

# Genetik varyasyon: SNP ve indeller

# Hatırlamakta fayda var...

- Alel
- Genotip
- Genetik belirteç

## Genetik çeşitlilik

- Mutasyon
- Rekombinasyon

Neden önemli??

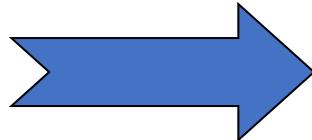


# Varyant ?

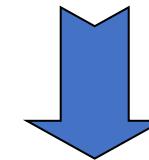
- Normal ya da yabancı tip
- Polimorfizm (ya da polimorfik)
- Mutasyon (ya da mutant)



Mutasyonla  
yeni varyantın  
Ortaya çıkması



nadir allele'in devamlılığı



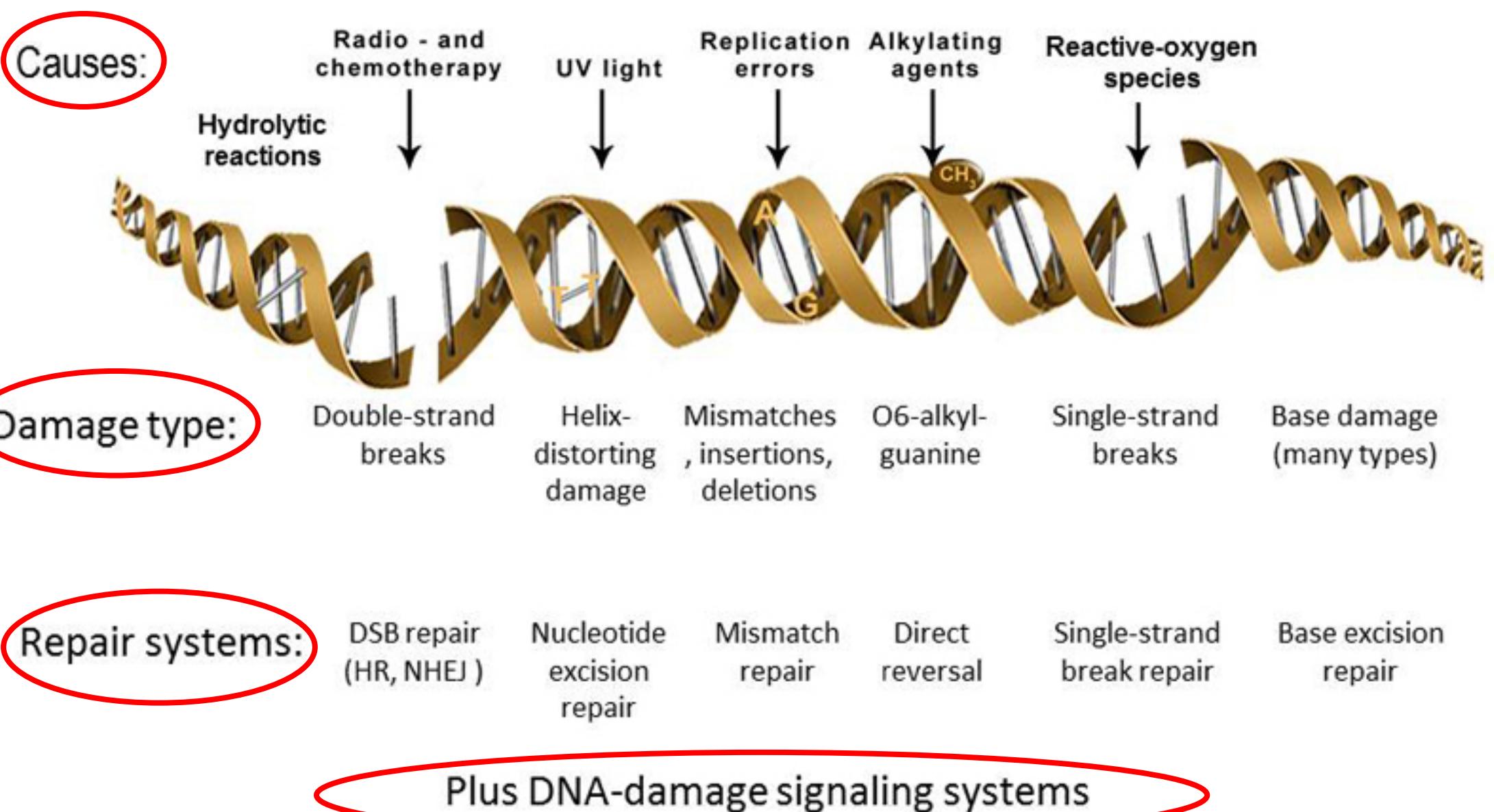
Populasyon genişlemesi ve  
allel frekansının  
artması



Yeni alel populasyonda  
yeni bir polimorfizm  
olarak **fikse olur**

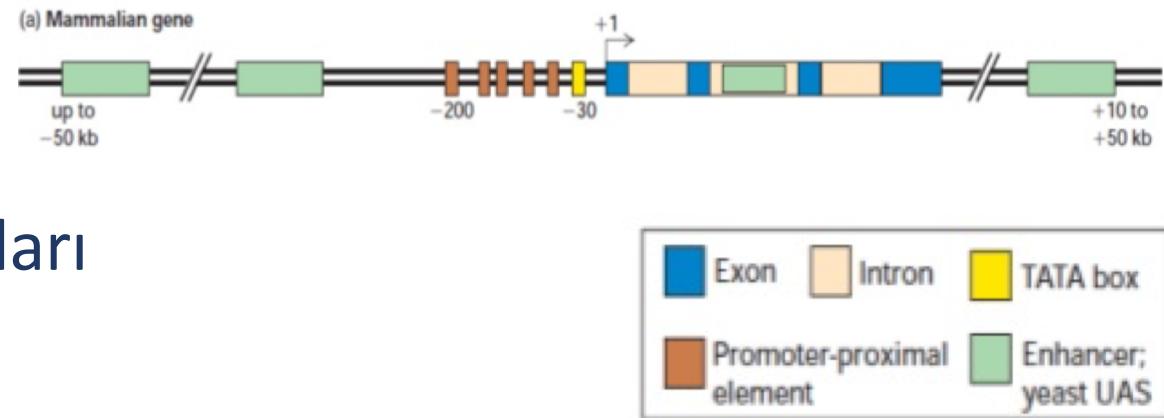
# Mutasyon tiplerinin sınıflandırılması

- Nedenine göre (spontan ya da uyarılmış)
- Yerleşimine göre
- İşlevine göre
- Uyum başarısına etkisine göre
- Büyüklüğüne göre



