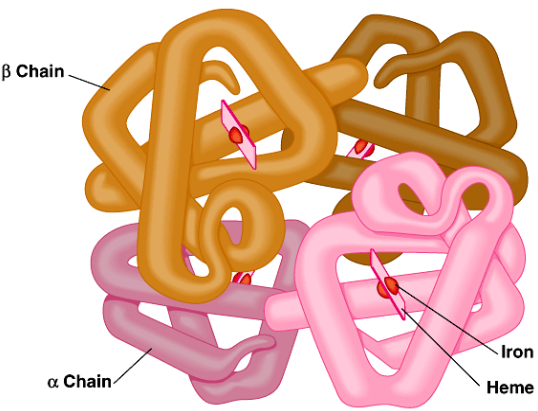
Using Bioinformatics in Medicine

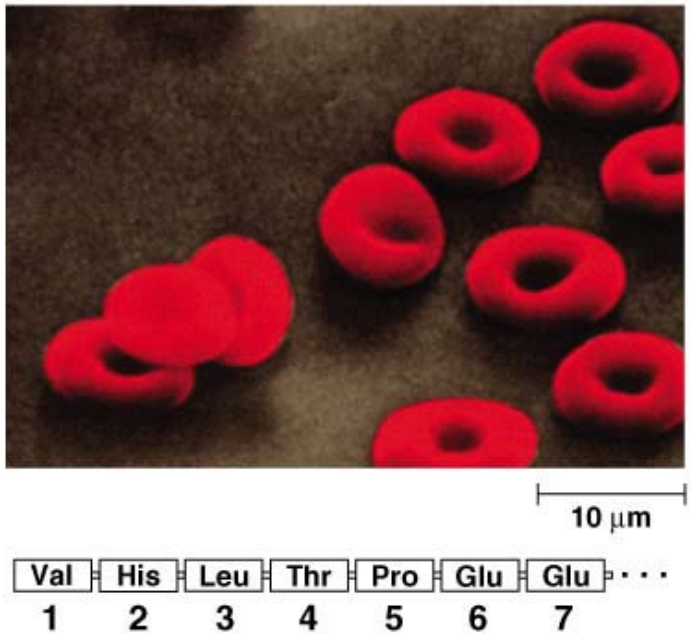
Sickle Cell Anemia & the Hemoglobin Gene



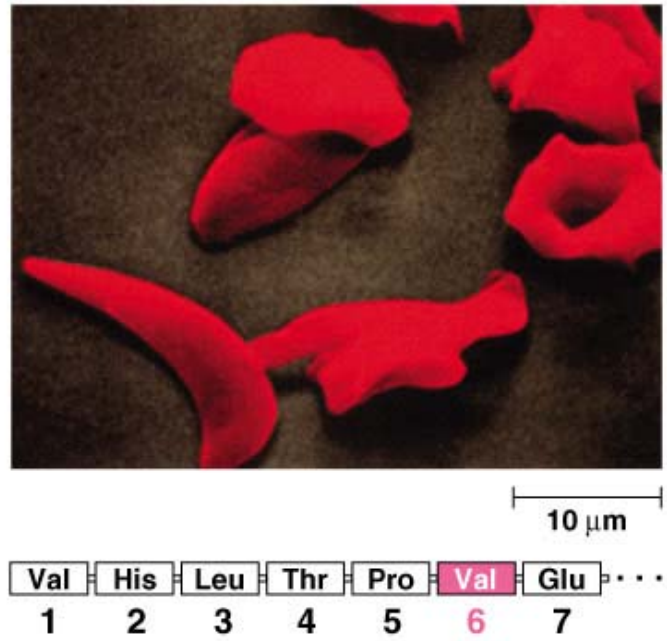
2004-2005

Sickle Cell Anemia

- **Most common genetic disease in US**
  - high incidence in African-Americans
  - affects red blood cells
  - potentially lethal



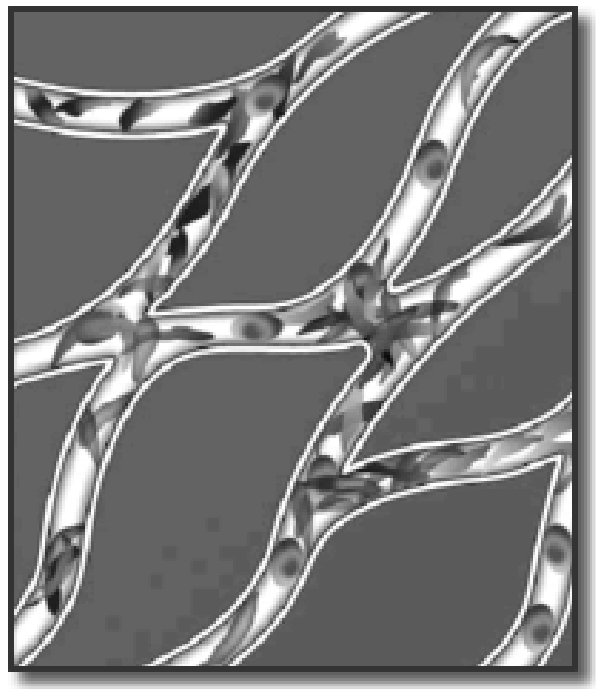
(a) Normal red blood cells and the primary structure of normal hemoglobin



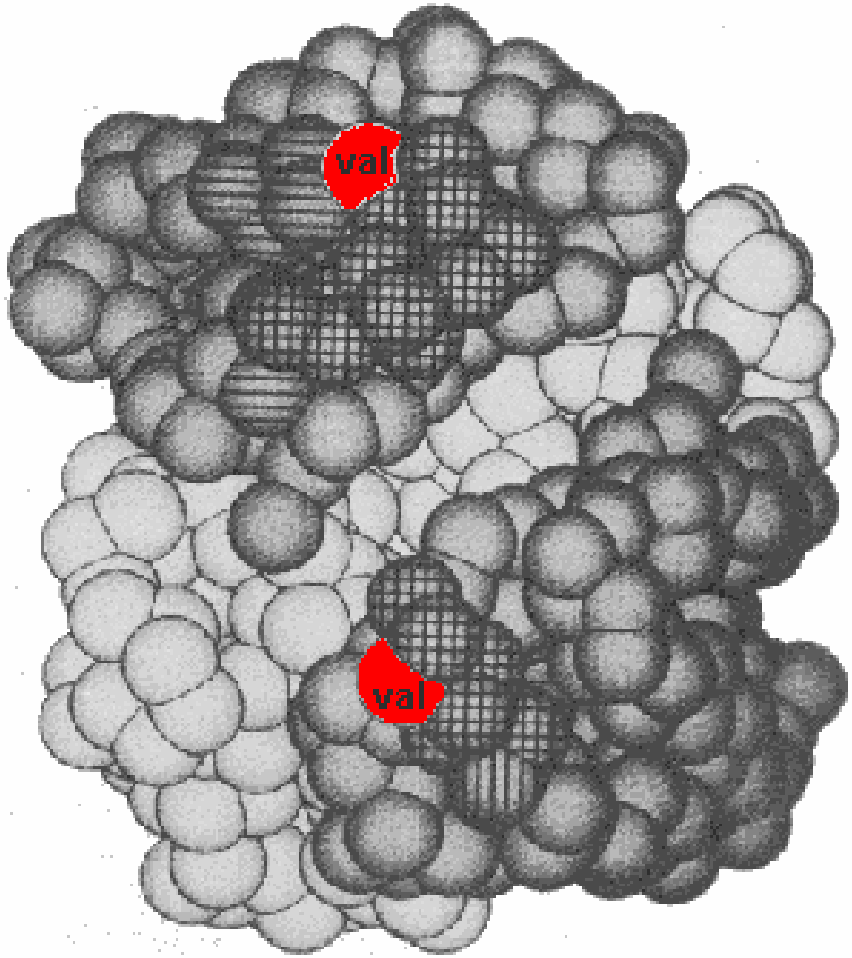
(b) Sickled red blood cells and the primary structure of sickle-cell hemoglobin

