How hereditary information is stored, passes on and changes is a fundamental problem in biology.

Three types of maps have been essential:

- Linkage maps of genes
- Banding pattern of chromosome
- DNA sequences

These maps lie on three very different types of data:

- Observed pattern of heredity; gene or lingake group transmission and recombination frequency.
- Identification of physical bands
- Physical sequence of nucleotides

High resolution maps

Are directly based on DNA sequence.

In particular any feature of DNA that vary among individual can serve as marker

Sequence Tagged Site

STS

Sequence Tagged Site

- Short tracts of DNA sequence (200 to 500 bp)

- a single occurrence in the human genome.

STSs are defined by PCR primer pairs, their location and base sequence are known and therefore they can be used as landmarks in genome mapping.

Sequence Tagged Site

Are useful:

1. For locating and orienting the mapping and/or sequence data reported from different laboratories;

2. As framework for the physical map of the human genome.

Primary database: dbSTS

http://www.ncbi.nlm.nih.gov/dbSTS/

Secondary database: UniSTS

http://www.ncbi.nlm.nih.gov/sites/entrez?db=unists

dbSTS
PRIMARY DATABASE

- is an NCBI resource that contains sequence data for short genomic landmark sequences or Sequence Tagged Sites
- STS sequences are incorporated into the STS Division of GenBank.
- STS are directly submitted to GenBank

GenBank Divisions

1. PRI - primate sequences
2. ROD - rodent sequences
3. MAM - other mammalian sequences
4. VRT - other vertebrate sequences
5. INV - invertebrate sequences
6. PLN - plant, fungal, and algal sequences
7. BCT - bacterial sequences
8. VRL - viral sequences
9. PHG - bacteriophage sequences
10. SYN - synthetic sequences
11. UNA - unannotated sequences

- Organized by taxonomy (sort of)
- Direct submissions (Sequin/Bankit)
- Accurate (~1 error per 10,000 bp)
- Well characterized

12. EST - EST sequences (expressed sequence tags)
13. PAT - patent sequences
14. STS - STS sequences (sequence tagged sites)
15. GSS - GSS sequences (genome survey sequences)
16. HTG - HTGS sequences (high throughp
17. HTC - HTC sequences (high throughpu
18. ENV - Environmental sampling sequenc
19. CON - Constructed sequences

- Organized by sequence type
- Batch submissions (ftp/email)
- Less accurate
- Poorly characterized

GenBank Divisions

- Organized by sequence type

12. EST - EST sequences (expressed sequence tags)
13. PAT - patent sequences

14. STS - STS sequences (sequence tagged sites)

15. GSS - GSS sequences (genome survey sequences)
16. HTG - HTGS sequences (high throughput genomic sequences)
17. HTC - HTC sequences (high throughput cDNA sequences)
18. ENV - Environmental sampling sequences
19. CON - Constructed sequences

UniSTS
DERIVATIVE DATABASE

Is a comprehensive non-redundant collection of sequence tagged sites (STSs) derived from STS-based maps and other experiments.

UniSTS

- Records are univocally defined by PCR primer pairs and are associated with additional information (genomic position, genes, and sequences).
- integrate marker and mapping data from different organisms and a variety of public resources.
- If two or more markers have different names but the same primer pair, a single STS record is presented and all the marker names are shown.

- A primer is a short, single-stranded DNA sequence used in the polymerase chain reaction (PCR) technique. In the PCR method, a pair of primers is used to hybridize with the sample DNA and define the region of the DNA that will be amplified. Primers are also referred to as oligonucleotides.
- Polymerase chain reaction (PCR) is a laboratory technique used to amplify DNA sequences. The method involves using short DNA sequences called primers to select the portion of the genome to be amplified. The temperature of the sample is repeatedly raised and lowered to help a DNA replication enzyme copy the target DNA sequence. The technique can produce a billion copies of the target sequence in just a few hours.

Search for
INPP1
(Inositolpolyphosphate 1 phosphatase

In *Entrez*

http://www.ncbi.nlm.nih.gov/gquery/?term=inpp1

NCBI

UniSTS
Integrating Markers and Maps

Search for inpp1 Preview Go Clear

Limits | Preview/Index | History | Clipboard | Details

Entrez UniSTS
Help
Query tips
Submit
Submit map
FTP site
Statistics

Related Sites
e-PCR
Map Viewer
Gene
UniGene
dbSNP
MGI
RGD
ZFIN

Genomic Biology

Bos taurus
Canis familiaris
Danio rerio
Homo sapiens
Mus musculus
Rattus novegicus
Sus scrofa

Field: **Gene Name**

- Enter terms and click Preview to see only the number of search results.
- To save search indefinitely, click query # and select Save in My NCBI.
- To combine searches use #search, e.g., #2 AND #3 or click query # for more options.

Search	Most Recent Queries
#15	Search **inpp1** Field: **Gene Name**
#14	Search **inpp1**

Add Term(s) to Query or View Index:

- Enter a term in the text box; use the pull-down menu to specify a search field.
- Click Preview to add terms to the query box and see the number of search results, or click Index to view terms within a field.

Preview Index

Click AND OR NOT to add a term to the query box

Field: **Gene Name**

Items 1 - 6 of 6 One page.

1:	UniSTS:37815	**WI-9155**
	Homo sapiens	chromosome 2, locus INPP1
	Macaca mulatta	chromosome 12, locus INPP1
	Found by e-PCR in sequences from Homo sapiens, Macaca mulatta and Papio anubis.	
2:	UniSTS:462280	**INPP1_3186**
	Homo sapiens	chromosome 2, locus INPP1
	Macaca mulatta	chromosome 12, locus INPP1
	Pan troglodytes	locus INPP1
	Pongo abelii	chromosome 2B
	Found by e-PCR in sequences from Homo sapiens, Macaca mulatta, Pan troglodytes, Papio anubis and Pongo abelii.	
3:	UniSTS:183336	**SHGC-145513**
	Callithrix jacchus	chromosome 6, locus INPP1
	Homo sapiens	chromosome 2, locus INPP1
	Pan troglodytes	chromosome 2B, locus INPP1
	Pongo abelii	chromosome 2B
	Found by e-PCR in sequences from Callithrix jacchus, Homo sapiens, Pan troglodytes, Papio anubis and Pongo abelii.	
4:	UniSTS:183335	**SHGC-145512**
	Homo sapiens	chromosome 2, locus INPP1
	Pan troglodytes	chromosome 2B, locus INPP1
	Pongo abelii	chromosome 2B
	Found by e-PCR in sequences from Homo sapiens, Pan troglodytes and Pongo abelii.	
5:	UniSTS:73930	**G16421**
	Callithrix jacchus	chromosome 6, locus INPP1
	Homo sapiens	chromosome 2, locus INPP1
	Macaca mulatta	chromosome 12, locus INPP1
	Pan troglodytes	chromosome 2B, locus INPP1

NCBI

UniSTS
Integrating Markers and Maps

PubMed Entrez BLAST OMIM

Search UniSTS for WI-9155 Go

Entrez UniSTS
Help
Query tips
Submit
Submit map
FTP site
Statistics

Related sites
e-PCR
Map Viewer
Gene
UniGene
dbSNP
GeneMap'99
MGD
ZFIN

Genomic biology
Bos taurus
Canis familiaris
Danio rerio
Homo sapiens
Mus musculus
Rattus novegicus
Sus scrofa

UniSTS:37815

WI-9155

Homo sapiens chromosome 2, locus INPP1
Macaca mulatta chromosome 12, locus INPP1

Found by e-PCR in sequences from *Homo sapiens*, *Macaca mulatta* and *Papio anubis*.

