

## Übung 4

a) AATCG

b) AACG

c) CG

d) AA

M = 1

MM = -1

G = -2

4.1)  $\rightarrow N \cdot (N-1)/2 = 6$  Paarweise Alignments:

a) AATCG  $p = \frac{1}{5} = 0,2$

b) AA-CG

b) AACG  $p = \frac{2}{4} = 0,5$

c) AA--

a) AATCG  $p = \frac{3}{5} = 0,6$

c) AA---

b) AACG  $p = \frac{2}{4} = 0,5$

d) --CG

a) AATCG  $p = \frac{3}{5} = 0,6$

d) ---CG

c) AA--  $p = \frac{4}{4} = 1$

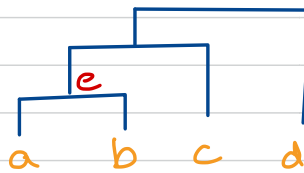
d) --CG

$\rightarrow$  Distanzmatrix:

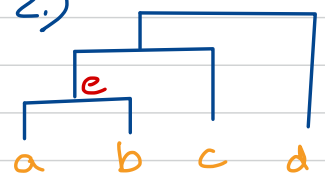
	a	b	c	d
a	-	0,2	0,6	0,6
b		-	0,5	0,5
c			-	1
d				-

$\rightarrow$  Guide Tree:

1.)



2.)



$\Rightarrow$  2 gleichwertige Varianten

	e	c	d
e	-	0,55	0,55
c		-	1
d			-

$$d(e,c) = \frac{1}{2 \cdot 1} (d(a,c) + d(b,c)) = \frac{1,1}{2} = 0,55$$

$$d(e,d) = \frac{1}{2 \cdot 1} (d(a,d) + d(b,d)) = \frac{1,1}{2} = 0,55$$

$\rightarrow$  MSA:

a) AATCG  
b) AA-CG  
c) --CG  
d) AA---

a) AATCG

b) AA-CG

d) AA---

c) ---CG

$\rightarrow$  Konsensus - Sequenz:

AA-CG

$$\rightarrow \text{Sum-of-Pairs-Score} = 2 + (-4) + (-4) + (-1) + (-1) + (-7) = -15$$

## 4.2) Vergleich der Sequenzen von Human, Maus, Ratte, Schimpanse, Rind, Katze

CLUSTAL O(1.2.4) multiple sequence alignment

```

sp|P02091|HBB1_RAT      MVHLTDAEKA AVNGLWGKVN PDDVGGEALGRLLVVPWTQRYFDSFGDLSSASAIMGNPK 60
sp|P02088|HBB1_MOUSE   MVHLTDAEKA AVSCLWGKVN SDEVGGEALGRLLVVPWTQRYFDSFGDLSSASAIMGNAK 60
sp|P07412|HBB_FELCA    -GFLTAE EKGLV NGLWGKVN DVGGEALGRLLVVPWTQRF FESFGDLSSADAIMSN AK 59
sp|P68871|HBB_HUMAN    MVHLTPEEKSA VTALWGKV NDEVGGEALGRLLVVPWTQRF FESFGDLSTPD AVMGNP K 60
sp|P68873|HBB_PANTR    MVHLTPEEKSA VTALWGKV NDEVGGEALGRLLVVPWTQRF FESFGDLSTPD AVMGNP K 60
sp|P02070|HBB_BOVIN    --MLTAE EKA AVTAFWGKV VDEVGGEALGRLLVVPWTQRF FESFGDLSTAD AVMMNPK 58
                        **  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

sp|P02091|HBB1_RAT      VKAHGKKVINAFNDGLKHL DNLKGTFAHLS ELHCDKLHVDPENFRLLGNMIVIVLGHHLG 120
sp|P02088|HBB1_MOUSE   VKAHGKKVITAFNDGLNHL DLSLKGTFASLSELHCDKLHVDPENFRLLGNMIVIVLGHHLG 120
sp|P07412|HBB_FELCA    VKAHGKKVLNSFSDGLKNID DLKGAFAKLSELHCDKLHVDPENFRLLGNVLCVLAHHFG 119
sp|P68871|HBB_HUMAN    VKAHGKKVLGAFSDGLAHL DNLKGTFA TLSELHCDKLHVDPENFRLLGNVLCVLAHHFG 120
sp|P68873|HBB_PANTR    VKAHGKKVLGAFSDGLAHL DNLKGTFA TLSELHCDKLHVDPENFRLLGNVLCVLAHHFG 120
sp|P02070|HBB_BOVIN    VKAHGKKVLD SFSNGMKHL DDLKGTFAALSELHCDKLHVDPENFKLLGNVLVVLARNFG 118
                        *****  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

sp|P02091|HBB1_RAT      KEFTPCAQA AFQKV VAGVASALAHKYH 147
sp|P02088|HBB1_MOUSE   KDFTPAAQA AFQKV VAGVATALAHKYH 147
sp|P07412|HBB_FELCA    HDFNPQVQA AFQKV VAGVANALAHKYH 146
sp|P68871|HBB_HUMAN    KEFTPPVQA AYQKV VAGVANALAHKYH 147
sp|P68873|HBB_PANTR    KEFTPPVQA AYQKV VAGVANALAHKYH 147
sp|P02070|HBB_BOVIN    KEFTPVLQA DFQKV VAGVANALAHRYH 145
                        :*: *  **  :*****  *****

```

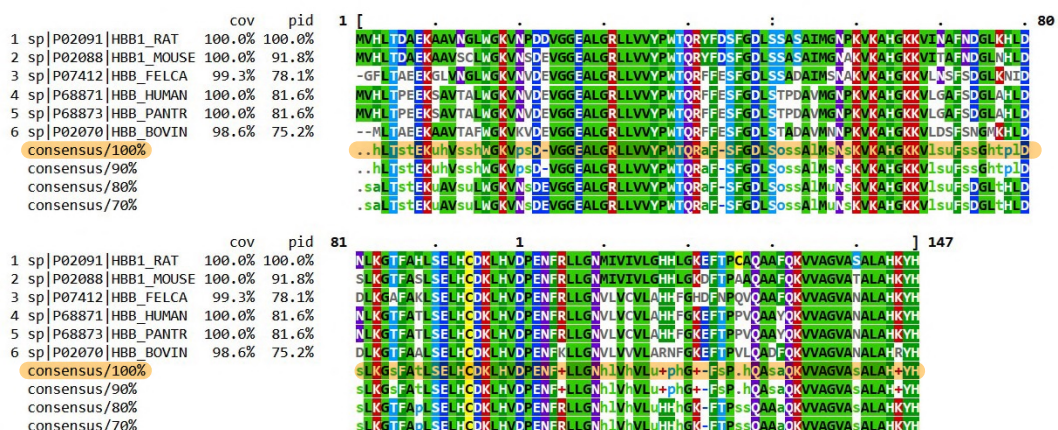
### Phylogenetic Tree

This is a Neighbour-joining tree without distance corrections.

Branch length: ☒ Cladogram ☐ Real



### Ansicht mit MView:



=> Ratte + Maus sind laut Phylogenetischem Baum nah verwandt

↗ auch anhand Sequenzen sichtbar

=> Konsensus: konservierte Aminosäuren (in jeder Sequenz vorhanden) : Groß

↗ viele vorhanden, nur wenige Mutationen in Hämoglobin Subunit beta

↗ wichtige Funktion in unterschiedlichsten Organismen (O<sub>2</sub>-Transport)