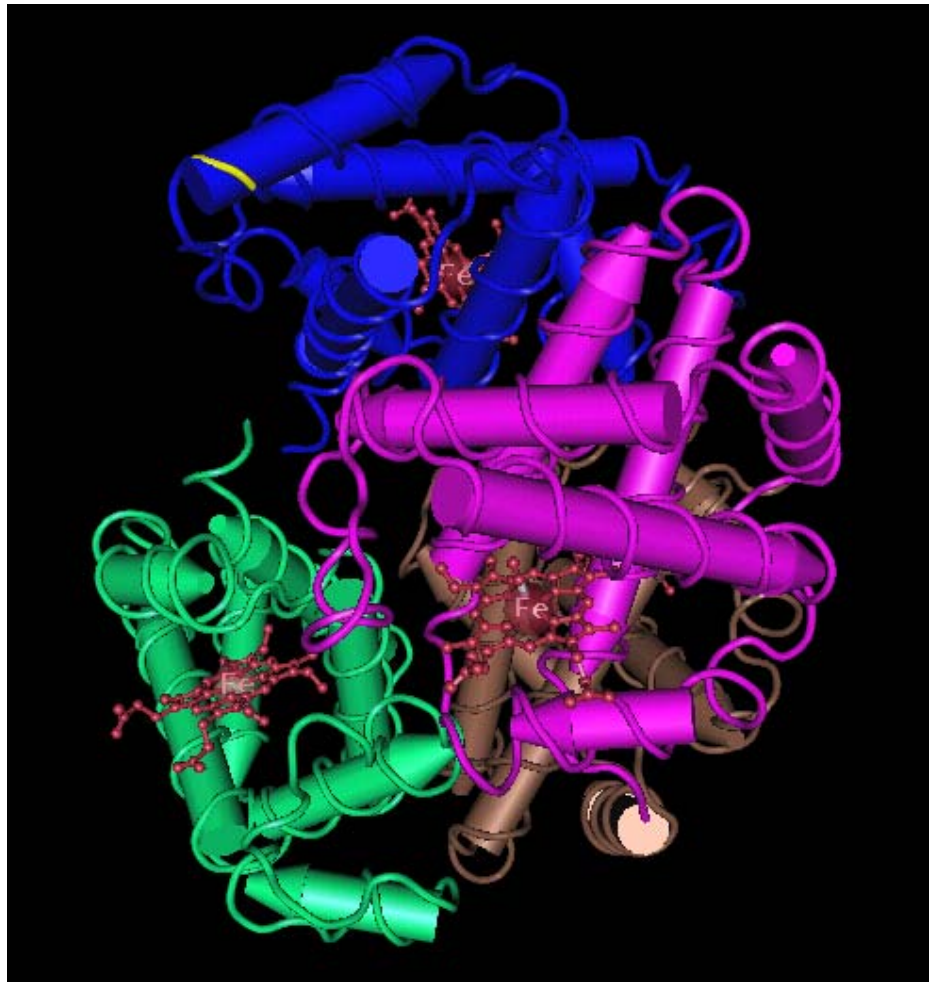
Symptoms

- Anemia
  - jaundice, fatigue, paleness, shortness of breath
- Hypoxia (low oxygen) & capillary damage
  - severe pain in organs & joints
  - retinal damage (blindness)
- Delayed growth
  - delayed puberty, stunted growth
- Infections
  - more susceptible
  - depressed immune
  - death from bacterial infections
- Stroke
  - blocked small blood vessels in brain
  - primarily in children

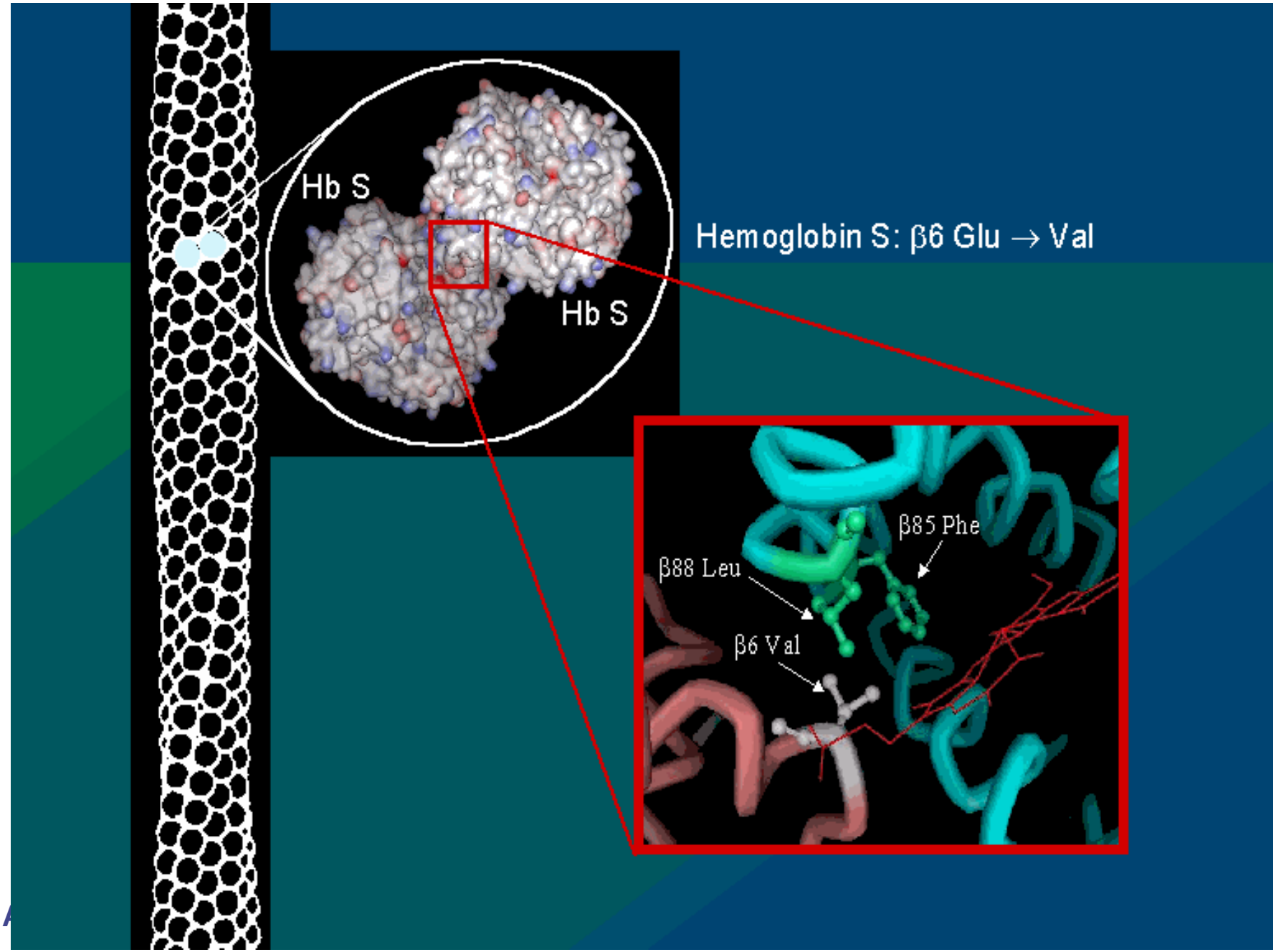


Sickle cell hemoglobin





mutant hemoglobin (Hb S)

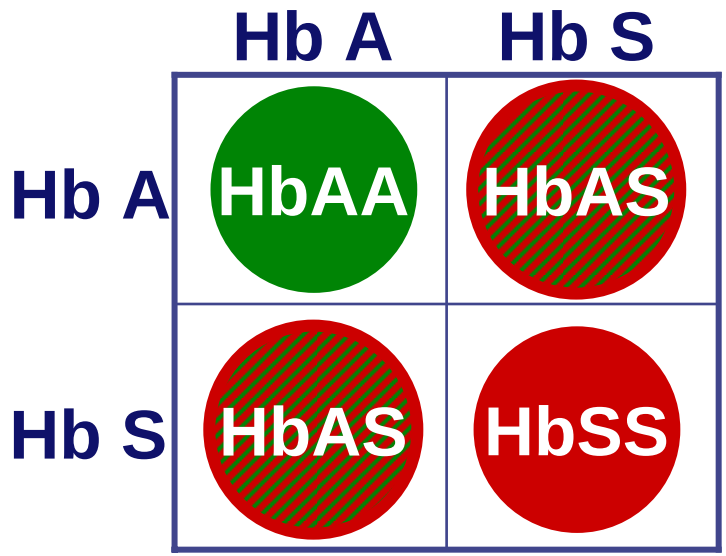

Hb S
Hb S
Hemoglobin S: β6 Glu → Val
β85 Phe
β88 Leu
β6 Val

Cell biology

- Hb S molecules stick together
  - form fibers
  - under low blood oxygen levels
  - distortion of cells from normal round to sickle shape

Genetics

- **Sickle cell mutation**
  - Hb S
  - changes 6th amino acid of β hemoglobin chain
  - normal glutamic acid → valine
- **Recessive allele**
  - heterozygote
    - Hb AS, normal, but carrier
  - homozygote recessive
    - Hb SS, sickle cell disease
  - 2 sickle cell carriers mate…
    - each child has 1/4 chance of having the disease