# Mutasyonun yerleşimi

- Promoterda
- Ekzonda
- Intronda
- Poliadenilasyon bölge mutasyonları
- 5`UTR mutasyonları
- Diğer düzenleyici bölge mutasyonları



# Mutasyonların gen işlevine etkisi

Geri mutasyon orijinal fenotipi geri getiren nokta mutasyonlarıdır

Letal mutasyon mutasyonu taşıyan canının ölümüne neden olur

İşlev kazandıran mutasyon yeni bir işlev kazandırarak gen ürününü değiştirir

İşlev kaybı mutasyonu genin işlevi azalır ya da tamamen kaybolur

Negatif baskın mutasyonlar değişikliğe uğramış gen ürünü yabanıl tip alelin tersi etki gösterir.

# Uyum başarısına etki

- Çoğu mutasyon nötrdür- fakat uzun vadede yararlı ya da zararlı etkileri olabilir
- Zararlı mutasyonlar yabanił hale döner ya da yararlı bir işlev kazandırabilir
- Bazı zararlı mutasyonlar ise hastalıklara neden olurlar

# Mutasyonlar (büyüklüğüne göre)

- **Büyük mutasyonlar**

Genom mutasyonu = kromozom sayısının değişmesi

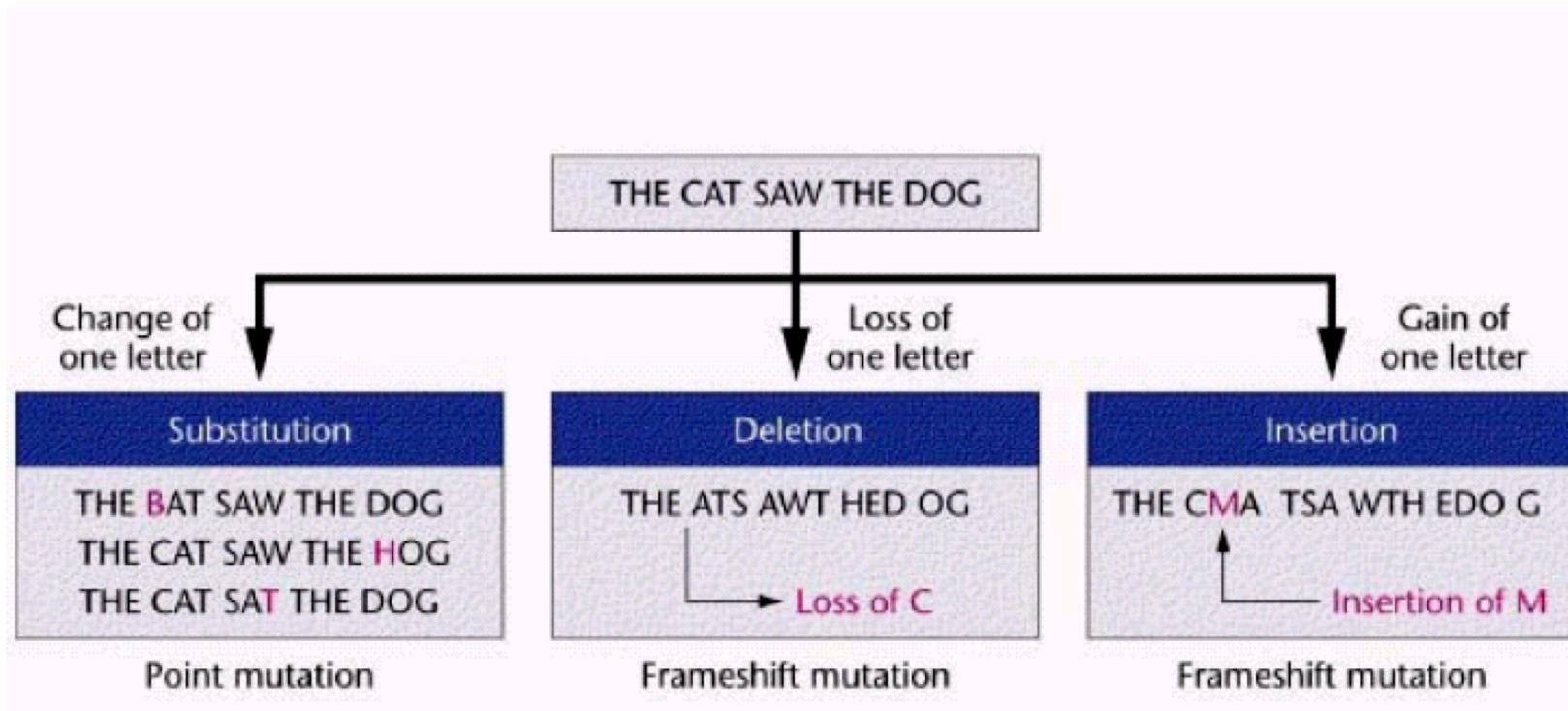
- **Orta büyülükte mutasyonlar**

Kromozom mutasyonları = kromozom yapısının değişmesi

- **Küçük mutasyonlar**

Gen mutasyonları = tek nükleotit değişikliklerinden tüm gene kadar

# DNA'nın büyüklüğünü etkileyen mutasyonlar

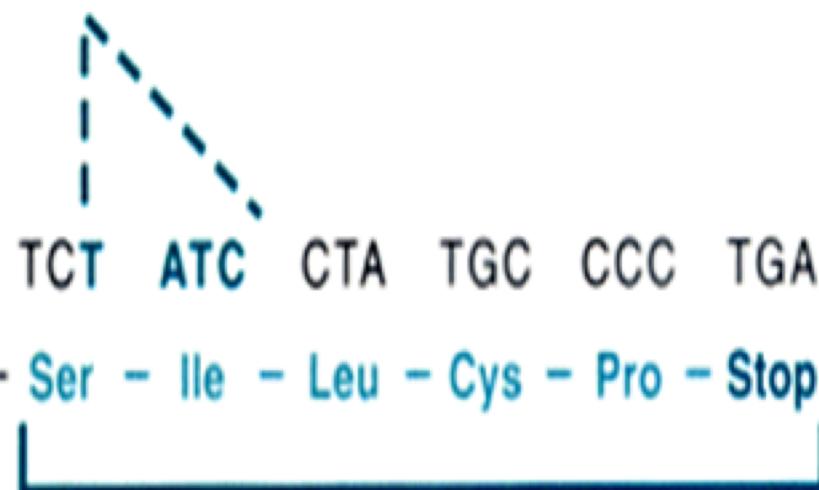


... - Arg - Ile - Ser - Tyr - Gly - Pro - Asp - ...

