\_L, SPIN\_R\_1, SPIN\_L\_1, DECLINEB, edate\_a, edate\_b, shareid. Table has 285 rows representing 285 individuals. The data dictionary is available at [ftp://ftp.ncbi.nlm.nih.gov/dbgap/NHLBI/SHARE/Framingham/phs000007.v1.p1-Framingham/substudies/HearingTest/phs000068.v1.p1-HearingTest/data-dictionary/phs000068.pht000158.v1.p1.heard0\\_22s.data\\_dict\\_2007\\_09\\_19.xml](ftp://ftp.ncbi.nlm.nih.gov/dbgap/NHLBI/SHARE/Framingham/phs000007.v1.p1-Framingham/substudies/HearingTest/phs000068.v1.p1-HearingTest/data-dictionary/phs000068.pht000158.v1.p1.heard0_22s.data_dict_2007_09_19.xml)

251 of 282

Refer to index number at the end of this report for a detailed description of each file.

phenotype-individual-traits

Cohort exam 1 - 7 GRU participants

6 NHLBI/SHARE/Framingham/phs000008v1/p1/phs000008.pht000009.v1.p1.c1.ex0\_7s.GRU.txt.gz

Embargo release after 2008-10-01

Filesize bytes

22 of 282

# Genotypes

Adobe Acrobat Professional - [Study\_Report phs000007 FHS v1 p1.pdf]

File Edit View Document Comments Tools Advanced Window Help

Note Tool Text Edits Stamp Tool Show

Select Object Data Tool 109%

Bookmarks Signatures Pages Model Tree Attachments Comments

**dbGaP Study Configuration Report**  
The Framingham Heart Study

dbGaP  
GENOTYPE and PHENOTYPE

genotype-calls-unfiltered-matrix-format  
phg000004: Affy 50K unfiltered genotype data (QC unfiltered). NPU participants. Matrix content = genotypes

259 NHLBI/SHARE/Framingham/phs000007v1/p1/phg000004v1/phg000004.FHS.genotype-calls.Affy50K.v1.p1.c2.NPU.unfiltered.matrixfmt.genotype.tar  
Embargo release after 2008-10-01  
Filesize 37 megabytes

genotype-calls-unfiltered-matrix-format  
phg000004: Affy 50K unfiltered genotype data (QC unfiltered). NPU participants. Matrix content = confidence scores

260 NHLBI/SHARE/Framingham/phs000007v1/p1/phg000004v1/phg000004.FHS.genotype-calls.Affy50K.v1.p1.c2.NPU.unfiltered.matrixfmt.confidence.tar  
Embargo release after 2008-10-01  
Filesize 957 megabytes

genotype-calls-unfiltered-matrix-format  
phg000004: Affy 50K unfiltered genotype data (QC unfiltered). NPU participants. Matrix content = allele 1 intensity

261 NHLBI/SHARE/Framingham/phs000007v1/p1/phg000004v1/phg000004.FHS.genotype-calls.Affy50K.v1.p1.c2.NPU.unfiltered.matrixfmt.allele1\_intensity.tar  
Embargo release after 2008-10-01  
Filesize 899 megabytes

genotype-calls-unfiltered-matrix-format  
phg000004: Affy 50K unfiltered genotype data (QC unfiltered). NPU participants. Matrix content = allele 2 intensity

262 NHLBI/SHARE/Framingham/phs000007v1/p1/phg000004v1/phg000004.FHS.genotype-calls.Affy50K.v1.p1.c2.NPU.unfiltered.matrixfmt.allele2\_intensity.tar  
Embargo release after 2008-10-01  
Filesize 895 megabytes

genotype-individual-information  
phg000005: Affy 100K genotypes - individual information

16 of 282

# NCBI



# Associations

Analysis: Genome-Wide Allelic Association of AMD Status in Illumina 100k Chip - Windows Internet Explorer

