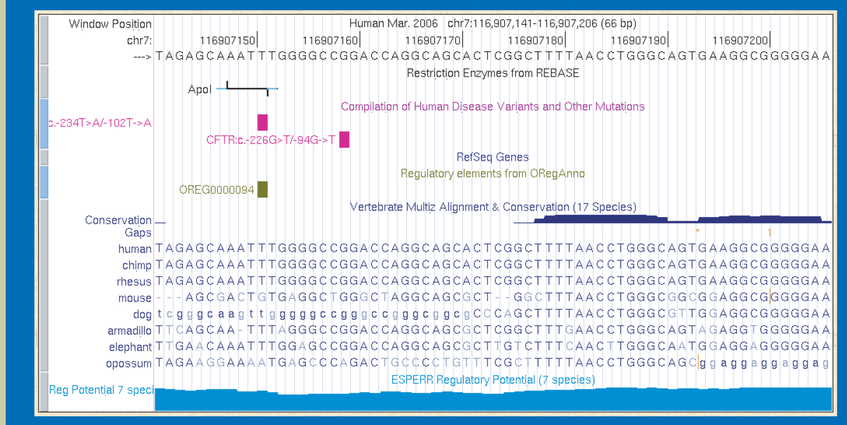
PHENCODE: CONNECTING GENOME AND PHENOTYPE

Giardine, B.[1], Riemer, C.[1], Trumbower, H.[2], Hsu, F.[2], Kent, W.J.[2], Hardison, R.C.[1], Miller, W.[1]
[1]Center for Comparative Genomics and Bioinformatics, Penn State University, University Park, PA;
[2]Center for Biomolecular Science and Engineering, University of California, Santa Cruz, CA

Acknowledgements:
Work was supported by NIH grants HG002238 (WM) and DK65806 (RH), and NHGRI grant 1P41HG02371 (WJK)

1. Example 1 follows links from the UCSC Genome Browser back to the data sources. In this region upstream of the CFTR gene, the Locus Variants track shows a substitution at the site of an ORegAnno regulatory element.



2. The details page for the Locus Variants track gives more details on the variant as well as a link to the source.

Compilation of Human Disease Variants and Other Mutations (CFMDB_712)

HGVS name: CFTR:c.-234T>A
Position: chr7:116713866-116713866
Band: 7q31.2
Genomic Size: 1
View DNA for this feature
source: LSDB; Cystic Fibrosis Mutation Database
location: not within known transcription unit
type: substitution

Common name:
-102T->A
External links:
CFMDB - 712
Type of mutation:
Promotor
RNA nucleotide change:
T to A at -102
Variation and Disease information related to gene locus:
OMIM title - CYSTIC FIBROSIS TRANSMEMBRANE CONDUCTANCE REGULATOR - 602421

View table schema

Data last updated: 2006-09-19

3. Details at CFMDB for this variant.

Nucleotide Change	T to A at -102
Exon	5' flanking
Consequence	regulatory mutation?
Original Report	This possible mutation(?) was found by DGGE then direct sequencing in the region upstream from the CFTR cap site. This change was not found in 200 normal alleles of our series. It was associated with the S549R(T->G) mutation in a CF patient who carries S945L on the other chromosome. It was also detected in another CF patient with genotype S549R(T->G)/[delta]F508, but no parental DNA was available at that time to further determine on which allele -102T->A is carried.
Contributors	Claustres M, Romey M C, Guittard C, Desgorges M, Carles S 1997-01-30
Institute	Institut de Biologie Montpelier
Updated Phenotypic Details	The mutation was found in 2 CF patients: -one male, 10 years old, diagnosed at 9 years, PS, FEV1 89%, sweat chloride 90 mmol/l. It was associated with the S549R(T->G) mutation in the CF patient who carries S945L on the other chromosome. -one female, 6 years old, diagnosed at 3,5 years of age. She is PS, has FEV1 89% and sweat chloride 122 mmol/l. Her genotype is S549R(T->G)/[delta]F508, but no parental DNA was available at that time to further determine on which allele -102T->A is carried. (pers. corr. Claustres)
Reference	Claustres et al. (NL#69)

4. The details page at UCSC for the ORegAnno track. For more detailed information follow the link to ORegAnno.

Regulatory elements from ORegAnno (OREG0000094)

ORegAnno ID: OREG0000094
Strand: +
Position: chr7:116907151-116907151
Band: 7q31.2
Genomic Size: 1
View DNA for this feature

Links:
ORegAnno - OREG0000094
PubMED - 10652351
Region type:
REGULATORY POLYMORPHISM
Gene:
CFTR
Entrez Gene - 1080
Transcription factor:
YY1
Entrez Gene - 7528
Evidence subtype(s):
Direct gel shift; Gel shift competition; Supershift; Transient transfection luciferase assay

View table schema

Data last updated: 2006-09-12

Aims:

- Connect genome data (evolutionary history, function) with phenotype and clinical data
- Facilitate better understanding of the associations between genotype and phenotype
- Generate novel hypotheses for pathological mechanisms

