

Single Nucleotide Polymorphism



- Concept
- Classification
- Search method
- Assay method
- Clinical application

Polymorphism

Single nucleotide polymorphism

Tandem repeat polymorphism

Insert/deletion polymorphism

Single Nucleotide Polymorphisms

DNA sequence variations
(A, T, C, or G)

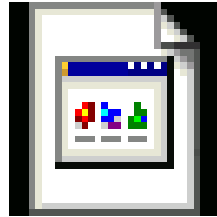
AAGGCTAA -> ATGGCTAA

Sickle cell anemia:

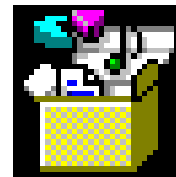
						GLU				
HBB	...	CAC	CTG	ACT	CCT	GTG	GAG	AAG	TCT	...
HBS	...	CAC	CTG	ACT	CCT	GAG	GAG	AAG	TCT	...
						VAL				

SNPs vs. Mutations

- ✿ by allele frequency
- ✿ A single base change, occurring at a frequency of $>1\%$
 - a single nucleotide polymorphism (SNP)
- ✿ When a single base change occurs at $<1\%$
 - a mutation.



DNA.rep.movie.swf



패키지

Recombination

Tandem Repeat Polymorphisms



A very common class of polymorphism
variable length of sequence motifs

- **Microsatellites or Short Tandem Repeat (STR)**

repeat unit: 1-6

(dinucleotide repeat: CACACACACAC )

- **Minisatellites**

repeat unit: 14-100

Insertion/Deletion Polymorphisms

Sequence repetitiveness (direct or inverted)
Predispose DNA to rearrangements

Example

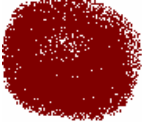
Association between coronary heart disease & a 287 bp

Indel Polymorphism

Located in intron 16 of the angiotensin converting enzyme
(ACE) (OMIM [106180](#))

Responsible for 50% of the inter individual variability of
plasma ACE concentration.

Estimated Numbers of SNP



0.3-1.0 kb average intervals
entire human genome - 3×10^7 bp
number: up to 5-10 million



short insertion/deletions
in between SNPs and VNTR

Polymorphisms & Disease Markers

- Very few of SNP
Direct impact on deleterious phenotype
- non-disease-causing polymorphisms
- Markers to identify & map other genes

Terminologies

Allele: Alternative form of a genetic locus
a single allele from each parent

Polymorphism: Difference in DNA sequence
among individuals.

Linkage Disequilibrium (LD): 2 alleles inherited together
more often than would be predicted

Haplotype : set of alleles on one particular chromosome
Each person has two haplotypes

SNP - Binary

ACGGGTACCGT

ACGGGGACCGT



haplotype.mov

From SNP to Haplotype

Phenotype

Black eye

Brown eye

Black eye

Green eye

Brown eye

Brown eye

GAT**A**TTTCGTAC**G**GA-T

GAT**G**TTTCGTACT**T**GA**A**T

GAT**A**TTTCGTAC**G**GA-T

GAT**A**TTTCGTAC**G**GA**A**T

GAT**G**TTTCGTACT**T**GA**A**T

GAT**G**TTTCGTACT**T**GA**A**T



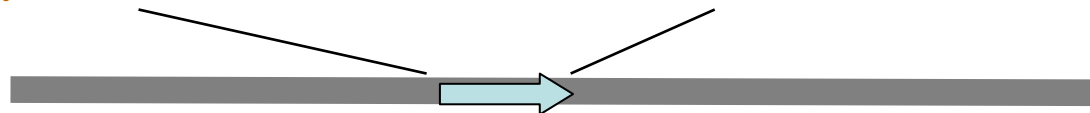
Haplotype

AG- 2/6 (Black)

GT**A** 3/6 (Brown)

AG**A** 1/6 (Green)

DNA



human genome is organized in blocks
~50 kb in size

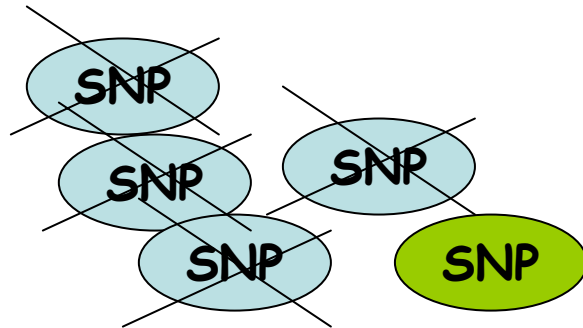
Haplotypes with high LD

regions with a high level of concomitant inheritance

Separated by regions of low LD

regions with a low level of concomitant inheritance

Life Cycle of SNPs & Mutations



Appearance of
New variants
By mutation



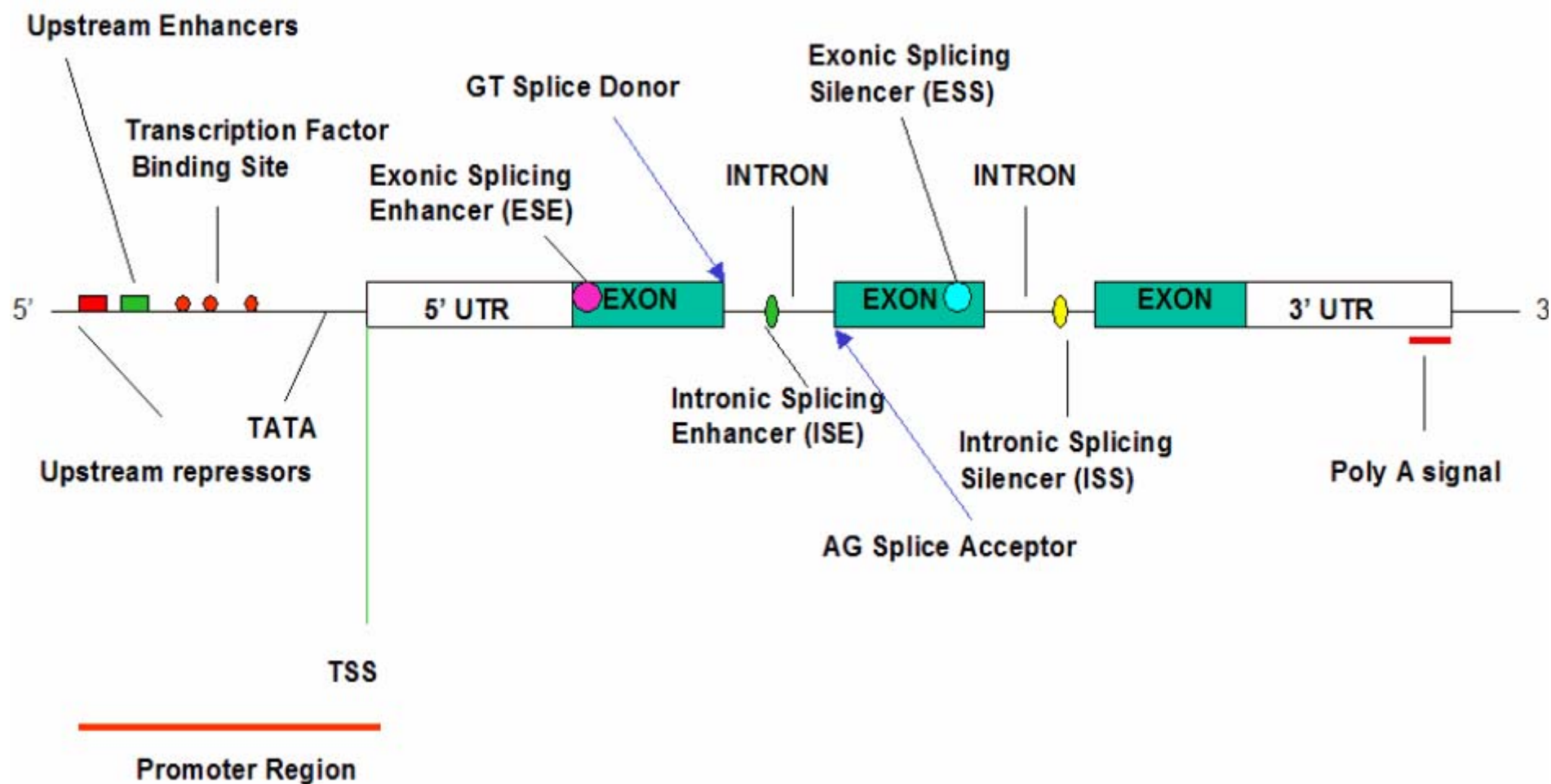
Survival of alleles



Increase of the allele
To a substantial
frequency

Species
Differentiation

Fixation of the allele
In a population



Classification of SNP

SNPs : intronic, exonic, promoter region

Coding SNPs

- Synonymous: no change in a.a
silent mutation
- Non-synonymous: a change in a.a.
conservative vs.
nonconservevative

Genetic Variations Databases



SNP Databases



dbSNP

<http://www.ncbi.nlm.nih.gov/SNP/index.html>



Human Genome Variation Database (HGVbase)

<http://hgvbase.cgb.ki.se/>



TSC: The SNP Consortium

<http://snp.cshl.org/>

dbSNP

<http://www.ncbi.nlm.nih.gov/SNP/index.html>

- ❁ Single-base nucleotide substitutions (**SNPs**)
- ❁ Small-scale multi-base deletions or insertions (deletion insertion polymorphisms- **DIPs**)
- ❁ Microsatellite repeat variations (short tandem repeats - **STRs**)

dbSNP Data Type

1. Submitted data from original observation

submitted SNPs (SS) with
ss# (ss 5586300)

