Hidden Markov Models and Bioinformatics

- Docentföreläsning

Marina Alexandersson

Göteborg, Dec 9, 2004

ITWM
Fraunhofer CHALMERS
Research Centre
Industrial Mathematics

2004-12-09

Some genetic history

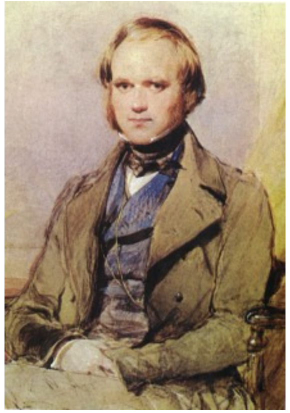
1850

1900

1950

2000

1859


Charles Darwin

1859: Origin of the Species

Some genetic history

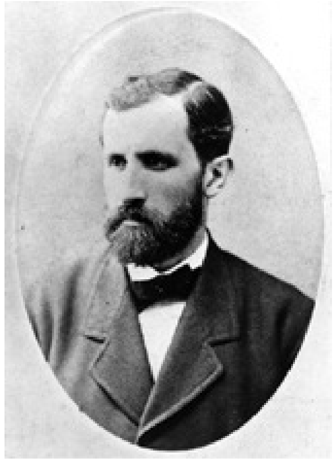
1850
1900
1950
2000
1865

Gregor Mendel

1865: Gregor Mendel’s peas

Some genetic history

1850
1900
1950
2000
1869



Courtesy of Herrn Courvoisier, Portrait-Sammlung, University of Basel.
Noncommercial, educational use only.