Normal HEXA allele     ... CGT ATA TCC TAT GCC CCT GAC ...

Tay-Sachs allele     ... CGT ATA TCT ATC CTA TGC CCC TGA C...

... - Arg - Ile - Ser - Ile - Leu - Cys - Pro - Stop



Altered reading frame

Four-base insertion in the hexosaminidase A gene in Tay-Sachs disease, leading to a frameshift mutation

# Nükleotit değişiklikleri

(a)



(b)

Silent mutation

Transversiyon

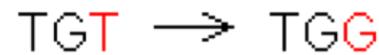
Pyr ↔ Pu

Missense mutation

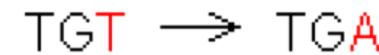
Nonsense mutation



Cys → Cys



Cys → Trp



Cys → Stop

# Missense mutasyon- orak hücreli anemi

## HBB Sequence in Normal Adult Hemoglobin (Hb A):

Nucleotide	CTG	ACT	CCT	GAG	GAG	AAG	TCT
Amino Acid	Leu	Thr	Pro	Glu	Glu	Lys	Ser
	3			6		9	

## HBB Sequence in Mutant Adult Hemoglobin (Hb S):

Nucleotide	CTG	ACT	CCT	GTG	GAG	AAG	TCT
Amino Acid	Leu	Thr	Pro	Val	Glu	Lys	Ser
	3			6		9	