2. Computed data

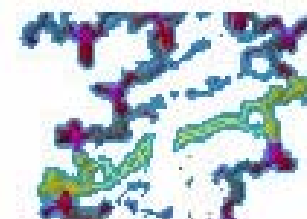
reference SNP Clusters (Ref SNP) with
rs# (rs 4986582)

Submitted SNP Details Page

- collection methods (assay technique)
- submitter information
 - contact data, individual submitter
- variation data
 - frequencies, genotypes



Single Nucleotide Polymorphism



PubMed Nucleotide Protein Genome Structure PopSet Taxonomy OMIM Books SNP

Search Entrez for

BUILD 126

GENERAL

[Contact Us](#)
[dbSNP Homepage](#)
[SNP Science Primer](#)
[Announcements](#)
[dbSNP Summary](#)
[FTP Download](#)
[Build History](#)

SNP SUBMISSION

[How to Submit](#)
[Handle Request](#)

DOCUMENTATION

[Build Release Note](#)
[Search FAQ Archive](#)

updated 2006

dbSNP Search Options

Entrez SNP

ID Numbers

Submission Info

Batch

Locus Info

Between
Markers

ANNOUNCEMENT

- 04/06/2006: [SNP FAQ Archive updated with 2006 contents](#)
- 03/06/2006: [dbSNP Mouse Build 126 is now available](#)
- 10/26/2005: [Accessioned Haplotype Content Now Available in dbSNP](#)
- 10/20/2005: [Schema Changes](#)
- 10/31/2005: [1 or 0 Based Mapping Position](#)

Search by IDs

Note: rs# and ss# must be prefixed with "rs" or "ss", respectively (i.e. rs25, ss25)

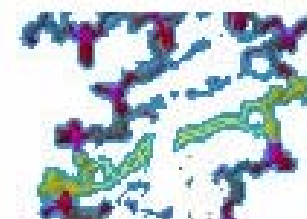
Reference cluster ID(rs#)

Search

Reset



Single Nucleotide Polymorphism



PubMed Nucleotide Protein Genome Structure PopSet Taxonomy OMIM Books SNP

Search Entrez for

BUILD 126

GENERAL

[Contact Us](#)
[dbSNP Homepage](#)
[SNP Science Primer](#)
[Announcements](#)
[dbSNP Summary](#)
[FTP Download](#)
[Build History](#)

SNP SUBMISSION

[How to Submit](#)
[Handle Request](#)

DOCUMENTATION

[Build Release Note](#)
[Search FAQ Archive](#)

updated 2006

dbSNP Search Options

Entrez SNP

ID Numbers

Submission Info

Batch

Locus Info

Between
Markers

ANNOUNCEMENT

- 04/06/2006: [SNP FAQ Archive updated with 2006 contents](#)
- 03/06/2006: [dbSNP Mouse Build 126 is now available](#)
- 10/26/2005: [Accessioned Haplotype Content Now Available in dbSNP](#)
- 10/20/2005: [Schema Changes](#)
- 10/31/2005: [1 or 0 Based Mapping Position](#)

Search by IDs

Note: rs# and ss# must be prefixed with "rs" or "ss", respectively (i.e. rs25, ss25)

dbSNP Search Result





PubMedNucleotideProteinGenomeStructurePopsetTaxonomySNP

Search for

☒ Limits [Preview/Index](#) [History](#) [Clipboard](#) [Details](#)

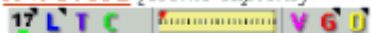
Limits: **coding nonsynon, Validated by cluster, Validated by frequency, Validated by submitter**

Display Show: Sort Send to

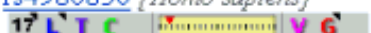
Items 1-17 of 17 One page.

☐ 1: [rs4986854](#) [*Homo sapiens*]

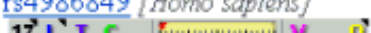

[Links](#)
[SNP500C](#)

☐ 2: [rs4986852](#) [*Homo sapiens*]


[Links](#)
[SNP500C](#)

☐ 3: [rs4986850](#) [*Homo sapiens*]


[Links](#)
[SNP500C](#)

☐ 4: [rs4986849](#) [*Homo sapiens*]


[Links](#)
[SNP500C](#)

☐ 5: [rs4986848](#) [*Homo sapiens*]


[Links](#)
[SNP500C](#)

☐ 6: [rs4986847](#) [*Homo sapiens*]


[Links](#)
[SNP500C](#)

☐ 7: [rs4986845](#) [*Homo sapiens*]


[Links](#)
[SNP500C](#)

☐ 8: [rs2227945](#) [*Homo sapiens*]


[Links](#)
[SNP500C](#)

NCBI

dbSNP BUILD 121

Entrez SNP

Search SNP

Search Mouse SNP

Entrez Batch Query

Entrez SNP Help

Search Fields

Programming Utilities

Batch Report

Graphic Summary

6 Detail Genotype

SNPLocusLink

10 MapViewer

OMIM

Examples

dbSNP Home Page

Overview

Entrez Help

General help

Limits

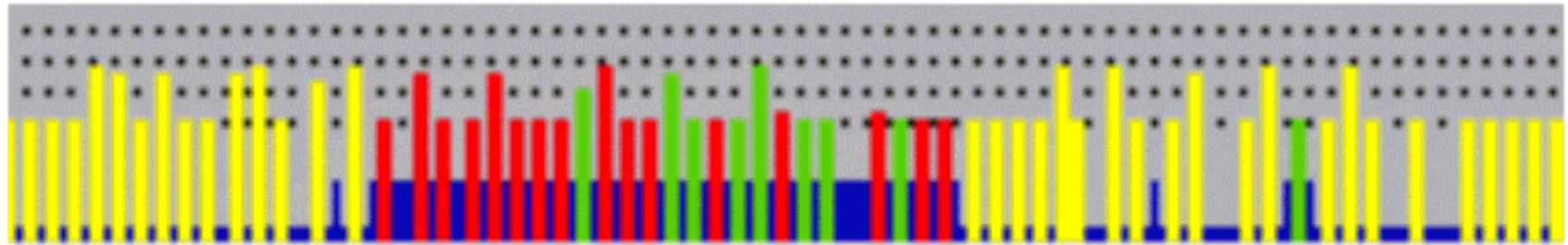
Preview/Index

History

Clipboard

Details

Other Services



Snp In Gene Model Legend:	
█	- Region exon
█	- Region intron
 	- snp: coding
 	- snp: Synonymous change
 	- snp: Nonsynonymous change
 	- snp: Untranslated region
 	- snp: Intron
 	- snp: Splice site
 	- snp: coding: synonymy unknown

Gene-Oriented SNP Visualization



Display of all known ref SNPs overlaid on gene structure

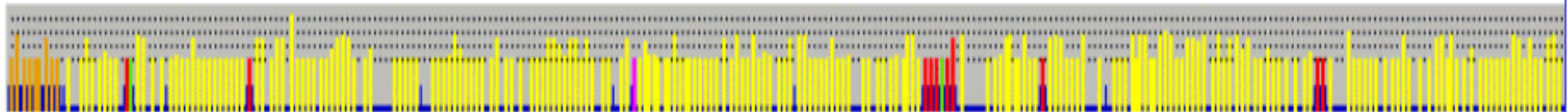
GeneView via analysis of contig annotation: [BRCA1](#) breast cancer 1, early onset

Click to see [\[all\]](#) [\[cSNP\]](#) [\[has frequency\]](#) [\[double hit\]](#) [\[haplotype tagged\]](#) variations associated with this gene.

view rs

☒ in gene region ☐ cSNP ☐ has frequency ☐ double hit ☐ haplotype tagged

	Contig	mrna	protein	mrna orientation	snp graph
gene model (contig mRNA transcript):	NT_010755	NM_007294	NP_009225	reverse	transcript on minus strand



GeneView for All SNPs

[RefSNP Summary Info](#)

Summary Information for refSNPs

The NCBI genome map viewer can be configured to supply additional information about elements in the master map by selecting **verbose mode** in the display settings dialog box.

When the "SNP" map is selected as the master map, the user is presented with a graphical summary of several properties of the SNP in a display that looks like this:

	Map	Gene	Het	Validation	Genotypes Available	Linkout Available
			0 100%	>80 >90 >95%		
rs999991		LTC				
rs999992		LTC				
rs999993		LTC				
rs999994		LTC				
rs999995		LTC				

In silico map results

The quality of the SNP is computed from mapping and validation data. The current set of map icons and their meanings are:

Map	Meaning
	Marker mapped to Unique position in the genome. Markers mapped to two contigs on the same chromosome will also receive this icon, on the assumption that the contigs might not have been joined during sequence assembly.
	Marker mapped to Multiple position in the genome. Markers mapped to 2 to 10 locations will receive this warning icon. Markers mapped to more than 10 locations will not be annotated on human genome sequence, nor will they be included in the genome map views.
	Markers that have been identified as a Withdrawn marker artifact will be flagged with this WITHDRAWN icon. Evidence for withdrawing the marker can be found in the dbSNP refSNP report for the marker.
	This marker is not Not mapped to any genomic location on.

Mouseover Text

MAP COUNTS: reports the number of hits to each of the following categories:

- Total number of chromosomes [CH]
- Total number of contigs [NT]
- Total distinct locations [TOT]

Map Viewer icon for SNPs

Gene region / functional classification

Markers placed on the sequence map are examined for neighboring gene features (gene, mRNA, CDS, and exon). Markers that satisfy the conditions below are considered a member of the set of markers associated with the gene, and each is assigned a functional classification depending on its particular location in the gene region:

Gene

Meaning

L

Locus

S: Any part of the marker position on sequence map is a 2kb interval 5' of the most 5' feature of gene (CDS, mRNA, gene), OR the marker position is within a 500 base interval 3' of the most 3' feature of the gene. Both strands of sequence are examined for gene features, so a marker can potentially be a variation on multiple genes at a single location.

