1

```
         1                                          10                                       20
α        val leu ser pro ala asp lys thr asn val lys ala ala trp gly lys val gly ala his ala gly glu tyr gly ala glu ala leu -----
β    val his leu thr pro glu glu lys ser ala val thr ala leu trp gly lys val asn val         asp glu val gly gly glu ala leu ------
     1                                  10                                  20

    30                                        40                                    50
α   glu arg met phe leu ser phe pro thr thr     lys thr tyr phe pro his phe asp leu ser his gly ser ala ----------
β   gly arg leu leu val val tyr pro trp thr gln arg phe phe glu ser phe gly asp leu ser thr pro asp ala val met gly asn pro ------
        30                                    40                                    50

                        60                                        70                                  80
α   gln val lys gly his gly lys lys val ala asp ala leu thr asn ala val ala his val asp asp met pro asn ala leu ser ala ------
β   lys val lys ala his gly lys lys val leu gly ala phe ser asp gly leu ala his leu asp asn leu lys gly thr phe ala thr -----
        60                                  70                                  80

                        90                                          100                                  110
α   leu ser asp leu his ala his lys leu arg val asp pro val asn phe lys leu leu ser his cys leu leu val thr leu ala ala his -----
β   leu ser glu leu his cys asp lys leu his val asp pro glu asn phe arg leu leu gly asn val leu val cys val leu ala his his -----
            90                                  100                                 110

                                120                               130                                 140 141
α   leu pro ala glu phe thr pro ala val his ala ser leu asp lys phe leu ala ser val ser thr val leu thr ser lys tyr arg
β   phe gly lys glu phe thr pro pro val gln ala ala tyr gln lys val val ala gly val ala asn ala leu ala his lys tyr his
            120                               130                                 140             146
```

2

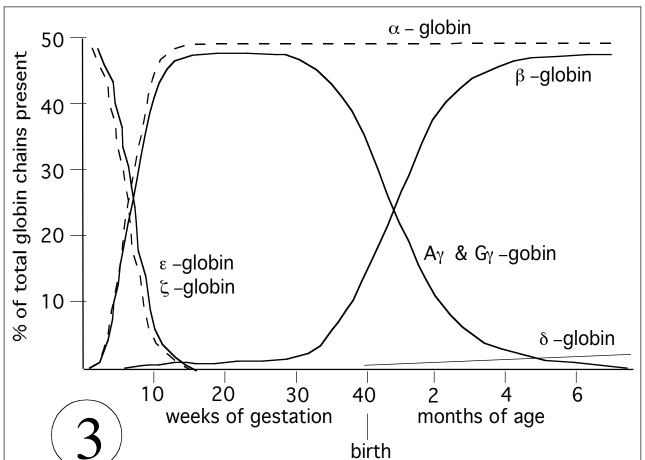
α– chain

	1	5	15	30	54	116
normal sequence:	val--	--ala--	--gly--	--glu--	--gln--	--glu
HbJ Toronto		--asp--				
HbJ Oxford			--asp--			
HbG Chineese				--gln--		
Hb Mexico					--glu--	
Hb Shimonoseki					--arg--	
HbO Indonesia						--lys--

β– chain

	1	2	6	7	26
normal sequence:	val--	--his--	--glu--	--glu--	--glu--
Hb Tokuchi		--try--			
HbC			--lys--		
HbS			--val--		
Hb San Jose				--gly--	
Hb Siriraj				--lys--	
HbE					--lys--

3



after Collins & Weissman,1984; page 331

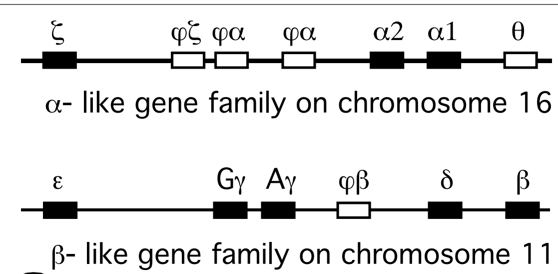
3a

embryonic (<8 weeks) $\zeta_2 \varepsilon_2$ $\zeta_2 \gamma_2$ $\alpha_2 \varepsilon_2$

fetal (3-9 months) $\alpha_2 \gamma_2$

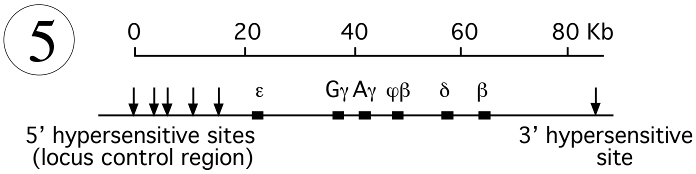
adult (from birth) $\alpha_2 \delta_2$ $\alpha_2 \beta_2$ after Genes VII

4



(after Genes VI, page 692)

5



(after Genes VI, page 871)

Alanine	ala	Leucine	leu
Arginine	arg	Lysine	lys
Asparagine	asn	Methionine	me
Aspartic Acid	asp	Phenylalanine	phe
Cysteine	cys	Proline	pro
Glutamine	gln	Serine	ser
Glutamic Acid	glu	Threonine	thr
Glycine	gly	Tryptophan	trp
Histidine	his	Tyrosine	tyr
Isoleucine	ile	Valine	val

