So far we have concentrated on the intrinsic properties of single sequences, such as CpG islands.
 3.
 4.
 5.
 6.
 7.

- 1. So far we have concentrated on the intrinsic properties of single sequences, such as CpG islands.
- 2. However functional biological sequences typically come in families, so what we are after is identifying the relationship of an individual sequences to a sequence family.
- 3.
 4.
 5.
 6.
 7.

- 1. So far we have concentrated on the intrinsic properties of single sequences, such as CpG islands.
- 2. However functional biological sequences typically come in families, so what we are after is identifying the relationship of an individual sequences to a sequence family.
- 3. A multiple sequence alignment can show how the sequences in a family relate to each other.
- 4. 5.

6.

- 1. So far we have concentrated on the intrinsic properties of single sequences, such as CpG islands.
- 2. However functional biological sequences typically come in families, so what we are after is identifying the relationship of an individual sequences to a sequence family.
- 3. A multiple sequence alignment can show how the sequences in a family relate to each other.
- 4. Some positions in a multiple sequence alignment are more conserved than others (e.g. helices as opposed to loop regions).
- 5.

6.

- 1. So far we have concentrated on the intrinsic properties of single sequences, such as CpG islands.
- 2. However functional biological sequences typically come in families, so what we are after is identifying the relationship of an individual sequences to a sequence family.
- 3. A multiple sequence alignment can show how the sequences in a family relate to each other.
- 4. Some positions in a multiple sequence alignment are more conserved than others (e.g. helices as opposed to loop regions).
- 5. Therefore it would be desirable, when identifying new sequence members to concentrate more on conserved features.
- 6.

- 1. So far we have concentrated on the intrinsic properties of single sequences, such as CpG islands.
- 2. However functional biological sequences typically come in families, so what we are after is identifying the relationship of an individual sequences to a sequence family.
- 3. A multiple sequence alignment can show how the sequences in a family relate to each other.
- 4. Some positions in a multiple sequence alignment are more conserved than others (e.g. helices as opposed to loop regions).
- 5. Therefore it would be desirable, when identifying new sequence members to concentrate more on conserved features.
- 6. For this task we will discuss a special type of HMM, well suited to modelling multiple alignments; we will call these *profile HMMs*.

- 1. So far we have concentrated on the intrinsic properties of single sequences, such as CpG islands.
- 2. However functional biological sequences typically come in families, so what we are after is identifying the relationship of an individual sequences to a sequence family.
- 3. A multiple sequence alignment can show how the sequences in a family relate to each other.
- 4. Some positions in a multiple sequence alignment are more conserved than others (e.g. helices as opposed to loop regions).
- 5. Therefore it would be desirable, when identifying new sequence members to concentrate more on conserved features.
- 6. For this task we will discuss a special type of HMM, well suited to modelling multiple alignments; we will call these *profile HMMs*.
- 7. Profile HMMs are the most popular application of HMMs in molecular biology.

A common feature of protein multiple sequence alignments is that gaps tend to line up with each other, leaving solid blocks with no insertions or deletions.

A common feature of protein multiple sequence alignments is that gaps tend to line up with each other, leaving solid blocks with no insertions or deletions.

A probabilistic model for a conserved region (block) would be to specify independent probabilities $e_i(\alpha)$ of observing amino acid α in position *i*. The probability of a new sequence *x*, is:

$$P(x \mid M) = \prod_{i=1}^{L} e_i(x_i)$$

where *L* is the length of the block.

A common feature of protein multiple sequence alignments is that gaps tend to line up with each other, leaving solid blocks with no insertions or deletions.

A probabilistic model for a conserved region (block) would be to specify independent probabilities $e_i(\alpha)$ of observing amino acid α in position *i*. The probability of a new sequence *x*, is:

$$P(x \mid M) = \prod_{i=1}^{L} e_i(x_i)$$

where *L* is the length of the block. To test for family membership we evaluate, as usual log-odds ratio:

$$S = \sum_{i=1}^{L} \log \frac{e_i(x_i)}{q_{x_i}}$$

A common feature of protein multiple sequence alignments is that gaps tend to line up with each other, leaving solid blocks with no insertions or deletions.

A probabilistic model for a conserved region (block) would be to specify independent probabilities $e_i(\alpha)$ of observing amino acid α in position *i*. The probability of a new sequence *x*, is:

$$P(x \mid M) = \prod_{i=1}^{L} e_i(x_i)$$

where *L* is the length of the block. To test for family membership we evaluate, as usual log-odds ratio:

$$S = \sum_{i=1}^{L} \log \frac{e_i(x_i)}{q_{x_i}}$$

The values $log(e_i(a)/q_a)$ behave like elements in a score matrix s(a, b) where the second index is position *i*, rather than amino acid *b*.

Such an approach is known as a *position specific score matrix* (PSSM).

Profile HMMs: αs and es

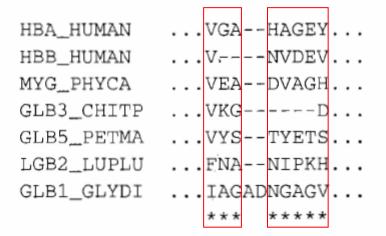
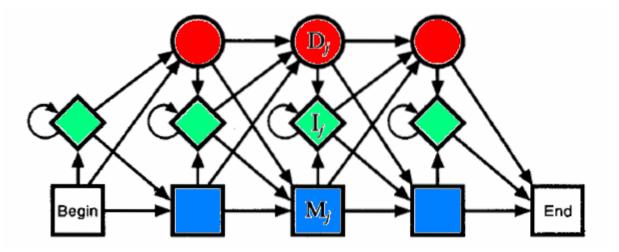


Figure 5.3 Ten columns from the multiple alignment of seven globin protein sequences shown in Figure 5.1. The starred columns are ones that will be treated as 'matches' in the profile HMM.

Let us assume that we use Laplace's rule to obtain parameters for an HMM corresponding to the alignment in Figure 5.3. Then $e_{M_1}(V) = 6/27$, $e_{M_1}(I) = e_{M_1}(F) = 2/27$, and $e_{M_1}(a) = 1/27$ for all residue types *a* other than V, I, F. Similarly, $a_{M_1M_2} = 7/10$, $a_{M_1D_2} = 2/10$ and $a_{M_1I_1} = 1/10$ (following column 1 there are six transitions from match to match, one transition to a delete state, in HBB_HUMAN, and no insertions). Figure 5.4 shows the complete set of parameters for the HMM

Profile HMMs: transition structure



MatchInsertionDeletion

Profile HMMs: Viterbi

Let $V_j^M(i)$ be the log-odds score of the best path matching subsequence $x_{1...i}$ to the submodel up to state *j*, ending with x_i being emitted by state *M*^{*i*} (match).

Similarly $V_j^{I}(i)$, $V_j^{D}(i)$ are the scores for state I_j (insertion) and D_j (deletion) respectively.

Then we can write:

$$V_{j}^{M}(i) = \log \frac{e_{M_{j}}(x_{i})}{q_{x_{i}}} + \max \begin{cases} V_{j-1}^{M}(i-1) + \log a_{M_{j-1}M_{j}}, \\ V_{j-1}^{I}(i-1) + \log a_{D_{j-1}M_{j}}, \\ V_{j}^{D}(i-1) + \log a_{D_{j-1}M_{j}}, \end{cases}$$

$$V_{j}^{I}(i) = \log \frac{e_{1_{j}}(x_{i})}{q_{x_{i}}} + \max \begin{cases} V_{j}^{M}(i-1) + \log a_{M_{j}1_{j}}, \\ V_{j}^{I}(i-1) + \log a_{M_{j}1_{j}}, \\ V_{j}^{D}(i-1) + \log a_{D_{j}I_{j}}, \\ V_{j}^{D}(i-1) + \log a_{D_{j}I_{j}}, \end{cases}$$

$$V_{j}^{D}(i) = \max \begin{cases} V_{j-1}^{M}(i) + \log a_{M_{j-1}D_{j}}, \\ V_{j-1}^{I}(i) + \log a_{D_{j-1}D_{j}}, \\ V_{j-1}^{D}(i) + \log a_{D_{j-1}D_{j}}. \end{cases}$$

Profile HMMs: Viterbi

Let $V_j^M(i)$ be the log-odds score of the best path matching subsequence $x_{1...i}$ to the submodel up to state *j*, ending with x_i being emitted by state *M*^{*i*} (match).

Similarly $V_j^{I}(i)$, $V_j^{D}(i)$ are the scores for state I_j (insertion) and D_j (deletion) respectively.

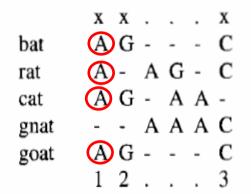
Then we can write:

$$V_{j}^{M}(i) = \log \frac{e_{M_{j}}(x_{i})}{q_{x_{i}}} + \max \begin{cases} V_{j-1}^{M}(i-1) + \log a_{M_{j-1}M_{j}}, \\ V_{j-1}^{I}(i-1) + \log a_{D_{j-1}M_{j}}, \\ V_{j}^{D}(i-1) + \log a_{D_{j-1}M_{j}}; \end{cases}$$

$$V_{j}^{I}(i) = \log \frac{e_{1j}(x_{i})}{q_{x_{i}}} + \max \begin{cases} V_{j}^{M}(i-1) + \log a_{M_{j}D_{j}}, \\ V_{j}^{I}(i-1) + \log a_{D_{j}D_{j}}, \\ V_{j}^{D}(i-1) + \log a_{D_{j}D_{j}}; \end{cases}$$

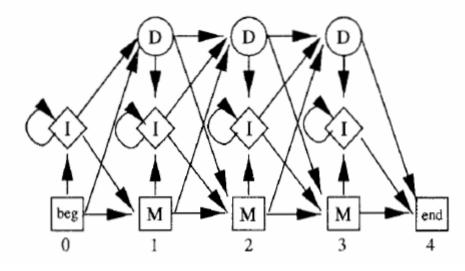
$$V_{j}^{D}(i) = \max \begin{cases} V_{j-1}^{M}(i) + \log a_{M_{j-1}D_{j}}, \\ V_{j-1}^{I}(i) + \log a_{D_{j-1}D_{j}}, \\ V_{j-1}^{D}(i) + \log a_{D_{j-1}D_{j}}. \end{cases}$$

Profile HMMs: model construction

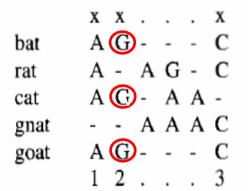


•

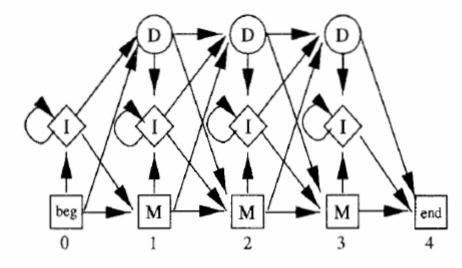
(b) Profile-HMM architecture:



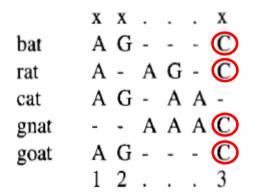
(c) Observed emission/transition counts									
		model position							
		0	1	2	3				
/	A	-	4	0	0				
match /	С	-	0	0	4				
emissions	G	-	0	3	0				
\ \	Т	-	0	0	0				
/	A	0	0	6	0				
insert /	С	0	0	0	0				
emissions	G	0	0	1	0				
	Т	0	0	0	0				
/	М-М	4	3	2	4				
/	M-D	1	1	0	0				
/	M-I	0	0	1	0				
state //	1-M	0	0	2	0				
transitions	1-D	0	0	1	0				
1	1-1	0	0	4	0				
\	D-M	-	0	0	1				
\	D-D	-	1	0	0				
1	D-l	-	0	2	0				



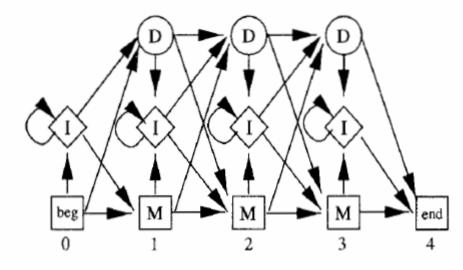
(b) Profile-HMM architecture:



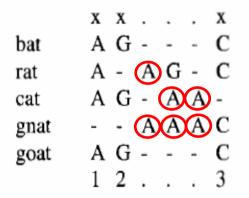
(c) Observed emission/transition counts model position $\frac{3}{0}$ -А С match 0 0 0 emissions G -Т -А 0 С insert emissions G Т $\frac{0}{4}$ M-M M-D M-I I-M state transitions 1-D 1-1 D-M -D-D -D-l -



(b) Profile-HMM architecture:

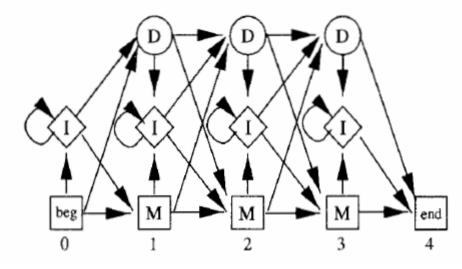


(c) Observed emission/transition counts model position $\frac{3}{0}$ 0 2 0 4 -А $\begin{pmatrix} 4 \\ 0 \\ 0 \\ \hline 0 \\ 0 \\ 0 \\ \end{pmatrix}$ С 0 3 match 0 emissions 0 G 0 0 Т -А 0 6 0 0 0 С 0 0 0 insert emissions G 0 1 0 0 Т 0 0 $\frac{0}{4}$ 4 M-M 0 M-D 1 1 0 0 M-I 0 0 1 0 2 0 I-M state transitions 1-D 0 0 1 0 0 0 1-1 0 4 0 0 D-M -D-D 1 0 0 -D-l 0 2 0 -



•

(b) Profile-HMM architecture:



(c) Observed emission/transition counts

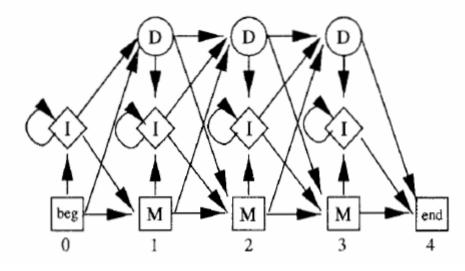
model position

			1		
		0	1	2	3
/	A	-	4	0	0
match /	С	-	0	0	4
emissions	G	-	0	3	0
	Т	-	0	0	0
/	A	0	0	6	0
insert /	С	0	0	0	0
emissions	G	0	0	1	0
	Т	0	0	0	0
,	M-M	4	3	2	4
/	M-D	1	1	0	0
/	M-I	0	0	1	0
state //	1-M	0	0	2	0
transitions	1-D	0	0	1	0
1	1-1	0	0	4	0
\	D-M	-	0	0	1
\	D-D	-	1	0	0
`	D-l	-	0	2	0

	Х	Х				Х
bat	А	G	-	-	-	С
rat	А	-	A	G	-	С
cat	А	G	-	A	A	-
gnat	-	-	А	А	A	С
goat	A	G	-	-	-	С
	1	2				3

•

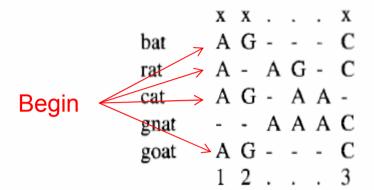
(b) Profile-HMM architecture:



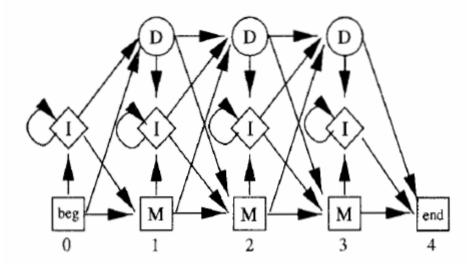
(c) Observed emission/transition counts

model position

			P		
		0	1	2	3
/	A	-	4	0	0
match /	С	-	0	0	4
emissions	G	-	0	3	0
\ \	Т	-	0	0	0
/	A	0	0	6	0
insert /	С	0	0	0	0
emissions	G	0	0		0
	Т	0	0	0	0
/	М-М	4	3	2	4
/	M-D	1	1	0	0
/	M-I	0	0	1	0
state //	1-M	0	0	2	0
transitions	1-D	0	0	1	0
1	1-1	0	0	4	0
\	D-M	-	0	0	1
\	D-D	-	1	0	0
1	D-l	-	0	2	0



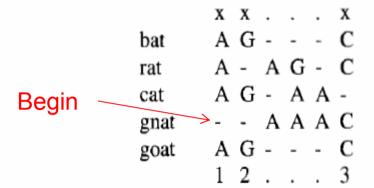
(b) Profile-HMM architecture:



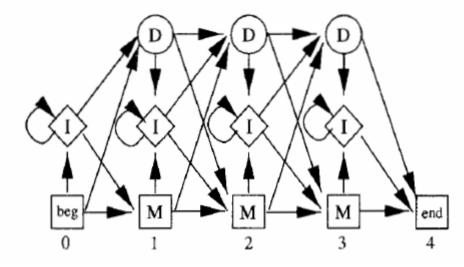
(c) Observed emission/transition counts								
		model position						
		0	1	2	3			
/	A	-	4	0	0			
match /	С	-	0	0	4			
emissions	G	-	0	3	0			
\backslash	Т	-	0	0	0			
/	A	0	0	6	0			
insert	С	0	0	0	0			
emissions	G	0	0	1	0			
$\langle \rangle$	Т	0	0	0	0			
/	М-М	4	3	2	4			
/	M-D	1	1	0	0			
/	M-I	0	0	1	0			
state //	1-M	0	0	2	0			
transitions	1-D	0	0	1	0			
1	1-1	0	0	4	0			
	D-M	-	0	0	1			
/	D-D	-	1	0	0			
\	D-l	-	0	2	0			

(c) Observed emission/transition counts

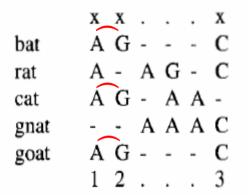
•



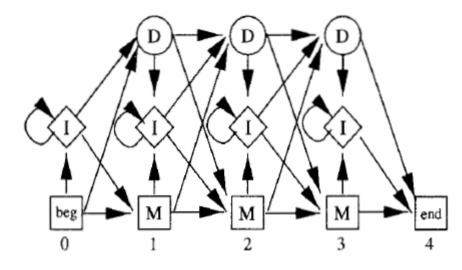
(b) Profile-HMM architecture:



(c) Observed emission/transition counts model position $\frac{3}{0}$ -А С match emissions G Т -А С insert emissions G Т $\frac{0}{4}$ M-M M-D M-I I-M state transitions 1-D 1-1 D-M -D-D -D-l -



(b) Profile-HMM architecture:

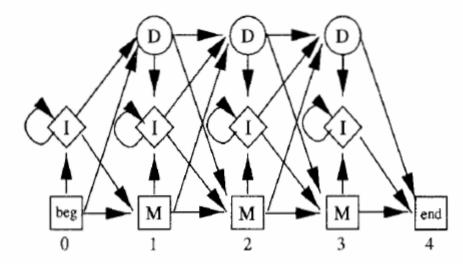


(c) Observed emission/transition counts model position $\frac{3}{0}$ -А С match emissions G Т -А С insert emissions G Т 2 $\frac{0}{4}$ M-M M-D M-I I-M state transitions 1-D $\frac{0}{1}$ 1-1 D-M -D-D -D-l -

	Х	Х				Х
bat	А	G	-	-	-	С
rat	Á	-	A	G	-	С
cat	А	G	-	A	A	-
gnat	-	-	A	А	A	С
goat	A	G	-	-	-	С
	1	2				3

•

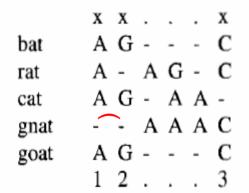
(b) Profile-HMM architecture:



(c) Observed emission/transition counts

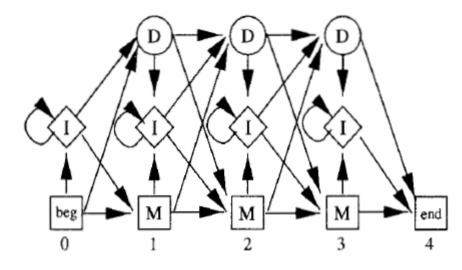
model position

			F		
		0	1	2	3
/	A	-	4	0	0
match /	С	-	0	0	4
emissions	G	-	0	3	0
	Т	-	0	0	0
/	A	0	0	6	0
insert /	С	0	0	0	0
emissions	G	0	0	1	0
	Т	0	0	0	0
/	M-M	4	3	2	4
/	M-D	1		0	0
/	M-I	0	0	1	0
state //	1-M	0	0	2	0
transitions	1-D	0	0	1	0
1	1-1	0	0	4	0
\	D-M	-	0	0	1
\	D-D	-	1	0	0
`	D-I	-	0	2	0



•

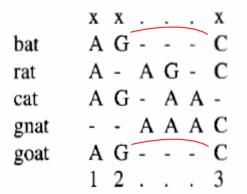
(b) Profile-HMM architecture:



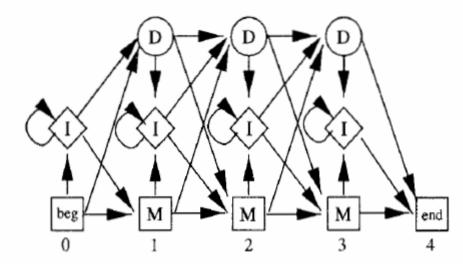
(c) Observed emission/transition counts

model position

			our p		-
		0	1	2	3
/	A	-	4	0	0
match /	С	-	0	0	4
emissions	G	-	0	3	0
	Т	-	0	0	0
/	A	0	0	6	0
insert /	С	0	0	0	0
emissions	G	0	0	1	0
	Т	0	0	0	0
/	М-М	4	3	2	4
/	M-D	1	1	0	0
/	M-I	0	0	1	0
state //	1-M	0	0	2	0
transitions	1-D	0	0	1	0
1	1-1	0	0	4	0
\	D-M	-	0	0	1
\	D-D	-	()	0	0
1	D-l	-	0	2	0



(b) Profile-HMM architecture:

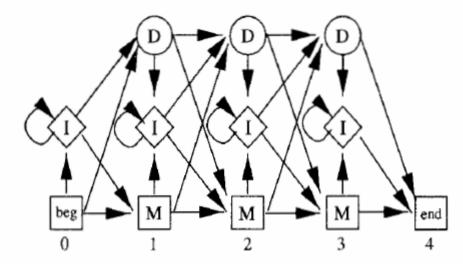


(c) Observed emission/transition counts model position $\frac{3}{0}$ -А С match emissions G Т -А С insert emissions G Т 4 (2)M-M M-D M-I I-M state transitions 1-D $\frac{0}{1}$ 1-1 -D-M D-D -D-l -

	Х	Х				Х
bat	А	G	-	-	-	С
rat	А	-	A	G	-	С
cat	А	G	-	A	A	-
gnat	-	-	A	А	A	С
goat	A	G	-	-	-	С
	1	2				3

•

(b) Profile-HMM architecture:



(c) Observed emission/transition counts

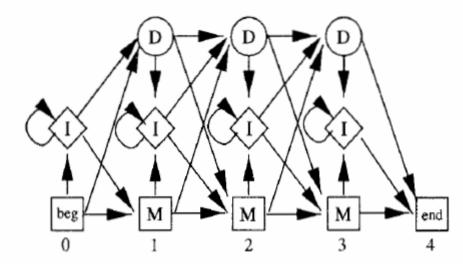
model position

			out p	0011011	
		0	1	2	3
/	A	-	4	0	0
match /	С	-	0	0	4
emissions	G	-	0	3	0
\ \	Т	-	0	0	0
/	A	0	0	6	0
insert /	С	0	0	0	0
emissions	G	0	0	1	0
	Т	0	0	0	0
,	M-M	4	3	2	4
/	M-D	1	1	0	0
/	M-I	0	0		0
state //	1-M	0	0	2	0
transitions	1-D	0	0	1	0
1	1-1	0	0	4	0
\	D-M	-	0	0	1
\	D-D	-	1	0	0
	D-l	-	0	2	0

	х	Х				х
bat	A	G	-	-	-	С
rat	А	-	A	G	-	С
cat	А	G	-	A	А	-
gnat	-	-	A	A	Á	С
goat	A	G	-	-	-	С
	1	2				3

•

(b) Profile-HMM architecture:

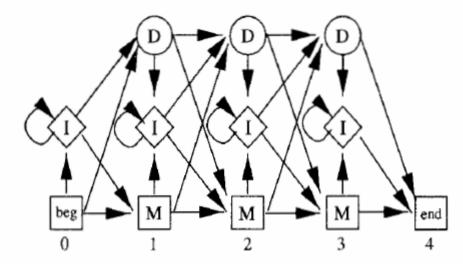


(c) Observed emission/transition counts model position А -С match emissions G -Т -А С insert emissions G Т $\frac{0}{4}$ M-M M-D M-I \bigcirc I-M state transitions 1-D $\frac{0}{1}$ 1-1 D-M -D-D -D-l -

	Х	Х				Х
bat	A	G	-	-	-	С
rat	А	-	A	G	-	С
cat	А	G	-	A	A	-
gnat	-	-	A	А	A	С
goat	A	G	-	-	-	С
	1	2				3

•

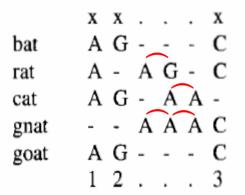
(b) Profile-HMM architecture:



(c) Observed emission/transition counts

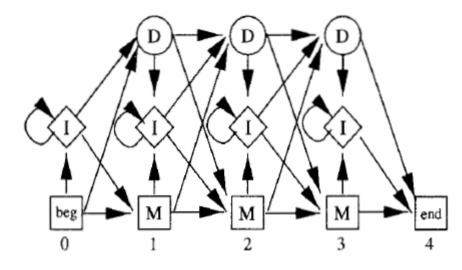
model position

		0		-	
		0	1	2	3
/	A	-	4	0	0
match /	С	-	0	0	4
emissions	G	-	0	3	0
	Т	-	0	0	0
/	A	0	0	6	0
insert /	С	0	0	0	0
emissions	G	0	0	1	0
	Т	0	0	0	0
/	M-M	4	3	2	4
/	M-D	1	1	0	0
/	M-I	0	0	1	0
state //	1-M	0	0	2	0
transitions	1-D	0	0	\bigcirc	0
1	1-1	0	0	4	0
	D-M	-	0	0	1
\	D-D	-	1	0	0
)	D-l	-	0	2	0
emissions	G T M-M M-D M-1 1-M 1-D 1-1 D-M D-D	0 0 4 1 0 0 0	0 0 3 1 0 0 0 0 0 0 1	$ \begin{array}{c} 1 \\ 0 \\ 2 \\ 0 \\ 1 \\ 2 \\ 1 \\ 4 \\ 0 \\ 0 \\ 0 \end{array} $	0 4 0 0 0 0 0 0 0 1 0



•

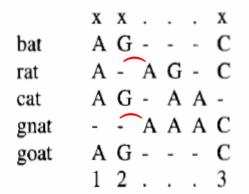
(b) Profile-HMM architecture:



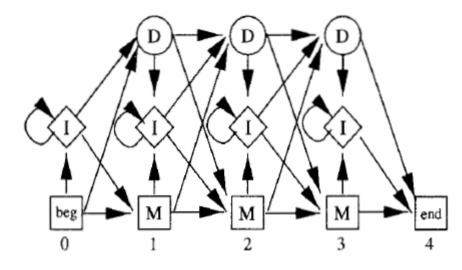
(c) Observed emission/transition counts

model position

		I			
		0	1	2	3
/	A	-	4	0	0
match /	С	-	0	0	4
emissions	G	-	0	3	0
	Т	-	0	0	0
/	A	0	0	6	0
insert /	С	0	0	0	0
emissions	G	0	0	1	0
	Т	0	0	0	0
,	М-М	4	3	2	4
/	M-D	1	1	0	0
/	M-I	0	0	1	0
state //	1-M	0	0	2	0
transitions	1-D	0	0	1	0
1	1-1	0	0	4	0
\	D-M	-	0	0	1
\	D-D	-	1	0	0
,	D-l	-	0	2	0

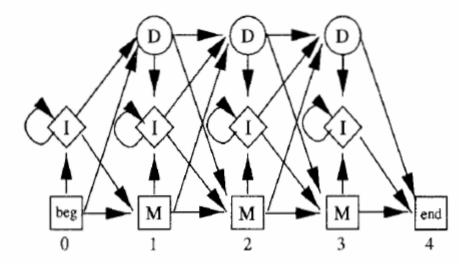


(b) Profile-HMM architecture:



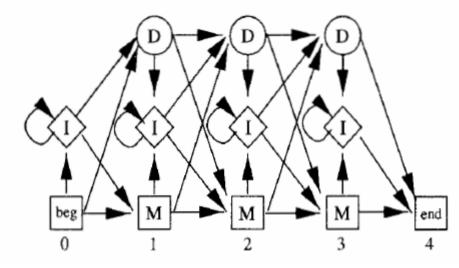
(c) Observed emission/transition counts model position 0 $\frac{3}{0}$ 2 0 4 -А С match 0 4 0 3 emissions 0 0 G 0 0 0 Т -А 0 0 0 6 С 0 0 0 0 insert 0 emissions G 0 1 0 Т 0 0 $\frac{0}{4}$ $\frac{0}{2}$ 3 4 M-M M-D 1 0 1 0 0 M-I 0 0 1 2 0 0 I-M state transitions 1-D 0 0 0 1 0 $\frac{0}{1}$ 1-1 0 4 0 0 D-M -D-D 0 0 1 -(2)D-l 0 0 -

(b) Profile-HMM architecture:



(c) Observed emission/transition counts						
		m	model position			
		0	1	2	3	
/	A	-	4	0	0	
match /	С	-	0	0	4	
emissions	G	-	0	3	0	
\sim	Т	-	0	0	0	
/	A	0	0	6	0	
insert /	С	0	0	0	0	
emissions	G	0	0	1	0	
	Т	0	0	0	0	
/	M-M	4	3	2	4	
/	M-D	1	1	0	0	
/	M-I	0	0	1	0	
state /	1-M	0	0	2	0	
transitions	1-D	0	0	1	0	
1	1-1	0	0	4	0	
\	D-M	-	0	0	1	
\	D-D	-	1	0	0	
````	D-l	-	0	2	0	

## (b) Profile-HMM architecture:



(c) Observed emission/transition counts					
		model position			
		0	1	2	3
/	A	-	4	0	0
match /	С	-	0	0	4
emissions	G	-	0	3	0
	Т	-	0	0	0
/	A	0	0	6	0
insert /	С	0	0	0	0
emissions	G	0	0	1	0
$\langle \rangle$	Т	0	0	0	0
/	M-M	4	3	2	4
/	M-D	1	1	0	0
/	M-I	0	0	1	0
state //	1-M	0	0	2	0
transitions	1-D	0	0	1	0
1	1-1	0	0	4	0
	D-M	-	0	0	
\	D-D	-	1	0	0
1	D-l	-	0	2	0

# **HMMER**

hmmalign - align sequences to a profile HMM SOURCE: http://hmmer.janelia.org/

# HMMER

hmmalign - align sequences to a profile HMM

hmmbuild - construct profile HMM(s) from multiple sequence alignment(s)

SOURCE: http://hmmer.janelia.org/



hmmalign - align sequences to a profile HMM

hmmbuild - construct profile HMM(s) from multiple sequence alignment(s)

hmmconvert - convert profile file to a HMMER format

hmmalign - align sequences to a profile HMM

hmmbuild - construct profile HMM(s) from multiple sequence alignment(s)

hmmconvert - convert profile file to a HMMER format

hmmemit - sample sequences from a profile HMM

hmmalign - align sequences to a profile HMM

hmmbuild - construct profile HMM(s) from multiple sequence alignment(s)

hmmconvert - convert profile file to a HMMER format

hmmemit - sample sequences from a profile HMM

hmmfetch - retrieve profile HMM(s) from a file

hmmalign - align sequences to a profile HMM

hmmbuild - construct profile HMM(s) from multiple sequence alignment(s)

hmmconvert - convert profile file to a HMMER format

hmmemit - sample sequences from a profile HMM

hmmfetch - retrieve profile HMM(s) from a file

hmmpress - prepare an HMM database for hmmscan

hmmalign - align sequences to a profile HMM

hmmbuild - construct profile HMM(s) from multiple sequence alignment(s)

hmmconvert - convert profile file to a HMMER format

hmmemit - sample sequences from a profile HMM

hmmfetch - retrieve profile HMM(s) from a file

hmmpress - prepare an HMM database for hmmscan

hmmscan - search sequence(s) against a profile database

hmmalign - align sequences to a profile HMM

hmmbuild - construct profile HMM(s) from multiple sequence alignment(s)

hmmconvert - convert profile file to a HMMER format

hmmemit - sample sequences from a profile HMM

hmmfetch - retrieve profile HMM(s) from a file

hmmpress - prepare an HMM database for hmmscan

hmmscan - search sequence(s) against a profile database

hmmsearch - search profile(s) against a sequence database

hmmalign - align sequences to a profile HMM

hmmbuild - construct profile HMM(s) from multiple sequence alignment(s)

hmmconvert - convert profile file to a HMMER format

hmmemit - sample sequences from a profile HMM

hmmfetch - retrieve profile HMM(s) from a file

hmmpress - prepare an HMM database for hmmscan

hmmscan - search sequence(s) against a profile database

hmmsearch - search profile(s) against a sequence database

hmmsim - collect score distributions on random sequences

hmmalign - align sequences to a profile HMM

hmmbuild - construct profile HMM(s) from multiple sequence alignment(s) hmmconvert - convert profile file to a HMMER format hmmemit - sample sequences from a profile HMM hmmfetch - retrieve profile HMM(s) from a file hmmpress - prepare an HMM database for hmmscan

hmmscan - search sequence(s) against a profile database

hmmsearch - search profile(s) against a sequence database

hmmsim - collect score distributions on random sequences

hmmstat - display summary statistics for a profile file

hmmalign - align sequences to a profile HMM

hmmbuild - construct profile HMM(s) from multiple sequence alignment(s) hmmconvert - convert profile file to a HMMER format hmmemit - sample sequences from a profile HMM hmmfetch - retrieve profile HMM(s) from a file hmmpress - prepare an HMM database for hmmscan hmmscan - search sequence(s) against a profile database hmmsearch - search profile(s) against a sequence database hmmsim - collect score distributions on random sequences hmmstat - display summary statistics for a profile file jackhmmer - iteratively search sequence(s) against a protein database

hmmalign - align sequences to a profile HMM

hmmbuild - construct profile HMM(s) from multiple sequence alignment(s) hmmconvert - convert profile file to a HMMER format hmmemit - sample sequences from a profile HMM hmmfetch - retrieve profile HMM(s) from a file hmmpress - prepare an HMM database for hmmscan hmmscan - search sequence(s) against a profile database hmmsearch - search profile(s) against a sequence database hmmsim - collect score distributions on random sequences hmmstat - display summary statistics for a profile file jackhmmer - iteratively search sequence(s) against a protein database

NAME fa LENG 45 ALPH DN RF nd CS nd MAP ye DATE Md NSEQ 10 EFFN 1. CKSUM 35 STATS LO	NA 5 5 5 5 5 7 5 7 7 5 7 7 7 5 7 7 7 5 7 7 7 5 7 7 7 5 7 7 5 7 7 5 7 7 5 7 7 5 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	-11.02	011 55 0.698				
	OCAL VITERE						
	CAL FORWAR						
HMM	A			т			
	m->m					d->m	d->d
COMPO	1.24690						
		1.38629					
	0.06034						*
1	0.38370					-	
		1.38629				1 000.01	0 40547
	0.06034 2.19189						0.4054/
2		1.38629				-	
	0.06034					1 09961	0 40547
	2.12548						0.1031/
		1.38629					
	0.06034					1.09861	0.40547
	1.82247						
		1.38629					
						1.09861	0.40547
5	2.12548						
	1.38629	1.38629	1.38629	1.38629			
	0.06034	3.53100	3.53100	1.46634	0.26236	1.09861	0.40547
6	2.12548	2.54191	0.34386	2.37646	6 -	-	
		1.38629					
	0.06034	3.53100	3.53100	1.46634	0.26236	1.09861	0.40547

$\begin{array}{c c c c c c c c c c c c c c c c c c c $	the number of match states in the model. Symbol alphabet type the number of sequences that the HMM was trained on	NAME fa: LENG 45 ALPH DN RF no CS no MAP ye: DATE Mo: NSEQ 10	A n Dec 51 215820 07035756	.2:34:59 2	011	4.6		
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$								
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$								
COMPO 1.24690 1.59679 1.28051 1.46032 1.38629 1.38629 1.38629 1.38629 0.06034 3.53100 3.53100 1.46634 0.26236 0.00000 * 1 0.38370 2.41911 2.07966 2.25673 1 1.38629 1.38629 1.38629 1.38629 0.06034 3.53100 3.53100 1.46634 0.26236 1.09861 0.40547 2 2.19189 1.97227 2.26310 0.43831 2 1.38629 1.38629 1.38629 1.38629 0.06034 3.53100 3.53100 1.46634 0.26236 1.09861 0.40547 3 2.12548 2.54191 0.34386 2.37646 3 1.38629 1.38629 1.38629 1.38629 0.06034 3.53100 3.53100 1.46634 0.26236 1.09861 0.40547 4 1.82247 2.39915 0.44485 2.23812 4		HMM	A	С	G	т		
$\begin{array}{cccccccccccccccccccccccccccccccccccc$			m->m	m->i	m->d	i->m	i->i d->m	d->d
$\begin{array}{cccccccccccccccccccccccccccccccccccc$		COMPO				1.46032		
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$								
$\begin{array}{cccccccccccccccccccccccccccccccccccc$								*
$\begin{array}{cccccccccccccccccccccccccccccccccccc$		1					1	
2 2.19189 1.97227 2.26310 0.43831 2 1.38629 1.38629 1.38629 1.38629 0.06034 3.53100 3.53100 1.46634 0.26236 1.09861 0.40547 3 2.12548 2.54191 0.34386 2.37646 3 1.38629 1.38629 1.38629 1.38629 0.06034 3.53100 3.53100 1.46634 0.26236 1.09861 0.40547 4 1.82247 2.39915 0.44485 2.23812 4								
1.38629 1.38629 1.38629 1.38629 0.06034 3.53100 3.53100 1.46634 0.26236 1.09861 0.40547 3 2.12548 2.54191 0.34386 2.37646 3 1.38629 1.38629 1.38629 1.38629 0.06034 3.53100 3.53100 1.46634 0.26236 1.09861 0.40547 4 1.82247 2.39915 0.44485 2.23812 4								0.40547
0.06034 3.53100 3.53100 1.46634 0.26236 1.09861 0.40547 3 2.12548 2.54191 0.34386 2.37646 3 1.38629 1.38629 1.38629 1.38629 0.06034 3.53100 3.53100 1.46634 0.26236 1.09861 0.40547 4 1.82247 2.39915 0.44485 2.23812 4		2					2	
3 2.12548 2.54191 0.34386 2.37646 3 1.38629 1.38629 1.38629 1.38629 0.06034 3.53100 3.53100 1.46634 0.26236 1.09861 0.40547 4 1.82247 2.39915 0.44485 2.23812 4								
1.38629 1.38629 1.38629 1.38629 0.06034 3.53100 3.53100 1.46634 0.26236 1.09861 0.40547 4 1.82247 2.39915 0.44485 2.23812 4								0.40547
0.06034 3.53100 3.53100 1.46634 0.26236 1.09861 0.40547 4 1.82247 2.39915 0.44485 2.23812 4		3					3	
4 1.82247 2.39915 0.44485 2.23812 4								
								0.40547
1.30029 1.30029 1.30029 1.30029		4					4	
0.06034 3.53100 3.53100 1.46634 0.26236 1.09861 0.40547							0 26226 1 08861	0 40547
5  2.12548  2.54191  0.34386  2.37646  5  -  -		5						0.40547
1.38629 1.38629 1.38629 1.38629		5					5	
0.06034 3.53100 3.53100 1.46634 0.26236 1.09861 0.40547							0.26236 1.09861	0.40547
6 2.12548 2.54191 0.34386 2.37646 6		6						0.1001/
1.38629 1.38629 1.38629 1.38629		Ŭ						