Primer Information

Forward primer: CTGAACTGTGAAACTGTTTCGG
Reverse primer: TGGAAAAATACTCCTCAAAAGAGG
PCR product size: 99-100 (bp), *Homo sapiens*
GenBank Accession: G07151 L08488

Homo sa

Name:
Also known as
Polymorphism

Cross Refer

Gene Ge
Sy
De
Po
UniGene Hs.

Mapping Information

WI-9155 Sequence Map: Chr.2|HuRef Map Viewer

Whitehead Institute for Biomedical Research

Each unique primer pair corresponds to one UniSTS ID that is unique and stable over time.

Entrez UniSTS
Help
Query tips
Submit
Submit map
FTP site
Statistics

Related sites
e-PCR
Map Viewer
Gene
UniGene
dbSNP
GeneMap'99
MGD
ZFIN

Genomic biology
Bos taurus
Canis familiaris
Danio rerio
Homo sapiens
Mus musculus
Rattus novegicus
Sus scrofa

UniSTS:37815 Links

WI-9155

Homo sapiens chromosome 2, locus INPP1
Macaca mulatta chromosome 12, locus INPP1

Found by e-PCR in sequences from *Homo sapiens*, *Macaca mulatta* and *Papio anubis*.

Primer Information

Forward primer:	CTGAACTGTGAAACTGTTTCGG
Reverse primer:	TGGAAAAATACTCCTCAAAAGAGG
PCR product size:	99-100 (bp), *Homo sapiens*
GenBank Accession:	G07151 L08488

Homo sapiens

Name: WI-9155
Also known as: G00-678-142 GDB:678142 INP
Polymorphism info:

Cross References

Gene	GeneID:	3628
	Symbol:	INPP1
	Description:	inositol polyphosphate-1-phosphatase
	Position:	2q32
UniGene	Hs.32309	Inositol polyphosphate-1-phosphatase

Mapping Information

WI-9155 Sequence Map: Chr.2|HuRef Map Viewer

GenBank records of *Homo sapiens*

	Symbol:	INPPT
	Description:	inositol polyphosphate-1-phosphatase
	Position:	2q32
UniGene	Hs.32309	Inositol polyphosphate-1-phosphatase

Mapping Information

WI-9155	Sequence Map:	Chr 2|HuRef-Primary_Assembly	Map Viewer
	Position:	183095975-183096074 (bp)	
WI-9155	Sequence Map:	Chr 2|Hs_Celera-Primary_Assembly	Map Viewer
	Position:	184831162-184831261 (bp)	
WI-9155	Sequence Map:	Chr 2|GRCh37.p2-Primary_Assembly	Map Viewer
	Position:	191236160-191236259 (bp)	
WI-9155	Whitehead-YAC Map:	Chr 2	Map Viewer
	Position:	721 (ordinal)	
	Reference Interval:	WC2.15	

Electronic PCR results

RefSeq mRNA (2)

NM_002194.3	1788 .. 1887 (100 bp)
NM_001128928.1	1932 .. 2031 (100 bp)

mRNA (4)

L08488.1	1558 .. 1657 (100 bp)
BC015496.1	1298 .. 1397 (100 bp)
AK093560.1	1890 .. 1989 (100 bp)
BC106006.1	1332 .. 1431 (100 bp)

Genomic RefSeqs (4)

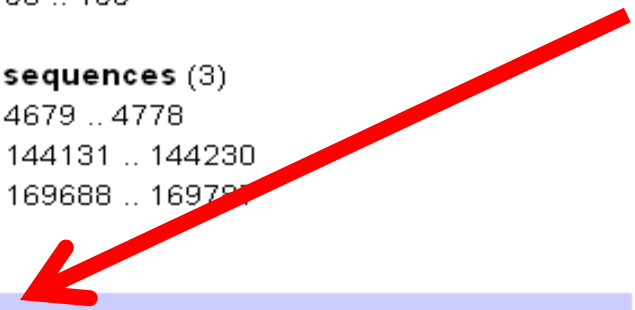
AC013698.3 60164 .. 60263

ESTs (5 of 98)[Show All Hits]

D25971.1	106 .. 205
T97209.1	30 .. 129
R01085.1	42 .. 141
H30231.1	90 .. 189
H52036.1	56 .. 155

Whole Genome Shotgun sequences (3)

AADD01026157.1	4679 .. 4778
AADC01022760.1	144131 .. 144230
AADB02002629.1	169688 .. 169781



Macaca mulatta

Name: WI-9155

Cross References

Gene GeneID: 703301
Symbol: INPP1
Description: inositol polyphosphate-1-phosphatase
Position:

Mapping Information

WI-9155 Sequence Map: Chr 12|NW_001098499.1 Map Viewer
Position: 7369-7469 (bp)

Questions or Comments?
Write to the NCBI Service Desk

Disclaimer Privacy statement

Link to MapViewer

http://www.ncbi.nlm.nih.gov/genome/sts/sts.cgi?uid=37815

Search | Find | Find in This View | Advanced Search

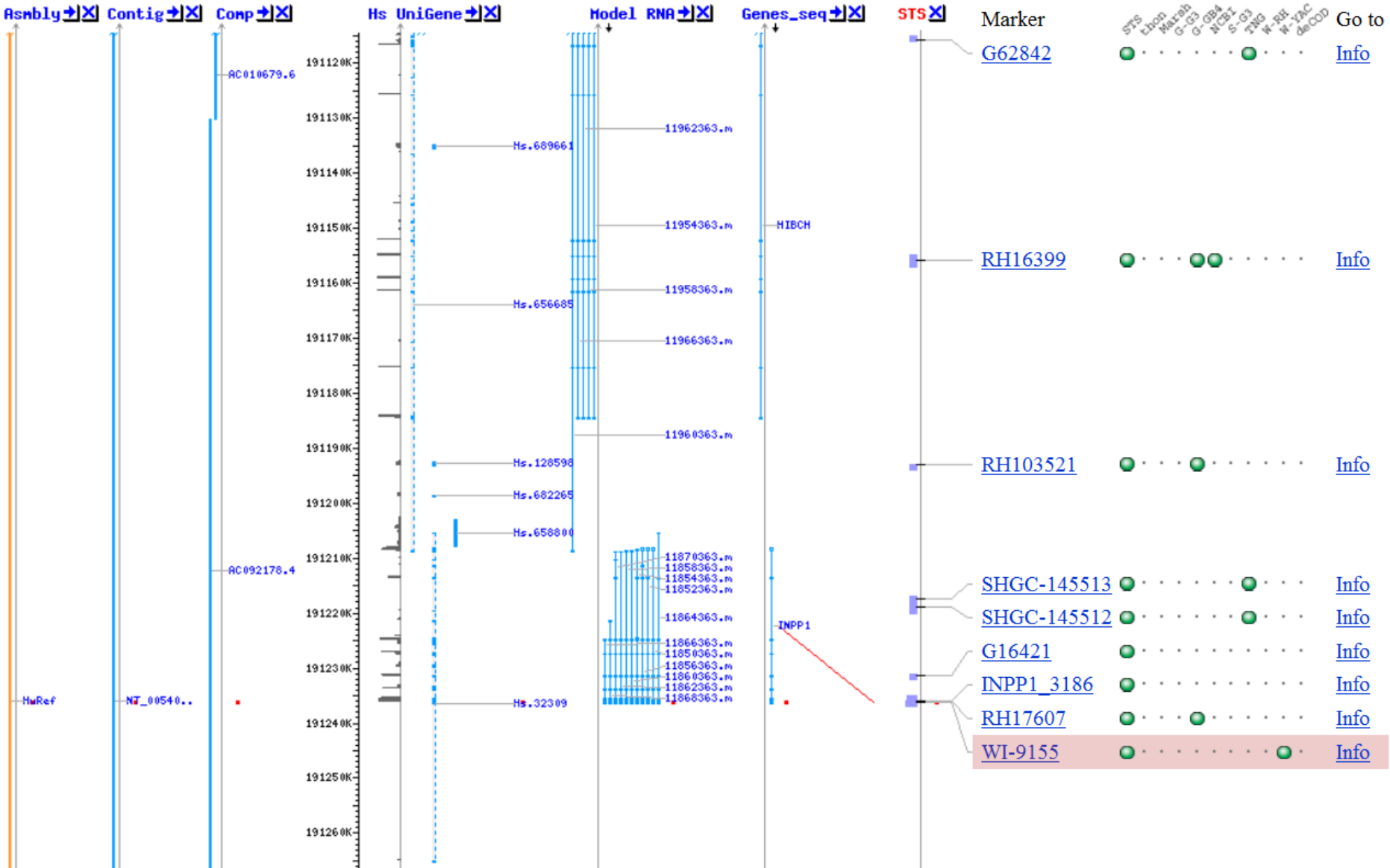
Homo sapiens (human) Build 37.3 (Current) BLAST human seq