T

Transcript

T: Any part of marker position overlaps with mRNA (overlaps with UTR/intron and mRNA feature is missing), BUT marker position is not within the coding region of the transcript.

C

Coding

C: Any part of the marker position overlaps with a sequence (CDS) region (or overlaps with exon region in the unlikely case an exon is annotated but CDS is missing).

L

Locus not within known gene locus

L: Marker position is not within 2kb of any annotated gene (as defined above) for any annotated gene.

T

Transcript not within known transcript region

T: Marker position is not within 500 bases of any annotated transcript for any annotated transcript.

C

Locus not within known coding region

C: Marker position is not within 500 bases of any annotated coding region for any annotated coding region.

Mouseover Text

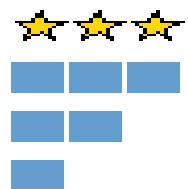
Mouseover on any letter will show the set of gene symbols for each respective functional category

Map Viewer Icon for SNPs

Quality Information

This scale distinguishes validated markers (★) from unvalidated ones (■). Viewed as a vertically stacked histogram, the width of each scale reports the likelihood that a marker is real as defined by the submitter. This likelihood is based on validation experiments conducted by the submitter for each batch of reported markers, where the success rate for a batch is defined as 1 minus the false positive rate. Scale probabilities are:

Validation Meaning



★ ★ ★ Marker has been validated by independent resequencing

■ ■ ■ Unvalidated marker - success rate >95%

■ ■ Unvalidated marker - success rate >90%

■ Unvalidated marker - success rate >80%

The success rate for the marker is reported. If several submissions have been clustered into a single refSNP, then the reported figure is the maximum success rate for the cluster.

Other Information

Success rate = $1 - \text{false positive rate}$

Meaning



Individual genotypes have been submitted to dbSNP for this marker.



Links have been established between the dbSNP marker and a submitter website or database. These resources can provide additional information about the quality of the marker or its phenotypic effects for the organism or gene.

Map Viewer Icon for SNPs

Marker heterozygosity

This 0 – 100% scale indicates the average heterozygosity as the range $\text{avg_het} \pm 2 [\text{SE}(\text{avg_het})]$, where $\text{SE}(\text{avg_het})$ is the standard error of the estimator. Thus the graph shows an approximate 95% confidence interval for the marker.

Het	Meaning
	No allele frequencies measures of observed heterozygosity were submitted for this marker.
	Average heterozygosity 0.26-0.30 % confidence interval (0.26 – 0.30). The estimate has a small standard error.
	Average heterozygosity 0.00-0.40 % confidence interval (0.00 – 0.40) The estimate has a large standard error.

Mouseover Text

Mouseover on the graph will report average heterozygosity and standard error in the form Het (S.E.). E.g. Het (S.E.)=0.32 (0.008)

Map Viewer Icon for SNPs

Validation Methods



validated by multiple, independent submissions to the refSNP cluster



validated by frequency or genotype data: minor alleles observed in at least two chromosomes.



validated by submitter confirmation



all alleles have been observed in at least two chromosomes apiece



validated by HapMap project

Map Viewer Display Option

Map Viewer - Microsoft Internet Explorer

Organism: Homo sapiens [Help](#)

Chromosome: 4 Region Shown: 9532963.50 9538038.50

Available Maps:

Org: human Assembly: ref

Rn_EST
Ssc_UniGene
Ssc_EST
Bt_UniGene
Bt_EST
Variation
-Cytogenic maps-
Ideogram
FISH Clone

ADD>>
<<REMOVE

Maps Displayed (left to right):

[] Gene_Cytogenetic
[] Hs_UniGene
[R] Gene
[] Variation

Change Assembly
Move UP
Move DOWN
Make Master/Mc
Toggle Ruler
([R] before map means 'rule')

More Options:

☐ Show Connections ☒ Verbose Mode

Compress Map: auto Auto Compress if > 350 px

Page Length: 30

Thumbnail View: ☒ default (ideogram) ☐ master

Apply Close

SNP: Genome View


NCBI

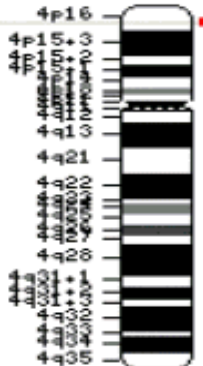

NCBI Map Viewer

[PubMed](#)
[Entrez](#)
[BLAST](#)
[OMIM](#)

[Search](#)

[Map Viewer Home](#)
[Map Viewer Help](#)
[Human Maps Help](#)
[FTP](#)
[Data As Table View](#)
[Maps & Options](#)
☒ **Compress Map**
 Region Shown:

☐ out
☒ zoom
☐ in



***Homo sapiens* build 34.3**
 Chromosome: [1](#) [2](#) [3](#) [[4](#)] [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [11](#) [12](#) [13](#) [14](#) [15](#) [16](#) [17](#) [18](#) [19](#) [20](#) [21](#) [22](#) [X](#) [Y](#)
 Query: DRD5 [\[clear\]](#)
 Master Map: Variation [View Summary](#)
 Region Displayed: 9,532K-9,538K bp [Download/View Sequence/Evidence](#)

Gen...	HsUniG	Genes...	Var...	Map	Gene	Het	Validation	Genotypes / Linkou
						0 100%	>80 >90 >95%	
		9533000			rs2006139	▼	LTC	inbredinbred
		9533100			rs12504432	▼	LTC	inbredinbred
		9533200			rs2076907	▼	LTC	inbredinbred ★
		9533300			rs2227839	▼	LTC	inbredinbred ★
		9533400			rs2227840	▼	LTC	inbredinbred ★
		9533500			rs2227848	▼	LTC	inbredinbred ★
		9533600			rs6282	▼	LTC	inbredinbred ★
		9533700			rs2227849	▼	LTC	inbredinbred ★
		9533800			rs2227841	▼	LTC	inbredinbred ★
		9533900			rs2227845	▼	LTC	inbredinbred ★
		9534000			rs2227846	▼	LTC	inbredinbred ★
		9534100			rs2227843	▼	LTC	inbredinbred ★
		9534200			rs2227852	▼	LTC	inbredinbred ★
		9534300			rs2227847	▼	LTC	inbredinbred ★
		9534400						
		9534500						
		9534600						
		9534700						
		9534800						
		9534900						
		9535000						
		9535100						
		9535200						
		9535300						

Online Mendelian Inheritance in Man (OMIM)

OMIM Allelic Variants

- OMIM : genetic disorders by all levels of mutation/variation
 - nucleotide substitutions
 - to large-scale chromosomal abnormalities.
- "Allelic Variants"
 - nucleotide substitutions;
 - small insertions and deletions (indels)
- 10-digit OMIM number (e.g., 141900.0003)
- 6 digit OMIM number of the parent locus
- + decimal point + 4-digit variant number.

Viewing the List of Allelic Variants

OMIM allelic variant number (4 digit)
name of associated disorder or condition
gene symbol
type/location of variant

Display Allelic Variants Show: 20 Send to Text

+305900
GLUCOSE-6-PHOSPHATE DEHYDROGENASE; G6PD

ALLELIC VARIANTS
(selected examples)

- [0001 G6PD A+](#) [G6PD, ASN126ASP]
- [0002 G6PD A-](#) [G6PD, VAL68MET, ASN126ASP]
- [0003 G6PD CHATHAM](#) [G6PD, ALA335THR]
- [0004 G6PD ILESHA](#) [G6PD, GLU156LYS]
- [0005 G6PD MAHIDOL](#) [G6PD, GLY163SER]
- [0006 G6PD MEDITERRANEAN](#) [G6PD, SER188PHE]
- [0007 G6PD METAPONTO](#) [G6PD, ASP58ASN]
- [0008 G6PD PORTICI](#) [G6PD, ARG393HIS]
- [0009 G6PD SANTIAGO DE CUBA](#) [G6PD, GLY447ARG]
- [0010 G6PD SEATTLE-LIKE](#) [G6PD, ASP282HIS]
- [0011 G6PD HARILAOU](#) [G6PD, PHE216LEU]

Viewing the List of Allelic Variants

OMIM allelic variant number (4 digit)
name of associated disorder or condition
gene symbol
type/location of variant

Display Allelic Variants Show: 20 Send to Text

+305900
GLUCOSE-6-PHOSPHATE DEHYDROGENASE; G6PD

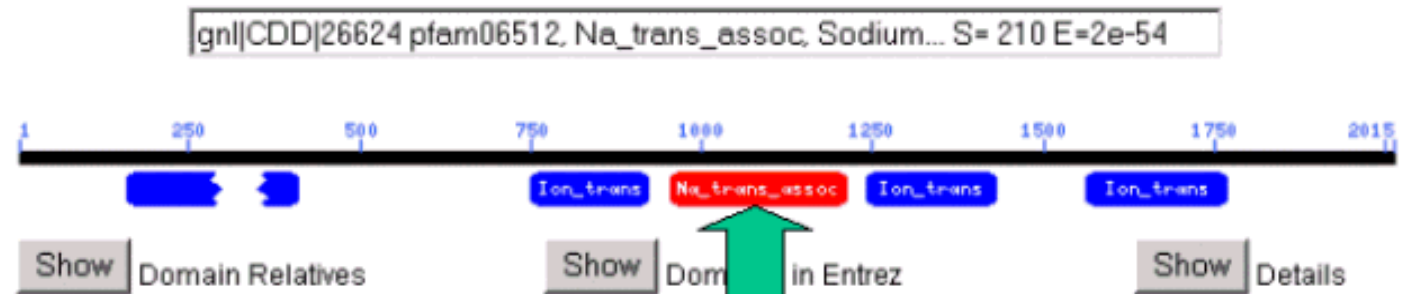
ALLELIC VARIANTS
(selected examples)