0.06034 3.53100 3.53100 1.46634 0.26236 1.09861 0.40547							0.26236 1.09861	0.40547

the number of match states in the model.	LENG 459 ALPH DNA RF no CS no MAP yes	ta2.aln	-					
the number of sequences that the HMM was trained on	NSEQ 10 EFFN 1.2 CKSUM 350 STATS LOC	AL MSV	-11.02	55 0.698				
the effective total number of sequences determined	STATS LOC. STATS LOC. HMM							
by hmmbuild during sequence weighting, for combining observed	СОМРО		m-≻i 1.59679 1.38629		i-≻m 1.46032 1.38629	i->i	d->m	d->d
counts with Dirichlet prior information in	1	0.38370	3.53100 2.41911 1.38629	2.07966	1.46634 2.25673 1.38629	0.26236 1	0.00000	*
parameterizing the model	2		3.53100 1.97227	3.53100	1.46634 0.43831 1.38629	0.26236 2		0.40547
	3	0.06034 2.12548	3.53100 2.54191	3.53100 0.34386	1.46634 2.37646	0.26236 3		0.40547
	4	0.06034 1.82247		3.53100 0.44485	1.38629 1.46634 2.23812	0.26236 4		0.40547
	5	1.38629 0.06034 2.12548	2.54191	3.53100 0.34386	1.38629 1.46634 2.37646	0.26236 5		0.40547
	6		3.53100 2.54191	1.38629 3.53100 0.34386	1.38629 1.46634 2.37646	0.26236 6	1.09861	0.40547
		1.38629 0.06034	1.38629 3.53100	1.38629 3.53100	1.38629 1.46634	0.26236	1.09861	0.40547

the number of match states in the model.	NAME fas LENG 459		rch 2010]					
Symbol alphabet type	ALPH DNA RF no CS no MAP yes	3						
the number of sequences that the HMM was trained on	NSEQ 10 EFFN 1.2 CKSUM 350 STATS LOO	215820 07035756 CAL MSV	2:34:59 2	55 0.698				
the effective total number of sequences determined		AL VITERB AL FORWAR A	D -5.13 C	06 0.698 G	46 T			
by hmmbuild during sequence weighting, for combining observed	Сомро	m->m 1.24690 1.38629	m->i 1.59679 1.38629	1.38629	i->m 1.46032 1.38629	i->i	d->m	d->d
counts with Dirichlet prior information in parameterizing the model	1		2.41911 1.38629	1.38629	1.46634 2.25673 1.38629	1 -	0.00000	*
This number is calculated	2	2.19189	3.53100 1.97227 1.38629	2.26310	1.46634 0.43831 1.38629	0.26236 2 -	1.09861	0.40547
from the training sequence data, and used in conjunctio	n 3	0.06034 2.12548 1.38629	3.53100 2.54191 1.38629		1.46634 2.37646 1.38629	0.26236 3 -	1.09861	0.40547
with the alignment map information to verify that a given alignment is indeed th	e 4	0.06034 1.82247 1.38629	3.53100 2.39915 1.38629	3.53100 0.44485 1.38629	1.46634 2.23812 1.38629	0.26236 4 -	1.09861	0.40547
alignment that the map is for		0.06034 2.12548 1.38629	3.53100 2.54191 1.38629	3.53100 0.34386 1.38629	1.46634 2.37646 1.38629	0.26236 5 -	1.09861	0.40547
	6	0.06034 2.12548 1.38629 0.06034	3.53100 2.54191 1.38629 3.53100	3.53100	1.46634 2.37646 1.38629	6 -	1.09861	0.40547
		0.00034	3.33100	3.33100	1.10034	0.20230	1.00001	0.1031/

the number of match states in the model.	HMMER NAME	3/b [3.0   1 fasta2.aln 459	March 2010]	colum	n annotatio	tion flag;.Reference n is picked up from a ent file's #=GC RF line	
Symbol alphabet type	ALPH RF CS MAP	DNA no no yes	10-04-50-0				
the number of sequences that the HMM was trained on	STAT\$	10 1.215820 3507035756 LOCAL MSV	-11.02	:55 0.698			
the effective total number of sequences determined	/ /	LOCAL VITE LOCAL FORW A	ARD -5.13 C	06 0.698 G	346 Т		
by hmmbuild during sequence weighting, for combining observed	сом	1.3862	9 1.38629	m->d 1.28051 1.38629	i->m 1.46032 1.38629	i->i d->m	d->d
counts with Dirichlet prior information in		0.0603 1 0.3837 1.3862	0 2.41911 9 1.38629	3.53100 2.07966 1.38629	1.46634 2.25673 1.38629	0.26236 0.00000 1	*
parameterizing the model This number is calculated		0.0603 2 2.1918 1.3862	9 1.97227	3.53100 2.26310 1.38629	1.46634 0.43831 1.38629	0.26236 1.09861 2	0.40547
from the training sequence data, and used in conjunction	n	3 2.1254	4 3.53100 8 2.54191 9 1.38629	3.53100 0.34386 1.38629	1.46634 2.37646 1.38629	0.26236 1.09861 3	0.40547
with the alignment map information to verify that a given alignment is indeed th	IE III		4 3.53100 7 2.39915	3.53100 0.44485 1.38629	1.46634 2.23812 1.38629	0.26236 1.09861 4	0.40547
alignment that the map is fo		0.0603 5 2.1254 1.3862	4 3.53100 8 2.54191	3.53100 0.34386 1.38629	1.46634 2.37646 1.38629	0.26236 1.09861 5	0.40547
		0.0603 6 2.1254 1.3862	4 3.53100 8 2.54191	3.53100 0.34386 1.38629	1.38629 1.46634 2.37646 1.38629	0.26236 1.09861 6	0.40547
		0.0603	4 3.53100	3.53100	1.46634	0.26236 1.09861	0.40547

the number of match states in the model.	NAME LENG	/b [3.0   Ma fasta2.aln 459	rch 2010]	colum Stocki	n annotation nolm alignm	tion flag;.Reference n is picked up from lent file's #=GC RF cture annotation fla	a Iine.
Symbol alphabet		DNA				nolm file's #=GC S	
type		no <del>&lt;</del>					
.,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,		yes					
the number of		Mon Dec 51 10	2:34:59 2	011			
sequences that the	~	1.215820					
HMM was trained on	CKSUM	3507035756					
		LOCAL MSV		55 0.698			
the effective total number		LOCAL VITERB LOCAL FORWAR					
of sequences determined	HMM	LOCAL FORWAR A	D -5.13 C	06 0.698 G	40 T		
by hmmbuild during	//	m->m	m->i	m->d	i->m	i->i d-3	>m d->d
sequence weighting, for	COMP		1.59679	1.28051	1.46032		
combining observed			1.38629	1.38629	1.38629		
counts with Dirichlet prior		0.06034 1 0.38370	3.53100 2.41911	3.53100 2.07966	1.46634 2.25673	0.26236 0.00	• 000
information in	· · · · ·	1.38629		1.38629	1.38629	1	
parameterizing the model		0.06034		3.53100	1.46634	0.26236 1.09	861 0.40547
	/ :	2 2.19189		2.26310	0.43831	2	
This number is calculated		1.38629		1.38629	1.38629		
from the training sequence		0.06034		3.53100 0.34386	1.46634	0.26236 1.09	861 0.40547
data, and used in conjunctio	<b>n</b> / `	3 2.12548 1.38629	2.54191	1.38629	2.37646 1.38629	3	
with the alignment map		0.06034			1.46634	0.26236 1.09	861 0.40547
information to verify that a		4 1.82247	2.39915	0.44485	2.23812	4	
given alignment is indeed th		1.38629	1.38629	1.38629	1.38629		
alignment that the map is for		0.06034	3.53100	3.53100	1.46634	0.26236 1.09	861 0.40547
		5 2.12548 1.38629	2.54191 1.38629	0.34386 1.38629	2.37646 1.38629	5	
		0.06034		3.53100	1.46634	0.26236 1.09	861 0.40547
		6 2.12548	2.54191	0.34386	2.37646	6	
		1.38629	1.38629	1.38629	1.38629		
		0.06034	3.53100	3.53100	1.46634	0.26236 1.09	861 0.40547

the number of match states in the model.	HMMER: NAME LENG	3/b [3.0   Ma fasta2.aln 459	arch 2010]	Sto	ference annota umn annotatior ockholm alignm	n is picked u ent file's #=	ip from a GC RF line.	
	ALPH	DNA			onsensus struc		· · · · · ·	
Symbol alphabet	RF CS	no		υμ	o from a Stockh	ioim nie s #•	=GC 55_00	is line
type	MAP	yes		M	ap annotation f	lag. The HN	/M/alignmen	t map
	DATE	Mon Dec 5	L2:34:59 2		notates each r			
the number of	NSEQ	10		ali	ignment columr	n from whic	h it came. It	can be
sequences that the	EFFN	1.215820 3507035756		us	sed for quickly r	mapping an	y subsequer	t HMM
HMM was trained on		LOCAL MSV	-11.02		ignment back to	o the origina	al multiple ali	gnment,
		LOCAL VITER			a the model			
the effective total number		LOCAL FORWAR			69846			
of sequences determined	HMM	A	C .	G	Т			
by hmmbuild during		m->m	m->i	m->c		i->i	d->m	d->d
sequence weighting, for	COM	PO 1.24690 1.38629	1.59679 1.38629	1.2805				
combining observed		0.06034	3.53100	3.5310		0.26236	0.00000	*
counts with Dirichlet prior		1 0.38370	2.41911	2.0796		1 -	-	
information in		1.38629	1.38629	1.3862		-		
parameterizing the model		0.06034	3.53100	3.5310	00 1.46634	0.26236	1.09861	0.40547
	/	2 2.19189	1.97227	2.2631	10 0.43831	2 -	-	
This number is calculated		1.38629	1.38629	1.3862				
from the training sequence		0.06034	3.53100	3.5310			1.09861	0.40547
data, and used in conjunctio	n /	3 2.12548	2.54191	0.3438		3 -	-	
with the alignment map		1.38629	1.38629	1.3862		0.00000	1 00061	0 405 47
information to verify that a		0.06034	3.53100 2.39915	3.5310		0.26236 4 -	1.09861	0.40547
given alignment is indeed the	<b>_</b>	1.38629	1.38629	1.3862		-		
alignment that the map is for		0.06034	3.53100	3.5310		0.26236	1.09861	0.40547
alignment that the map is for		5 2.12548	2.54191	0.3438		5 -		
		1.38629	1.38629	1.3862	29 1.38629			
		0.06034	3.53100	3.5310				0.40547
		6 2.12548	2.54191	0.3438		6 -	-	
		1.38629	1.38629	1.3862				
		0.06034	3.53100	3.5310	00 1.46634	0.26236	1.09861	0.40547

Statistical parameters needed for E-value calculations.

			rch 2010]					
		ta2.aln						
LENG								
ALPH								
RF								
CS								
MAP				~ ~ ~				
		Dec 51	2:34:59 2	011				
NSEQ		15000						
EFFN		7035756						
			_11_02	55 0 690	46			
gmamg	LOCZ	AL MOV	-11.02 I -12.62	01 0.698	46			
			D -5.13					
			c		т			
						i->i	d->m	d->d
COM			1.59679					
			1.38629					
						0.26236	0.00000	*
			2.41911					
		1.38629	1.38629	1.38629	1.38629			
		0.06034	3.53100	3.53100	1.46634	0.26236	1.09861	0.40547
	2	2.19189	1.97227	2.26310	0.43831	2 -	-	
			1.38629					
							1.09861	0.40547
			2.54191				-	
			1.38629					
							1.09861	0.40547
			2.39915				-	
			1.38629					
							1.09861	0.40547
			2.54191				-	
			1.38629				1 00001	0 40547
			2.54191				1.09861	0.4034/