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 MT

Query: (sts[obj_type] AND (37815[sts_id])) [clear]

Master Map: STS Summary of Maps Maps &

Region Displayed: **191,115K-191,358K bp** Download/View Sequence/F

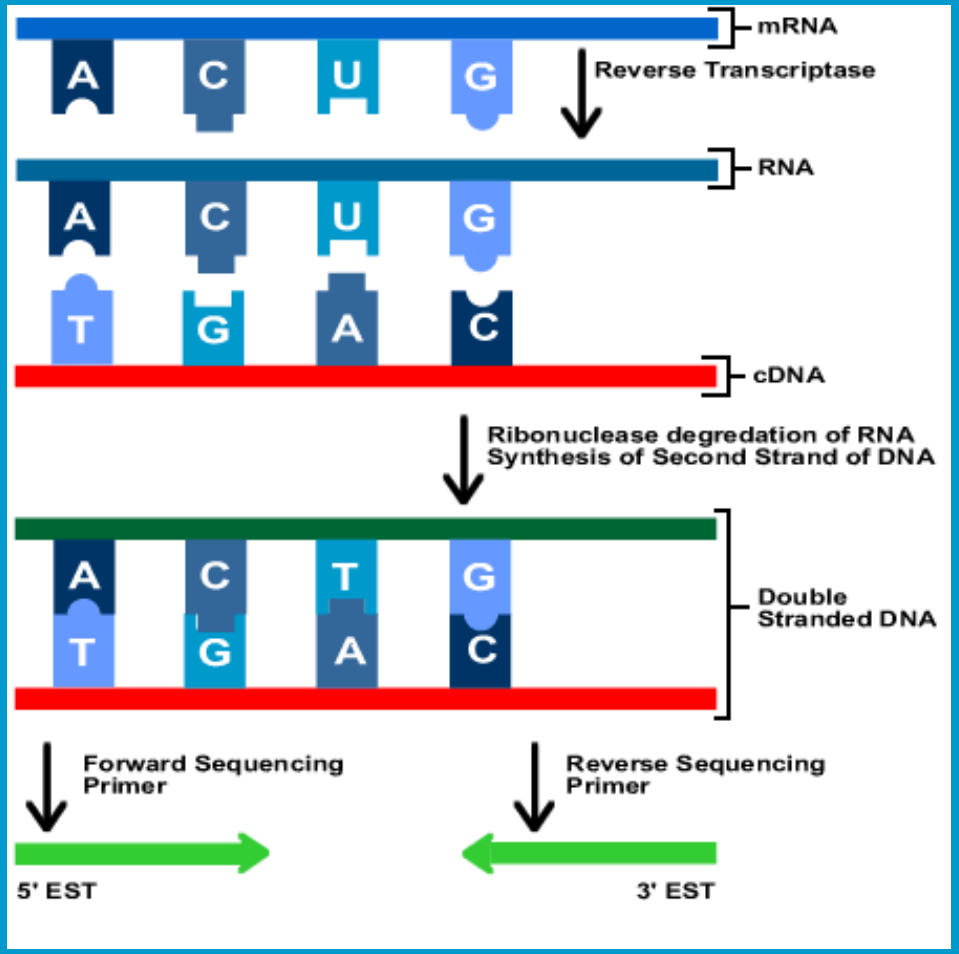


Marker	STS	Chon	Marsh	G-G3	G-GB4	NCBI	S-G3	TNG	W-RH	W-YAC	deCOD	Go to
G62842	●						●					Info
RH16399	●			●	●							Info
RH103521	●			●								Info
SHGC-145513	●						●					Info
SHGC-145512	●						●					Info
G16421	●											Info
INPP1_3186	●											Info
RH17607	●			●								Info
WI-9155	●										●	Info

EST

"Expressed Sequence Tags"

- is a type of STS
- a short sub-sequence of a transcribed cDNA sequence.



ESTs are small DNA sequences (usually 200 to 500 nucleotides long) that are generated by sequencing either one or both ends of an expressed gene.

The identification of ESTs has proceeded rapidly, with approximately 65,9 million ESTs now available in public databases

EST PROPERTIES

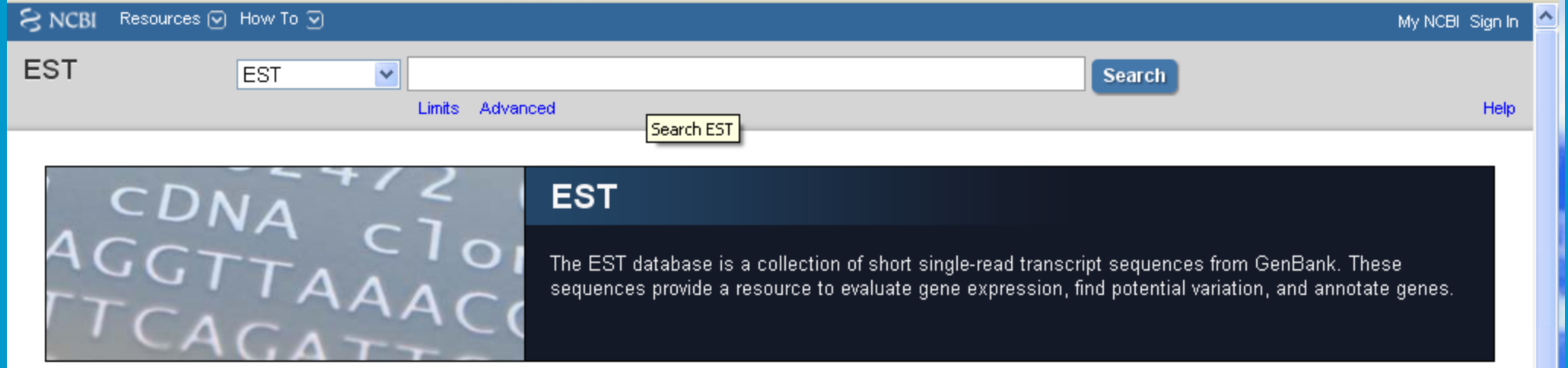
- the ESTs represent portions of expressed genes. The sequence contains only exons of the gene, spliced together to form the sequence encoding for the protein.
- ESTs can be mapped to specific chromosome locations using physical mapping techniques (such as radiation hybrid mapping or FISH).
  - Alternatively, if the genome of the organism that originated the EST has been sequenced, the EST sequence can be aligned to that genome.