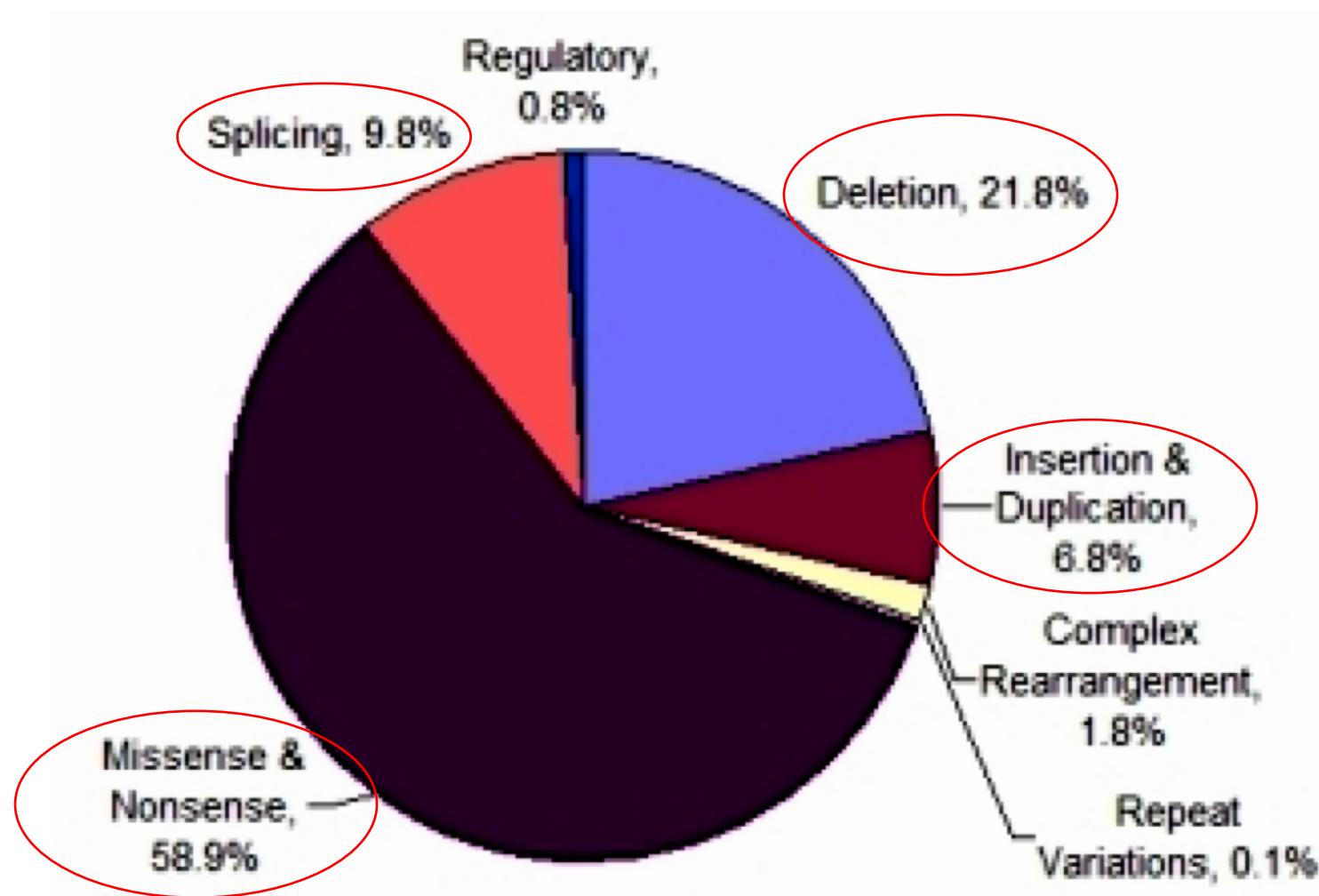


Normal red blood cell



Sickled red blood cell

# Hastalığa neden olan mutasyonların siklikları



# Polimorfizm farklı seviyelerde görülür

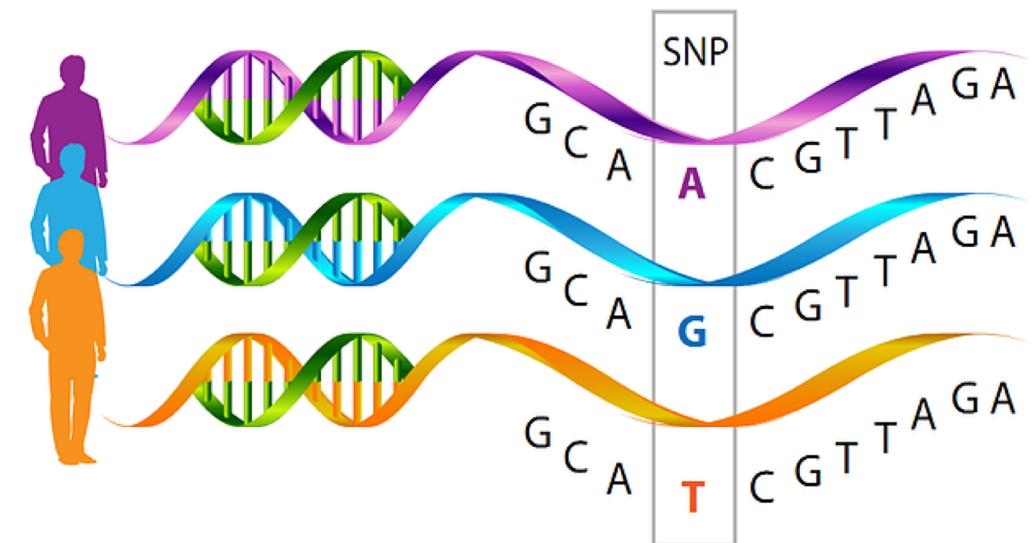
- Fenotip polimorfizmi
- Protein polimorfizmi (immunoglobulinler, AB0 kan grupları)
- DNA polimorfizmi

# Genetik polimorfizmler

- Kromozomal
- Ardarda gelen tekrar dizileri
- Satelit dna (perisentrik heterokromatin)
- Minisatelit (VNTR, telomer)
- Microsatelit (STR)
- SNP

# SNP

- Polimorfizmin en basit ve yaygın görülen tipidir
- İnsan genomunda her 1000 bazda bir SNP bulunur
- Çoğu SNP hücresel işlevler üzerine etkisizdir