- [0001 G6PD A+](#) [G6PD, ASN126ASP]
- [0002 G6PD A-](#) [G6PD, VAL68MET, ASN126ASP]
- [0003 G6PD CHATHAM](#) [G6PD, ALA335THR]
- [0004 G6PD ILESHA](#) [G6PD, GLU156LYS]
- [0005 G6PD MAHIDOL](#) [G6PD, GLY163SER]
- [0006 G6PD MEDITERRANEAN](#) [G6PD, SER188PHE]
- [0007 G6PD METAPONTO](#) [G6PD, ASP58ASN]
- [0008 G6PD PORTICI](#) [G6PD, ARG393HIS]
- [0009 G6PD SANTIAGO DE CUBA](#) [G6PD, GLY447ARG]
- [0010 G6PD SEATTLE-LIKE](#) [G6PD, ASP282HIS]
- [0011 G6PD HARILAOU](#) [G6PD, PHE216LEU]

Functional Analysis of Polymorphisms

Sequence Alignment



		190	200	210	220	230	240		
		*....*....*....*....*....*....							
consensus	104	QPSE--E--SQFEDLPTE-TKR--LTSPDDES	VETEVR	SE	E-DLNL	TIQDSRKK	----	149	
query	1071	EESS--K--QESQPVSGGpEAP--PDSRTWSQVS	ATAS	SE	A-EASASQAD	WRQQ	----	1117	
gi 4033376	1071	EESS--K--QESQPVSG-WPRgpPDSRTWSQVS	ATAS	SE	A-EASASQAD	WRQQ	----	1118	
gi 13774490	1062	QEED--E--ENSLGTEEE-SSK--QTPEDS	-----	SE	G-ST	-----	-----	1090	
gi 116452	1074	EESS--K--QESQVSGG-HEP--YQEPRAWSQV	SETT	SS	E-AGASTS	QADWQQ	----	1119	
gi 1209467	1023	WQEE--DpkGQQEQLPQV-QKC--ENHQAARSP	ASHSS	SE	DlAPYLGES	WKRKDs	spqvp	1076	
gi 116449	1054	VPIAvGE--SDFENLNTE	-----	EF	SE	S-ELE	----	ESKEK	1083
gi 27263190	1087	VPIAvGE--SDFENLNTE	-----	DF	SE	S-DLE	----	ESKEK	1116
gi 3087876	918	PIAS--E--ESDLEMPTE-E	-----	ETDA	SE	E-DIKKPL	QPL	-----	950
gi 14165234	309	-----	-----	DDESVDL	ST	E-ENMIK	LKGA--	L	328

S1102 Human	RTWSQVSATASSEAEASASQA
Mouse	RAWSQVSETTSSEAEASTSQA
Rat	RAWSQVSETTSSEAGASTSQA

Amino Acid Comparison Text View

[NCBI Amino Acid Explorer](#)

Serine

Amino Acid Explorer

[Course Main Page](#)

[Description of Displayed Data](#)

Compare

S - Ser

to

Y - Tyr

using

text

Serine vs Tyrosine

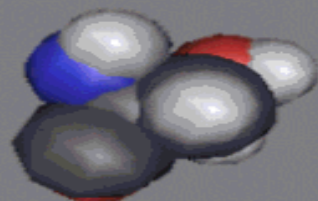
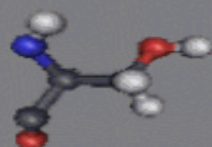
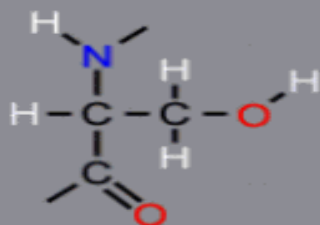
Feature	Serine	Tyrosine
Synopsis	Serine is the simplest hydroxyl amino acid, being equivalent to alanine with the hydroxyl replacing one of the methyl protons. The hydroxyl group not only gives serine hydrogen bonding potential, but enables it to serve as a substrate to both kinases and glycosyltransferases. Serine is also the third most abundant amino acid in proteins.	Tyrosine is both an aromatic and a hydroxyl amino acid, being equivalent to phenylalanine with the para position of the phenyl group substituted with a hydroxyl. This phenolic hydroxyl endows tyrosine with unique chemical properties, and allows it to serve as both a catalyst and a substrate to the vast array of kinases specific for this residue.
Side chain flexibility	Low	Moderate
Interaction modes	H-bonds, van der Waals	H-bonds, aromatic stacking, van der Waals
Potential side chain H-bonds	3	3
Residue molecular weight	87	163
Isoelectric point	5.7	5.7
Hydrophobicity	0.601	0.714
Standard codon(s)	UCN,AGY	UAY
Properties	Polar Hydroxyl Nonessential	Polar Hydroxyl Aromatic Nonessential

Amino Acid Comparison Graphic View

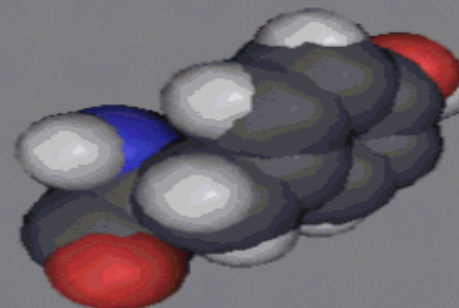
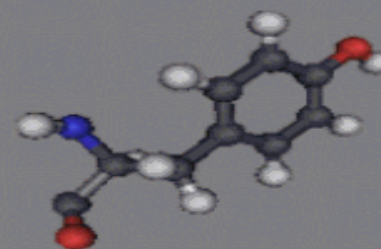
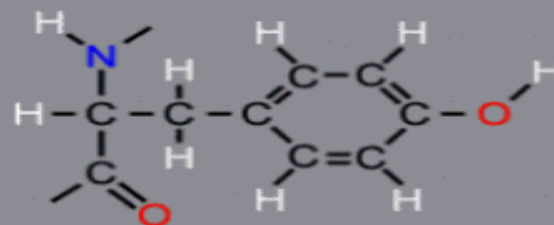
[NCBI Amino Acid Explorer](#)

Serine vs Tyrosine

[Click to display with Cn3D](#)



[Click to display with Cn3D](#)