EST FUNCTIONS

- The human set of discovered genes includes thousands of genes based solely on EST evidence.
- ESTs have become a tool to refine the predicted transcripts for those genes, which leads to the prediction of their protein products and ultimately their function.
- Moreover, the situation in which ESTs are obtained (tissue, organ, disease state) gives information on the conditions in which the gene is expressed.
- ESTs contain enough information to permit the design of precise probes that then can be used to determine the gene expression.

dbEST

dbEST is a division of GenBank that contains sequence data and other information on the "Expressed Sequence Tags" from a number of organisms.



The EST database is a collection of short single-read transcript sequences from GenBank. These sequences provide a resource to evaluate gene expression, find potential variation, and annotate genes.

You are here: NCBI > DNA & RNA > Database of Expressed Sequence Tags (dbEST)

Write to the Help Desk

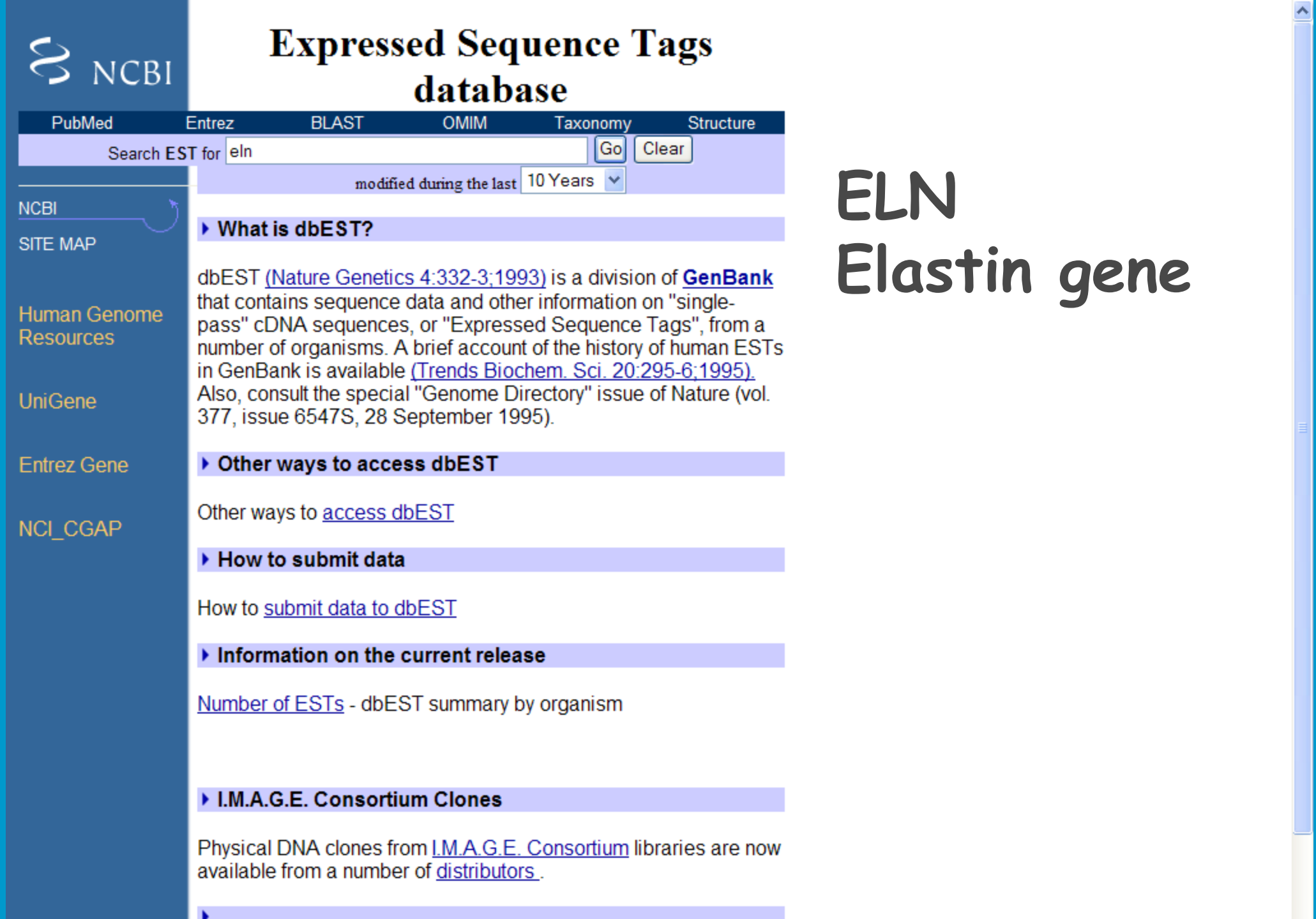
GETTING STARTED
NCBI Education
NCBI Help Manual
NCBI Handbook
Training & Tutorials

RESOURCES
Chemicals & Bioassays
Data & Software
DNA & RNA
Domains & Structures
Genes & Expression
Genetics & Medicine
Genomes & Maps

POPULAR
PubMed
Nucleotide
BLAST
PubMed Central
Gene
Bookshelf
Protein

FEATURED
GenBank
Reference Sequences
Map Viewer
Genome Projects
Human Genome
Mouse Genome
Influenza Virus

NCBI INFORMATION
About NCBI
Research at NCBI
NCBI Newsletter
NCBI FTP Site
NCBI on Facebook
NCBI on Twitter
NCBI on YouTube



ELN
Elastin gene

Display Settings: EST Send to:

TC119860 Human endothelial cells, large insert, pCMV expression library Homo sapiens cDNA clone TC119860 5- similar to Homo sapiens elastin (supravalvular aortic stenosis, Williams-Beuren syndrome) (ELN), mRNA sequence

GenBank: DN999424.1

GenBank FASTA

IDENTIFIERS

dbEST Id: 28903320
EST name: TC119860
GenBank Acc: DN999424
GenBank gi: 66259251

CLONE INFO
Clone Id: TC119860 (5')
Source: OriGene Technologies, Inc. (www.origene.com)
Id as DNA: TC119860
Id in host: TC119860
DNA type: cDNA

PRIMERS
Sequencing: pCMV6 5prime forward vector primer, OriGene Technologies Inc.
PolyA Tail: Unknown

SEQUENCE

GCACGAGGCCGAGATGGCGGGTCTGACGGCGGCGGCCCCGCGGCCCGGAGTCCTCCTGCT
CCTGCTGTCCATCCTCCACCCCTCTCGGCCTGGAGGGGTCCCTGGGGCCATTCCTGGTGG
AGTTCCTGGAGGAGTCTTTTATCCAGGGGCTGGTCTCGGAGCCCTTGGAGGAGGAGCGCT
GGGGCCTGGAGGCAAACCTCTTAAGCCAGTTCCCGGAGGGCTTGCGGGTGCTGGCCTTGG
GGCAGGGCTCGGCGCCTTCCCCGCAGTTACCTTTCCGGGGGCTCTGGTGCCTGGTGGAGT

Analyze this sequence
Run BLAST
Pick Primers

Reference sequence information
RefSeq alternative splicing
See 5 reference mRNA sequence splice variants for the ELN gene.