	0		1.38629					
							1.09861	0 40547
		0.00034	5.55100	5.55100	1.10034	0.20230	1.00001	0.1001/

Statistical parameters needed for E-value calculations. Flags the start of the main model section	LENG 459 ALPH DNA RF no CS no MAP yes DATE Mon NSEQ 10 EFFN 1.23 CKSUM 350	Dec 5 1 5820 7035756	2:34:59 2					
$\sim$	STATS LOCA		-11.02					
	STATS LOCA							
	HMM	A	c	G. 010	т			
		m->m	m->i	m->d	i->m	i->i	d->m	d->d
	COMPO	1.24690	1.59679	1.28051	1.46032			
			1.38629		1.38629			
			3.53100		1.46634			*
	1		2.41911		2.25673	1	-	
			1.38629		1.38629			
			3.53100			0.26236		0.40547
	2		1.97227		0.43831	2	-	
			1.38629 3.53100		1.38629	0.26236	1 09961	0 40547
	3		2.54191		2.37646	3		0.1031/
			1.38629		1.38629	5		
			3.53100		1.46634	0.26236	1.09861	0.40547
	4		2.39915		2.23812	4		
		1.38629	1.38629	1.38629	1.38629			
		0.06034	3.53100	3.53100	1.46634	0.26236	1.09861	0.40547
	5	2.12548	2.54191		2.37646	5	-	
			1.38629		1.38629			
			3.53100		1.46634			0.40547
	6		2.54191		2.37646	6	-	
			1.38629		1.38629	0.00000	1 000.00	0 40545
		0.06034	3.53100	3.53100	1.46634	0.26236	1.09861	0.40547

Statistical parameters needed for E-value calculations.	NAME fas LENG 459 ALPH DNA RF no CS no	7	rch 2010]						
	MAP yes DATE Mor	3 1 Dec 51	2:34:59 2	011					
Flags the start of the	NSEQ 10								
main model section	EFFN 1.2 CKSUM 350	215820							
	STATS LOC		-11.02	55 0.698	46				
		CAL VITERB							
design of the second state		CAL FORWAR							
these are the model's	HMM	A	С	G	т				
overall average match		m−>m			i->m	i->i	d->m	d->d	
state emission	COMPO		1.59679	1.28051	1.46032				
probabilities, which are		1.38629 0.06034	1.38629 3.53100	1.38629 3.53100	1.38629 1.46634	0.26236	0 00000	*	
used as a background	1	0.38370			2.25673	0.26236		-	
residue composition in the	1		1.38629		1.38629	1			
"filter null" model.			3.53100			0.26236	1.09861	0.40547	
	2	2.19189	1.97227	2.26310	0.43831	2 -	-		
			1.38629		1.38629				
	_		3.53100		1.46634		1.09861	0.40547	
	3	2.12548			2.37646	3 -	-		
			1.38629 3.53100		1.38629 1.46634	0 26236	1.09861	0.40547	
	4	1.82247			2.23812	4 -		0.1031/	
	-		1.38629	1.38629	1.38629	-			
		0.06034	3.53100	3.53100	1.46634	0.26236	1.09861	0.40547	
	5	2.12548	2.54191		2.37646	5 -	-		
			1.38629	1.38629	1.38629				
	-	0.06034		3.53100	1.46634		1.09861	0.40547	
	6	2.12548 1.38629	2.54191 1.38629	0.34386 1.38629	2.37646 1.38629	6 -	-		
		0.06034	3.53100	3.53100		0.26236	1.09861	0.40547	
		0.00001	3.00100	5.00100	2		1.05001		

	ains information f d as model node							
proba conta	begin state is mu abilities. The first ains the transition a seven numbers	line is the ins from the	insert 0 emi	ssions. The	second line	•		
$B \rightarrow$	M1, $B \rightarrow I_0, B \rightarrow$	D1; $I_0 \rightarrow N$						
trans	a 0.0 and a '*', b itions from the no og 0 = $-\infty$ = '*'.							
		m->m	m->i	m->d	i->m	i->i	d->m	d->d
	COMPO	1.24690	1.38629	1.28051	1.38629			
		0.06034	3.53100	3.53100	1.46634	0.26236	0.00000	*
	1	0.38370	2.41911	2.07966	2.25673	1 -	-	
		1.38629	1.38629	1.38629	1.38629			
		0.06034	3.53100	3.53100	1.46634		1.09861	0.40547
	2	2.19189 1.38629	1.97227 1.38629	2.26310 1.38629	0.43831 1.38629	2 -	-	
		0.06034	3.53100	3.53100	1.46634	0 26236	1.09861	0.40547
	3	2.12548	2.54191	0.34386	2.37646	3 -		0.1001/
		1.38629	1.38629	1.38629	1.38629			
		0.06034	3.53100	3.53100	1.46634	0.26236	1.09861	0.40547
	4	1.82247	2.39915	0.44485	2.23812	4 -	-	
		1.38629	1.38629	1.38629	1.38629			
	-	0.06034	3.53100	3.53100	1.46634		1.09861	0.40547
	5	2.12548 1.38629	2.54191	0.34386	2.37646	5 -	-	
		0.06034	1.38629 3.53100	1.38629 3.53100	1.38629 1.46634	0 26236	1.09861	0.40547
	6	2.12548		0.34386	2.37646	6 -		0.1001/
	-	1.38629		1.38629	1.38629	-		
		0.06034	3.53100	3.53100	1.46634	0.26236	1.09861	0.40547

	NAME LENG ALPH RF CS	fas 459	sta2.aln	urch 2010]					
			Dec 51	2:34:59 2	011				
		10	215820						
		3507035756							
				-11.02					
				SI -12.62					
	STATS HMM	LOC	AL FORWAR A	ир -5.13 с	06 0.698 G	146 T			
	IIIMM		m->m	m->i	m->d	i->m	i->i	d->m	d->d
	COME	0		1.59679					
				1.38629	1.38629				
				3.53100		1.46634	0.26236 0	.00000	*
Line 1: Match emission line	e	1		2.41911 1.38629	2.07966 1.38629		1		
			1.38629 0.06034	3.53100			0.26236 1	09861	0 40547
		2	2.19189	1.97227	2.26310	0.43831	2	.00001	0.1001/
Line 2: Insert emission line			1.38629	1.38629	1.38629	1.38629			
			0.06034	3.53100	3.53100		0.26236 1	.09861	0.40547
		3	2.12548	2.54191	0.34386	2.37646	3		
Line 3: State transition line			1.38629 0.06034	1.38629 3.53100	1.38629 3.53100	1.38629 1.46634	0.26236 1	09961	0 40547
		4	1.82247	2.39915	0.44485	2.23812	4	.05001	0.40347
		-	1.38629	1.38629	1.38629	1.38629	-		
			0.06034	3.53100	3.53100		0.26236 1	.09861	0.40547
		5		2.54191	0.34386		5		
				1.38629	1.38629		0.00000 1	00061	0 40547
		6		3.53100 2.54191		1.46634 2.37646	0.26236 1	.09861	0.4054/
		0		1.38629	1.38629		0		
				3.53100		1.46634	0.26236 1	.09861	0.40547

	NAME fas LENG 459 ALPH DNA RF no CS no MAP yes DATE Mon NSEQ 10 EFFN 1.3 CKSUM 350 STATS LOO STATS LOO STATS LOO	A n Dec 51 215820 07035756	2:34:59 2 -11.02 I -12.62	55 0.698 01 0.698	46 46		
	HMM		m->i	_	r i->m	i->i d->m	d->d
	COMPO	110010010010010010010010010010010010010	m-⊭1 1.59679				<b> </b>
	COMPU	1.38629	1.38629	1.38629	1.38629		
		0.06034	3.53100	3.53100		0.26236 0.00000	*
Line 1. Motok emission line	1		2.41911	2.07966	2.25673	1	
Line 1: Match emission line		1.38629	1.38629	1.38629	1.38629		
		0.06034	3.53100			0.26236 1.09861	0.40547
	2	2.19189	1.97227	2.26310	0.43831	2	
Line 2: Insert emission line		1.38629	1.38629	1.38629	1.38629		
		0.06034	3.53100	3.53100	1.46634	0.26236 1.09861	0.40547
		2.12548	2.54191	0.34386	2.37646	3	101012010101010122202020202020
Line 3: State transition line		1.38629	1.38629	1.38629	1.38629		
		0.06034	3.53100	3.53100	1.46634	0.26236 1.09861	0.40547
	4	1.82247	2.39915	0.44485	2.23812	4	
		1.38629	1.38629	1.38629	1.38629		
	-	0.06034	3.53100	3.53100		0.26236 1.09861	0.40547
	5	2.12548	2.54191	0.34386	2.37646	5	
		1.38629 0.06034	1.38629	1.38629 3.53100	1.38629	0 06006 1 00061	0 40547
	6	2.12548	3.53100 2.54191		1.46634 2.37646	0.26236 1.09861	0.4034/
	0		1.38629	1.38629	1.38629	0	
		0.06034	3.53100	3.53100		0.26236 1.09861	0.40547
		0.00001	3.00100	0.00100	1.10001	0.20200 1.00001	0.1001/

	NAME LENG ALPH RF CS MAP DATE NSEQ EFFN	fas 459 DNA no no yes Mon 10 1.2				th th	e MAP annotation e RF annotation e CS annotation for this node	
				-11.02	55 0.698	46		
			AL VITERB					
		FOC	AL FORWAR					
	HMM		A m−>m	C m->i	G	T i−>m	i->i d->m	d->d
	COMP	0		1.59679				
	COM	Č		1.38629	1.38629			
				3.53100		1.46634	0.26230 0.00000	*
Line 1: Match emission line	2	1	0.38370	2.41911	2.07966	2.25673	1	
				1.38629	1.38629			
				3.53100			0.26236 1.09861	0.40547
		2	2.19189	1.97227	2.26310	0.43831	2	
Line 2: Insert emission line			1.38629	1.38629	1.38629	1.38629		
			0.06034	3.53100	3.53100	1.46634	0.26236 1.09861	0,4054/
		8	1.38629	1.38629	1.38629	1.38629	3	
Line 3: State transition line			0.06034	3.53100	3.53100	1.46634	0.26236 1.09861	0 40547
		4	1.82247	2.39915	0.44485	2.23812	4	0.1001/
		-	1.38629	1.38629	1.38629	1.38629	-	
				3.53100	3.53100	1.46634	0.26236 1.09861	0.40547
		5	2.12548	2.54191	0.34386	2.37646	5	
			1.38629	1.38629	1.38629	1.38629		
				3.53100	3.53100	1.46634	0.26236 1.09861	0.40547
		6		2.54191	0.34386	2.37646	6	
				1.38629	1.38629	1.38629		
			0.06034	3.53100	3.53100	1.46634	0.26236 1.09861	0.40547

Query: fasta2.aln [M=459] Scores for complete sequences (score includes all domains): --- full sequence --- --- best 1 domain --- -#dom-E-value score bias E-value score bias exp N Sequence Description

0 6700.2 173.8 2e-176 5	574.1 5.1 1	12.5 12	selected bases
-------------------------	-------------	---------	----------------

Domain annotation for each sequence (and alignments):

>> selected bases

#		score	bias	c-Evalue	i-Evalue	hmmfrom	hmm to		alifrom	ali to		envfrom	env to		acc
1	!	573.1	5.2	4e-176	4e-176	1	459	[]	271076	271534		271076	271534		1.00
2	!	573.0	5.0	4.3e-176	4.3e-176	1	459	[]	446265	446723		446265	446723		1.00
3	!	568.7	5.6	8.5e-175	8.5e-175	1	459	[]	507385	507843		507385	507843		1.00
4	!	574.1	5.1	2e-176	2e-176	1	459	[]	657246	657704		657246	657704		1.00
5	!	574.1	5.1	2e-176	2e-176	1	459	[]	864041	864499		864041	864499		1.00
6	!	574.1	5.1	2e-176	2e-176	1	459	[]	1082125	1082583		1082125	1082583		1.00
7	!	574.1	5.1	2e-176	2e-176	1	459	[]	1085337	1085795		1085337	1085795		1.00
8	!	566.2	4.5	4.9e-174	4.9e-174	1	459	[]	1178594	1179052		1178594	1179052		1.00
9	!	574.1	5.1	2e-176	2e-176	1	459	[]	1235316	1235774		1235316	1235774		1.00
10	!	570.5	4.9	2.5e-175	2.5e-175	1	459	[]	1367421	1367879		1367421	1367879		1.00
11	!	566.2	4.5	4.9e-174	4.9e-174	1	459	[]	3122530	3122988		3122530	3122988		1.00
12	1	481.9	4.3	1.9e-148	1.9e-148	1	459	[]	3551516	3551974	• •	3551516	3551974	•••	1.00

#### score: The bit score for this domain.