Mutation Database

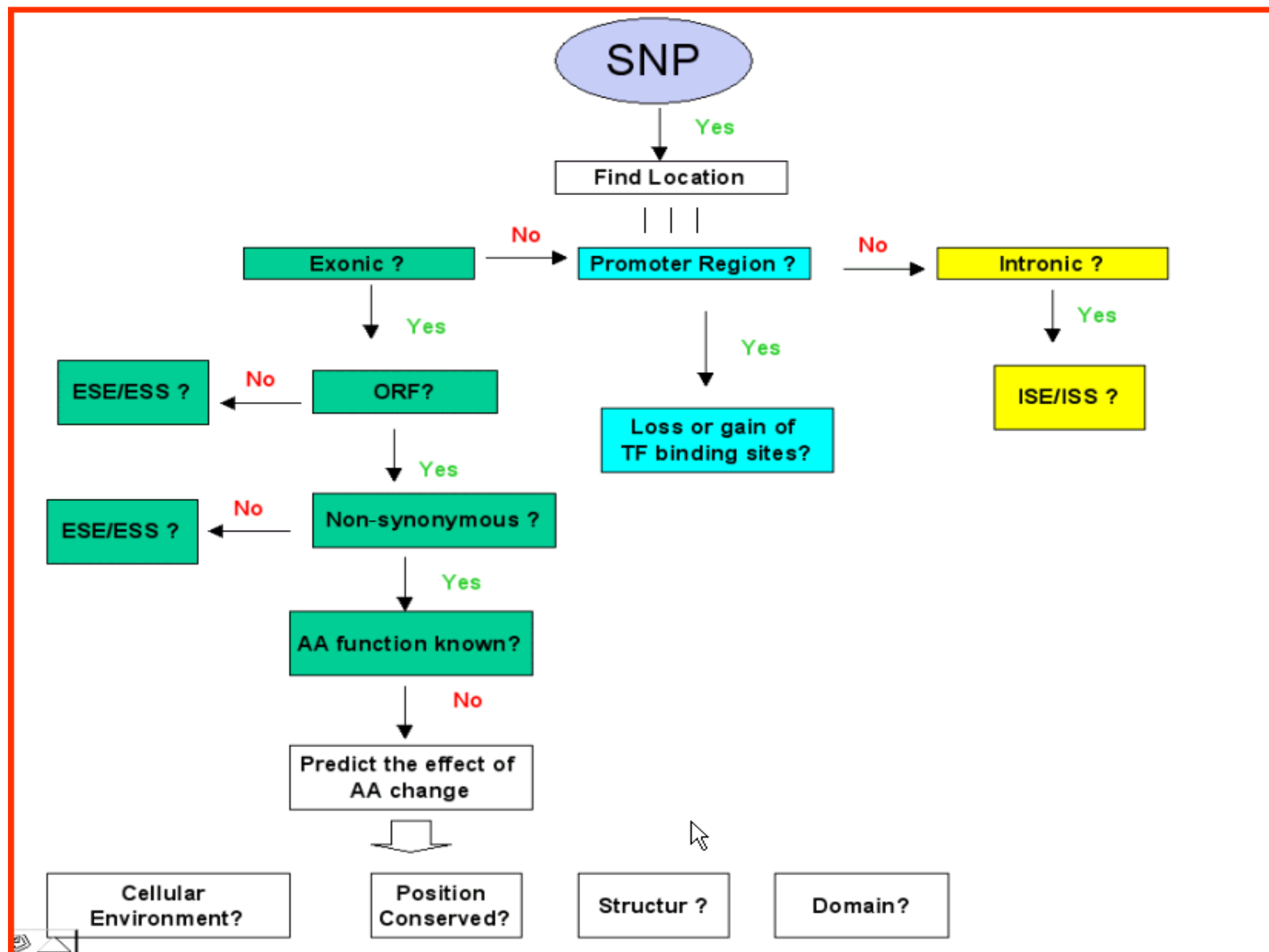


- Nucleic Acids Research Database Issue

http://nar.oupjournals.org/content/vol32/suppl_1/

- HUGO Mutation Database Initiative

<http://www2.ebi.ac.uk/mutations/cotton/dblist/dblist.html>



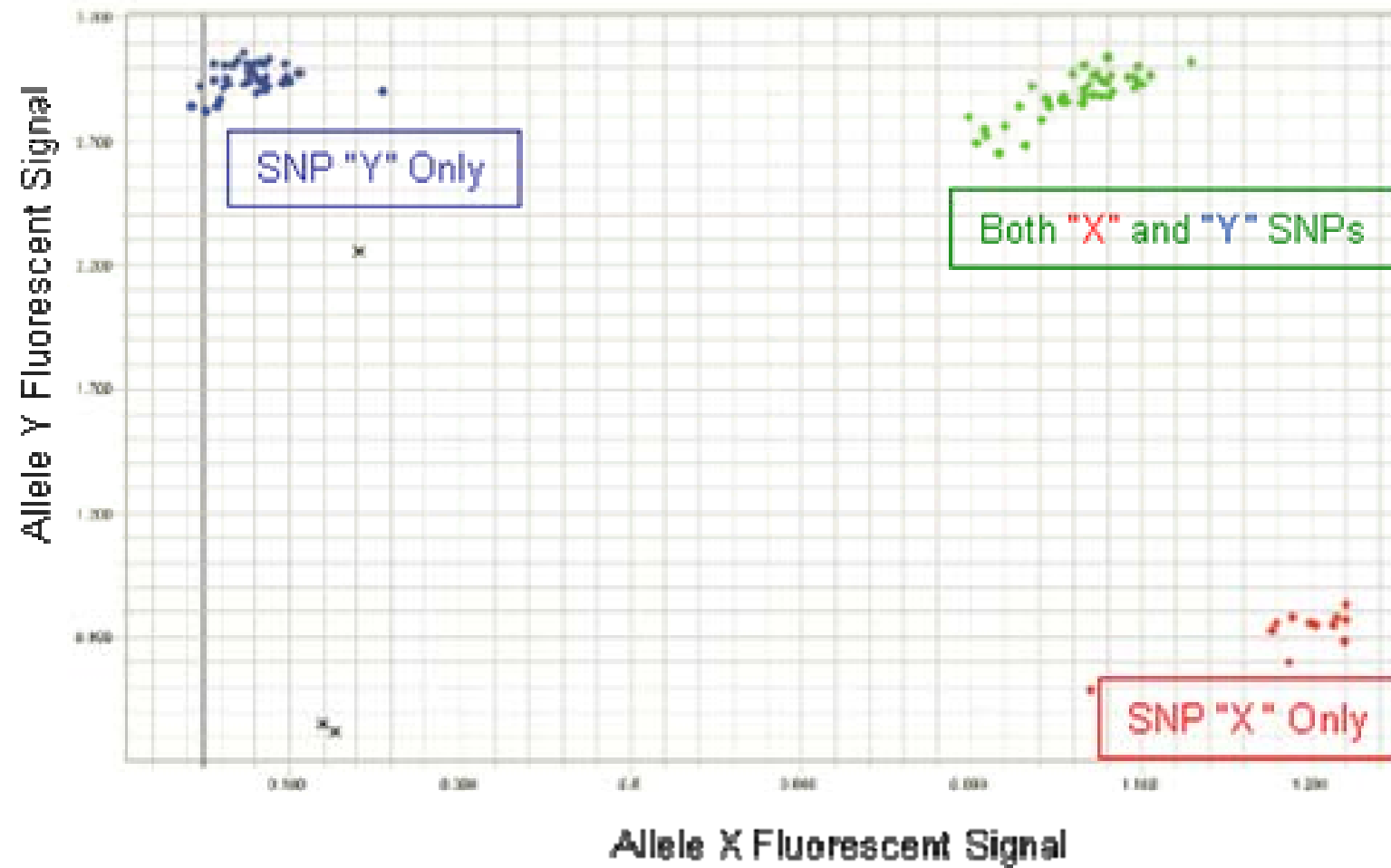
SNP class:

☐ het	variation has unknown sequence composition, but is observed to be heterozygous
☐ in del	insertion deletion polymorphism, deletions represented by '-' in allele string
☐ microsat	microsatellite / simple sequence repeat
☐ mixed	
☐ mnp	multiple nucleotide polymorphism (all alleles same length where length>1)
☐ named	allele sequences defined by name tag instead of raw sequence, e.g. (Alu)/-
☐ no variation	submission reports invariant region in surveyed sequence
☐ snp	true single nucleotide polymorphism

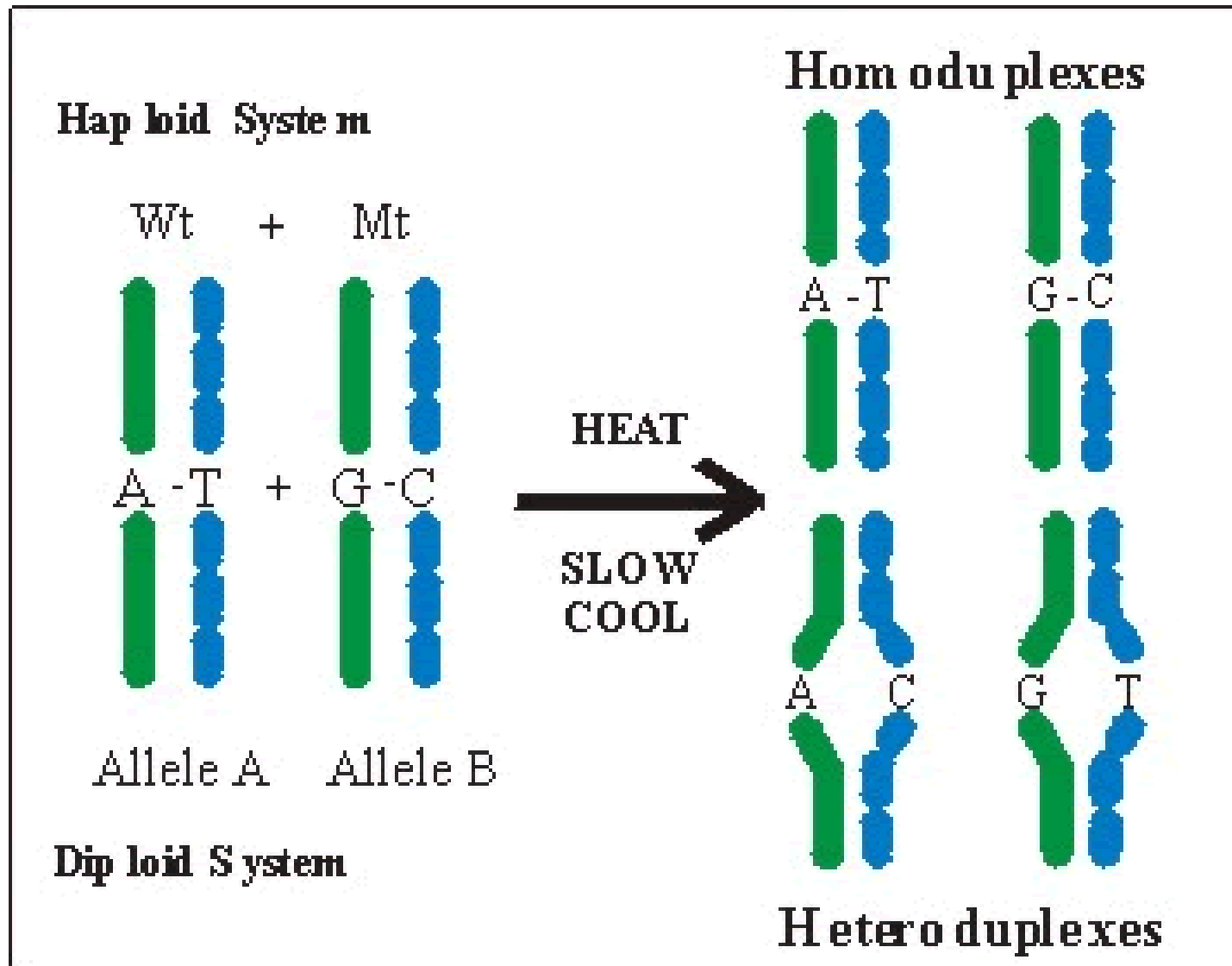
Method class

☐ computed	variation was mined from sequence alignment with software
☐ dhplc	Denaturing High Pressure Liquid Chromatography used to detect SNP
☐ hybridize	hybridization method (e.g. chip) was used to assay for variation
☐ other	other method used to detect variation
☐ rflp	variation in enzyme restriction site used to detect variation
☐ sequence	samples were sequenced and resulting alignment used to define variation
☐ sscp	single stranded conformational polymorphism used to detect variation
☐ unknown	