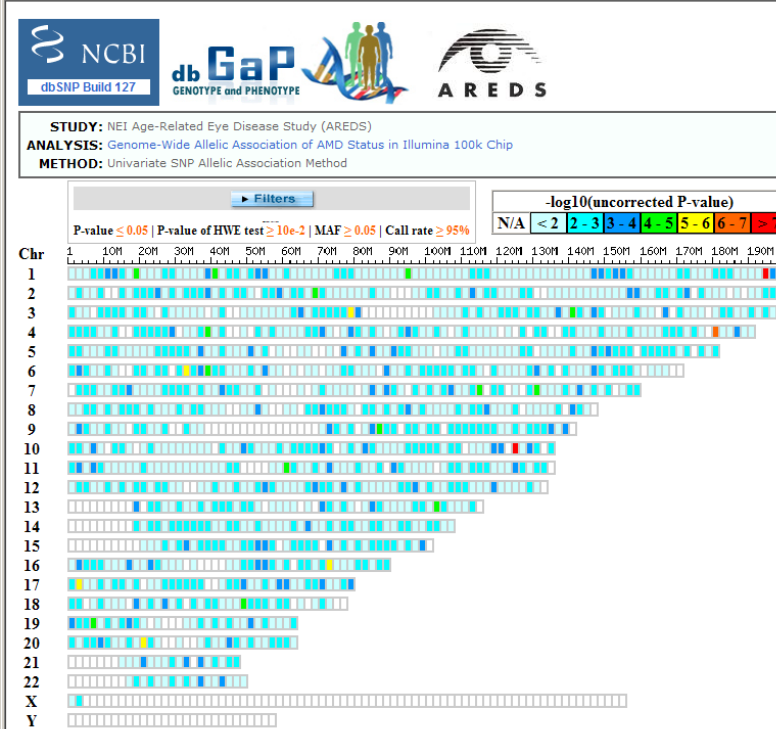
http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/analysis.cgi?id=pha000001

GaP Chromosome Browser - Windows Internet Explorer

http://www.ncbi.nlm.nih.gov/SNP/GaP.cgi?rm=plot&frame&test\_id=1&chr=1&from=194000000&to=196000000&method\_id=

GaP Genome View - Windows Internet Explorer

http://www.ncbi.nlm.nih.gov/SNP/GaP.cgi?rm=genomeTraitView&test\_id=1



dbSNP Build 127 dbGaP GENOTYPE and PHENOTYPE AREDS

STUDY: NEI Age-Related Eye Disease Study (AREDS)