6

```
1                                                                ttagacctcaccctgtggagccaca

2   ccctagggcagttggccaatctactcccaggagcagggagggcaggagcgggctggcataaaagtcagggca

3   ggaccatctattgcttACATTTGCTTCTACGACACAACTGTGTTCTAGCAACCTCAAACAGACACC

    met val his leu thr pro glu glu lys ser ala val thr ala leu trp gly lys val asn
4   ATG GTG CAC CTG ACT CCT GAG GAG AAG TCT GCC GTT ACT GCC CTG TGG GGC AAG GTG AAC

    val asp glu val gly gly glu ala leu gly ar
5   GTG GAT GAA GTT GGT GGT GAG GCC CTG GGC  AGGTTGGTATCAAGGTTACAAGACAGGTTT

6   AAGGAGACCAATAGAAACTGGGCATGTGGAGACAGAGAAGACTCTTGGGTTTCTGATAGGCACTGACTCTCTC

                                          g leu leu val val tyr pro trp thr gln arg
7   TGCCTATTGGTCTATTTT CCCACCCTTAGG CTG CTG GTG GTC TAC CCT TGG ACC CAG AGG

    phe phe glu ser phe gly asp leu ser thr pro asp ala val met gly asn pro lys val
8   TTC TTT GAG TCC TTT GGG GAT CTG TCC ACT CCT GAT GCT GTT ATG GGC AAC CCT AAG GTG

    lys ala his gly lys lys val leu gly ala phe ser asp gly leu ala his leu asp asn
9   AAG GCT CAT GGC AAG AAA GTG CTC GGT GCC TTT AGT GAT GGC CTG GCT CAC CTG GAC AAC

    leu lys gly thr phe ala thr leu ser glu leu his cys asp lys leu his val asp pro
10  CTC AAG GGC ACC TTT GCC ACA CTG AGT GAG CTG CAC TGT GAC AAG CTG CAC GTG GAT CCT

    glu asn phe arg
11  GAG AAC TTC AGG GTG AGTCTATGGGACCCTTGATGTTTTCTTTCCCCTTCTTTTCTATGGTTAAGTTCAT

12  GTCATAGGAAGGGGAGAAGTAACAGGGTACAGTTTAGAATGGGAAACAGACGAATGATTGCATCAGTGTGGAAG
13  TCTCAGGATCGTTTTAGTTTCTTTTATTTGCTGTTCATAACAATTGTTTTCTTTTGTTTAATTCTTGCTTTCTT
14  TTTTTTTCTTCTCCGCAATTTTTACTATTATACTTAATGCCTTAACATTGTGTATAACAAAAGGAAATATCTCT
15  GAGATACATTAAGTAACTTAAAAAAAAACTTTACACAGTCTGCCTAGTACATTACTATTTGGAATATATGTGTG
16  CTTATTTGCATATTCATAATCTCCCTACTTTATTTTCTTTTATTTTTAATTGATACATAATCATTATACATATT
17  TATGGGTTAAAGTGTAATGTTTTAATATGTGTACACATATTGACCAAATCAGGGTAATTTTGCATTTGTAATTT
18  TAAAAAATGCTTTCTTCTTTTAATATACTTTTTTGTTTATCTTATTTCTAATACTTTCCCTAATCTCTTTCTTT
19  CAGGGCAATAATGATACAATGTATCATGCCTCTTTGCACCATTCTAAAGAATAACAGTGATAATTTCTGGGTTA
20  AGGCAATAGCAATATTTCTGCATATAAATATTTCTGCATATAAATTGTAACTGATGTAAGAGGTTTCATATTGC
21  TAATAGCAGCTACAATCCAGCTACCATTCTGCTTTTATTTTATGGTTGGGATAAGGCTGGATTATTCTGAGTCC

                                                                leu leu gly asn val leu
22  AAGCTAGGCCCTTTTGCTAATCATGTTCATACCTCTTATCTTCCTCCCACAGCTCCTG GGC AAC GTG CTG

    val cys val leu ala his his phe gly lys glu phe thr pro pro val gln ala ala tyr
23  GTC TGT GTG CTG GCC CAT CAC TTT GGC AAA GAA TTC ACC CCA CCA GTG CAG GCT GCC TAT

    gln lys val val ala gly val ala asn ala leu ala his lys thr his
24  CAG AAA GTG GTG GCT GGT GTG GCT AAT GCC CTG GCC CAC AAG TAT CAC TAAGCTCGCTTT

25  CTTGCTGTCCAATTTCTATTAAAGGTTCCTTTGTTCCCTAAGTCCAACTACTAAACTGGGGGATATTATGAAG
26  GGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTGC•AATGATGTATTTAAATTATTTCT
27  GAATATTTTACTAAAAAGGGAATgtgggaggtcagtgcatttaaaacataaagaaatgaagagctagttcaa
28  accttgggaaaatacactatatcttaaactccatgaaagaaggtgaggctgcaaacagctaatgcacattggcaac
29  agccctgatgcctatgccttattcatccctcagatgctaaaggattcaagtagaggcttgatttggaggttaaagttt
30  atgcttgtattttacattacttattgttttagctgtcctcatgaatgtcttttcacacccatttgcttatcctgcatctc
31  tcagccttgactccactcagtt
```

Nucleotide sequence of beta-globin gene

non-transcribed strand shown

5' ..... 3'

lower case (a) -- not transcribed

upper case (A) -- transcribed

5' leader to ATG;
translated bases with amino acid designations;
introns between translated bases;
intron splice sites: at 5' end first base of intron underlined, at 3' end first base of next exon underlined

First base transcribed = +1
Base to the left = -1

ccaat = cat box;
ata = tata box
double underline = other promoters

underline at 3' end = sequences required for poly-A site and ? termination of transcription

dot = attachment site poly A

References

sequence and promoters from Collins and Weissman; Prog. Nucleic Acid Res. and Mol. Biol. v.31; 1984

3' details from Singer and Berg, 1991

For sequence details see : http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?cmd=Retrieve&db=nucleotide&list_uids=455025&dopt=GenBank&term=HUMHBB&qty=1

Diagram of some important regions of the beta-globin gene (not to scale)

7