ESTs for the ELN gene
This EST is one of 1019 sequences matched to ELN: Elastin.

More about the ELN gene
This gene encodes a protein that is one of the two components of elastic fibers. The encoded protein is rich in hydrophobic amino acids such...
Also Known As: FLJ38671, FLJ43523, SVA...

Homologs of the ELN gene
The ELN gene is conserved in chimpanzee, cow, mouse, and rat.

All links from this record
Taxonomy
Map Viewer

Nucleotide

Alphabet of Life

Search: Nucleotide Limits Advanced search Help

Search Clear

Display Settings: Summary, Sorted by Default order

Send to:

Results: 5

1. Homo sapiens elastin (ELN), transcript variant 2, mRNA
   GI:126352699
   GenBank FASTA Graphics Related Sequences
2. Homo sapiens elastin (ELN), transcript variant 5, mRNA
   GI:126352607
   GenBank FASTA Graphics Related Sequences
3. Homo sapiens elastin (ELN), transcript variant 4, mRNA
   GI:126352445
   GenBank FASTA Graphics Related Sequences
4. Homo sapiens elastin (ELN), transcript variant 1, mRNA
   GI:126352439
   GenBank FASTA Graphics Related Sequences
5. Homo sapiens elastin (ELN), transcript variant 3, mRNA
   GI:126352321
   GenBank FASTA Graphics Related Sequences

Display Settings: Summary, Sorted by Default order

Send to:

Filter your results:

All (5)

Bacteria (0)

INSDC (GenBank) (0)

mRNA (5)

RefSeq (5)

Analyze these sequer

Run BLAST

Find related data

Database: Select

Find items

Recent activity

Single nucleotide polymorphism (SNPs)

A Single Nucleotide Polymorphism (SNP) is a small genetic change that occurs within a DNA sequence.

SNP variation occurs when a change, insertion or deletion affects a single nucleotide.

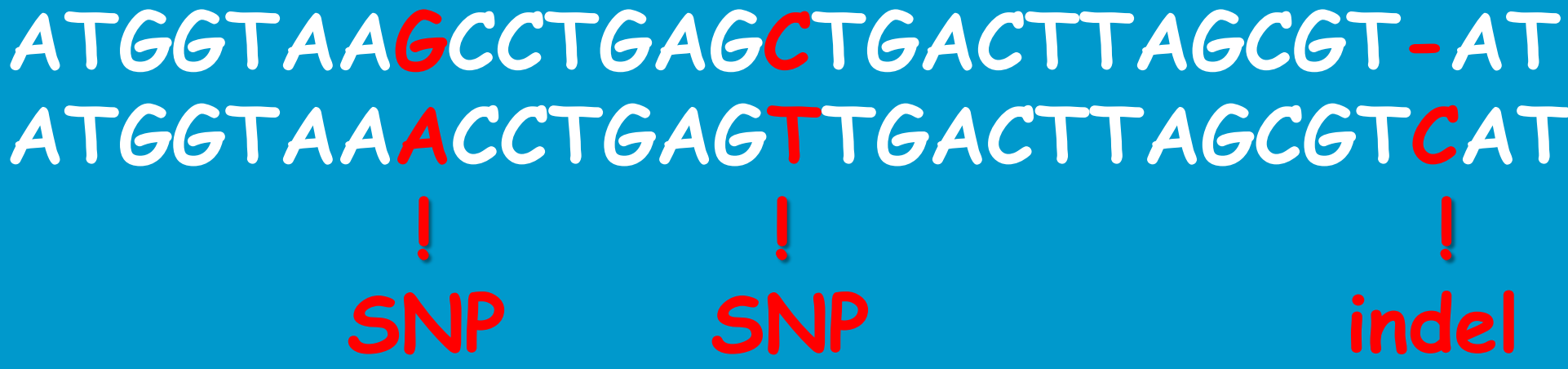
Also short deletions can be included among SNPs

SNP variation occurs when a single nucleotide, such as an *A*, replaces one of the other three nucleotide letters—*C*, *G*, or *T*.

AAGGTTA to ATGGTTA

Two of every three SNPs involve

cytosine (C) → thymine (T).



SNPs result from replication errors and DNA damage

GENOMIC SNP FREQUENCY

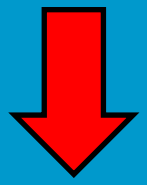
- from 1 every 100 to 300 bases along the 3-billion-base human genome.
- Because only about 3 - 5 % of a person's DNA sequence codes for proteins, most SNPs are found outside of "coding sequences".

SNP VARIATION AMONG PERSONS

- 99.9% of one individual DNA sequences is identical to that of another person.
- Of the 0.1% difference, over 80% is imputable to single nucleotide variation.
- For a variation to be considered polymorphic (SNP), it must occur in at least 1% of the population.

SNPS OCCUR FREQUENTLY THROUGHOUT THE GENOME,

TEND TO BE RELATIVELY STABLE,



SERVE AS BIOLOGICAL MARKERS.

SNP pattern helps in disease association studies

- in Mendelian diseases (e.g. cystic fibrosis, Huntington's disease)

  - Genetic mutations causes the disease
  - Environmental variation usually irrelevant
  - Usually rare
  - Occurs in isolated pedigrees

SNP pattern helps in disease association studies

- Complex diseases (schizophrenia, hypertension, diabetes)

  - Genetic variants increases the risk of disease
  - Environmental variation usually important
  - Often common
  - Occurs in general population

Many independent mutations can be associated with increased predisposition to a disease.

SNPs found within a coding sequence

- can directly alter the biological function of a protein.
  - Mutations from a sense codon to a stop codon
  - Deletion or insertion causing phase shift in translation
  - Amino acid substitutions affecting the protein function
- do *not* alter the biological function of a protein.
  - Mutations to synonymous codons
  - Substitutions not affecting the protein function

SNPs, prevalently outside the coding sequences, have no effect on cell function

But it has been hypothetized that

- can account for differences in the risk of disease or response to drugs.
- can help to identify the multiple genes associated with or conferring predisposition to complex pathologies (cancer, diabetes, vascular disease, and some forms of mental illness).

NCBI plays a major role in facilitating the identification and cataloging of SNPs through creation and maintenance of the public SNP database (dbSNP)

dbSNP

is a public database of single nucleotide polymorphisms (SNPs).

The data can be from any species, and from any part of a particular genome.