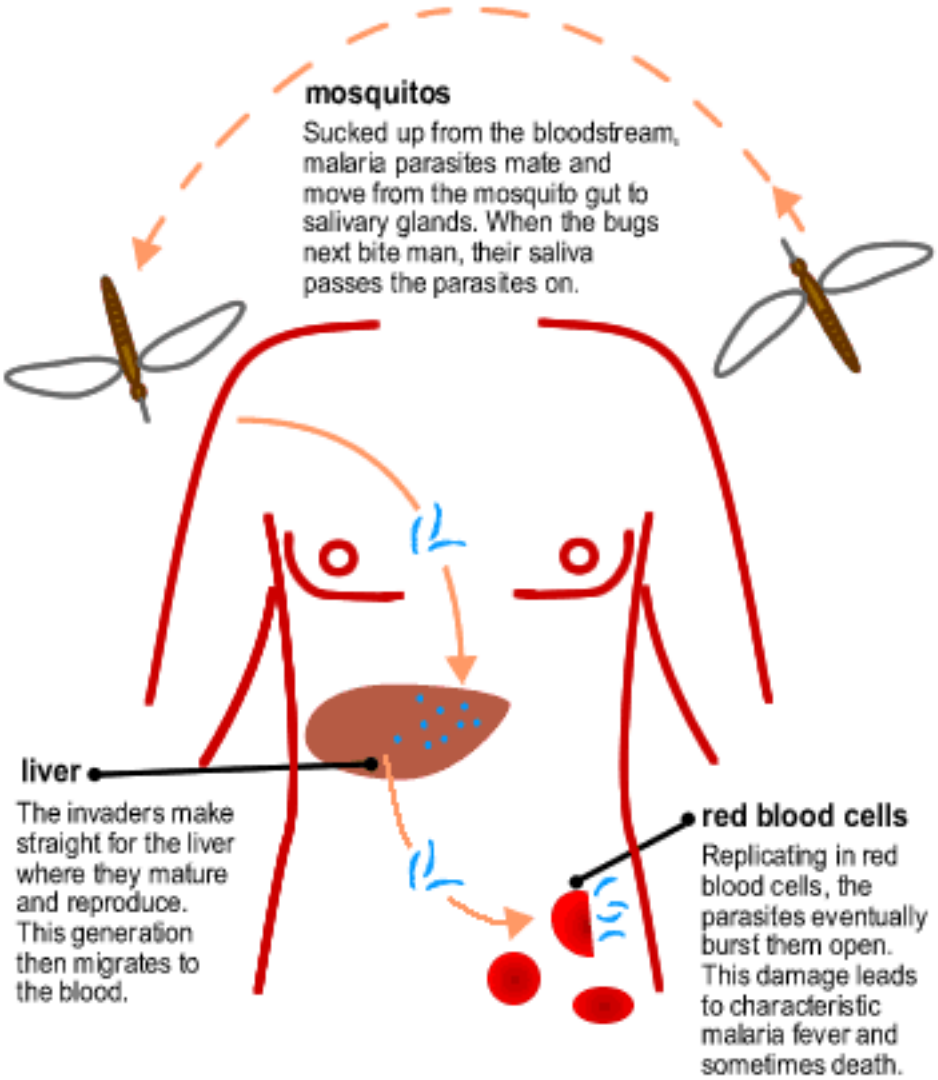
Prevalence in U.S.

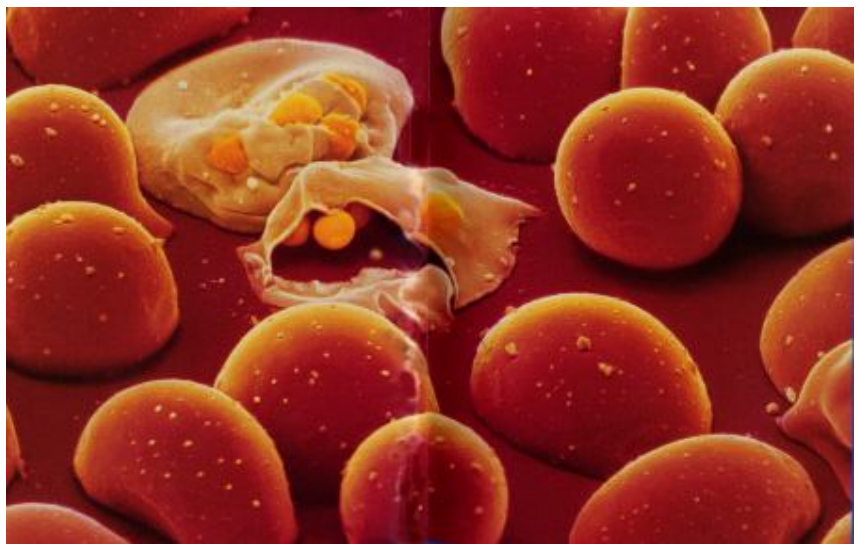
- Carriers
  - ~2 million Americans carry sickle cell trait
  - 1 in 14 African-Americans
- Disease
  - ~72,000 Americans have disease
  - ~1 in every 700 African-American babies born in U.S. has sickle cell disease

The Malaria Connection

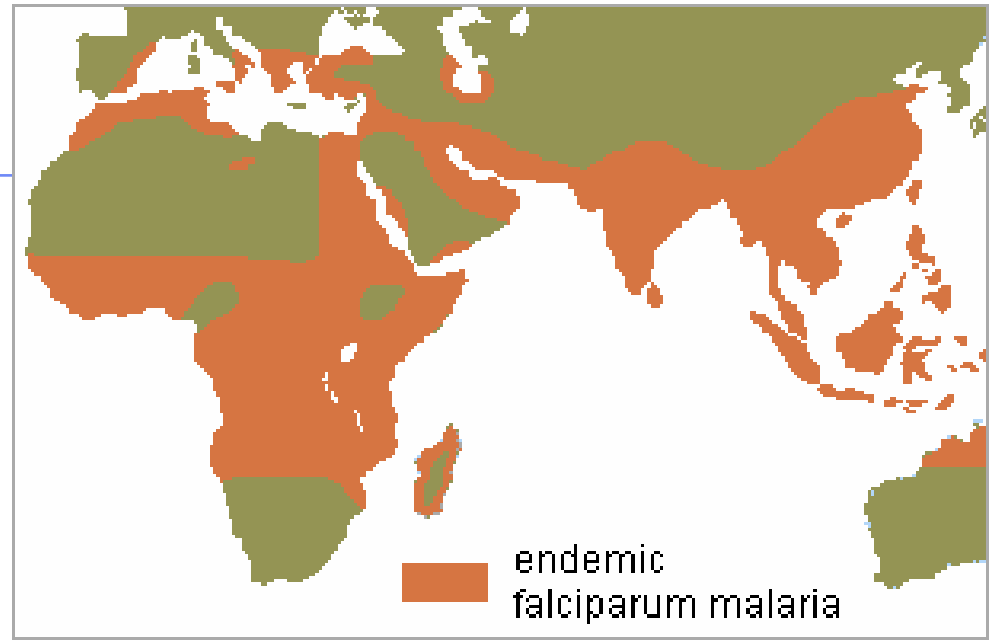
- Sickle cell disease is surprisingly common for a potentially lethal genetic disease
- Heterozygote advantage
  - heterozygotes are tolerant of malaria infection & do not suffer symptoms of sickle cell disease

Malaria

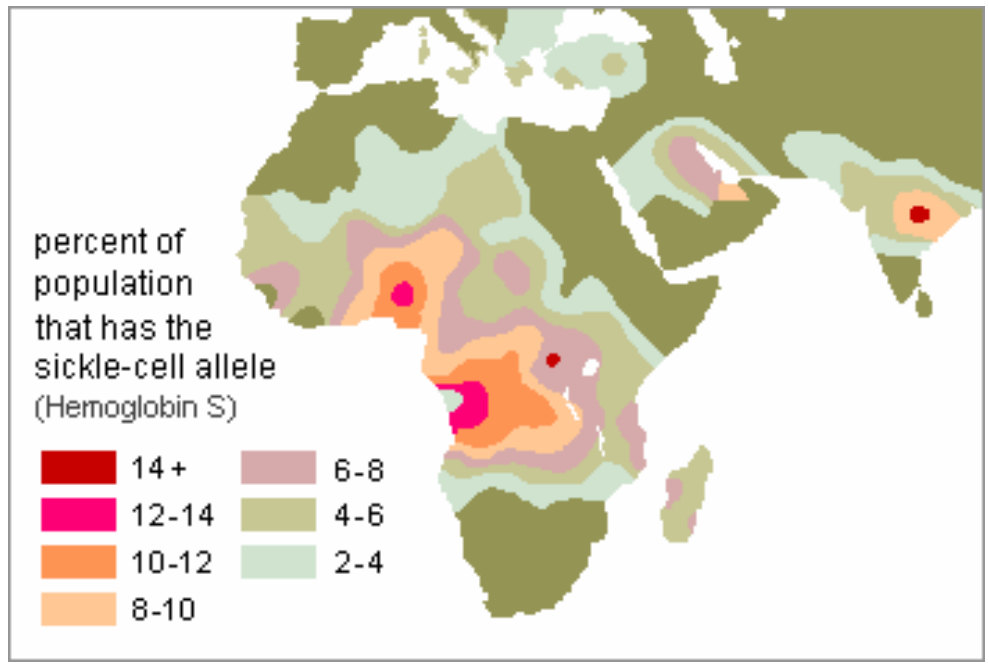

mosquitos
Sucked up from the bloodstream, malaria parasites mate and move from the mosquito gut to salivary glands. When the bugs next bite man, their saliva passes the parasites on.
liver
The invaders make straight for the liver where they mature and reproduce. This generation then migrates to the blood.
red blood cells
Replicating in red blood cells, the parasites eventually burst them open. This damage leads to characteristic malaria fever and sometimes death.