Chromosome 1 From 194000000 To 196000000

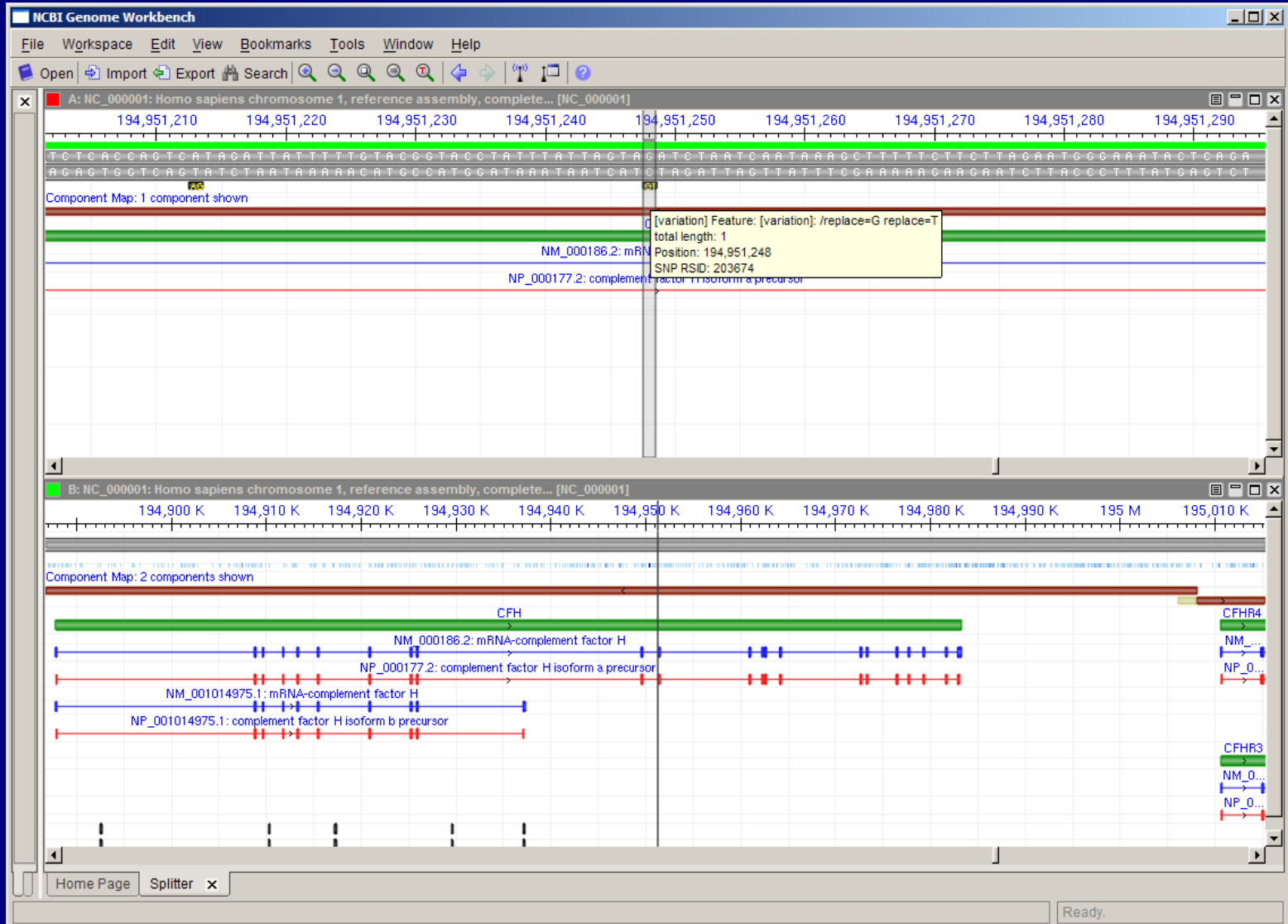
Traits/Analysis:  Genome-Wide All

Maps:  MapViewer

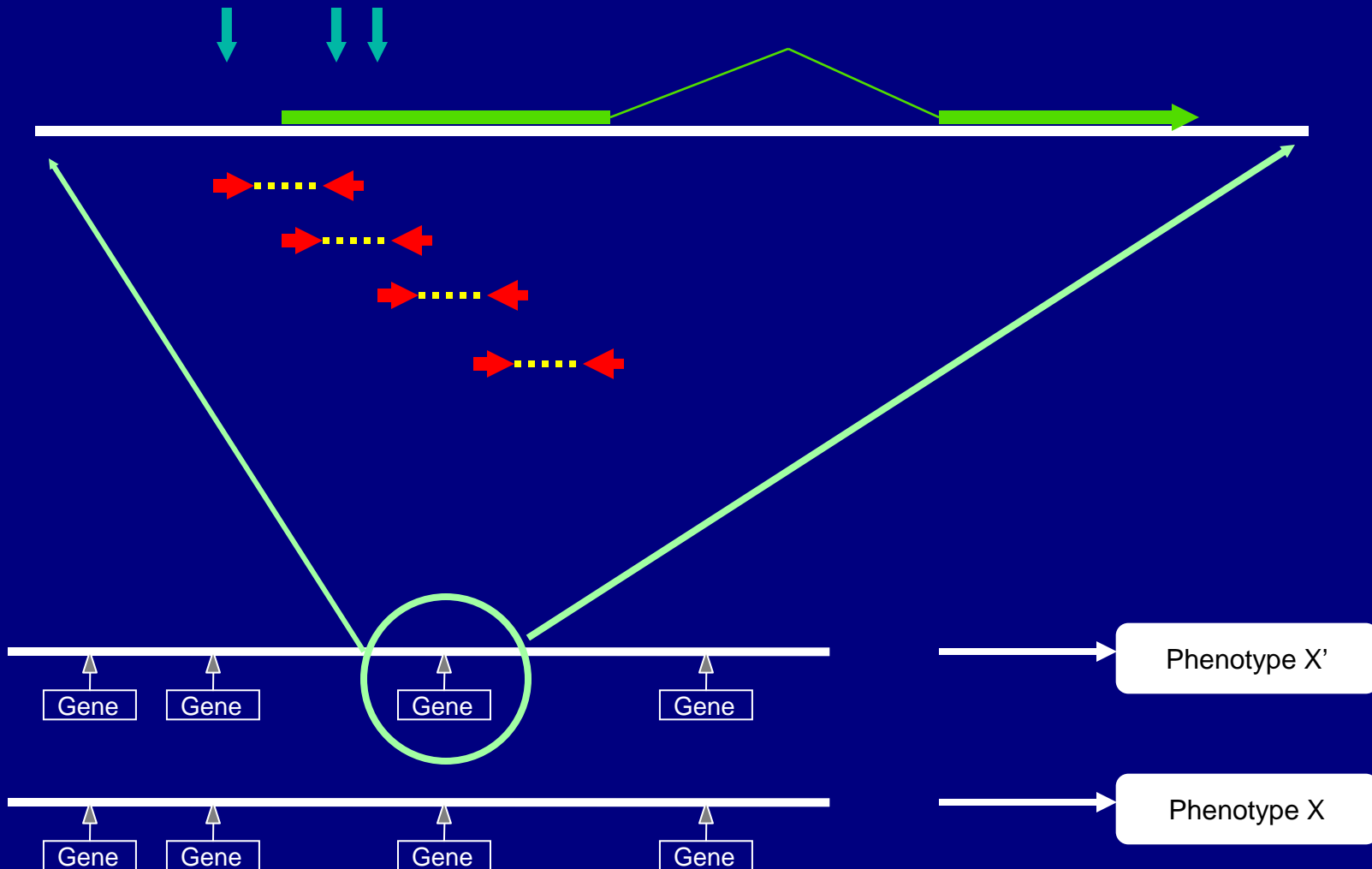
Links	Marker ID	Chr	Position	P-value	P-value Rank	MAF (control)	MAF (case)	HWE(control)	HWE(case)
	rs2989966	1	194013113	0.13	29/62	0.44	0.49	0.38	0.0
	rs921516	1	194121260	0.09	27/62	0.47	0.48	0.20	0.8
	rs1451912	1	194133832	0.55	46/62	0.42	0.44	0.38	0.6
	rs10494738	1	194178191	0.34	41/62	0.17	0.15	0.08	0.8
	rs12032769	1	194200833	0.78	56/62	0.39	0.40	0.55	0.0
	rs12402403	1	194325083	0.89	59/62	0.49	0.48	0.31	0.4