Hastalığa dirençli populasyon



Hastalığa yatkın populasyon

Her birey binlerce SNP için tarandı

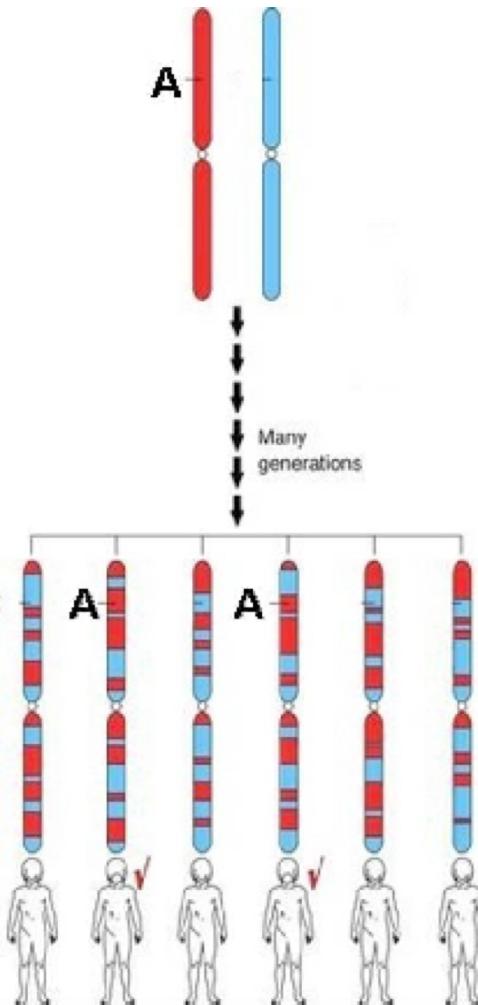
ATG**A**TTATAG

*geneX*

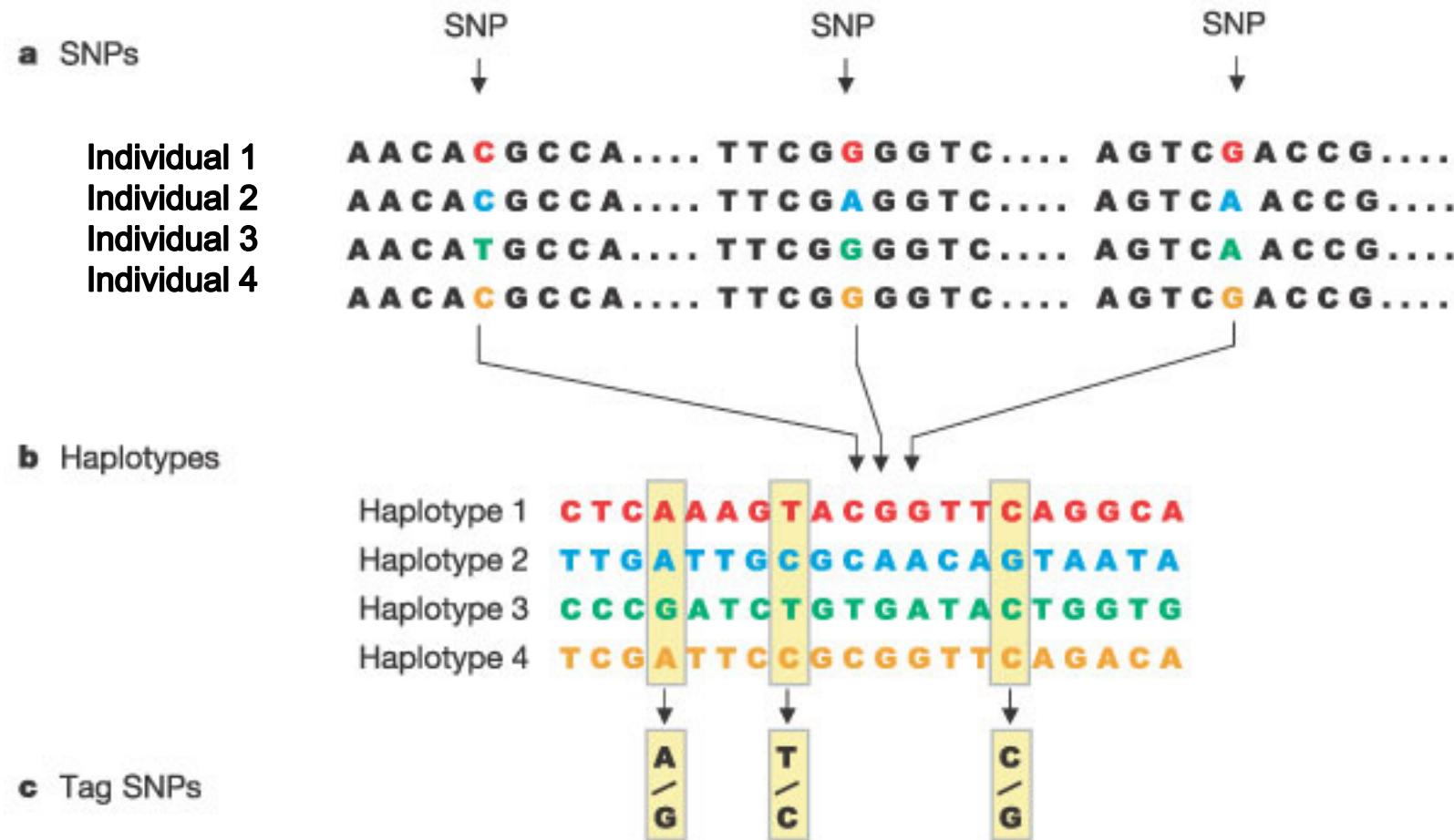
ATG**T**TTATAG

Dirençli bireylerin hepsi *geneX*'de 4. pozisyonda 'A' içerirken, hastalığa yatkın olanlar 'T' içerir

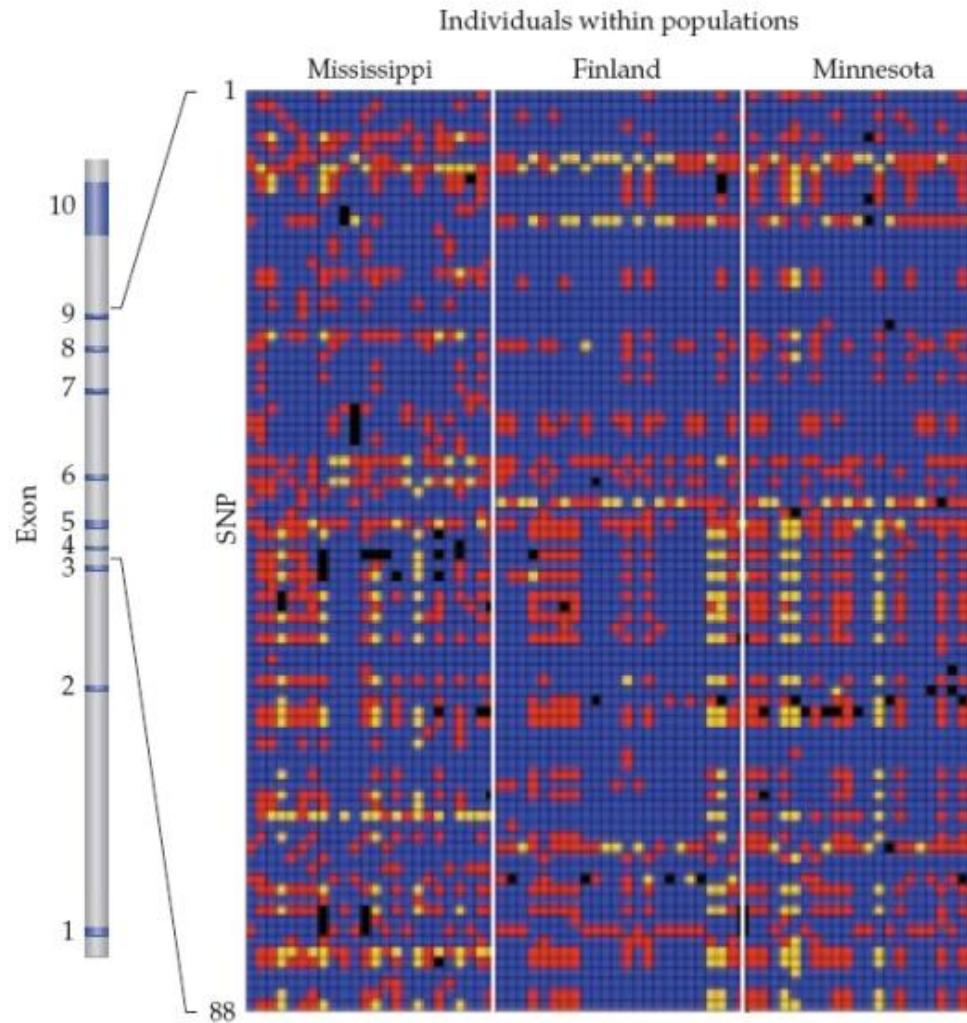
# Haplɔtiplerin orijini



# haplotip



# Farklı populasyonlarda SNP taramaları



# SNP yerine haplotip kullanmanın faydaları

- Bilgi daha fazladır
- Gen işlevi birden fazla SNP'den etkileniyor olabilir
- Yanlış pozitiflik oranının daha düşük olması
- Genotipleme hatalarının büyük ölçüde elimine edilmesi

# International HapMap Project

<http://www.hapmap.org/>



**International HapMap Project**

[Home](#) | [About the Project](#) | [Data](#) | [Publications](#) | [Tutorial](#)

[中文](#) | [English](#) | [Français](#) | [日本語](#) | [Yoruba](#)

The International HapMap Project is a partnership of scientists and funding agencies from Canada, China, Japan, Nigeria, the United Kingdom and the United States to develop a public resource that will help researchers find genes associated with human disease and response to pharmaceuticals. See "[About the International HapMap Project](#)" for more information.