Friedrich Miescher

1869: DNA isolated

Some genetic history

1850

1900

1950

2000

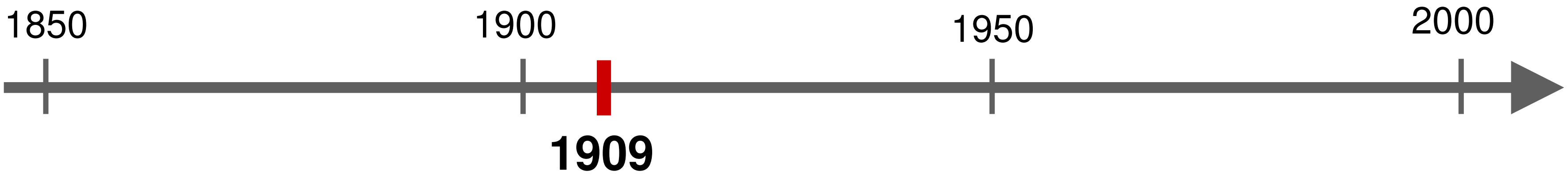
1902

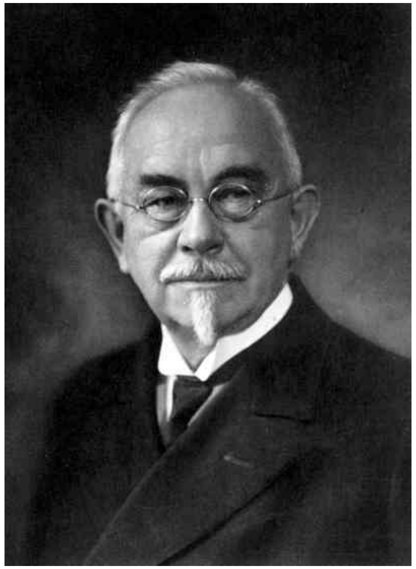
Walter Sutton

Theodor Boveri

1902: Chromosome theory of heredity

Some genetic history

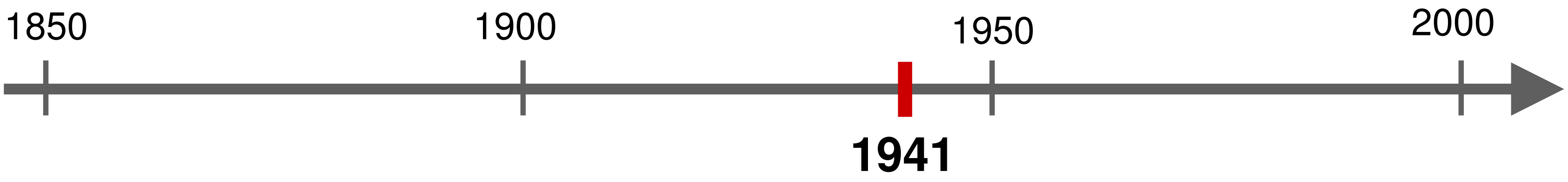




Willhem Johannsen

1909: the word ‘gene’ introduced

Some genetic history



George Beadle

Edward Tatum

1941: One gene, one enzyme

Some genetic history

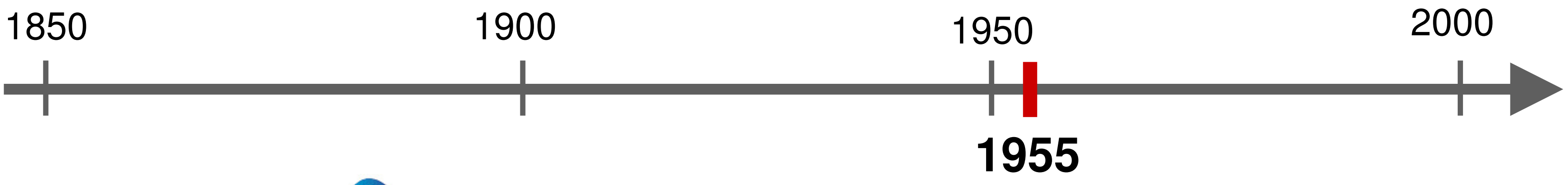
1850
1900
1950
2000
1953

James Watson

Francis Crick

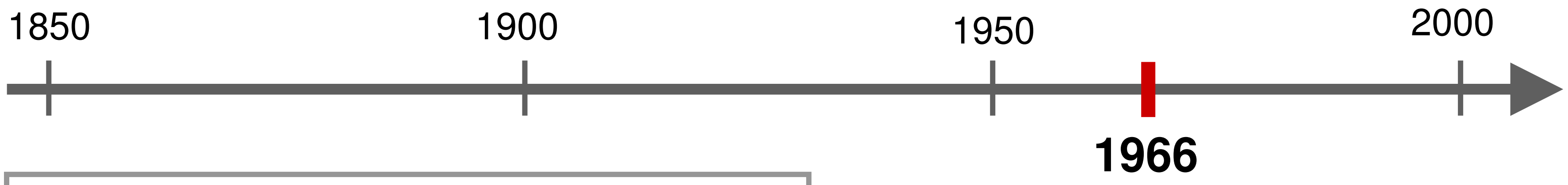
1953: The DNA double helix

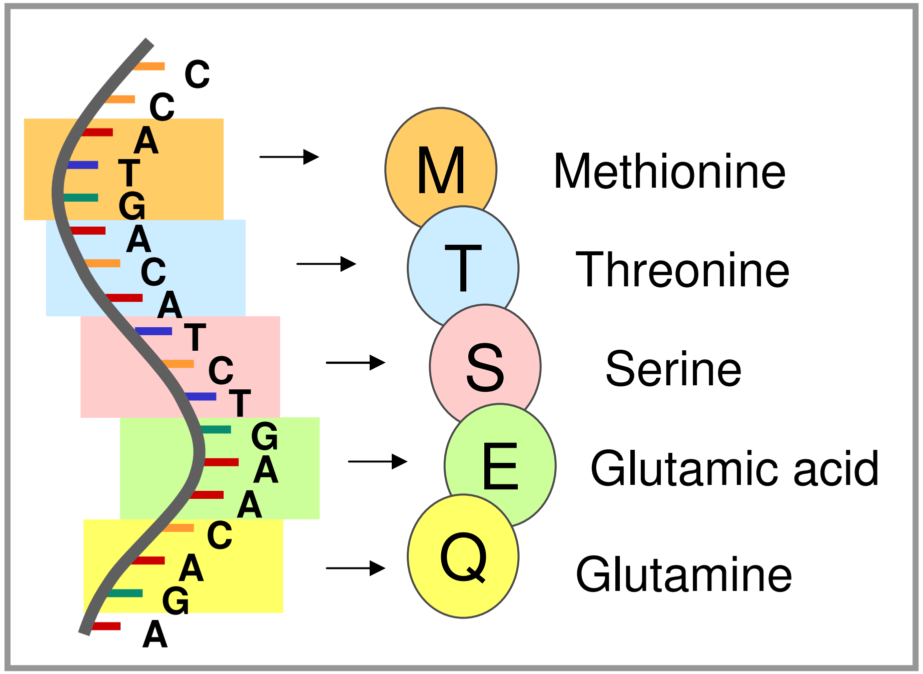
Some genetic history



1955: 46 human chromosomes

Some genetic history





1966: Genetic code cracked

Some genetic history