Prevalence of Malaria



Prevalence of Sickle Cell Anemia



[~sickle cell movie~]

Public health

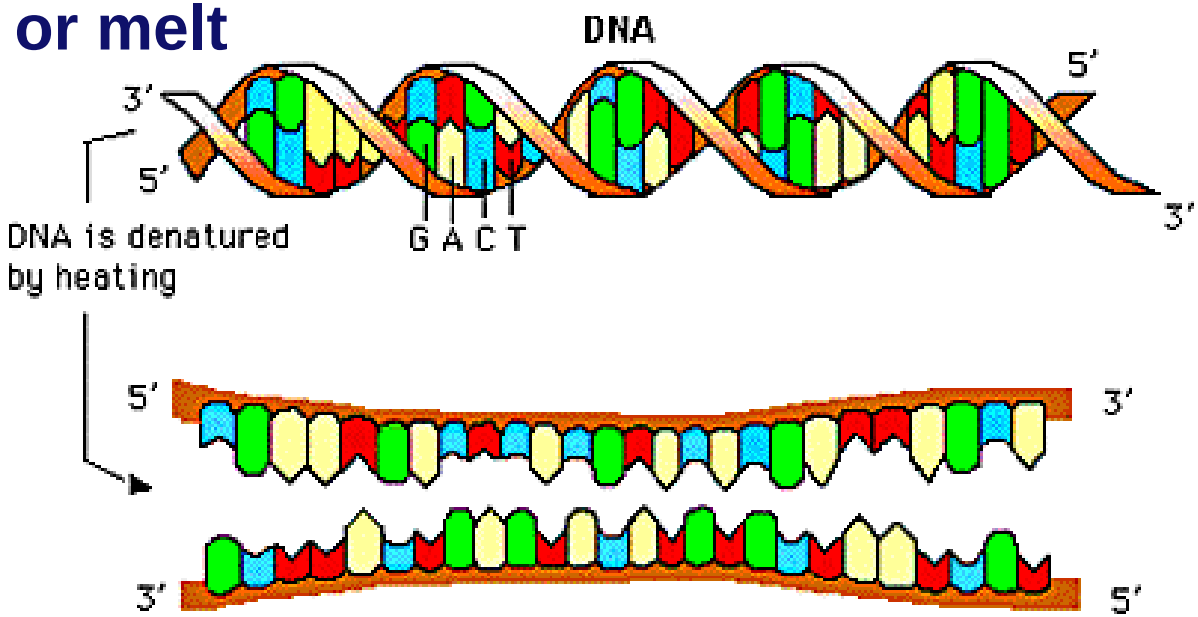
- Many carriers of this mutant allele are not aware that they have it
  - at risk of having children with the disease
- DNA test for sickle cell allele would benefit public health
  - genetic counseling
  - pre-natal testing

Your Assignment

- Develop a simple inexpensive DNA test for sickle cell allele
  - develop DNA probe
    - test for presence of sickle cell mutation
  - use bioinformatics tools
    - online databases of DNA sequences
      - UCSC Genome Browser
    - probe design tool
      - Primer3

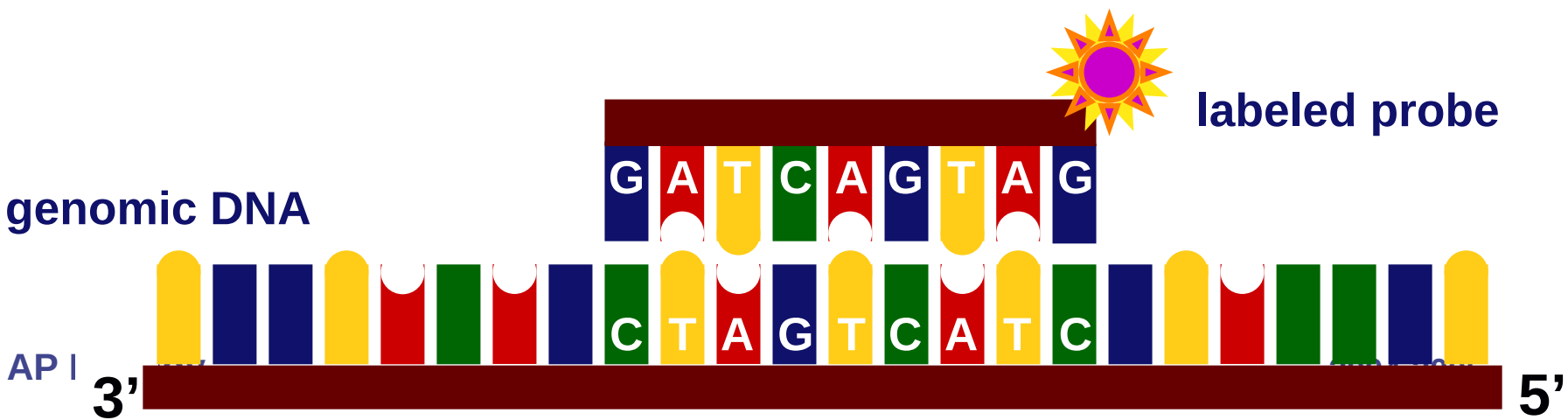
DNA review

- DNA double helix
  - A–T, C–G
  - base pair bonds can be broken by heating to 100°C
    - separate strands
    - denature, or melt



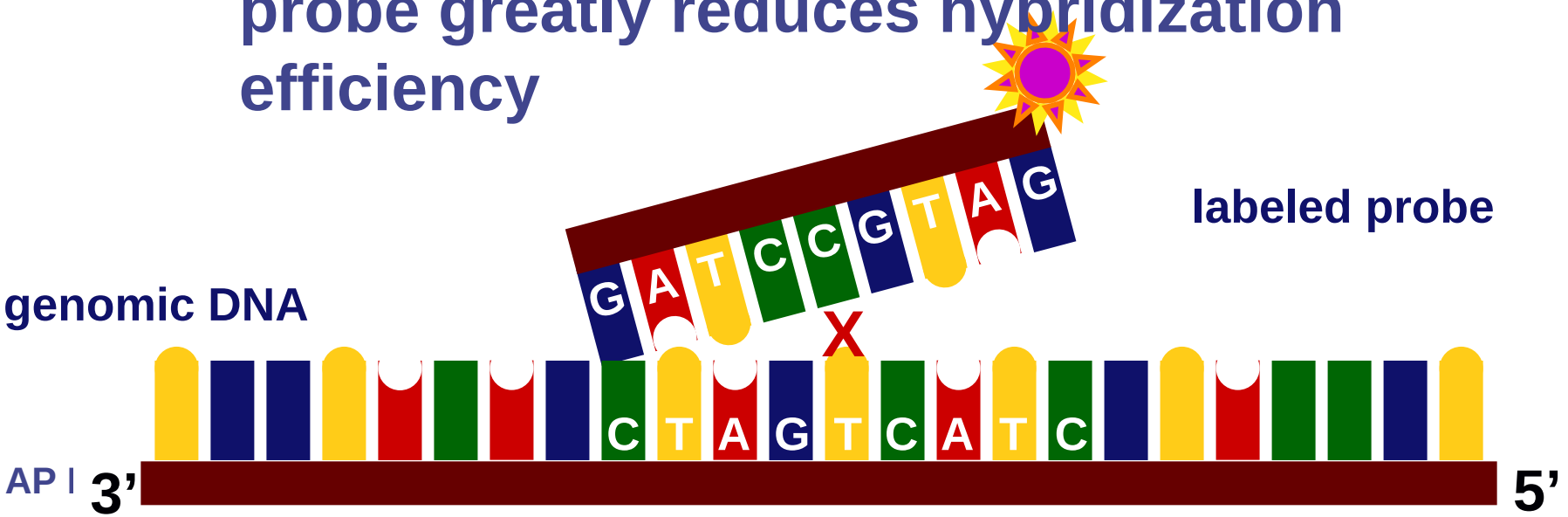
DNA probes

- **Probe**
  - short, single stranded DNA molecule
  - mix with denatured DNA
- **DNA Hybridization**
  - probe bonds to complementary DNA sequence
- **Label**
  - probe is labeled for easy detection



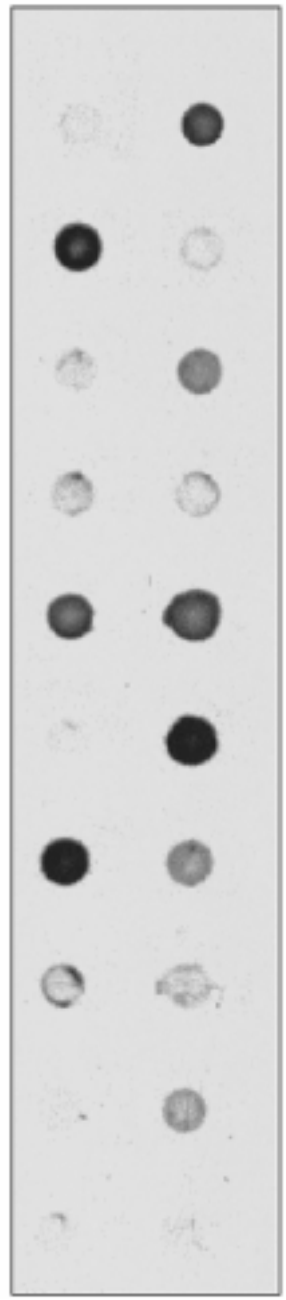
Designing Probes

- **Allele specific probes**
  - probes require matched sequences
  - can detect single base differences in alleles
  - single mis-matched base near middle of probe greatly reduces hybridization efficiency