```
hmmsearch :: search profile(s) against a sequence database
HMMER 3.0 (March 2010); http://hmmer.org/
Copyright (C) 2010 Howard Hughes Medical Institute.
Freely distributed under the GNU General Public License (GPLv3).
query HMM file:
 is200.hmm
target seguence database: St test.dna
Query:
 fasta2.aln [M=459]
Scores for complete sequences (score includes all domains):
 --- full sequence --- --- best 1 domain --- -#dom-
 E-value score bias E-value score bias exp N Sequence Description
 0 6700.2 173.8 2e-176 574.1 5.1 12.5 12 selected bases
Domain annotation for each sequence (and alignments):
>> selected, bases
 #
 score bias c-Evalue i-Evalue hmmfrom hmm to
 alifrom ali to envfrom env to
 _____ ____

 ----- -----
 1 ! 573.1 5.2 4e-176
 4e-176
 459 [] 271076 271534 .. 271076 271534 .. 1.00
 1
 2 ! 573.0 5.0 4.3e-176 4.3e-176 1
3 ! 568.7 5.6 8.5e-175 8.5e-175 1
 459 [] 446265 446723 .. 446265 446723 .. 1.00
 1 459 [] 507385 507843 .. 507385 507843 .. 1.00
 4 ! 574.1 5.1 2e-176 2e-176
 1 459 [] 657246 657704 .. 657246 657704 .. 1.00
 5 ! 574.1 5.1 2e-176 2e-176
 1
 459 [] 864041 864499 .. 864041 864499 .. 1.00
 6 ! 574.1 5.1 2e-176 2e-176
 1
 459 [] 1082125 1082583 .. 1082125 1082583 .. 1.00
 7 ! 574.1 5.1 2e-176
 2e-176
 1 459 [] 1085337 1085795 .. 1085337 1085795 .. 1.00
 8 ! 566.2 4.5 4.9e-174 4.9e-174
 1
 459 [] 1178594 1179052 .. 1178594 1179052 .. 1.00
 9 ! 574.1 5.1 2e-176 2e-176
 1 459 [] 1235316 1235774 .. 1235316 1235774 .. 1.00
 10 ! 570.5 4.9 2.5e-175 2.5e-175
 1 459 [] 1367421 1367879 .. 1367421 1367879 .. 1.00
 1 459 [] 3122530 3122988 .. 3122530 3122988 .. 1.00
 11 ! 566.2 4.5 4.9e-174 4.9e-174
 1 459 [] 3551516 3551974 .. 3551516 3551974 .. 1.00
 12 ! 481.9 4.3 1.9e-148 1.9e-148
```

acc

____

score: The bit score for this The biased composition domain. (null2) score correction that was applied to the domain bit # hmmsearch :: search profile(s) SCORe. # HMMER 3.0 (March 2010); http://hmmer.org/ # Copyright (C) 2010 Howard Hughes Medical Institute. # Freely distributed under the GNU General Public License (GPLv3). # query HMM file: is200.hmm # target seguence database: St/test.dna fasta2.aln [M=459] Ouerv: Scores for complete sequences (score includes all domains): --- full sequence ------ best 1 domain ----#dom-E-value score bias Z-value score bias exp N Sequence Description ----- -----0 6700.2 173.8 / 2e-176 574.1 5.1 12.5 12 selected bases Domain annotation for each sequence (and alignments): >> selected, bases, # score bias c-Evalue i-Evalue hmmfrom hmm to alifrom ali to envfrom env to acc _____ ____ _____ ___ _____ _ ____ ____ 1 ! 573.1 5.2 4e-176 459 [] 271076 271534 .. 271076 271534 .. 1.00 4e-176 1 2 ! 573.0 5.0 4.3e-176 4.3e-176 459 [] 446265 446723 .. 446265 446723 .. 1.00 1 3 ! 568.7 5.6 8.5e-175 8.5e-175 1 459 [] 507385 507843 .. 507385 507843 .. 1.00 4 ! 574.1 5.1 2e-176 2e-176 1 459 [] 657246 657704 .. 657246 657704 .. 1.00 5 ! 574.1 5.1 2e-176 1 2e-176 459 [] 864041 864499 .. 864041 864499 .. 1.00 6 ! 574.1 5.1 2e-176 2e-176 1 459 [] 1082125 1082583 .. 1082125 1082583 .. 1.00 7 ! 574.1 5.1 2e-176 2e-176 1 459 [] 1085337 1085795 .. 1085337 1085795 .. 1.00 8 ! 566.2 4.5 4.9e-174 4.9e-174 1 459 [] 1178594 1179052 .. 1178594 1179052 .. 1.00 1 459 [] 1235316 1235774 .. 1235316 1235774 .. 1.00 9 ! 574.1 5.1 2e-176 2e-176 10 ! 570.5 4.9 2.5e-175 2.5e-175 1 459 [] 1367421 1367879 .. 1367421 1367879 .. 1.00 1 459 [] 3122530 3122988 .. 3122530 3122988 .. 1.00 11 ! 566.2 4.5 4.9e-174 4.9e-174 1 459 [] 3551516 3551974 .. 3551516 3551974 .. 1.00 12 ! 481.9 4.3 1.9e-148 1.9e-148

score: domair	The bit s า.	core fo	r this	(null2) sc		ection that			permissiv	e m	nal E-valu		ble	;
<pre># HMMER # Copyr # Freel # # query</pre>	8 3.0 (M right (C .y distr  y HMM fi	arch 2 ) 2010 ibuted  le:	Howard H under th	acoro	rg/ al Insti al Publi  .hmm			GPL⊽3) 	calculated space tha Evalue. T uses the r	The I on In th he o num	r domain conditiona a smaller ne indeper conditiona ber of tar orting thre	r search ndent al E-value gets that	e is	
 E-⊽ 	<pre>#</pre>													
	annotat cted, b		/	equence (and	alignme	nts):								
#		V.	_	e i-Evalue H	hmmfrom	hmm to		alifrom	ali to		envfrom	env to		acc
1 ! 2 ! 3 ! 5 ! 6 ! 7 ! 8 ! 9 ! 10 ! 11 !	573.1 573.0 568.7 574.1 574.1 574.1 574.1 574.1 566.2 574.1 570.5 566.2	5.1 4.9 4.5	2e-176 2.5e-175 4.9e-174	4.3e-176 8.5e-175 2e-176 2e-176 2e-176 2e-176 4.9e-174 2e-176 2.5e-175 4.9e-174	1 1 1 1 1 1 1 1 1	459 459 459 459 459		446265 507385 657246 864041 1082125 1085337 1178594 1235316 1367421 3122530	446723 507843 657704 864499 1082583 1085795 1179052 1235774 1367879 3122988	· · · · · · · · · · · · · · · · · · ·	271076 446265 507385 657246 864041 1082125 1085337 1178594 1235316 1367421 3122530	446723 507843 657704 864499 1082583 1085795 1179052 1235774 1367879 3122988	· · · · · · · · · · · · ·	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
12 !	481.9	4.3	1.9e-148	1.9e-148	1	459	[]	3551516	3551974	••	3551516	3551974	••	1.00

domain # hmmse # HMMEF	earch :: 2 3.0 (Ma	searc rch 2	h profile 010); http	(null2) s was app score.		ction tha domain		]	permissi [,] this parti may be.	ve m culai The	nal E-valu easure of domain conditiona a smaller	how relia		
<pre># Freel # # query # targe #</pre>	ly distri  / HMM fil st sequen 	buted  e: ce da 	under the	/	al Publi		3e ( 	(GPLv3)	uses t pass t d	alue omp eceiv oma	ndepende that the s arison wo red if this in envelop ding any c	equence/ uld have were the pe found i	profi only n it,	le
Query:		sta2. lete	-	9] (score ind	ludes al	1 domain	).				ent measu			
	full seq			best 1 do						- U	articular d			
E-v	value sc	ore	bias 🗵	-value sco	ore bias	exp	N	Sequen	-		endent E-			
				- 176 574	1 1 5 1,	10 5	10		1   []	ump	er or large		laiud	રા ા
	0 670	0.2 1	73.8 2	2e-176 574	1.1 5.1	12.5	12	select	ou su.	atab	er of targe ase.		laryc	÷L
							12	select	ou su.		-		laryc	÷l
	annotati	on fo		2e-176 574 quence (and			12	select	ou su.		-		laryc	÷L
	annotati ected, ba	on fo ses/	y each see	quence (and	ł alignme	nts):	12	-	d	atab	ase.	_		÷L
	annotati ected, ba	on fo ses/			ł alignme	nts):	12	-	ou su.	atab	-	_	larye	acc
	annotati ected, ba	on fo ses/	y each see	quence (and	ł alignme	nts): hmm to		alifrom	ali to	atab	envfrom	env to		acc
>> sele # 	annotati ected ba score	on fo ses bias	r each seo c-Evalue	quence (and i-Evalue	d alignme hmmfrom	nts): hmm to  459	[]	alifrom  271076	ali to 271534	atab	envfrom 271076	env to  271534		acc  1.00
>> sele #  1 !	annotati ected ba score 	on fo ses bias  5.2	c-Evalue 4e-176	uence (and i-Evalue  4e-176	d alignme hmmfrom  1	nts): hmm to  459 459	[]	alifrom  271076 446265	ali to	atab	envfrom 271076 446265	env to		acc  1.00 1.00
>> sele #  1 ! 2 !	annotati ected ba score 	on fo ses bias  5.2 5.0	r each seo c-Evalue  4e-176 4.3e-176	uence (and i-Evalue 4e-176 4.3e-176	d alignme hmmfrom  1 1	nts): hmm to  459 459	[]	alifrom 271076 446265 507385	ali to 271534 446723	atab	envfrom 271076 446265 507385	env to  271534 446723		acc  1.00 1.00 1.00
>> sele #  1 ! 2 ! 3 !	annotati score 573.1 573.0 568.7	on fo ses bias  5.2 5.0 5.6	r each seo c-Evalue 4e-176 4.3e-176 8.5e-175	uence (and i-Evalue 4e-176 4.3e-176 8.5e-175	d alignme hmmfrom 1 1	nts): hmm to  459 459 459 459	[] [] []	alifrom 271076 446265 507385	ali to 271534 446723 507843 657704	atab	envfrom 271076 446265 507385 657246	env to 271534 446723 507843		acc 1.00 1.00 1.00 1.00
>> sele #  1 ! 2 ! 3 ! 4 !	annotati score 573.1 573.0 568.7 574.1	on fo ses bias  5.2 5.0 5.6 5.1	each seo c-Evalue 4e-176 4.3e-176 8.5e-175 2e-176	uence (and i-Evalue 4e-176 4.3e-176 8.5e-175 2e-176	d alignme hmmfrom  1 1 1 1	nts): hmm to  459 459 459 459	[] [] []	alifrom 271076 446265 507385 657246 864041	ali to 271534 446723 507843 657704 864499	atab	envfrom 271076 446265 507385 657246	env to 271534 446723 507843 657704 864499		acc 1.00 1.00 1.00 1.00 1.00
>> sele #  1 ! 2 ! 3 ! 4 ! 5 !	annotati score 573.1 573.0 568.7 574.1 574.1	on fo ses bias  5.2 5.0 5.6 5.1 5.1	each seo c-Evalue 4e-176 4.3e-176 8.5e-175 2e-176 2e-176	uence (and i-Evalue 4e-176 4.3e-176 8.5e-175 2e-176 2e-176	d alignme hmmfrom  1 1 1 1 1	nts): hmm to 459 459 459 459 459	[] [] [] [] [] []	alifrom 271076 446265 507385 657246 864041 1082125 1085337	ali to 271534 446723 507843 657704 864499 1082583 1085795	atab	envfrom 271076 446265 507385 657246 864041 1082125 1085337	env to 271534 446723 507843 657704 864499 1082583 1085795		acc 1.00 1.00 1.00 1.00 1.00 1.00
>> sele #  1 ! 2 ! 3 ! 4 ! 5 ! 6 !	annotati score 573.1 573.0 568.7 574.1 574.1 574.1	on fo ses bias  5.2 5.0 5.6 5.1 5.1 5.1 5.1 4.5	each seo c-Evalue 4e-176 4.3e-176 8.5e-175 2e-176 2e-176 2e-176	uence (and i-Evalue 4e-176 4.3e-176 8.5e-175 2e-176 2e-176 2e-176	d alignme hmmfrom  1 1 1 1 1 1	nts): hmm to 459 459 459 459 459 459	[] [] [] [] [] []	alifrom 271076 446265 507385 657246 864041 1082125 1085337	ali to 271534 446723 507843 657704 864499 1082583 1085795	atab	envfrom 271076 446265 507385 657246 864041 1082125	env to 271534 446723 507843 657704 864499 1082583 1085795		acc 1.00 1.00 1.00 1.00 1.00 1.00