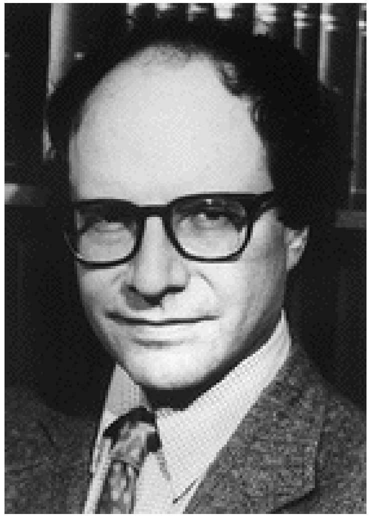
1850

1900

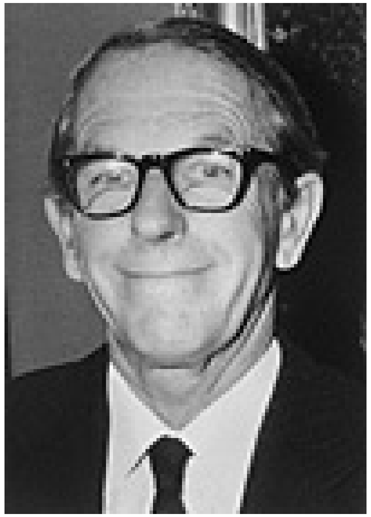
1950

2000

1977



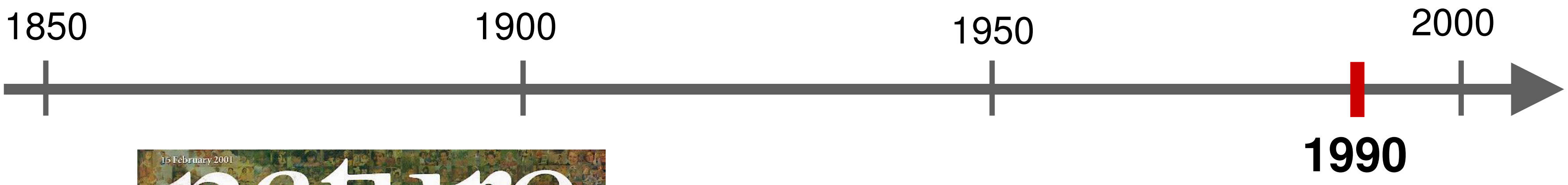
Walter Gilbert



Frederick Sanger

1977: DNA sequencing

Some genetic history



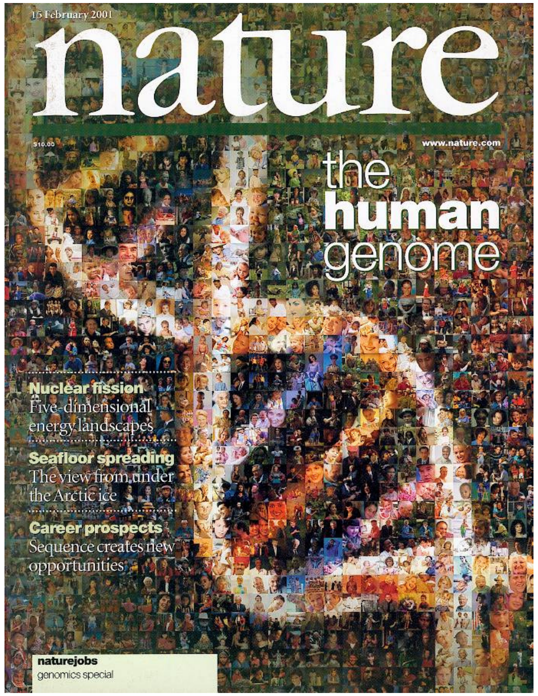
1850

1900

1950

2000

1990



15 February 2001

nature

$10.00

www.nature.com

the **human** genome

Nuclear fission
Five-dimensional energy landscapes

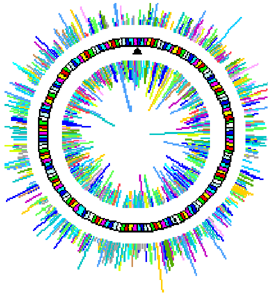
Seafloor spreading
The view from under the Arctic ice

Career prospects
Sequence creates new opportunities

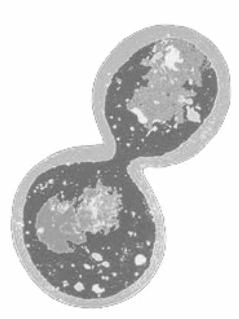
naturejobs
genomics special

1990: Human Genome Project

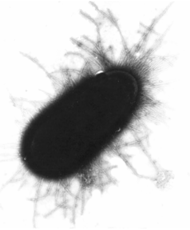
Sequencing history



1995: Two microbial genomes (1.8, 0.6)