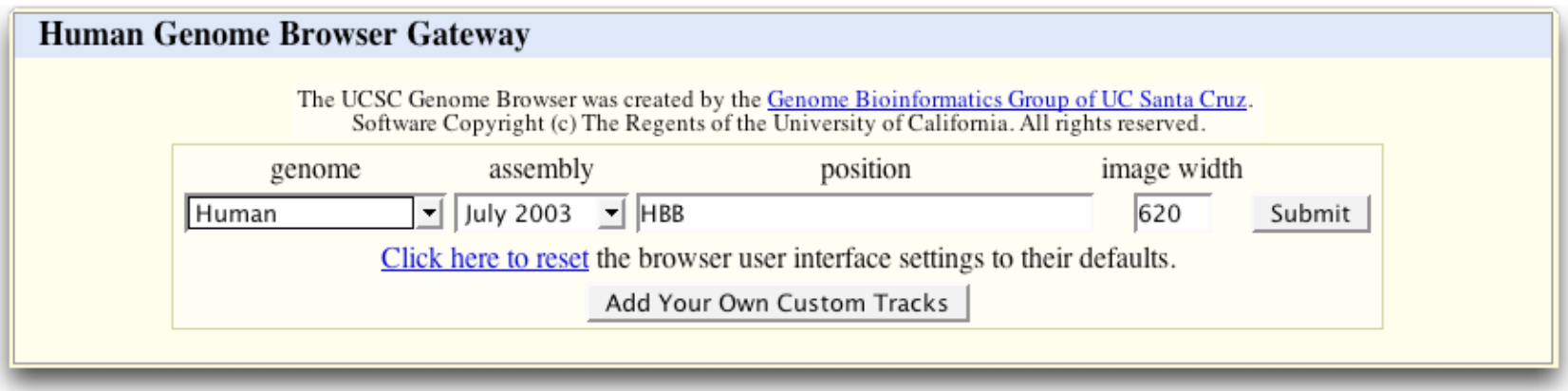
Dot blot

- Genomic DNA
  - denature DNA
  - bind DNA from cells on filter paper
- DNA hybridization
  - wash probe over filter paper
  - if complementary sequence present, probe binds to genomic DNA
  - expose on X-ray film
    - dark spots show bound probe



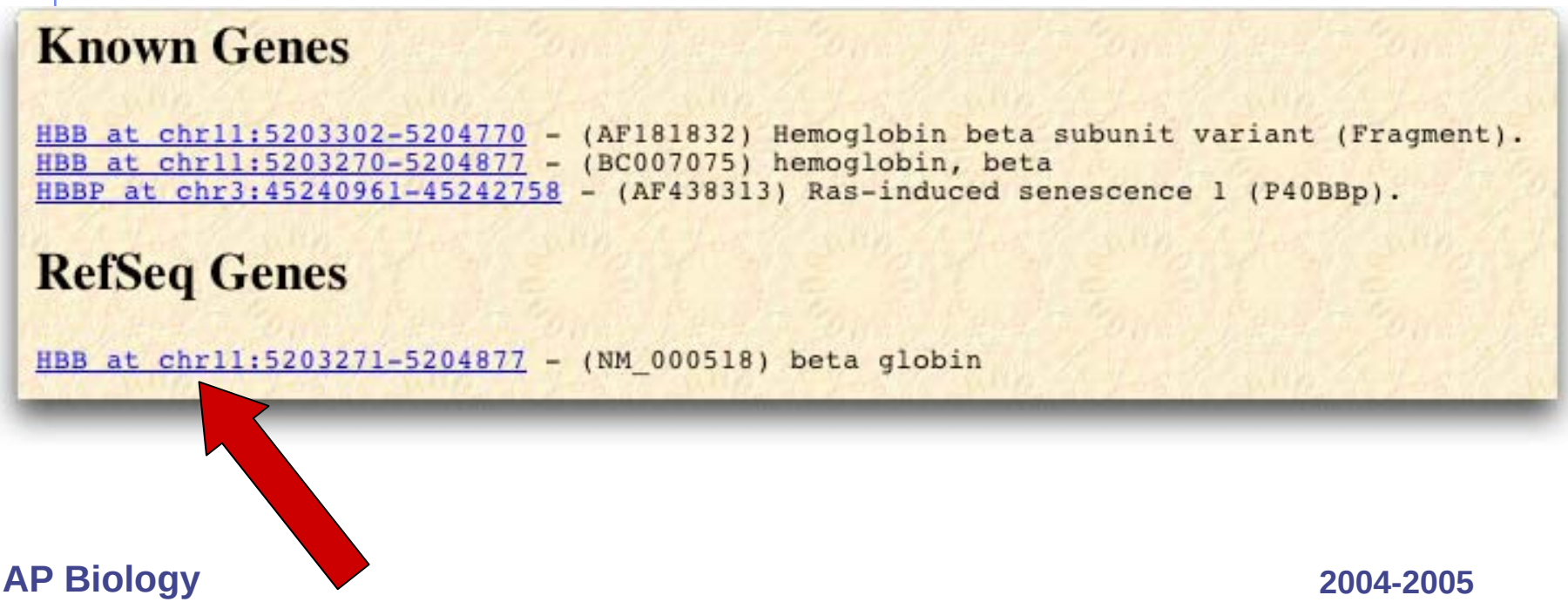
Get hemoglobin sequence

- UCSC Genome Browser
  - human genome database
  - http://genome.ucsc.edu/
    - UCSC Genome Browser home page
    - click on link to Genome Browser
    - in genome pulldown menu, choose "Human"
    - for position text box, type "HBB" (hemoglobin β)
    - hit "submit"

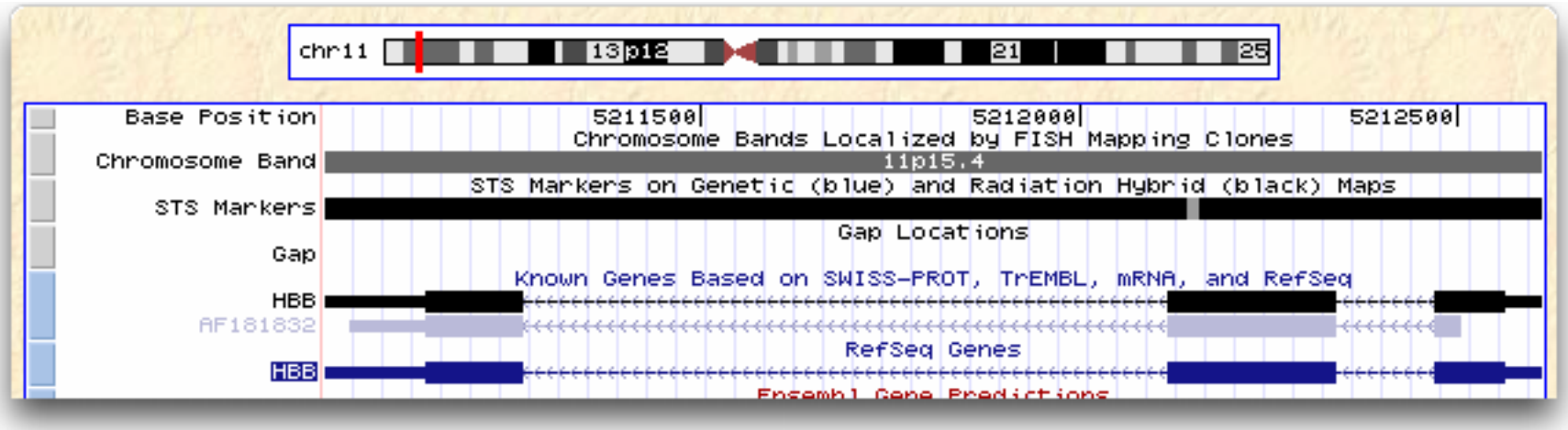


Genome Browser Results

- Listing of genes & sequences in database
  - Click on “RefSeq” gene for HBB (NM_000518)