## Project Information

- [About the Project](#)
- [HapMap Publications](#)
- [HapMap Tutorial](#)
- [HapMap Mailing List](#)
- [HapMap Project Participants](#)

## Project Data

- [HapMap Genome Browser release #28 \(Phases 1, 2 & 3 - merged genotypes & frequencies\)](#)
- [HapMap3 Genome Browser release #3 \(Phase 3 - genotypes & frequencies\)](#)
- [HapMap Genome Browser release #27 \(Phase 1, 2 & 3 - merged genotypes & frequencies\)](#)
- [HapMap3 Genome Browser release #2 \(Phase 3 - genotypes, frequencies & LD\)](#)
- [HapMap Genome Browser release #24 \(Phase 1 & 2 - full dataset\)](#)
- [GWAs Karyogram](#)
- [HapMart](#)
- [HapMap FTP](#)
- [Bulk Data Download](#)
- [Data Freezes for Publication](#)
- [ENCODE Project](#)
- [Guidelines For Data Use](#)

## News

- 2011-01-19: HapMap phase II recombination rate on GRCh37**

The liftover of the HapMap II genetic map from human genome build b35 to GRCh37 is available. Data is [available for bulk download](#).

- 2010-08-18: HapMap Public Release #28**

Genotypes and frequency data in hapmap format are now available for data in merged HapMap phases I+II+III release #28 (NCBI build 36, dbSNP b126). Data is [available for bulk download](#) and also [available for browsing](#). Click here to read the latest [release notes](#).

- 2010-05-28: HapMap3 Public Release #3**

Genotypes and frequency data in hapmap format are now available for data in HapMap phase 3 release #3 (NCBI build 36, dbSNP b126). Data is [available for bulk download](#) and also [available for browsing](#). Click here to read the latest [release notes](#).

- 2010-05-28: HapMap3 CNV Genotypes**

Copy Number Variation genotypes for HapMap phase samples are [available for bulk download](#).

- 2009-12-10: Corrected HapMap3 phased haplotypes available for chromosome X**

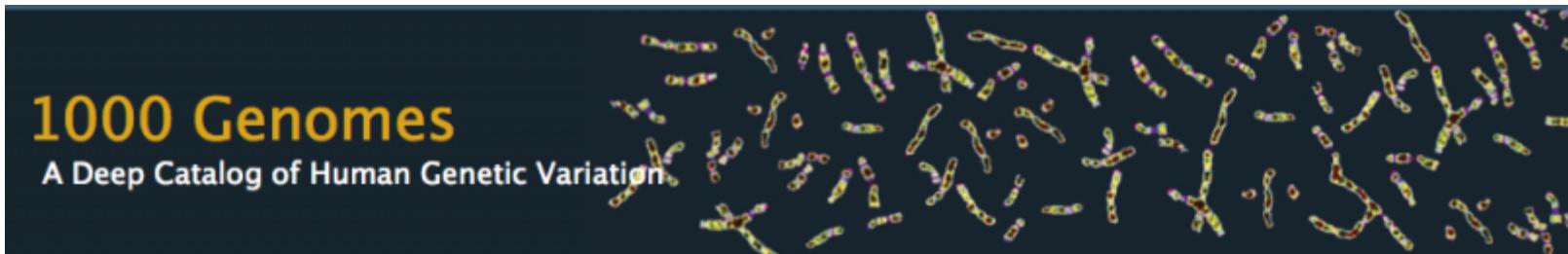
Phased haplotypes for consensus HapMap3 release 2 data for chromosome X has been corrected and the new data are now [available for bulk download](#). Sorry for any inconvenience this might have caused.

- 2009-12-02: HapMap3 phased haplotypes available for chromosome X**

Phased haplotypes for consensus HapMap3 release 2 data has been phased for chromosome X and are now available for bulk download. [Update: The downloading was disabled because several users have found that there are repeating data in some of the chrX phasing data files. The data source is being contacted and the downloading will be enabled as soon as the problem is cleared.]

# Thousand Genomes Project

<http://www.1000genomes.org/>



**1000 Genomes**  
A Deep Catalog of Human Genetic Variation

Home   About   Data   Analysis   Participants   Contact   Browser   Wiki   FTP search    Search

**LATEST ANNOUNCEMENTS**

WEDNESDAY OCTOBER 12, 2011  
**October 2011 Integrated Variant Set release #ICHG2011**  
This [October 2011](#) release represents an integrated set of variant calls and phased genotypes including SNPs, short INDELs and Deletions based on low coverage and exome sequencing data across 1092 individuals.  
Our [FAQ](#) contains instructions on how to get [smaller subsections](#) of these files  
Data access links: [EBI](#) / [NCBI](#)  
Link to additional information:[README file](#)

---

THURSDAY JUNE 23, 2011  
**June 2011 Data Release**  
Genotypes for 1094 individuals for the [May 2011 SNP calls](#) from the 20101123 sequence and alignment release of the 1000 genomes project has now been made. This release is based on the GRCh37 assembly of the human genome and is released in the format [VCF 4.0](#).  
Our [FAQ](#) contains instructions on how to get [smaller subsections](#) of these files

**NAVIGATION**

- [Frequently Asked Questions](#)

**LINKS**

-  [All Project Announcements](#)
-  [Sample and Project Information](#)
-  [Media Archive](#)
-  [Download the 1000 Genomes Pilot Paper](#)

# 10,000 Genomes Project Evolutionary Biology

<http://www.genome.gov/>

The screenshot shows the homepage of the NHGRI website. The header features the NHGRI logo and navigation links for Home, About NHGRI, Newsroom, and Staff. Below the header is a banner with links for Research, Grants, Health, Policy & Ethics, Educational Resources, and Careers & Training. The main content area includes a news item about peer review changes, a video interview with Drs. Green and Felsenfeld, a story on the 10K Genomes Project, a scientist proposal for a genome zoo, and a horse genome analysis. There are also sections for the Human Genome Project, Newsroom, and a Calendar.

genome.gov  
National Human Genome Research Institute  
National Institutes of Health