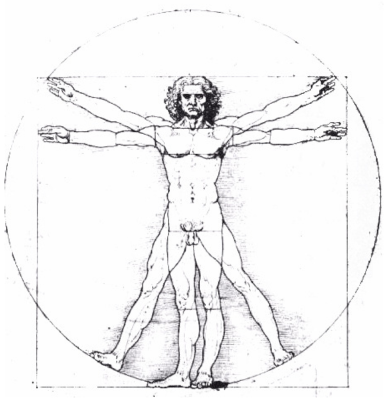


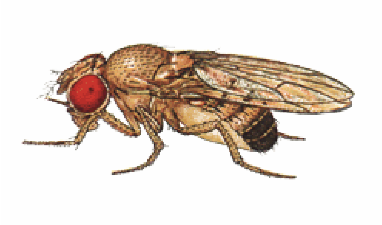
1996: Saccharomyces cerevisiae (12)



1997: E. coli (4.6)

1998: C. elegans (97)





2000: Drosophila melanogaster (180)

2003: Homo sapiens (3,200)

More to come...

Fraunhofer ITWM

CHALMERS
Research Centre
Industrial Mathematics

Whole genome analysis

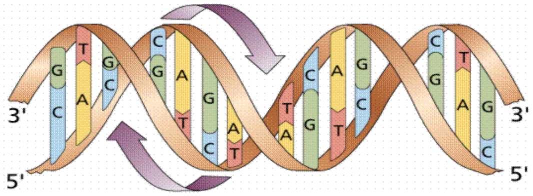
- Gene finding
- Sequence alignment
- Regulatory region discovery

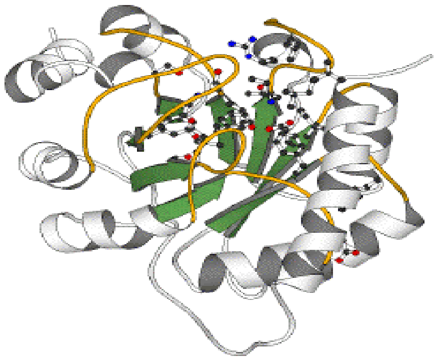
Whole genome analysis - Why?

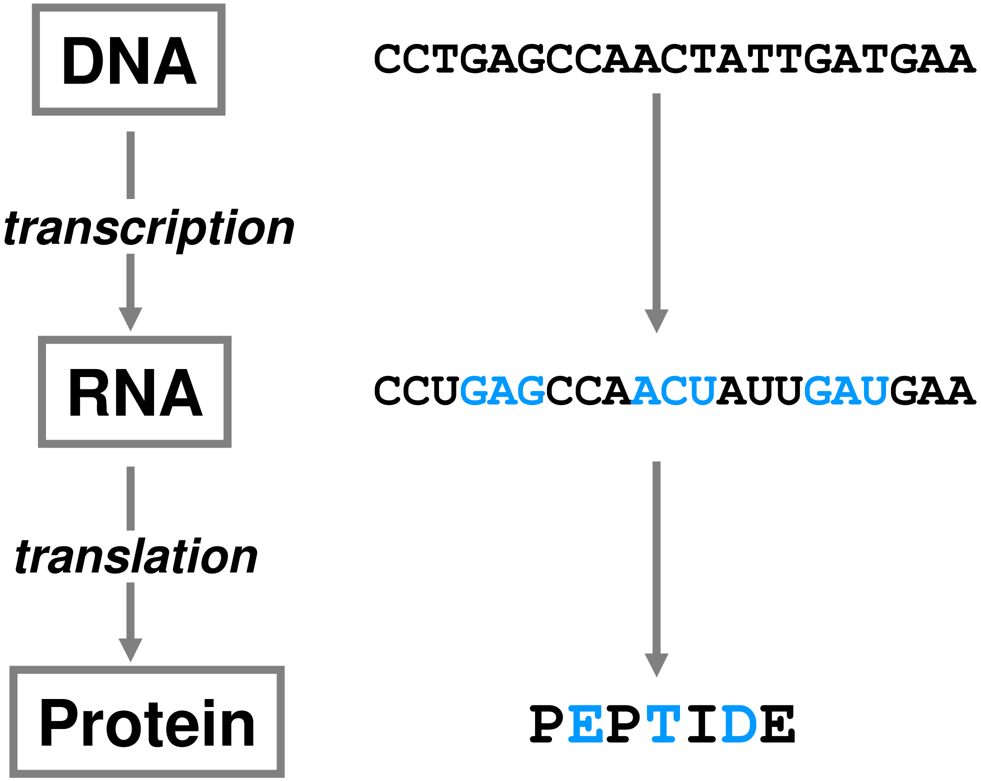
- Complete gene and protein sets
- Primary sequence of all genes
- Sequence relationships between genes and proteins
- Function of new proteins
- Transcriptional level of all genes
- Understanding methabolic pathways
- Trace disease genes
- ...