>> sele #  1 ! 2 ! 3 ! 4 ! 5 ! 6 ! 7 ! 8 ! 9 !	annotati score 573.1 573.0 568.7 574.1 574.1 574.1 574.1 574.1 574.1 574.1 574.1	on fo ses bias  5.2 5.0 5.6 5.1 5.1 5.1 5.1 4.5 5.1	r each seo c-Evalue 4e-176 4.3e-176 8.5e-175 2e-176 2e-176 2e-176 4.9e-176	uence (and i-Evalue 4e-176 4.3e-176 8.5e-175 2e-176 2e-176 2e-176 2e-176	d alignme hmmfrom  1 1 1 1 1 1 1 1 1 1	nts): hmm to 459 459 459 459 459 459 459	[] [] [] [] [] [] []	alifrom 271076 446265 507385 657246 864041 1082125 1085337 1178594 1235316	ali to 271534 446723 507843 657704 864499 1082583 1085795 1179052 1235774	atab	envfrom 271076 446265 507385 657246 864041 1082125 1085337 1178594 1235316	env to 271534 446723 507843 657704 864499 1082583 1085795 1179052 1235774		acc 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.
>> sele #  1 ! 2 ! 3 ! 4 ! 5 ! 6 ! 7 ! 8 ! 9 ! 10 !	annotati score 573.1 573.0 568.7 574.1 574.1 574.1 574.1 574.1 574.1 574.1 574.1 574.1 574.1 574.1	on fo ses 5.2 5.0 5.6 5.1 5.1 5.1 5.1 4.5 5.1 4.5 5.1 4.9	each seo c-Evalue 4e-176 4.3e-176 8.5e-175 2e-176 2e-176 2e-176 2e-176 4.9e-174	uence (and i-Evalue 4e-176 4.3e-176 8.5e-175 2e-176 2e-176 2e-176 2e-176 4.9e-174	d alignme hmmfrom  1 1 1 1 1 1 1 1 1 1	nts): hmm to 459 459 459 459 459 459 459 459	[] [] [] [] [] [] [] []	alifrom 271076 446265 507385 657246 864041 1082125 1085337 1178594 1235316 1367421	ali to 271534 446723 507843 657704 864499 1082583 1085795 1179052 1235774 1367879	atab	envfrom 271076 446265 507385 657246 864041 1082125 1085337 1178594 1235316 1367421	env to 271534 446723 507843 657704 864499 1082583 1085795 1179052 1235774 1367879		acc 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.
>> sele #  1 ! 2 ! 3 ! 4 ! 5 ! 6 ! 7 ! 8 ! 9 !	annotati score 573.1 573.0 568.7 574.1 574.1 574.1 574.1 574.1 574.1 574.1 574.1	on fo ses bias  5.2 5.0 5.6 5.1 5.1 5.1 5.1 4.5 5.1	r each seo c-Evalue 4e-176 4.3e-176 8.5e-175 2e-176 2e-176 2e-176 4.9e-176	uence (and i-Evalue 4e-176 4.3e-176 8.5e-175 2e-176 2e-176 2e-176 4.9e-176	d alignme hmmfrom  1 1 1 1 1 1 1 1 1 1	nts): hmm to 459 459 459 459 459 459 459 459 459 459	[] [] [] [] [] [] [] []	alifrom  271076 446265 507385 657246 864041 1082125 1085337 1178594 1235316 1367421 3122530	ali to 271534 446723 507843 657704 864499 1082583 1085795 1179052 1235774 1367879 3122988	atab	envfrom 271076 446265 507385 657246 864041 1082125 1085337 1178594 1235316	env to 271534 446723 507843 657704 864499 1082583 1085795 1179052 1235774 1367879 3122988		acc 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.

<pre>score: The bit score for th domain. # hmmsearch :: search # HMMER 3.0 (March 201</pre>	(null2) s was app profile(s) score. 0); http://hmmer.		pern this , may	<ul> <li>"conditional E-valumissive measure of particular domain</li> <li>be. The conditional condititatical conditional condition</li></ul>	f how reliable al E-value is
<pre># Copyright (C) 2010 H # Freely distributed u #</pre>	nder the GNU Gener is20 base: St to n [M=459] quences (score in as E-value sco .8 2e-176 57	ral Public License 0.hmm est.dna 	uses pass	lue The Independent s t s t comparison wo received if this domain envelop excluding any o stringent measu this particular d independent E-	were the only
Domain annotation for >> selected, bases,	each sequence (and	d alignments):			
# score bias c	-Evalue i-Evalue	hmmfrom hmm to	alifrom al	li to envfrom	env to acc
	4e-176 4e-176 .3e-176 4.3e-176	1 459 []	446265 44		446723 1 00
3 ! 568.7 5.6 8 4 ! 574.1 5.1	.5e-175 8.5e-175 2e-176 2e-176			07843 507385 57704 657246	
5 ! 574.1 5.1	2e-176 2e-176		864041 86	54499 864041	
6 ! 574.1 5.1	2e-176 2e-176			32583 1082125	
7 ! 574.1 5.1 8 ! 566.2 4.5 4	2e-176 2e-176 .9e-174 4.9e-174			35795 1085337 79052 1178594	
9 ! 574.1 5.1	2e-176 2e-176			35774 1235316	the everall alignment is
10 ! 570.5 4.9 2	.5e-175 2.5e-175			57879 1367421	
	.9e-174 4.9e-174 .9e-148 1.9e-148			22988 3122530 51974 3551516	

#	tp://hmmer.org/ Hughes Medical Institute. he GNU General Public License (G is200.hmm	uses t value that the sequence/profile
# target sequence database: #	St_test.dna	pass t received if this were the only domain envelope found in it,
full sequence	s (score includes all domains): best 1 domain#dom-	excluding any others. This is a stringent measure of how reliable this particular domain may be. The independent E-value uses the total
Domain annotation for each so >> selected, bases,	equence (and alignments):	
	e i-Evalue hmmfrom hmm to a	alifrom ali to envfrom env to acc
3 ! 568.7 5.6 8.5e-17 4 ! 574.1 5.1 2e-17 5 ! 574.1 5.1 2e-17 6 ! 574.1 5.1 2e-17 7 ! 574.1 5.1 2e-17 8 ! 566.2 4.5 4.9e-17 9 ! 574.1 5.1 2e-17	6 $4.3e-176$ $1$ $459$ $[]$ $5$ $8.5e-175$ $1$ $459$ $[]$ $6$ $2e-176$ $1$ $459$ $[]$ $6$ $2e-176$ $1$ $459$ $[]$ $6$ $2e-176$ $1$ $459$ $[]$ $1$ $6$ $2e-175$ $1$ $459$ $[]$ $1$ $6$ $2 e-175$ $1$ $459$ $[]$ $1$ $6$ $2 e-175$ $1$ $459$ $[]$ $1$ $4$ $4.9e-174$ $1$ $459$ $1$ $1$ <td>507385 507843 507388 The mean posterior</td>	507385 507843 507388 The mean posterior
	of L residues. subsequence	umbered 1L for a sequence . The envelope defines a e for which their is substantial ass supporting a homologous

<pre>score: The bit score for this domain. # hmmsearch :: search profile(s) # HMMER 3.0 (March 2010); http://domain.</pre>		The "conditional E-value", a permissive measure of how reliable this particular domain may be. The conditional E-value is calculated on a smaller search
<pre># Copyright (C) 2010 Howard Hugh # Freely distributed under the ( #</pre>	hes Medical Institute. GNU General Public License (GPLv3 is200.hmm St_test.dna 	<ul> <li>space</li> <li>Evalue</li> <li>uses t</li> <li>pass t</li> <li>The "independent E-value", the E-value value that the sequence/profile</li> <li>comparison would have</li> <li>received if this were the only</li> <li>domain envelope found in it,</li> <li>excluding any others. This is a</li> <li>stringent measure of how reliable</li> <li>this particular domain may be. The</li> <li>independent E-value uses the total</li> <li>number of targets in the target</li> <li>database.</li> </ul>
1 ! 573.1 5.2 4e-176 2 ! 573.0 5.0 4.3e-176 3 ! 568.7 5.6 8.5e-175 4 ! 574.1 5.1 2e-176 5 ! 574.1 5.1 2e-176	i-Evalue hmmfrom hmm to alifr 4e-176 1 459 [] 2710 4.3e-176 1 458 [] 4462 8.5e-175 1 459 [] 5073 2e-176 1 459 [] 6572 2e-176 1 459 [] 8640	65       446723        1       00         85       507843        507388        1       00         46       657704        657246       The mean posterior         41       864499        864041       probability of aligned
9 ! 574.1 5.1 2e-176 10 ! 570.5 4.9 2.5e-175	2e-176       1       459 [] 108533         4.9e-174       1       459 [] 11785         2e-176       1       459 [] 12353         2.5e-175       1       459 [] 136743         4.9e-174       1       459 [] 136743         1       459 [] 136743       1         4.9e-174       1       459 [] 136743         1       459 [] 136744       1         4.9e-174       1       459 [] 136744         1       159 [] 136744       1         1       459 [] 136744       1         1       159 [] 136744       1         1       159 [] 136744       1         1       159 [] 136744       1         1       159 [] 136744       1         4.9e-174       1       459 [] 136744         1       159 [] 136744       1         1       159 [] 136744       1         1       159 [] 136744       1         1       159 [] 136744       1         1       159 [] 136744       1         1       159 [] 136744       1         1       159 [] 136744       1         1       169 [] 160 [] 160 [] 160 [] 160 [] 160 [] 160 [] 160 [] 160 []	$\begin{array}{c} 25  1082583  \ldots  1082125 \\ 37  1085795  \ldots  1085337 \\ 94  1179052  \ldots  1178594 \\ 16  1235774  \ldots  1235316 \\ 21  1367879  \ldots  1367421 \\ 21  235774  \ldots  1367421 \\ 21  235789  \ldots  1367421 \\ 21  235789  \ldots  1.00 \\ 21  23598  \ldots  1.00 \\ 31  31  31678  31  31678  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  3168  $

<pre>score: The bit score for this domain. # hmmsearch :: search profile(s # HMMER 3.0 (March 2010); http:: # Copyright (C) 2010 Howard Hugh # Freely distributed under the compared to the second sec</pre>	//hmmer.org/ hes Medical Institute.		of how reliable n onal E-value is ller search ndent E-value", the E-
E-value score bias E-value	score includes all domains): best 1 domain#dom- alue score bias exp N Se	equence Des	e sequence/profile would have is were the only elope found in it, y others. This is a asure of how reliable r domain may be. The E-value uses the total rgets in the target
1 ! 573.1 5.2 4e-176 2 ! 573.0 5.0 4.3e-176	i-Evalue hmmfrom hmm to ali 4e-176 1 459 [] 27 4.3e-176 1 458 [] 44	from ali to envfr 1076 271534 2710 16265 446723 4462	76 27/1534 1.00 65 446723 1.00
4 ! 574.1 5.1 20-176 5 ! 574.1 5 1 20-176 The start/end of the alignment of this domain with respect to the profile	2e-176       1       459       []       65         2e-176       1       459       []       86         2e-176       1       459       []       108         2e-176       1       459       []       108         2e-176       1       459       []       108         .9e-174       1       459       []       117         .2e-176       1       459       []       123	07385       507843       5073         07246       657704       6572         54041       864499       8640         32125       1082583       10821         35337       1085795       10853         78594       1179052       11785         35316       1235774       12353	The mean posterior probability of aligned residues in the alignment; a measure of how reliable the overall alignment is
10 ! 570.5 4.9 2.5e-175 11 ! 566.2 4.5 4.9e-174 12 ! 481 The start/end of the a of this domain with res the sequence	lignment sequence, numb spect to of L residues. Th subsequence fo		he .974 1.00