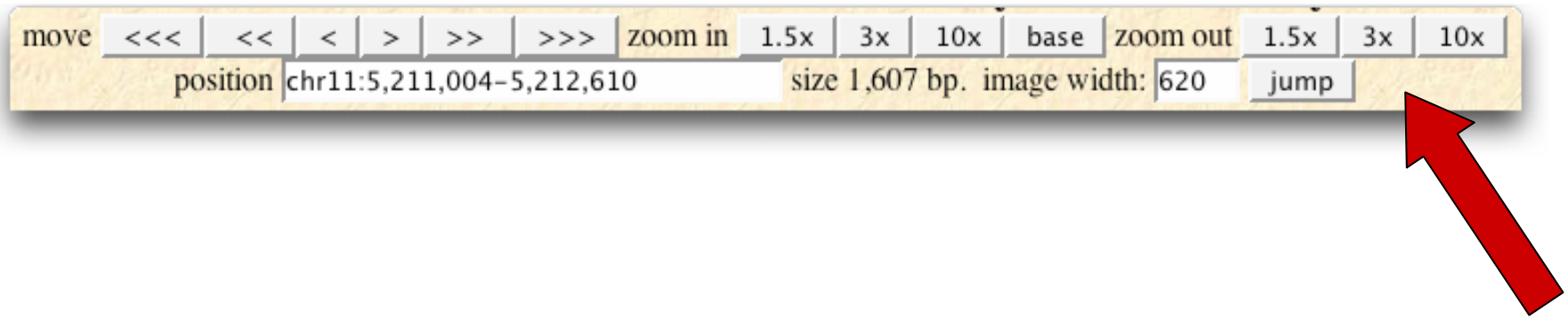
Chromosome view



- **Position of HBB in genome**
  - **at base 5.2 million on chromosome 11**

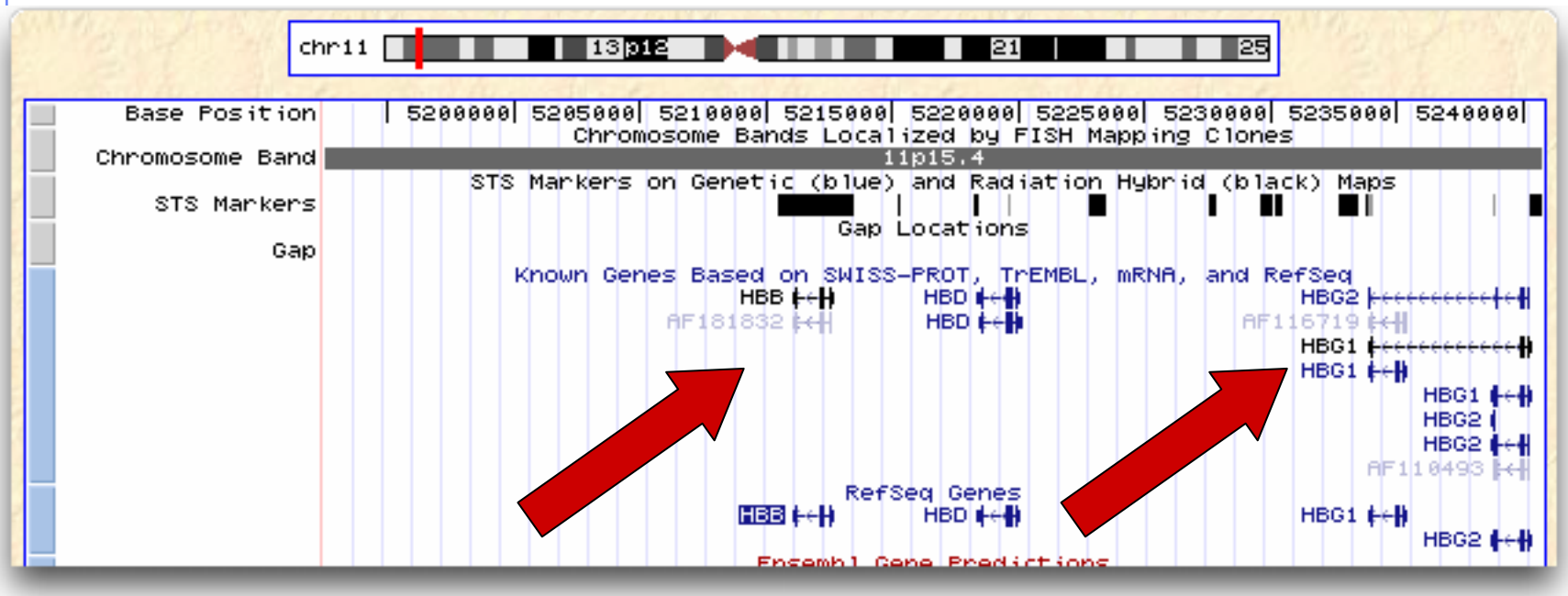
Change view of chromosome

- **Move & zoom tools**
  - **zoom out ~30x to see more of chromosome 11**



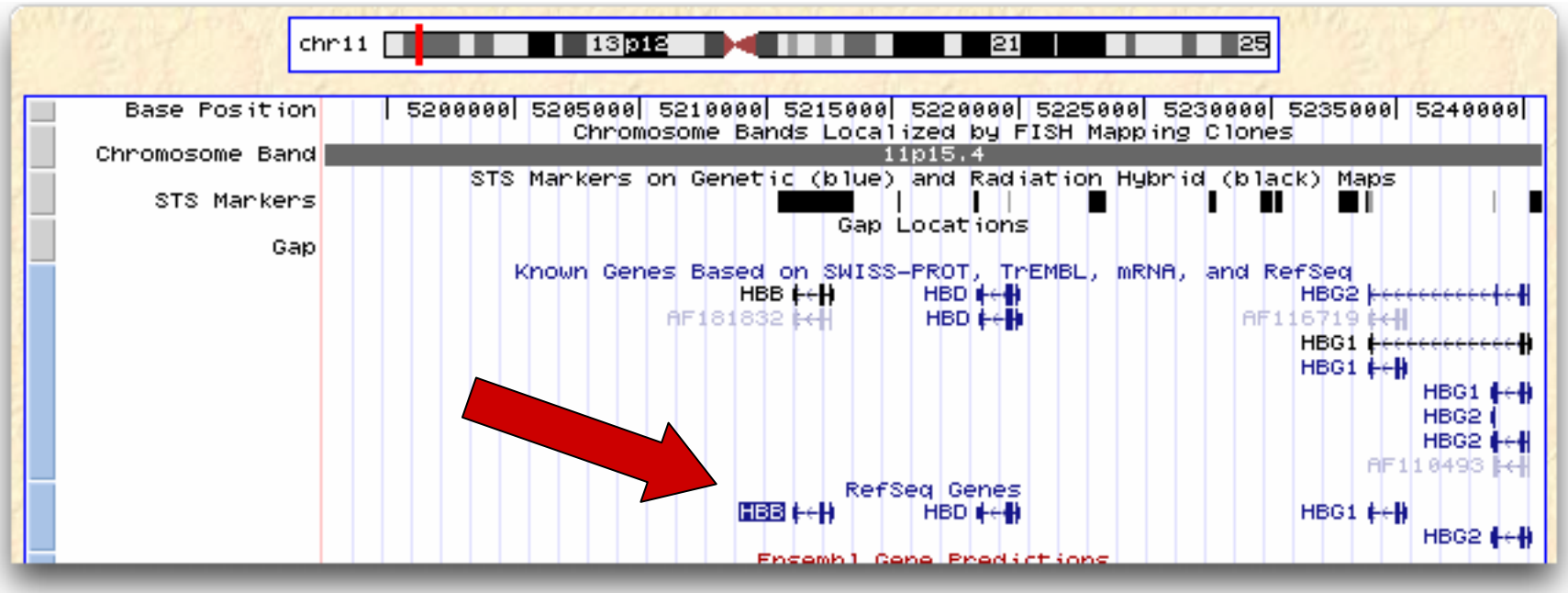
More Hb genes

- Cluster of hemoglobin genes on chromosome 11
  - HBD, HBG1, HBG2 & HBE1
  - what are these genes?



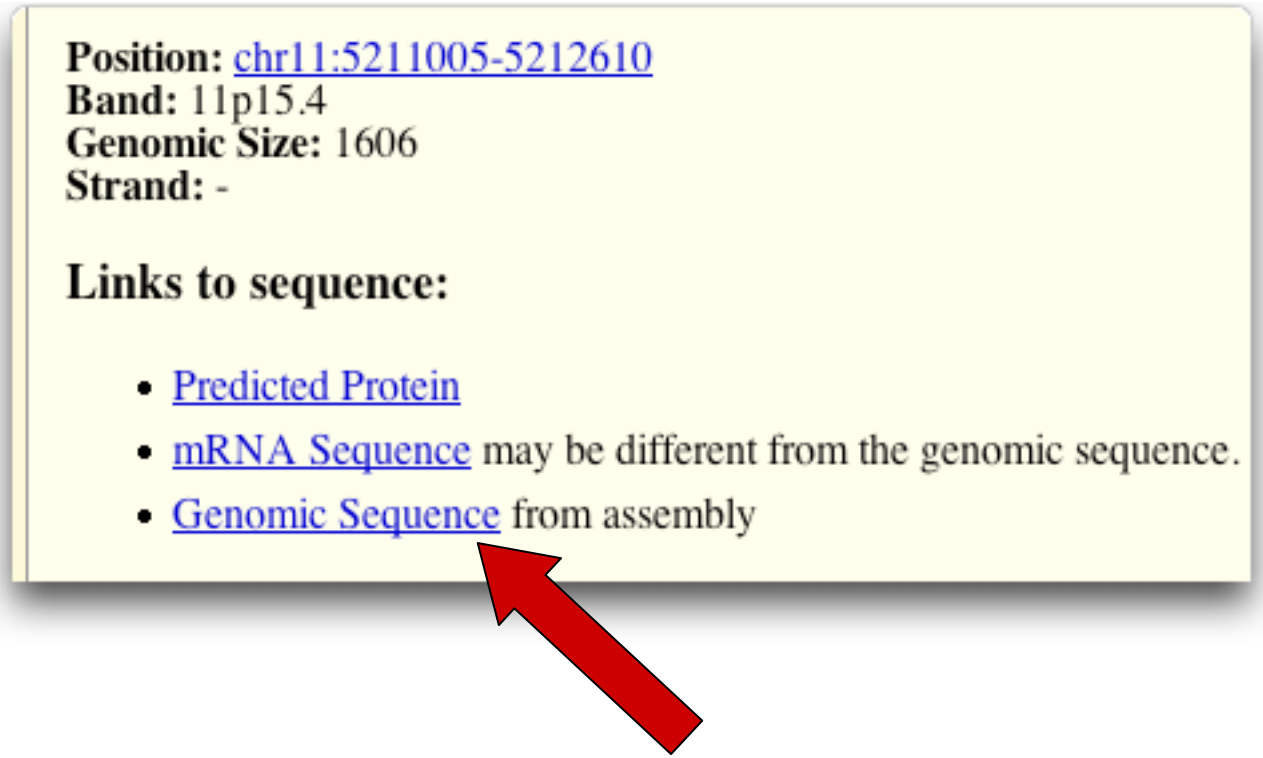
A

Get the DNA sequence



- **Click on the HBB RefSeq gene**
  - **HBB RefSeq summary page**

HBB RefSeq gene summary page



- **Click on “Genomic Sequence from assembly”**

Formatting the sequence

- Sequence Formatting Options
  - “exons in upper case, everything else in lower case”
  - hit “submit”
- Genomic DNA
  - lower case = introns
    - spliced out of mRNA before translation
  - upper case = exons
    - translated into polypeptide chain