Single Nucleotide Polymorphisms (SNPs)

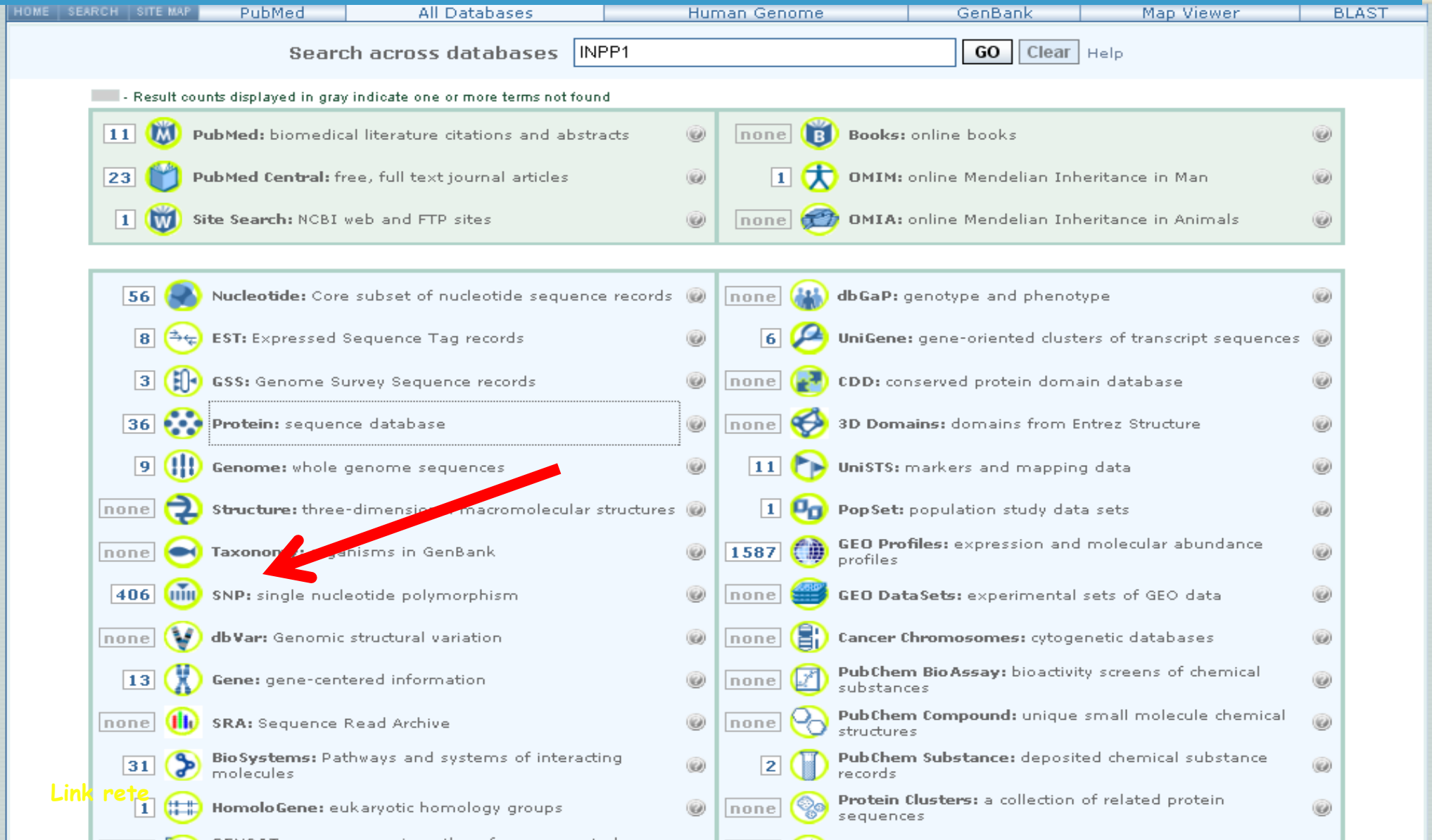
- Exist at defined positions
  - can be used for gene mapping, defining population structure, and performing functional studies.
- Report the sequence information around the polymorphism, the specific experimental conditions, and frequency information by population or individual genotype.

dbSNP

contains several classes of genetic variation:

- Single Nucleotide Polymorphism (SNP)
- Deletion/Insertion Polymorphism (DIP)*
- Microsatellite or Short Tandem Repeat (STR)
- Multi-Nucleotide Polymorphism (MNP)

Search for SNPs associated with INPP1 (inositol polyphosphate-1-phosphatase) gene



Link rete

NCBI

ENTREZ SNP
Single Nucleotide Polymorphism

My NCBI
[Sign In] [Reg

All Databases | PubMed | Nucleotide | Protein | Genome | Structure | OMIM | PMC | Journals | Books

Search SNP for [ ] Go Clear

Limits | Preview/Index | History | Clipboard | Details

NCBI
dbSNP BUILD 132

Entrez SNP
Search SNP
Search Mouse SNP
Common Query Filters
Entrez Batch Query
SNP Link Datamodel

My NCBI
My NCBI help

Entrez SNP Help
Searchable FAQ
Search Fields
Programming Utilities
Batch Report
Legend
Examples
dbSNP Home Page
Overview

Entrez Help
General help
Limits
Preview/Index
History
Clipboard
Details

Other Services
Citation Matcher
Clinical Queries

Limit your search by any of the following criteria.

Organism CLEAR
- Homo sapiens
- Agelaius phoeniceus
- Alectoris
- Alectoris chukar
- Alectoris rufa
- Allium cepa
- Amaranthus caudatus
- Anopheles funestus
- Anopheles gambiae

Chromosomes CLEAR
- 1
- 2
- 2a
- 2b
- 3
- 4
- 5
- 6
- 7

Chromosome Range CLEAR
From:
To:

Map Weight CLEAR
- 1
- 2
- 3-10
- 10+

Function Class CLEAR
- coding nonsynonymous
  - nonsense
  - missense
  - frame shift
- intron
- coding synonymous
- locus region
- mrna utr

SNP Class CLEAR
- het
- in del
- microsatellite
- mixed
- mnp
- named
- no variation
- snp

Method Class CLEAR
- computed
- dhplc
- hybridize
- other
- rflp
- sequence
- sscp
- unknown

Validation Status CLEAR
- by-cluster
- by-frequency
- by-submitter
- by-2hit-2allele
- no-info

Variation Allele CLEAR
- A
- C
- G
- T
- M
- R
- W
- S
- Y

Annotation CLEAR
- Clinical/LSDB Submissions
- nucleotide
- OMIM

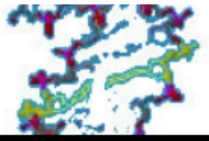
Heterozygosity CLEAR
- 0-10
- 40-50
- 10-20

Success Rate CLEAR
- 80-85
- 85-90
- 90-95

limits

NCBI

dbSNP
Short Genetic Variations



All Databases | PubMed | Nucleotide | Protein | Genome | Structure | OMIM | PMC | Journals | Books

Search SNP for INPP1 Go Clear Save Search

Limits | Preview/Index | History | Clipboard | Details

Display Graphic Summary Show 20 Sort By Send to

All: 915 | Cited in PubMed: 0 | Clinical/LSDB Submissions: 0 | Human: 585

Items 1 - 20 of 915

Page 1 of 46 Next

1: rs69127709 *[Equus caballus]* Links

GCTGARCCCAGGGTGGGTGCTTCATC[A/G]GTGGCAGTGATGAGCATCATTATCC

18 MapView No VarVu No PubMed No Gene Not on mRNA No 3D No OMIM

2: rs69127708 *[Equus caballus]* Links

TTGCATGATCCCATGACTTTCGCTGA[A/G]CCCAGGGTGGGTGCTTCATCRGTGG

18 MapView No VarVu No PubMed No Gene Not on mRNA No 3D No OMIM

3: rs137398285 *[Bos taurus]* Links

CTATTAAAGTAGTGCTCTTTCCTTCA[-/C]CCTTTGGGAGAAGGAAGAAGTGAAA

2 MapView No VarVu No PubMed No Gene Not on mRNA No 3D No OMIM

4: rs137239878 *[Bos taurus]* Links

TCTTTTTTTTGGTGGGGGGGGGGTGG[G/T]CAACAACAAAAATACCATCTAACAT

2 MapView No VarVu No PubMed No Gene Not on mRNA No 3D No OMIM

5: rs136981326 *[Bos taurus]* Links

TAACATTTCTCAGTGTTTAACTTACT[A/C]TATACGCTGGATAGTTTCTTTCCAG

Recent activity

INPP1 (915)