URLs:

genome.ucsc.edu
www.bx.psu.edu

5. ORegAnno record details.
Reference: J Biol Chem. 2000 Feb 4;275(5):3561-7

Record Details: Record Evidence

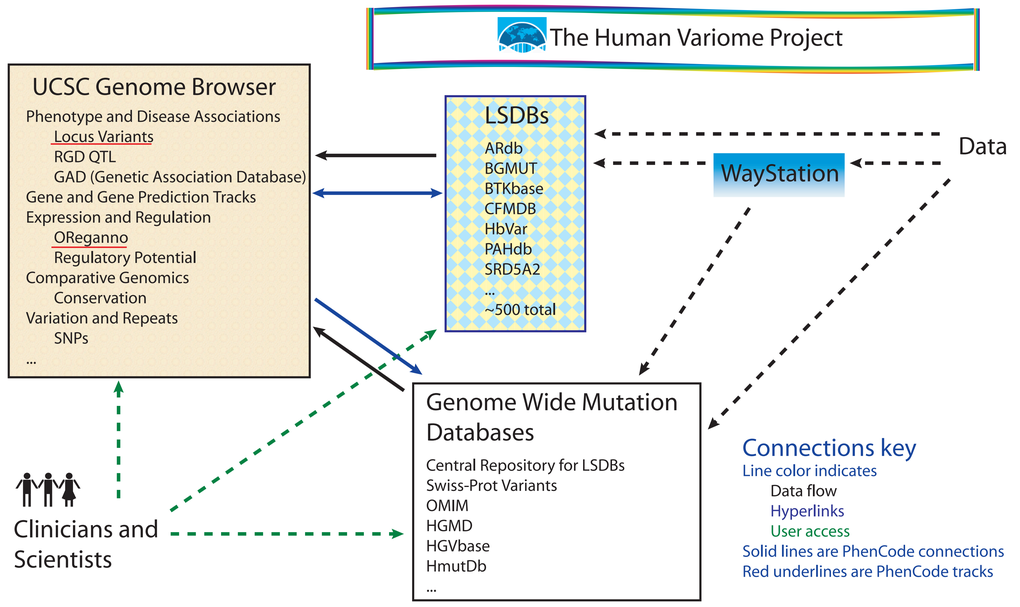
Evidence class: **UNKNOWN** (OREGEC00000)
Evidence type: **Reporter gene assay** (OREGET00002)
Evidence subtype: **Transient transfection luciferase assay** (OREGES00004)
Evidence comment: Cell lines: Calu-3, CF-PAC1, and HeLa. A allele had increased promoter activity by 60%, 66%, and 45% in the respective cell types.

Evidence class: **Transcription regulator** (OREGEC00002)
Evidence type: **Electrophoretic mobility shift assay (EMSA)** (OREGET00001)
Evidence subtype: **Direct gel shift** (OREGES00001)
Evidence comment: Oligonucleotides assayed harboring various cis-acting consensus sequences. Minimal promoter may act with SRF-like protein.

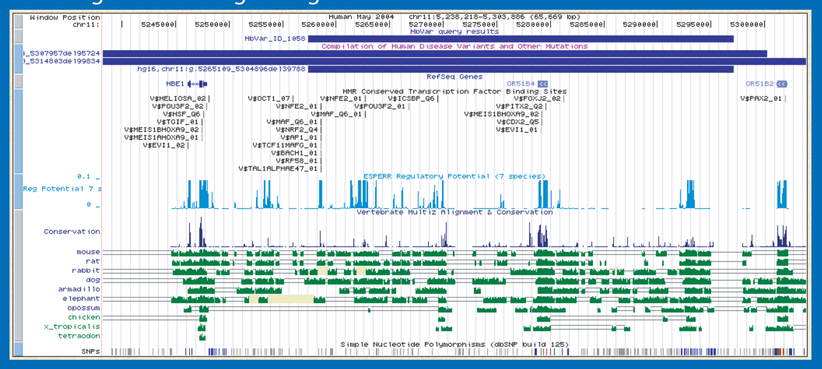
Evidence class: **Transcription regulator** (OREGEC00002)
Evidence type: **Electrophoretic mobility shift assay (EMSA)** (OREGET00001)
Evidence subtype: **Gel shift competition** (OREGES00003)
Evidence comment: Tested SRF-1 and YY1 protein. Factors could both occupy variant binding region.

Evidence class: **Transcription regulator** (OREGEC00002)
Evidence type: **Electrophoretic mobility shift assay (EMSA)** (OREGET00001)
Evidence subtype: **Supershift** (OREGES00002)
Evidence comment: Supershift against transcription factors binding YY1 and CArG sites. Supershifted YY1 not SRF1

PhenCode Connections



7. Eight variants are found and can be viewed in the UCSC Genome Browser. Zooming in on one deletion, we find that it contains several conserved transcription factor binding sites and segments with high regulatory potential. This deletion removes the major distal enhancer for the gene encoding beta-globin.



6. Example 2 shows connections from a locus specific database to the UCSC Genome Browser. Starting at HbVar, we search for all variants with an alpha-globin/beta-globin ratio >= 2.5, which is characteristic of beta-thalassemia.

HbVar: A database of Human Hemoglobin Variants and Thalassemias

Query Page

Name:
Category: Any
Type of Thalassemia: beta0, beta+, beta (0 or + unclear)
Chain: Agamma Ggamma alpha1 alpha2 beta delta zeta1 zeta2
Location: 3' UTR 5' UTR exon intron not within known transcription unit unknown

Mutation data
Contact
Haplotype
Hematology

Select one: Heterozygote, Homozygote, Simple Heterozygote only, Simple Homozygote only
chronic, compensated, congenital, few
Anemia, Anisocytosis, Basophilic stipplings, Cyanosis
Laboratory findings
Select one: Heterozygote, Homozygote, Simple Heterozygote only, Simple Homozygote only
AgammaT, AgammaX, Biosynthesis: alpha/beta ratio, Ggamma
Range from 2.5 to
L/L, fL, fraction, g/dL