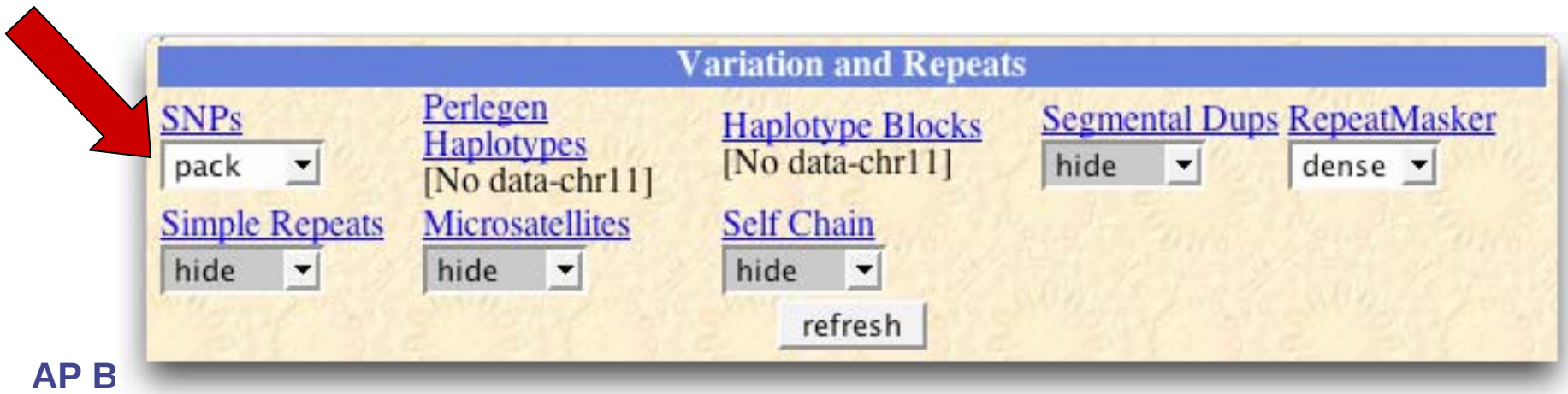
HBB DNA sequence

>hg16_refGene_NM_000518 range=chr11:5211005-5212610 5'pad=0 3'pad=0 revComp=TRUE
ACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACC
ATGGTGCATCTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGG
CAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCAGgttggtat
caaggttacaagacaggtttaaggagaccaatagaaactgggcatgtgga
gacagagaagactcttgggtttctgataggcactgactctctctgcctat
tggtctattttcccacccttagGCTGCTGGTGGTCTACCCTTGGACCCAG
AGGTTCTTTGAGTCCTTTGGGGATCTGTCCACTCCTGATGCTGTTATGGG
CAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGGTGCCTTTAGTG
ATGGCCTGGCTCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGT
GAGCTGCACTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGGgtgag
tctatgggacgcttgatgttttctttccccttcttttctatggttaagtt
catgtcataggaaggggataagtaacagggtacagtttagaatgggaaac
agacgaatgattgcatcagtgtggaagtctcaggatcgttttagtttctt
ttatttgctgttcataacaattgttttcttttgtttaattcttgctttct
ttttttttcttctccgcaatttttactattatacttaatgccttaacatt
gtgtataacaaaaggaaatatctctgagatacattaagtaacttaaaaaa
aaactttacacagtctgcctagtacattactatttggaatatatgtgtgc
ttatttgcatattcataatctccctactttatttcttttatttttaatt
gatacataatcattatacatatttatgggttaaagtgtaatgttttaata
tgtgtacacatattgaccaaatcagggtaattttgcatttgtaatttaa
aaaatgctttcttctttaatatactttttgtttatcttatttctaata
ctttccctaatctctttctttcagggcaataatgatacaatgtatcatgc
ctctttgcaccattctaaagaataacagtgataatttctgggttaaggca
atagcaatatctctgcatataaatatttctgcatataaattgtaactgat

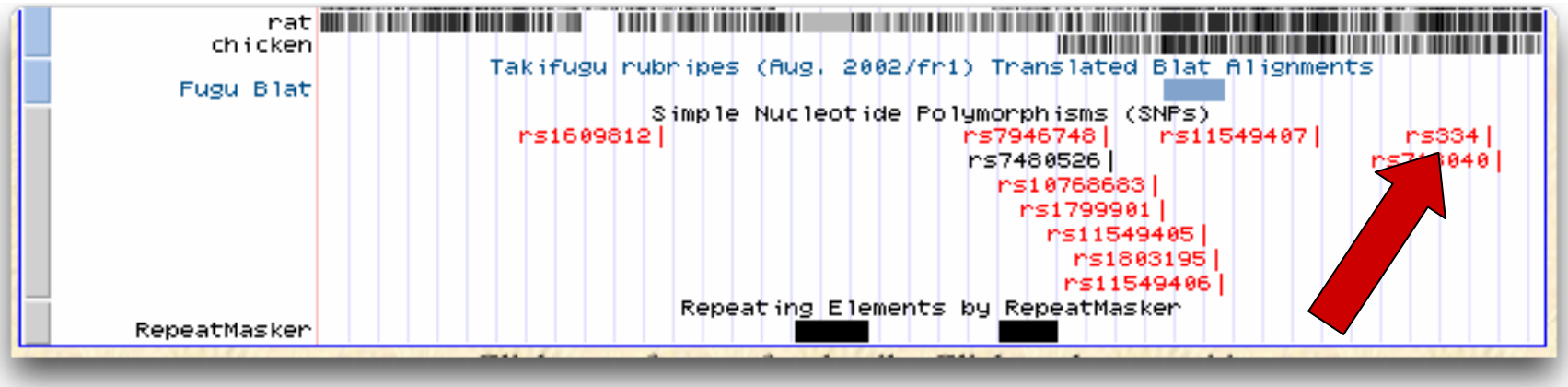
- first 50 bases are untranslated “leader” sequence
- actual protein coding sequence starts at base 51
  - starting with letters ATG

Get the mutant sequence

- **Sickle cell mutation**
  - single base mutation
  - 6th amino acid: **glutamic acid → valine**
  - need DNA sequence to design probe
- **SNPs**
  - single nucleotide polymorphisms
  - “variations and repeats” section: **pack**



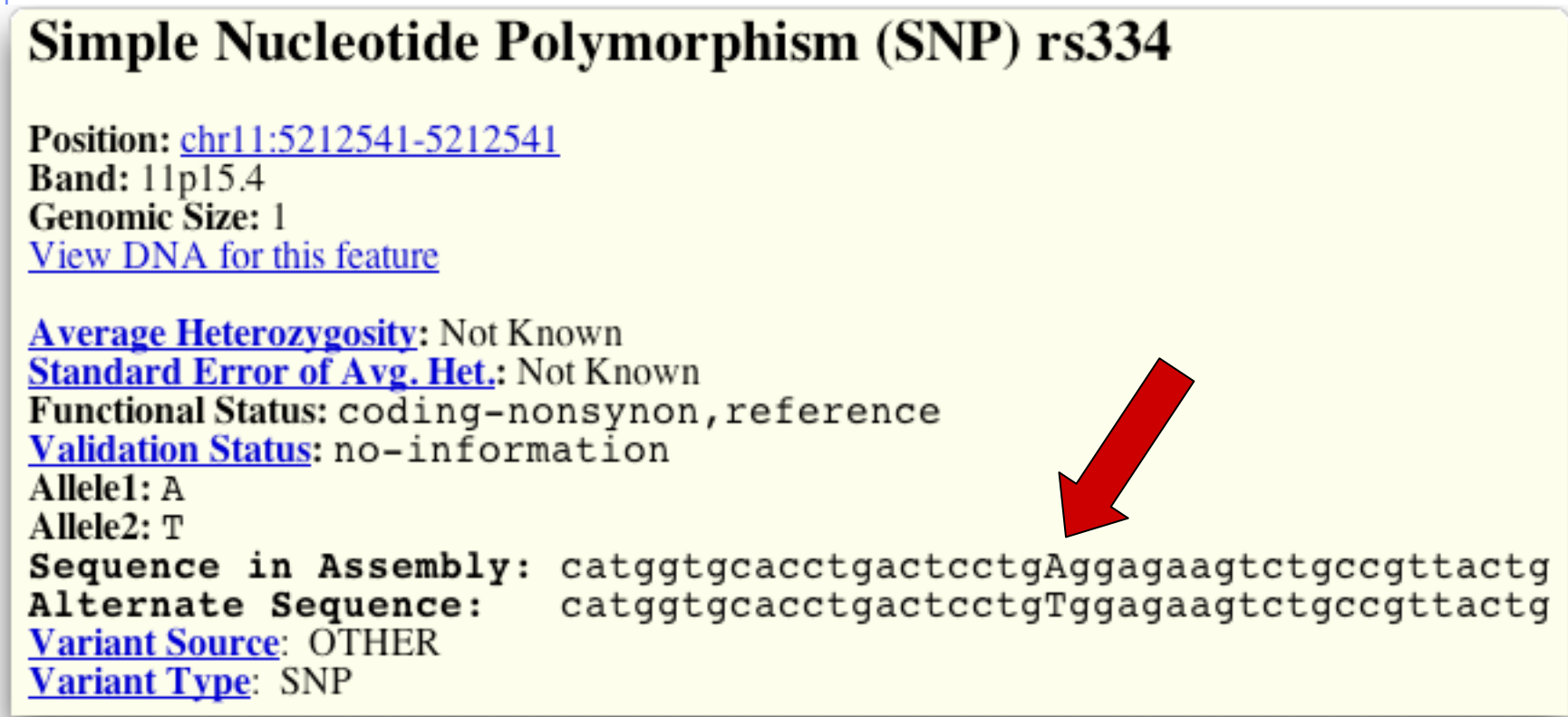
SNPs of HBB gene



- several SNPs of HBB gene
  - need mutation in exon
  - near beginning of HBB protein
  - rs334 = Hb S mutation

rs334 Hb S sickle cell mutation

- "Sequence in Assembly" = normal sequence
- "Alternate Sequence" = sickle cell sequence