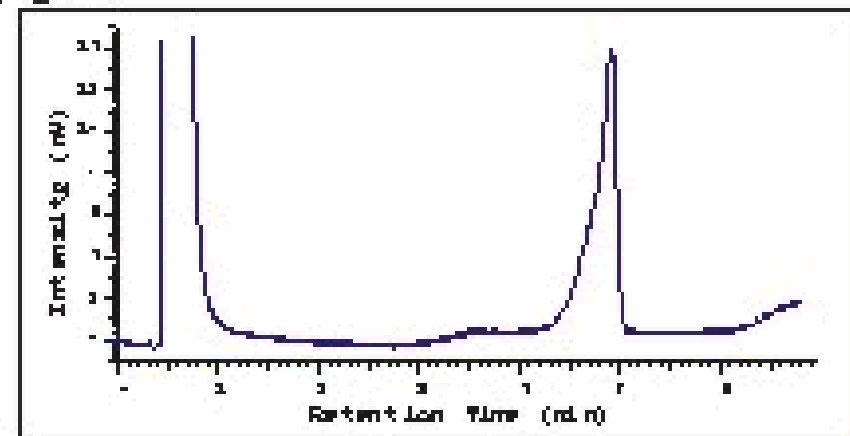
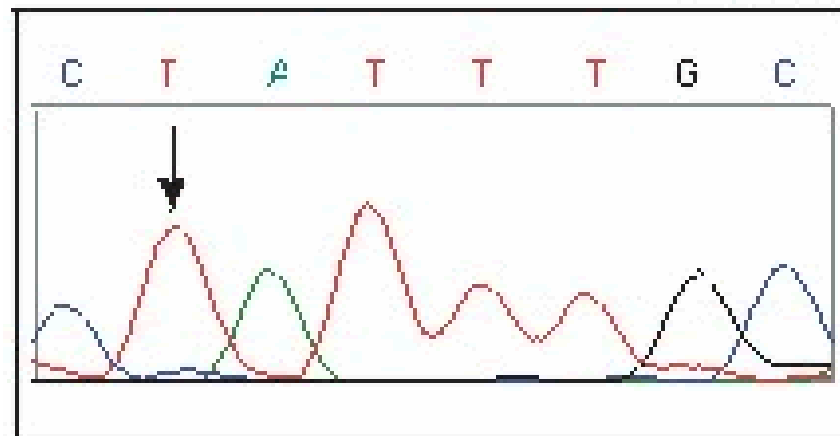
TaqMan Assay



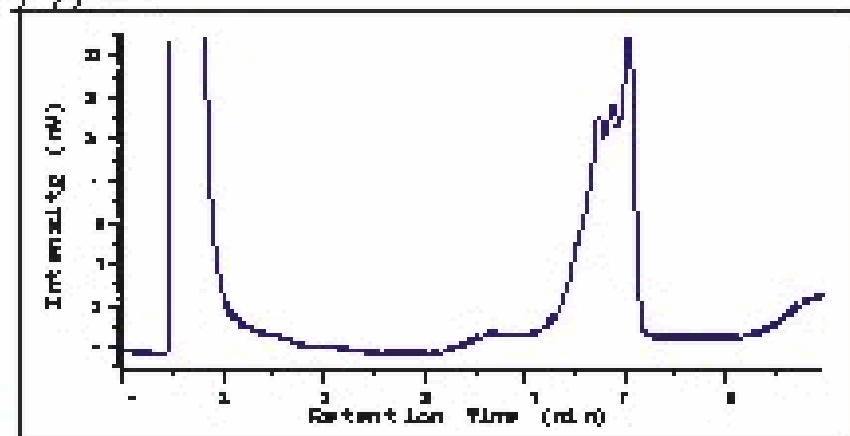
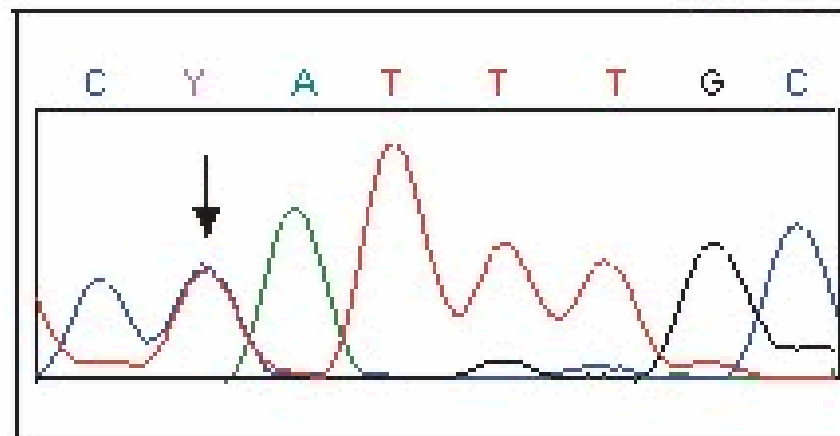
Denaturing High Performance Liquid Chromatography (DHPLC)



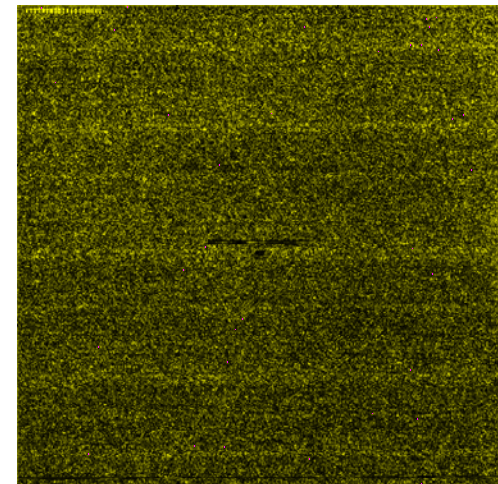
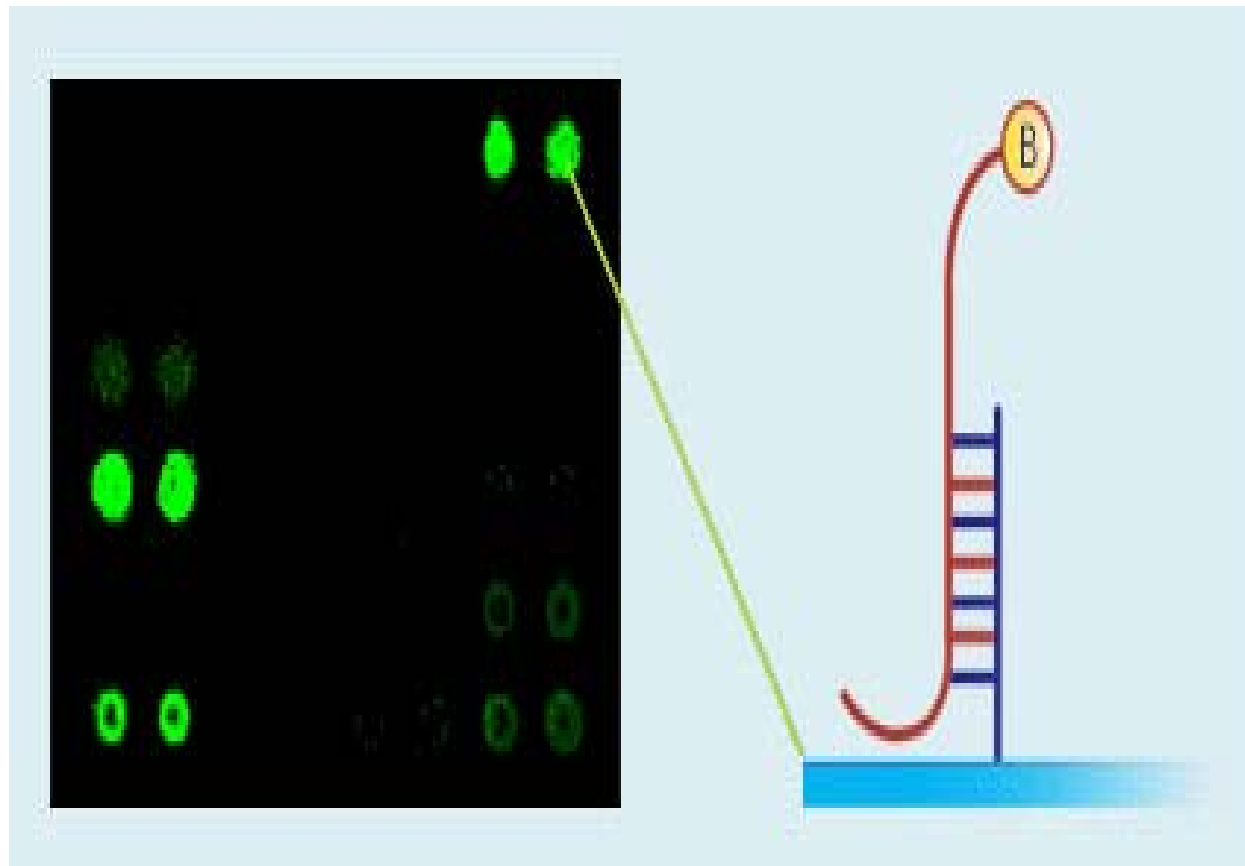
Homozygote



Heterozygote

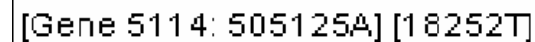


SNP Microarray Chip



Affymetrix Probe Layout

- Two alleles, A & B
- PerfectMatch
(Signal)
MisMatch
(Background)



Model method: PM/MM difference model



Pharmacogenomics



Inherited variations in genes
@ drug response



Predict
whether a good or bad response
to a drug



pharmacogenomic.swf

Chemosensitivity & Genomics

phase I enzymes

activation or inactivation of the drug

Quinone oxidoreductase

(NQO1, NADH)

carcinogenic quinones -> less toxic hydroxyquinones

inactivate the antileukemic agent **doxorubicin**

Lower activity **GSTP1 1578A > G** (*GSTP1*B*)
a.a. change Val105Ile

- associated with **higher etoposide clearance** in African Americans treated with steroids

GeneCard for protein-coding **RUNX1**
GC21M035081

runt-related transcription factor 1 (acute myeloid leukemia 1; aml1 oncogene)

Symbol approved by the [HUGO Gene Nomenclature Committee \(HGNC\) database](#)
 (Previous symbols: **AML1**, **CBFA2**)

Services

Jump to Section...