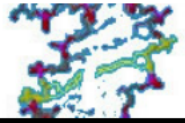
INPP1 (21)

Pan troglodytes

Gene Help: Integ
Genomes in the

NCBI

dbSNP
Short Genetic Variations



My NCBI
[Sign In] [Regis

All Databases | PubMed | Nucleotide | Protein | Genome | Structure | OMIM | PMC | Journals

Search SNP for INPP1 Go Clear Save Search

Homo sapiens

Limits | Preview/Index | History | Clipboard | Details

Display Graphic Summary Show 20 Sort By Send to

All: 915 | Cited in PubMed: 0 | Clinical/LSDB Submissions: 0 | Human: 585

Items 1 - 20 of 585

Page 1 of 30 Next

1: rs193290039 *[Homo sapiens]* Links

AAAATTTTTTAGTCTCCTTCATGCAC[C/T]ACATATTAAGATGCAGTTAGTTCTA

2 MapView No VarVu No PubMed GeneView Not on mRNA No 3D No OMIM

HGVS Names: [*NC_000002.11:g.191228383C>T*] [*NM_001128928.1:c.265+970C>T*] [*NM_002194.3:c.265+970C>T*]

2: rs193178379 *[Homo sapiens]* Links

GATGGAGGAGGAGTCAGAAATGTCTC[A/C]TGAAGGAGCCAACAGTTGAAATGAT

2 MapView No VarVu No PubMed GeneView Not on mRNA No 3D No OMIM

HGVS Names: [*NC_000002.11:g.191216000C>A*] [*NM_001128928.1:c.-65+2243C>A*] [*NM_002194.3:c.-65+7313C>A*]

3: rs193134347 *[Homo sapiens]* Links

CTTTTAAGAATCATGTTTAAATTGTT[C/T]CTTGATTTTGCCATCATAGATATGA

2 MapView No VarVu No PubMed GeneView Not on mRNA No 3D No OMIM

HGVS Names: [*NC_000002.11:g.191207186T>C*]

4: rs192963836 *[Homo sapiens]* Links

TCAAGCAATCCTCCTATCTCAGCTTC[C/T]CAAGTAGCTGGGACCCAAGTAGGTG

2 MapView No VarVu No PubMed GeneView Not on mRNA No 3D No OMIM

HGVS Names: [*NC_000002.11:g.191232401C>T*] [*NM_001128928.1:c.466+780C>T*] [*NM_002194.3:c.466+780C>T*]

5: rs192815424 *[Homo sapiens]* Links

Recent activity

Turn Off Cle

INPP1 (915)

INPP1 (21)

Pan troglodytes

Gene Help: Integrated Access to Genes of Genomes in the Reference Sequence Collectio

» See mor

Graphic Summary Help - Mozilla Firefox

File Modifica Visualizza Cronologia Segnalibri Strumenti Aiuto

http://www.ncbi.nlm.nih.gov/corehtml/query/Snp/EntrezSNPlegend.html

Più visitati Come iniziare Ultime notizie

Homo sapiens cystic fibrosis... | NCBI: All Resources | **Graphic Summary Help**

Graphic Summary :

- 10 MapView — Mapped to chromosome shown with map weight 1 (single green bar), linkout to MapViewer
- 10 MapView — Mapped to chromosome shown with map weight greater than 1 (two or more green bar)
- M no Map — Mapped to multiple chromosomes
- ? MapView — Unknown, not on chromosome
- GeneView — SNP in locus region, linkout to Gene View in dbSNP
- SeqView — SNP in coding region (Non-synonymous)
- SeqView — SNP in coding region (synonymous)
- SeqView — SNP in other mRNA regions (intron, UTR, etc.)
- Not on mRNA — SNP not on mRNA
- Protein 3D — Structure neighbor available (Cn3D), linkout to structure mapping summary
- OMIM — linkout to Omim record
- V — Validated
- G — Genotype data available

Actual percentage (1-100) heterozygosity indicated by the red arrow (ie. 9%)and actual success rate indicated by the blue arrow (ie. 95%).

Reference SNP clusters
rs[NCBI SNP ID]

Reference SNP cluster are created by NCBI during periodic 'builds' of the database.

Define a non-redundant set of markers that are used for annotation of reference genome sequence and integration with other NCBI resources.

Novel SNP submission at new positions in genome sequence

NCBI | ssASSAY ID

refers to an individual submission record and will instantiate a new refSNP cluster.

New submissions that match existing data will be merged into an existing refSNP cluster rs[NCBI SNP ID].

where 'rs' and ss are always lower case.

Reference SNP reports (rs) and submitter reports (ss) have different identifiers in dbSNP

When two submitted SNP records refer to the same location in the genome, Reference SNP records will provide a summary list of submitter records in dbSNP and a list of external resource and database links.