MapViewer showing variation and gene tracks for the region. Genes shown include ROP2, CFH3, CFH4, CFH5, F13B, GDF7, LOC127411, LOC127413, LOC127414, LOC127415, LOC127416, LOC127417, LOC127418, LOC127419, LOC127420, LOC127421, LOC127422, LOC127423, LOC127424, LOC127425, LOC127426, LOC127427, LOC127428, LOC127429, LOC127430, LOC127431, LOC127432, LOC127433, LOC127434, LOC127435, LOC127436, LOC127437, LOC127438, LOC127439, LOC127440, LOC127441, LOC127442, LOC127443, LOC127444, LOC127445, LOC127446, LOC127447, LOC127448, LOC127449, LOC127450, LOC127451, LOC127452, LOC127453, LOC127454, LOC127455, LOC127456, LOC127457, LOC127458, LOC127459, LOC127460, LOC127461, LOC127462, LOC127463, LOC127464, LOC127465, LOC127466, LOC127467, LOC127468, LOC127469, LOC127470, LOC127471, LOC127472, LOC127473, LOC127474, LOC127475, LOC127476, LOC127477, LOC127478, LOC127479, LOC127480, LOC127481, LOC127482, LOC127483, LOC127484, LOC127485, LOC127486, LOC127487, LOC127488, LOC127489, LOC127490, LOC127491, LOC127492, LOC127493, LOC127494, LOC127495, LOC127496, LOC127497, LOC127498, LOC127499, LOC127500.