Home | About NHGRI | Newsroom | Staff

Research Grants Health Policy & Ethics Educational Resources Careers & Training

**Research**  
Intramural Research  
Extramural Research  
Population Genomics

**Grants**  
Funding Opportunities  
Active Grants Database  
Minority Funding

**Health**  
Genetic Disorders FAQ  
Clinical Research FAQ  
GARD Center

**Policy & Ethics**  
Genetic Discrimination  
Genetic Testing  
Legislative Database

**Educational Resources**  
Fact Sheets  
National DNA Day  
Talking Glossary

**Careers & Training**  
Research Training  
Educational Opps  
Working at NHGRI

**Newsroom**

**About NHGRI**

**Enhancing Peer Review Application Changes**  
Read about shortened and restructured application forms and instructions

**10K Genomes Project to Create Vertebrate Genome Zoo**  
As DNA sequencing costs decline, a group of research biologists have pondered the possibility of reading evolution's notebooks in the genomes of thousands of vertebrate species. Using a workshop to organize their ideas, the group proposed The 10K Genomes Project in a paper published Nov. 5 in the *Journal of Heredity*. Co-authors include NHGRI's Eric Green M.D., Ph.D. and Adam Felsenfeld, Ph.D.  
**Scientists propose a "genome zoo" of 10,000 vertebrate species**  
Press release from the University of California at Santa Cruz

**Hay! First Analysis of Horse Genome**  
An international research team, supported in part by NHGRI, has found that the genome of the domestic horse, *Equus caballus*, has a structure remarkably similar to our own genome. Published in the Nov. 6 issue of *Science*, the team's landmark analysis also sheds new light on the centromere evolution.

**Horse genome sequence and analysis published in Science**  
**From NHGRI's Talking Glossary of Genetic Terms: Centromeres**

**Recovery Act at NHGRI**

**The Human Genome Project**  
Read the story behind the Human Genome Project

**Newsroom** Help

**Study Conclusively Ties Rare Disease Gene to Parkinson's**  
October 21, 2009

**NHGRI Launches Improved Online Talking Glossary of Genetic Terms**  
October 20, 2009

**NIH Funds Four Centers of Excellence in Genomic Science**  
September 28, 2009

**Calendar** Help

**DIR Seminar Series**  
Rhoda Alani, M.D.  
November 19, 2009

**Evolutionary conserved pathways suppress**

# dbSNP at NCBI

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

The screenshot shows the NCBI dbSNP homepage. At the top, there's a navigation bar with links to PubMed, Nucleotide, Protein, Genome, Structure, PopSet, Taxonomy, OMIM, Books, and SNP. Below the navigation bar is a search bar with the placeholder "Search for SNP on NCBI Reference Assembly". To the right of the search bar is a small graphic of a DNA helix.

A yellow announcement box on the left says "BUILD 131" and "Have a question about dbSNP? Try searching the SNP FAQ Archive!". It also includes a "Go" button. To the right of the announcement is a yellow "ANNOUNCEMENT" box with the text "03/25/2010: RELEASE: NCBI dbSNP Build 131 for Human" and a link to "dbSNP Build 131:human\_9606 is now available".

The main content area has several sections:

- Search by IDs on All Assemblies:** A note says "Note: rs# and ss# must be prefixed with "rs" or "ss", respectively (i.e. rs25, ss25)". It includes a search form with fields for "ID:" and "Reference cluster ID(rs#)" with a dropdown menu, and buttons for "Search" and "Reset".
- Submission Information:** A list of links:
  - By Submitter
  - New Submitted Batches
  - Method
  - Population
  - Publication
- Batch:** A list of methods:
  - Enter List
    - NCBI Assay ID(ss)
    - Reference SNP ID(rs)
    - Local SNP ID
  - Upload List
    - NCBI Assay ID(ss)
    - Reference SNP ID(rs)
    - Local SNP ID
- Locus Information:** A link to "Batch Query Help".

# Human β-Hemoglobin Gene SNPs

[http://www.ncbi.nlm.nih.gov/SNP/snp\\_ref.cgi?locusId=3043](http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?locusId=3043)

NCBI Single Nucleotide Polymorphism

PubMed Nucleotide Protein Genome Structure PopSet Taxonomy OMIM Books SNP

Search for SNP on NCBI Reference Assembly

Search Entrez SNP for Go

BUILD 131  
Have a question about dbSNP? Try searching the SNP FAQ Archive!

GENERAL HUMAN VARIATION  
Search, Annotate, Submit [FAQ](#)  
Annotate and Submit Batch Data with Clinical Impact [FAQ](#)

SNP SUBMISSION DOCUMENTATION  
SEARCH RELATED SITES

SNP linked to Gene HBB(geneID:3043) Via Contig Annotation

Send rs# on all gene models to Batch Query Download all rs# to file. [Genotype](#) [VarView](#) [Help](#)

Gene Model (mRNA alignment) information from genome sequence

Total gene model (contig mRNA transcript): 1  
mRNA transcript protein mRNA orientation Contig Contig Label List SNP

NM\_000518.4 plus strand NP\_000509.1 forward NT\_009237.18 GRCh37 < currently shown

Include clinically associated  in gene region  cSNP  has frequency  double hit [refresh](#)

gene model	Contig Label	Contig	mRNA	protein	mRNA orientation	transcript	snp.count	
(contig mRNA transcript):	GRCh37	NT_009237.18	NM_000518.4	NP_000509.1	forward	plus strand	22, coding	
5246854 465	rs41605449	N.D.		Yes	frame shift		1	139
					frame shift	AGC [S]	1	139
					contig reference	GCTA [AN]	1	139
5246870 452	rs113062294	N.D.		Yes	synonymous	C [V]	3	134
					contig reference	G [V]	3	134
5246875 430	rs41511744	N.D.		Yes	frame shift		2	127
					frame shift	(15bp) [QAYYQ]	2	127
					contig reference	(17bp) [VQAYYQ]	2	127
5246883 439	rs111645880	N.D.		Yes	missense	T [V]	2	130
					contig reference	C [A]	2	130
5246968 365	rs71811954	N.D.			frame shift		3	105
					frame shift	(15bp)	3	105
					contig reference	G [R]	3	105
5247600 309	rs41539866	N.D.			frame shift		1	87
					frame shift	(17bp)	1	87
					contig reference	G	1	87
5247655 317	rs11549405	N.D.		Yes	synonymous	C [L]	3	89
					contig reference	G [L]	3	89

# $\beta$ -Hemoglobin Gene SNP rs111645889

[http://www.ncbi.nlm.nih.gov/SNP/snp\\_ref.cgi?rs=111645889](http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=111645889)

Reference SNP(refSNP) Cluster Report: rs111645889

RefSNP		Allele			HGVS Names		
Organism: human ( <i>Homo sapiens</i> )		Variation Class:	SNP: single nucleotide polymorphism				
Molecule Type: Genomic		RefSNP Alleles:	C/T				
Created/Updated in build: 132/132		Ancestral Allele:	C				
Map to Genome Build: 37.1		Clinical Association:	unknown				