Simple Nucleotide Polymorphism (SNP) rs334

Position: chr11:5212541-5212541
Band: 11p15.4
Genomic Size: 1
View DNA for this feature

Average Heterozygosity: Not Known
Standard Error of Avg. Het.: Not Known
Functional Status: coding-nonsynon,reference
Validation Status: no-information
Allele1: A
Allele2: T
Sequence in Assembly: catggtgcacctgactcctgAggagaagtctgccgttactg
Alternate Sequence: catggtgcacctgactcctgTggagaagtctgccgttactg
Variant Source: OTHER
Variant Type: SNP

A

Align Hb A & Hb S sequences

- Line up sequences

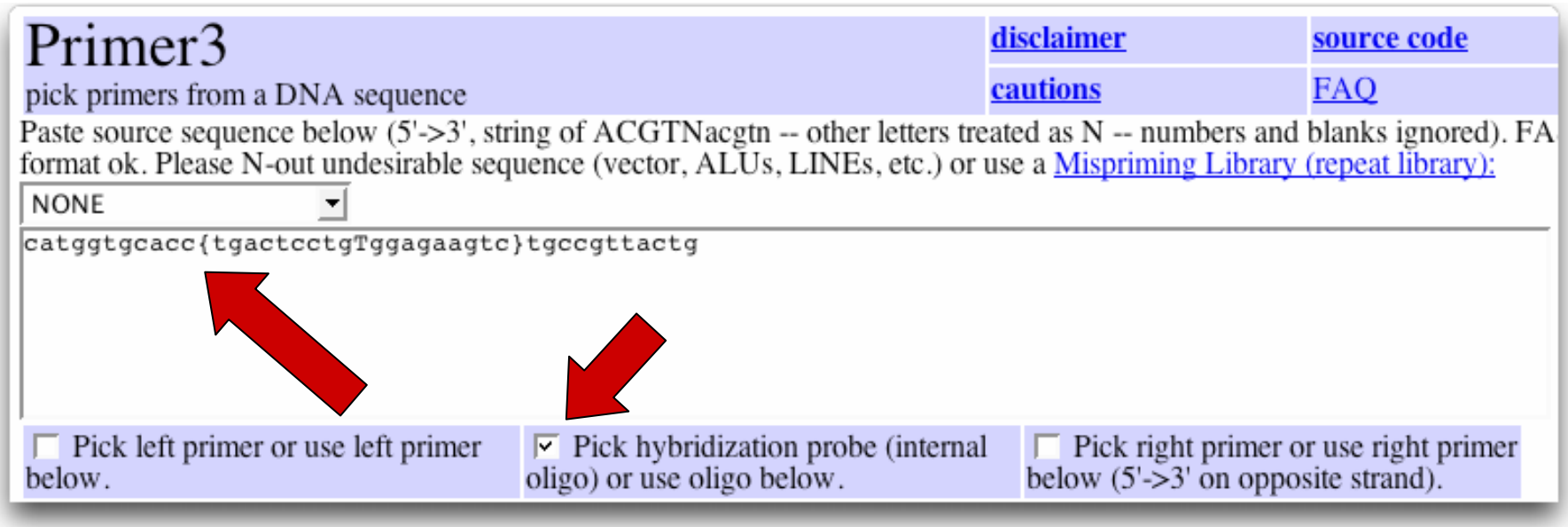
```
Normal:catggtgcacctgactcctgAggagaagtctgccgttactg
  HBB:
ATGGTGCATCTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGG
Mutant:    catggtgcacctgactcctgTggagaagtctgccgttactg
```

  - sequence fragment is enough to design DNA probes for normal & mutant sequences

Designing the probe

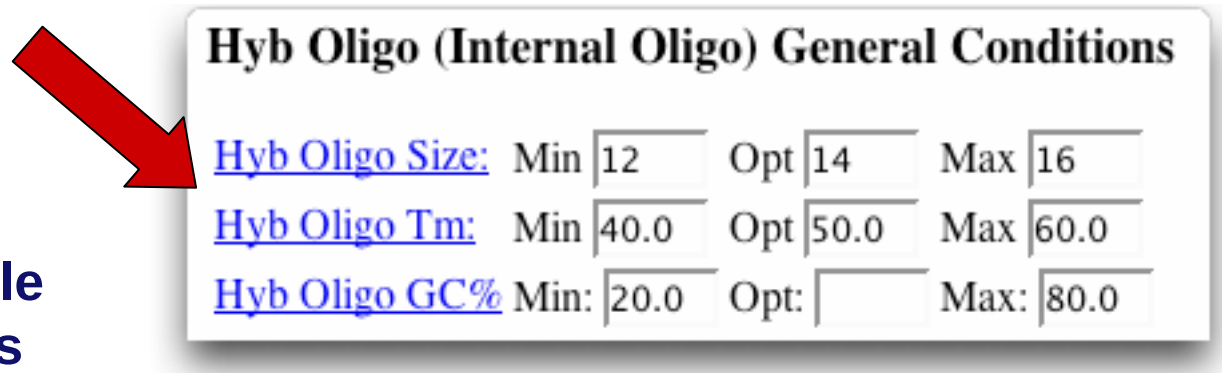
- **Primer3**
  - **free on Web from MIT**

    http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi
  - **powerful tool for primer design**
    - **paste in sequence fragment**



Allele specific probes

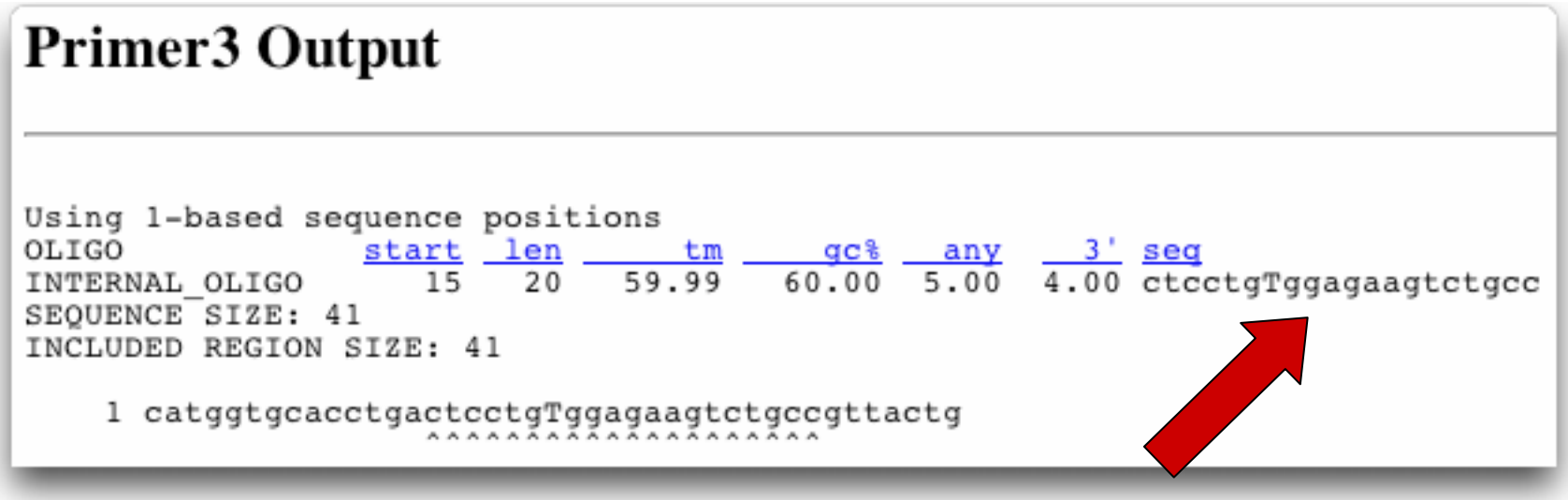
- Need 2 probes
  - normal allele probe
  - sickle cell allele probe
  - choose hybridization probes
- Customize probes
  - 12-16 bases
  - 40°-60°C



longer probes are stable at higher temperatures

Your probes…

- Ready to order!



Primer3 Output

```
Using 1-based sequence positions
OLIGO              start  len      tm     gc%   any   3'  seq
INTERNAL_OLIGO        15   20   59.99   60.00  5.00  4.00 ctcctgTggagaagtctgcc
SEQUENCE SIZE: 41
INCLUDED REGION SIZE: 41

    1 catggtgcacctgactcctgTggagaagtctgccgttactg
                          ^^^^^^^^^^^^^^^^^^^^
```

- Place an order at your local DNA lab!