Aliases and Descriptions

(According to ¹HGNC, ²Entrez Gene, ³UniProt/Swiss-Prot, ⁴UniProt/TrEMBL, ⁵GDB, ⁶OMIM, ⁸Nature:405:311-319 and ⁷CroW21, and/or ⁷GeneLoc)

[About This Section](#)

Aliases

AML1 ^{2,3,5,6}
AML1-EVI-1 ²
 AMLCR1 ^{1,2,5}
 CBFA2 ^{2,3,5,6}
 EVI-1 ²
 PEBP2A2 ^{1,2,5}
 PEBP2aB ²

Descriptions

Acute myeloid leukemia 1 protein ³
 CBF-alpha 2 ³
 Core-binding factor, alpha 2 subunit ³
 Oncogene AML-1 ³
 PEA2-alpha B ³
 PEBP2-alpha B ³
 Polyomavirus enhancer binding protein 2 alpha B subunit ³
 Runt-related transcription factor 1 ³
 SL3-3 enhancer factor 1 alpha B subunit ³
 SL3/AKV core-binding factor alpha B subunit ³
 acute myeloid leukemia 1 protein (oncogene AML-1), core-binding factor, alpha subunit ⁸
 core binding factor, runt domain, alpha subunit 2 (acute myeloid leukemia 1; **aml1** oncogene) ⁵

Genomic Location

(According to [GeneLoc](#) and/or [HGNC](#), and/or [Entrez Gene \(NCBI build 35\)](#), and/or [miRBase](#), Genomic Views According to [UCSC](#) and [Ensembl](#), Whole Chromosome Sequence According to [Nature](#) (cited Here with Permission):[405,311-319](#) and [CroVW21](#))
[About This Section](#)

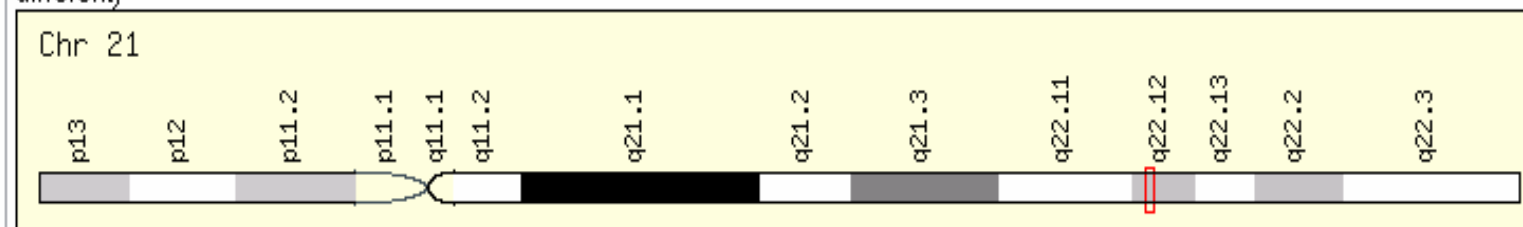
Jump to Section...

Chromosome: **21**

Entrez Gene cytogenetic band: [21q22.3](#) Ensembl cytogenetic band: [21q22.12](#)

Nature(405: 311-319) cytogenetic band: **21q22.12**

Gene in genomic location: bands according to Ensembl, locations according to [GeneLoc](#) (and/or [Entrez Gene](#) and/or [Ensembl](#) different)



[GeneLoc gene densities for chromosome 21](#)

[\(about GC identifiers\)](#)

GC21M035081:

Start:

End:

Size:

Orientation:

[GeneLoc](#)

[Nature:405,311-319](#)

35,081,971 bp from pter 21,770,223 bp from centromere

35,343,511 bp from pter 21,837,636 bp from centromere

261,540 bases

67,414 bases

minus strand

minus strand

RefSeq genomic assemblies:

[NC_000021.7](#) [NT_011512.10](#) [NT_086913.1](#)

Whole chromosome sequencing:

cDNA sequence:

[D43967](#)

genomic clones:

[PPQ140K16](#) to [P499A22](#)

Genomic View.

[UCSC Golden Path with GeneCards custom track](#)

UniProt/Swiss-Prot: [RUNX1_HUMAN_Q01196](#) (See protein sequence)

- **Size:** 453 amino acids; 48797 Da
- **Subunit:** Heterodimer with CBFB. RUNX1 binds DNA as a monomer and through the Runt domain. DNA-binding is increased by heterodimerization. Isoform AML-1L can neither bind DNA nor heterodimerize. Interacts with TLE1 and THOC4. Interacts with ELF1, ELF2 and SPI1. Interacts via its Runt domain with the ELF4 N-terminal region. Interaction with ELF2 isoform 2 (NERF-1a) may act to repress RUNX1-mediated transactivation. Interacts with MYST3 and MYST4.
- **Subcellular location:** Nucleus.
- **Caution:** The fusion of **AML1** with EAP in T-MDS induces a change of reading frame in the latter resulting in 17 AA unrelated to those of EAP.
- **3D structures:** PDB ids [1CMO \(3D\)](#) [1CO1 \(3D\)](#) [1E50 \(3D\)](#) [1H9D \(3D\)](#) [1LJM \(3D\)](#)

Proteins

According to ¹[UniProt](#),
and/or [Ensembl](#),
Phosphorylation sites
according to ²[Phosphosite](#),
Seq according to [NCBI](#),
3 rendering according to
[3](#), Ontologies according
to [Gene Ontology](#)
[nsortium](#) 2006-02-01.,
bodies by [Cell Signaling](#)
[hnology](#) and/or [Abcam](#))
[About This Section](#)

Up to Section... ▾

Post-translational modifications:



- Phosphorylated in its C-terminus upon IL-6 treatment. Phosphorylation enhances interaction with MYST3.¹
- View phosphorylation sites using [PhosphoSite](#)².

REFSEQ proteins (2 alternative transcripts):

[NP_001001890.1](#) [NP_001745.2](#)

ENSEMBL proteins:

[ENSP00000300305](#) [ENSP00000351123](#) [ENSP00000340690](#) [ENSP00000349079](#) [ENSP00000319459](#) [ENSP00000342892](#)

1 Gene Ontology (GO) cellular component term (links to tree view):

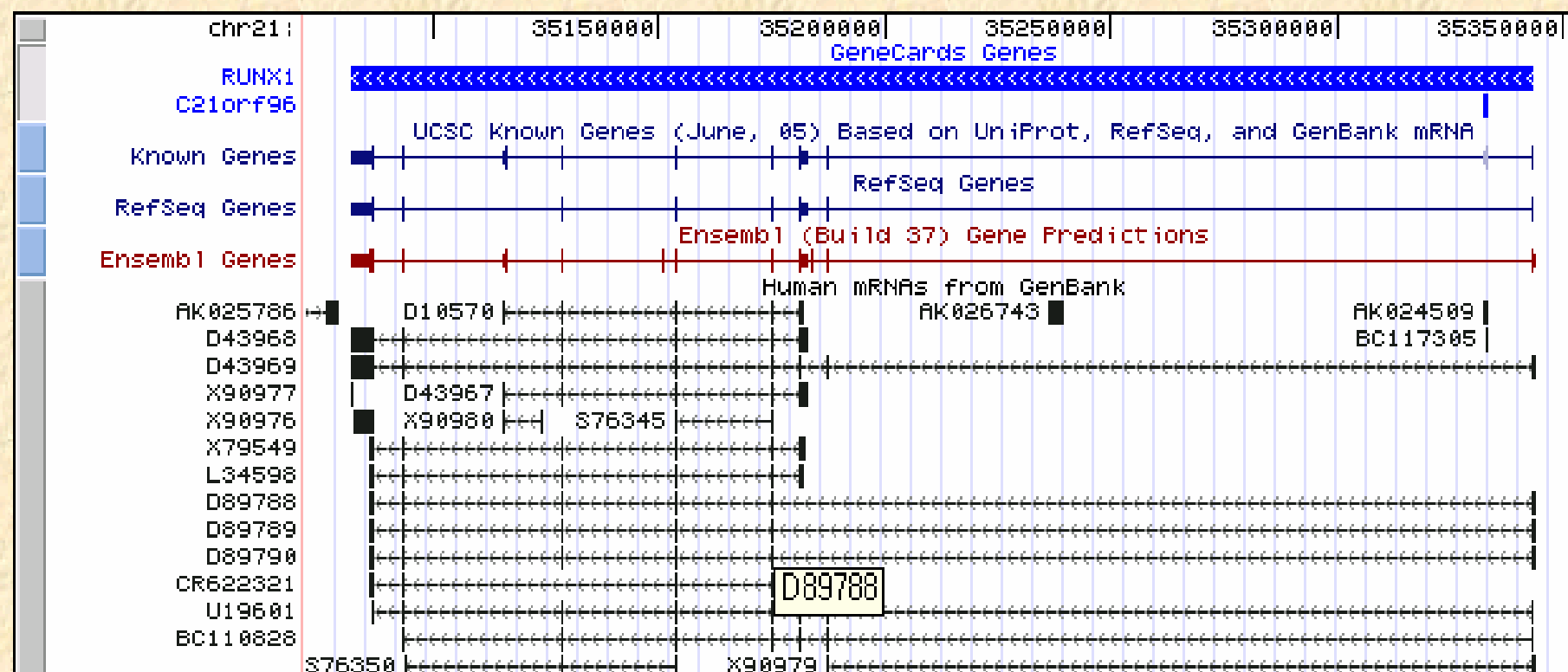
[GO:0005634](#) nucleus

UCSC Genome Browser on Human May 2004 Assembly

move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x

position/search chr21:35,071,971-35,353,511 jump clear size 281,541 bp. configure

chr21 (q22.12) p13 p12 p11.2 21q21.1 21.2 q21.3 q22.11 22.2 21q22.3



profiles by [Cell Signaling Technology](#) and/or [Abcam](#)
[About This Section](#)

Jump to Section...