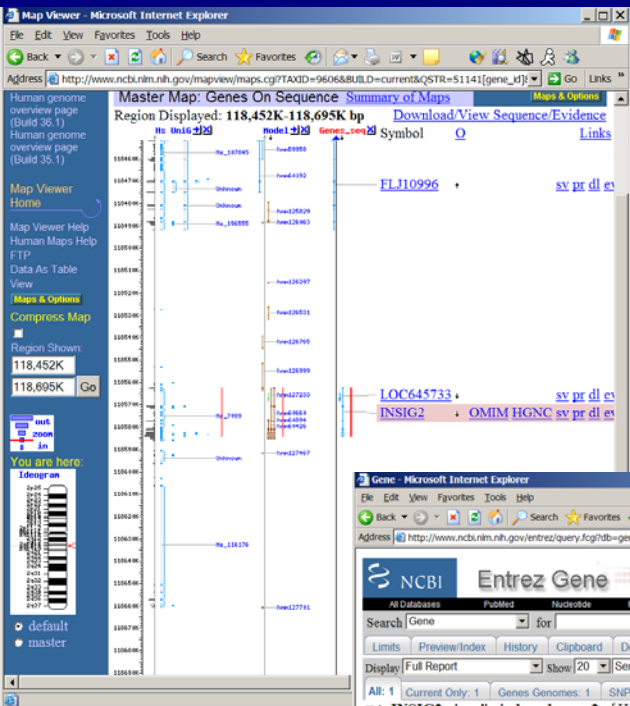
# Associations to the Basepair



# The Medical Sequencing Project – Finding the Causative Mutations



# Closing the Loop



Entrez GEO - Gene Expression Omnibus - Microsoft Internet Explorer  
 Address: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene&cmd=Display&dopt=gene\_geo&from\_uid=51141  
 Search: GEO Profiles for  
 Display: Summary Show 20  
 All: 221  
 Items 1 - 20 of 221  
 1: GDS501 record | GPL371 27849 [Homo Annotation: INSIG2: insulin induced gene 2 (Homo sapiens)]  
 Reporter: AL080184  
 Experiment: Inflammatory cytokine effect on gene expression array-based log

Entrez PubMed - Microsoft Internet Explorer  
 Address: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list\_uids=110101  
 Search: PubMed for  
 Display: Abstract Show 20 Sort by  
 All: 1 Review: 0  
 1: Science. 2006 Apr 14;312(5771):279-83.  
 A common genetic variant is associated with adult and childhood obesity.

Entrez Gene - Microsoft Internet Explorer  
 Address: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene&cmd=Retrieve&dopt=full\_report&list\_uids=51141  
 Search: Gene for  
 Display: Full Report Show 20  
 All: 1 Current Only: 1 Genes Genomes: 1 SNP GeneView: 1  
 1: INSIG2 insulin induced gene 2 [Homo sapiens]  
 GeneID: 51141 Primary source: HGNC:20452 updated 19-Apr-2006  
**Summary**  
**Official Symbol:** INSIG2 **and Name:** insulin induced gene 2 provided by HUGO Gene Nomenclature Committee  
**See related:** HPRD:16361, MIM:608660  
**Gene type:** protein coding  
**Gene name:** INSIG2  
**Gene description:** insulin induced gene 2  
**RefSeq status:** Reviewed  
**Organism:** *Homo sapiens*  
**Lineage:** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorhini; Catarrhini; Hominoidea; Homo  
**Gene aliases:** MGC26273  
**Summary:** The protein encoded by this gene is highly similar to the protein product encoded by gene INSIG1. Both INSIG1 protein and this protein are endoplasmic reticulum proteins that block the processing of sterol regulatory element binding proteins (SREBPs) by binding to SREBP cleavage-activating protein (SCAP), and thus prevent SCAP from escorting SREBPs to the Golgi.  
**Genomic regions, transcripts, and products**  
**RefSeq below**

Entrez PubMed  
 Overview  
 Help | FAQ  
 Tutorials  
 News/Noteworthy  
 E-Utilities

PubMed  
 Services  
 Journals Database  
 MeSH Database  
 Single Citation  
 Matcher  
 Batch Citation  
 Matcher  
 Clinical Queries  
 Special Queries  
 LinkOut  
 My NCBI

Related  
 Resources  
 Order Documents  
 NLM Mobile  
 NLM Catalog  
 NLM Gateway  
 TOXNET  
 Consumer Health  
 Clinical Alerts

**A common genetic variant is associated with adult and childhood obesity.**

**Herbert A, Gerry NP, McQueen MB, Heid IM, Pfeuffer A, Illig T, Wichmann HE, Meitinger T, Hunter D, Hu FB, Colditz G, Hinney A, Hebebrand J, Koberwitz K, Zhu X, Cooper R, Ardlie K, Lyon H, Hirschhorn JN, Laird NM, Lenburg ME, Lange C, Christman MF.**

Department of Genetics and Genomics, Boston University Medical School, E613, 715 Albany Street, Boston, MA 02118, USA. aherbert@bu.edu

Obesity is a heritable trait and a risk factor for many common diseases such as type 2 diabetes, heart disease, and hypertension. We used a dense whole-genome scan of DNA samples from the Framingham Heart Study participants to identify a common genetic variant near the INSIG2 gene associated with obesity. We have replicated the finding in four separate samples composed of individuals of Western European ancestry, African Americans, and children. The obesity-predisposing genotype is present in 10% of individuals. Our study suggests that common genetic polymorphisms are important determinants of obesity.