SNP Details are organized in the following sections:

[GeneView](#) [Map](#) [Submission](#) [Fasta](#) [Resource](#) [Diversity](#) [Validation](#)

Integrated Maps (Hint: click on 'Chr Pos' or 'Contig Pos' column value to see variation in NCBI sequence viewer)

Genome Build	Chr	Chr Pos	Contig	Contig Pos	SNP to Chr	Contig allele	Contig to Chr	Group term	Group label
37.1	11	<a href="#">5365554</a>	<a href="#">NW_925006.1</a>	<a href="#">865878</a>	-	G	+	Celera	Celera
37.1	11	<a href="#">4906056</a>	<a href="#">NW_001838021.1</a>	<a href="#">875431</a>	-	G	+	HuRef	HuRef
37.1	11	<a href="#">5246883</a>	<a href="#">NT_009237.18</a>	<a href="#">5186883</a>	-	G	+	GRCh37	GRCh37

GeneView

GeneView via analysis of contig annotation: [HBB](#) hemoglobin, beta

View more variation on this gene (click to hide).

Include clinically associated:  in gene region  cSNP  has frequency  double hit [Go](#)

Assembly	SNP to Chr	Chr	Chr position	Contig	Contig position	Allele
GRCh37	-	11	5246883	<a href="#">NT_009237.18</a>	<a href="#">5186883</a>	G

RefSeqGene	Gene (ID)	SNP to RefSeqGene	Position	Allele
<a href="#">NG_000007.3</a>	<a href="#">HBB (3043)</a>	+	<a href="#">71963</a>	C

Function	mRNA				Protein		
	SNP to mRNA	Accession	Position	Allele change	Accession	Position	Residue change
missense	+	<a href="#">NM_000518.4</a>	<a href="#">439</a>	<a href="#">GCC → GTC</a>	<a href="#">NP_000509.1</a>	<a href="#">130</a>	<a href="#">A [Ala] → V [Val]</a>

# $\beta$ -Hemoglobin SNP Variation Viewer

<http://www.ncbi.nlm.nih.gov/sites/varvu?gene=3043>

HBB Variation Viewer

Gene	HBB; hemoglobin, beta	Gene Reference Sequences	<a href="#">NG_000007.3</a> genomic <a href="#">NM_000518.4</a> transcript <a href="#">NP_000509.1</a> protein variation locations are based on the
Description	beta globin chain   hemoglobin beta chain   hemoglobin subunit beta Also known as: CD113t-C, beta-globin		
Species	<a href="#">Homo sapiens</a>	Links	<a href="#">HGMD</a> , <a href="#">Panther</a> , <a href="#">Gene</a> , <a href="#">OMIM</a>
Cyto	11p15.5		

Observed Variation									Displaying results 1 - 20 of 51		
Var Cl...	Genomic	Transcript	Protein	Clinical interpretation	...	...	Freq	Pub...	MIM AI Var	rs id	
SNC	g.70459C>G	c.-136C>G		pathologic	6	6			<a href="#">141900.0375</a>	<a href="#">rs33994806</a>	
SNC	g.70517A>G	c.-78A>G		pathologic	6	6			<a href="#">141900.03...</a>	<a href="#">rs33931746</a>	
SNC	g.70545A>C	c.-50A>C			3	3			<a href="#">141900.0387</a>	<a href="#">rs34305195</a>	
DIP	g.70554delT	c.-41delT			2	2				<a href="#">rs35352549</a>	
SNC	g.70584C>T	c.-31C>T			1	1				<a href="#">rs63750628</a>	
SNC	g.70566G>A	c.-29G>A			3	3				<a href="#">rs34704828</a>	
SNC	g.70577C>G	c.-18C>G			2	2				<a href="#">rs34135787</a>	
DIP	g.70584_705...	c.-11_-8del4			2	2				<a href="#">rs34196559</a>	
SNC	g.70595A>G	c.1A>G	p.Met1Val		2	2				<a href="#">rs34563000</a>	
SNC	g.70596T>C   ...	c.2T>C   c.2T>G	p.Met1Thr   p....		5	5			<a href="#">141900.034...</a>	<a href="#">rs33941849</a>	
SNC	g.70597G>A   ...	c.3G>A   c.3G...	p.Met1Ile   p....		6	6			<a href="#">141900.043</a>	<a href="#">rs33930702</a>	
DIP	g.70598delG	c.4delG	p.Val2fx		1	1			<a href="#">141900.0419</a>	<a href="#">rs63750475</a>	
SNC	g.70598G>A   ...	c.4G>A   c.4G>T	p.Val2Met   p....		4	4			<a href="#">141900.026...</a>	<a href="#">rs33958358</a>	
SNC	g.70602A>C   ...	c.8A>C   c.8A...	p.His3Pro   p....		8	8			<a href="#">141900.006...</a>	<a href="#">rs33983205</a>	
DIP	g.70602_706...	c.8_9insC			1	1			<a href="#">141900.0503</a>	<a href="#">rs63750898</a>	
DIP	g.70603_706...	c.9_10insT			1	1				<a href="#">rs34058656</a>	
SNC	g.70603T>G   ...	c.9T>G   c.9T...	p.His3Gln   p....		33	33	Q		<a href="#">141900.0206</a>	<a href="#">rs713040</a>	
SNC	g.70604C>G	c.10C>G	p.Leu4Val		2	2				<a href="#">rs34126315</a>	
SNC	g.70605T>A	c.11T>A	p.Leu4Gln		1	1				<a href="#">rs63750720</a>	
SNC	g.70608C>A	c.14C>A	p.Thr5Asn		1	1				<a href="#">rs63750605</a>	

# Mutasyonların evrimsel önemi

- Eğer mutasyon olmasaydı evrim olmazdı
- ancak,
  - çok fazla mutasyon zararlı
  - çok az mutasyon ise etkisi az
  - ortaya çıkan varyasyon ile uyum arasında bir denge olmalı

Yakın akraba canlılar arasında yüksek oranda eşleşen DNA dizilerinin varlığı geçmişteki ayrılma noktasını ve yeni çevreye uyum sırasındaki adaptasyonu gösterir