[NP_001001890.1](#) [NP_001745.2](#)

ENSEMBL proteins:

[ENSP00000300305](#) [ENSP00000351123](#) [ENSP00000340690](#) [ENSP00000349079](#) [ENSP00000319459](#) [ENSP00000342892](#)

1 Gene Ontology (GO) cellular component term (links to tree view):

[GO:0005634](#) nucleus

Antibodies for RUNX1:



[Search Cell Signaling Technology for Antibodies and Assays.](#)



Antibodies from Abcam ([Runx1/AML1](#)), each with their [Abpromise](#)SM.

2 InterPro domains/families:

[IPR012346](#) P53_RUNT_DNA_bd

[IPR000040](#) AML1_Runt

[Graphical View of Domain Structure for UniProt Entry Q01196](#)

ProtoNet protein and cluster: [Q01196](#)

1 Blocks protein family: [IPB000040](#) Acute myeloid leukemia 1 protein signature

UniProt/Swiss-Prot: [RUNX1_HUMAN, Q01196](#)

- **Domain:** A proline/serine/threonine rich region at the C-terminus is necessary for transcriptional activation of target genes.
- **Similarity:** Contains 1 Runt domain.

Protein Domains/ Families

According to [InterPro](#),
[ProtoNet](#), [UniProt](#), and/or
[BLOCKS](#)

[About This Section](#)

Jump to Section...

Product Pathways - Lymphocyte Signaling

Product Pathways

Chromatin Regulation

MAPK Signaling

Apoptosis/Caspase

Akt Signaling

Translational Control

Ca, cAMP & Lipid Signaling

Cell Cycle/Checkpoint

Jak/Stat Pathway

NF-kappaB Signaling

TGF-beta/SMAD Signaling

Lymphocyte Signaling

Neuroscience

Tyrosine Kinase/Adaptors

Cytoskeletal Signaling

Glucose Metabolism

Wnt/Hedgehog Signaling

Nuclear Receptor

Protein Folding & Stability

Other Signaling Pathways

AML1 Antibody

◀ back to
category



more info



application
references



FAQs



download
datasheet

Catalog #	Size	Price (USA)
# 4334	100 microliters (10 Western mini-blots)	\$150.00

add to
cart

in
stock

Species Cross-Reactivity	MW (kDa)	Applications	Source
H Mk	50	W IF-IC F	Rabbit
H=Human Mk=Monkey			
W=Western Blot IF-IC=Immunofluorescence (Immunocytochemistry) F=Flow Cytometry			

Specificity / Sensitivity

Drug Discovery Tools

Screening Technologies

Active Kinases & Kits

PathScan ELISA

Antibody Technologies

Rabbit Monoclonals

Motif Antibodies

PhosphoScan Proteomics

Flow Cytometry

Immunohistochemistry

Immunofluorescence

siRNA

Protein Classes

Kinases

Phosphatases

Transcription Factors

Growth Factors & Cytokines

Signaling

Sign-in

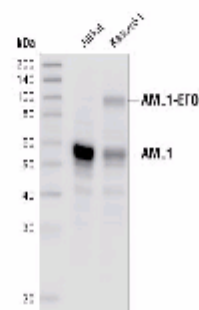
New User

Lost Password

protein.

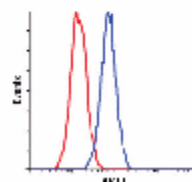
Click on any image below to Zoom

Western Blotting



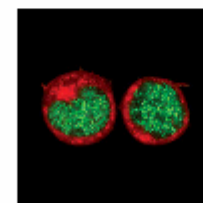
Western blot analysis of extracts from Jurkat and Kasumi-1 cells, using AML1 Antibody.

Flow Cytometry



Flow cytometric analysis of Jurkat cells, using AML1 Antibody (blue) compared to a nonspecific negative control antibody (red).

IF-IC



Confocal immunofluorescent image of Jurkat cells labeled with AML1 Antibody (green). Mitochondria have been labeled with MitoTracker Red CMXRos.

Source / Purification

(F) 편집(E) 보기(V) 즐겨찾기(A) 도구(I) 도움말(H)

뒤로 ◀ ▶ ⌂ ★ 즐겨찾기 📁 👤

(D) http://www.dsi.univ-paris5.fr/genatlas/

이동 연결 »

GENATLAS

Université René Descartes - Paris

20861 genes / 3883 phenotypes / 61367 citations

release Fri 2 Jun 2006

contact help

GENATLAS : GENE Database

omim	sequences	swissprot	citations	locuslink	source
HGNC	genelynx	genecards	Ensembl	Unigene	linkage

FLASH GENE

Symbol	MPO	<i>last update : 11/03/2006</i>
HGNC name	myeloperoxidase	
HGNC id	7218	
Corresponding disease	MPO	
Location	17q22-q23.1	
EC.number	1.11.1.7	

DNA	RNA	EXP/sub-loc	PROTEIN	PATHOLOGY
---------------------	---------------------	-----------------------------	-------------------------	---------------------------

DNA

TYPE	functioning gene
SPECIAL FEATURE	
text	tail to tail with LPO
STRUCTURE	11,08 kb 12 Exon(s)

present in the contig : [NT_010783](#) of Genbank in reverse/complement

DNA size	11.08 Kb	mRNA size	3215 bp	12 exons	NM_000250	see the exons
9782029		9793108				

인터넷

시작



3 Window...

5 Internet ...

3 Microso...

알씨

A 漢

오후 9
금요일
2006-06

present in the contig : [NT_010783](#) of Genbank in reverse/complement

DNA size 11.08 Kb mRNA size 3215 bp 12 exons [NM_000250](#) [see the exons](#)
9782029 1 2 3 4 5 6 7 8 9 10 11 12 9793108

UniGene Blastn

[see Grailexp prediction](#)

[see Genscan prediction](#)

[see Genomic Sequence](#) ■ Excellent ■ Good ■ Marginal

[10 Kb 5' upstream gene genomic sequence study](#)

MAPPING cloned Y linked status confirmed

mode in situ hybridation, recombinant DNA, somatic cell hybrid

Map cen - [MPO](#) [MPO](#) - [LPO](#) - [EPX](#) - qter

Authors Sakamaki (00)

Physical map

MPO Physical Map

■ Genatlas link

■ Locuslink link

5'→3'

EPX
LOC
HCHDVC

Atlas of Genetics and Cytogenetics in Oncology and Haematology



[Aim](#) [Editorial Board](#) [Authors](#) [Grants](#) [Citation](#) [Copyright](#)

How to contribute: [Instructions to Authors](#)

ENTITIES: by THEME

[GENES](#)
[LEUKAEMIAS](#)
[SOLID TUMOURS](#)
[CANCER PRONE DISEASES](#)

ENTITIES: by CHROMOSOME

[X](#) [Y](#) [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [11](#) [12](#) [13](#) [14](#) [15](#) [16](#) [17](#) [18](#) [19](#) [20](#) [21](#) [22](#) [Not Assigned](#)

REVIEWS and CASE REPORTS

[DEEP INSIGHTS](#)
[CASE REPORTS](#) (see also a list of [Rare Translocations](#) and [Authors Instructions](#))
11p15/NUP98 Project: [Authors Instructions](#)

PORTAL

[DATABASES AND LINKS](#)

(F) 편집(E) 보기(V) 즐겨찾기(A) 도구(T) 도움말(H)

뒤로 ◀ ▶ 앞으로 즐겨찾기

(D) http://canceratlas.spis.co.uk/login.aspx

이동 연결 »

Atlas
Cancer



Home My Slides Help

SEARCH

Head and Neck Cancers

Gynecologic Cancers

Lung Cancer

Upper Gastrointestinal Cancers

Lower Gastrointestinal Tract
Cancers

Leukemia

Lymphoma

Sarcomas

Breast Cancer

Genitourinary Cancers

Skin Cancer

Neuro-oncology

Welcome to Atlas of Cancer

Username

Password

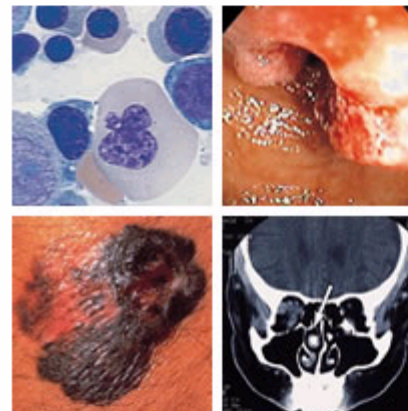
LOGIN

Forgotten your login details?

[Click here](#)

New User?

[Register here](#)



Privacy Statement Feedback

© [Current Medicine Group Ltd.](#) 2006

시작



3 Window...

5 Internet ...

3 Microso...

알씨

A 漢

오후 9
금요일
2006-06