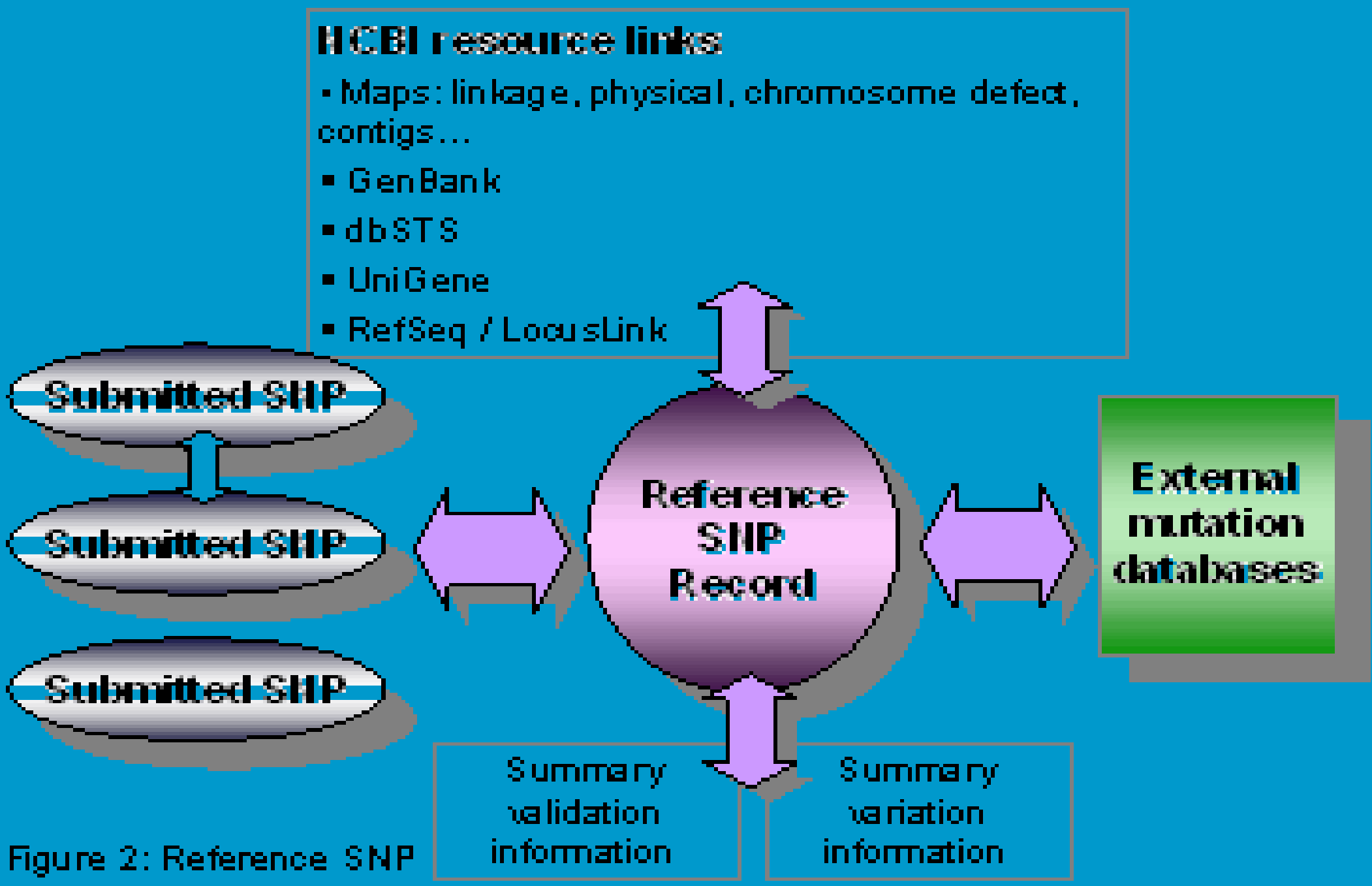


Figure 2: Reference SNP

SNPs are indexed by two different accession numbers in dbSNP

- the NCBI | ssASSAY ID form which refer to an individual submission record
- the NCBI | rsSNP ID form which refers to the reference SNP cluster record and all associated records.
  - where 'rs' and 'ss' are always lower case.

WITHIN REFERENCE SNP CLUSTERS (rs)

MUTATION DATA: information on the specific alleles and the flanking sequence that surrounds the mutation.
COLLECTION METHODS: Descriptions of the assay technique used to type the SNP.
SUBMITTER DATA: Contact information is maintained for the individual submitter. Bibliographic data for unpublished or in-press citations are recorded.
VARIATION DATA: frequency formation provided by population, and genotype information provided for individuals. Populations are defined by the submitter. Individuals may be sub-classified by population or sample frame.

In dbSNP record the sequence data consists of three elements:

- The sequence 5' to the site of mutation
- The mutation itself
- The sequence 3' to the site of mutation

For the 5' and 3' sides are reported 25 bases.

The standard IUPAC ambiguity characters are permitted to identify regions of known variation.

a slash ('/'), to denote the alternative alleles

IUPAC

International Union of Pure and Applied Chemistry

Search

The International Union of Pure and Applied Chemistry (IUPAC) serves to advance the worldwide aspects of the chemical sciences and to contribute to the application of chemistry in the service of Human kind. As a scientific, international, non-governmental and objective body, IUPAC can address many global issues involving the chemical sciences.

IUPAC nucleotide code	Base
A	Adenine
C	Cytosine
G	Guanine
T (or U)	Thymine (or Uracil)
R	A or G
Y	C or T
S	G or C
W	A or T
K	G or T
M	A or C
B	C or G or T
D	A or G or T
H	A or C or T
V	A or C or G
N	any base
. or -	gap

Search for SNP: rs61734584.

http://www.ncbi.nlm.nih.gov/sites/entrez?db=snp

Name the alleles for this SNP.

What contig does this SNP map to?

What gene does this SNP map to?

In which mRNA and Protein refseq records is this SNP?

Are there population diversity data?

More than 2.8 million of SNPs are already known and described in NCBI dbSNP.

The major source of this public SNP catalog was done by

The SNP Consortium (TSC)

a collaborative genomics effort of major pharmaceutical companies, the Wellcome Trust and academic centers.

The SNP Consortium (TSC)

was established in 1999 and the initial goal was to discover 300 000 SNPs in two years, but at the end of 2001, 1.4 million SNPs had been released into the public domain.

The human genome is thought to contain at least 10 million SNPs, about one in every 300 bases.

Theoretically, researchers could hunt for genes using a map listing all 10 million SNPs, but there are major practical drawbacks to that approach.

Often SNPs are 'conserved' across the genome, in patterns called 'haplotype blocks'

- Sets of SNPs that are close together and tend to be inherited together.
- A set of associated SNP alleles in a region of a chromosome is called a "haplotype".
- Most chromosome regions have only a few common haplotypes (each with a frequency of at least 5%), which account for most of the variation from person to person in a population.

SNPs and Haplotypes

- SNP: Single Nucleotide Polymorphism
- Haplotype: A set of closely linked genetic markers present on one chromosome which tend to be inherited together (not easily separable by recombination).

G——G——A——C——A

Set of SNP polymorphisms: a SNP haplotype

The International HapMap Project

is a collaboration among scientists from various countries to identify and catalog genetic similarities and differences in humans.

The goal is to develop a haplotype map of the human genome, which will describe the common patterns of human DNA sequence variation.

The Project officially started with a meeting on October 27-29, 2002

About the HapMap

What is the HapMap?
Origins of Haplotypes
Health Benefits
Populations Sampled
Ethical Issues
Consent Forms
Community Advisory Groups(CAG)
Data Release Policy
Guidelines For Data Use
Guidelines For Referring to HapMap Populations

Project Information

Home
Project Data
HapMap Mailing List
HapMap Project Participants

Useful Links

HapMap Project Press Release
NHGRI HapMap Page
NCBI Variation Database (dbSNP)
Japanese SNP Database (JSNP)

About the HapMap

The International HapMap Project is a multi-country effort to identify and catalog genetic similarities and differences in human beings. Using the information in the HapMap, researchers will be able to find genes that affect health, disease, and individual responses to medications and environmental factors. The Project is a collaboration among scientists and funding agencies from Japan, the United Kingdom, Canada, China, Nigeria, and the United States. [See **Participating Groups** and **Initial Planning Groups**.] All of the information generated by the Project will be released into the public domain.

The goal of the International HapMap Project is to compare the genetic sequences of different individuals to identify chromosomal regions where genetic variants are shared. [**See What is the HapMap?**] By making this information freely available, the Project will help biomedical researchers find genes involved in disease and responses to therapeutic drugs. [**See How Will the HapMap Benefit Human Health?**] In the initial phase of the Project, genetic data are being gathered from **four populations** with African, Asian, and European ancestry. Ongoing interactions with members of these populations are addressing potential **ethical issues** and providing valuable experience in conducting research with identified populations.

Public and private organizations in six countries are participating in the International HapMap Project. Data generated by the Project can be **downloaded** with minimal constraints. [**See Data Release Policies**.] The Project officially started with a meeting in October 2002 (**http://genome.gov/10005336**) and is expected to take about three years.

Please send questions and comments on website to hapmap-help@ncbi.nlm.nih.gov

The HapMap is a catalog of common genetic variants that occur in human beings. It describes what these variants are, where they occur in our DNA, and how they are distributed among people within populations and among populations in different parts of the world. The Project is designed to provide information that researchers can use to link genetic variants to the risk for specific illnesses, which will lead to new methods of preventing, diagnosing, and treating disease.