Gene finding

Gene expression







DNA

transcription

RNA

translation

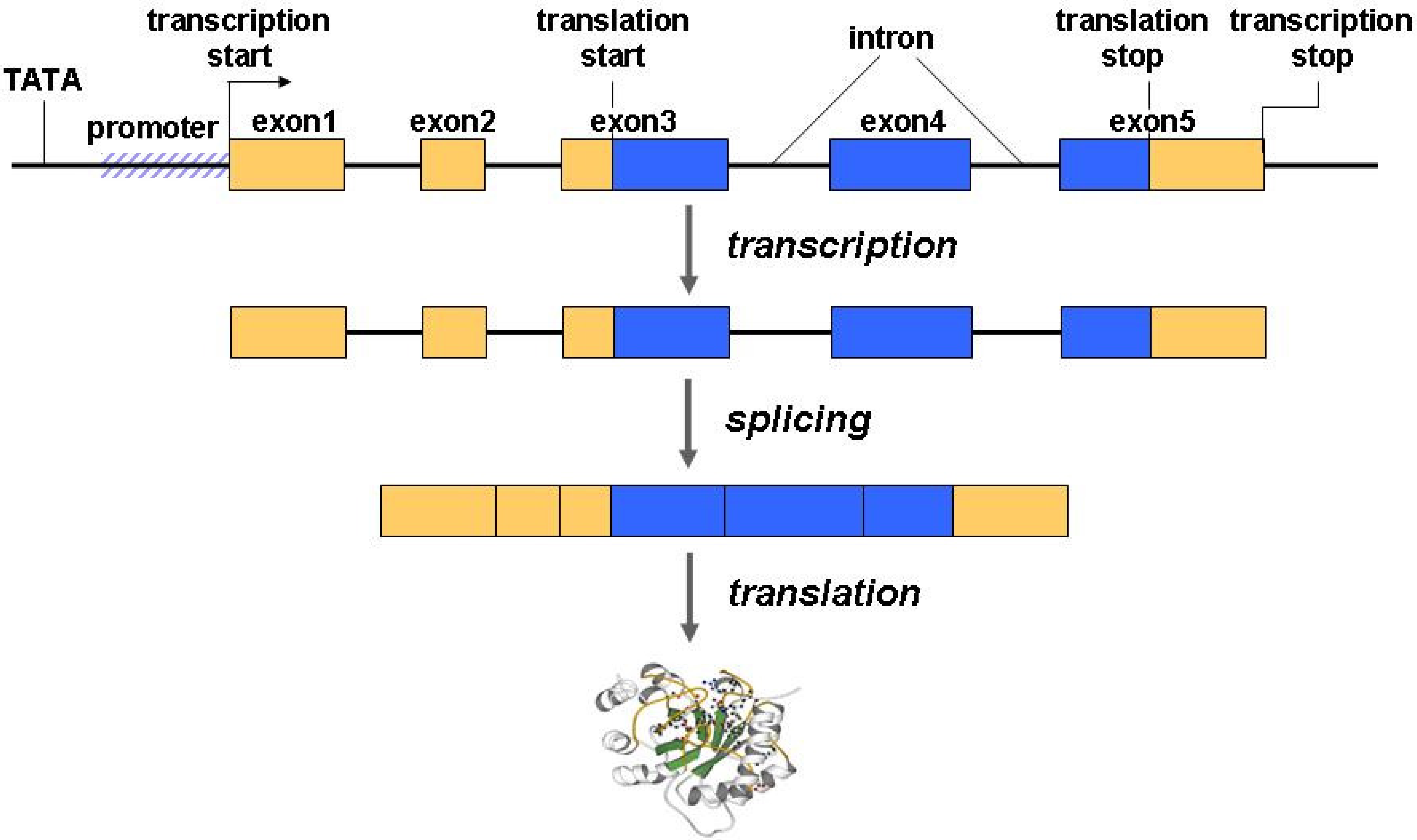
Protein

CCTGAGCCAACTATTGATGAA

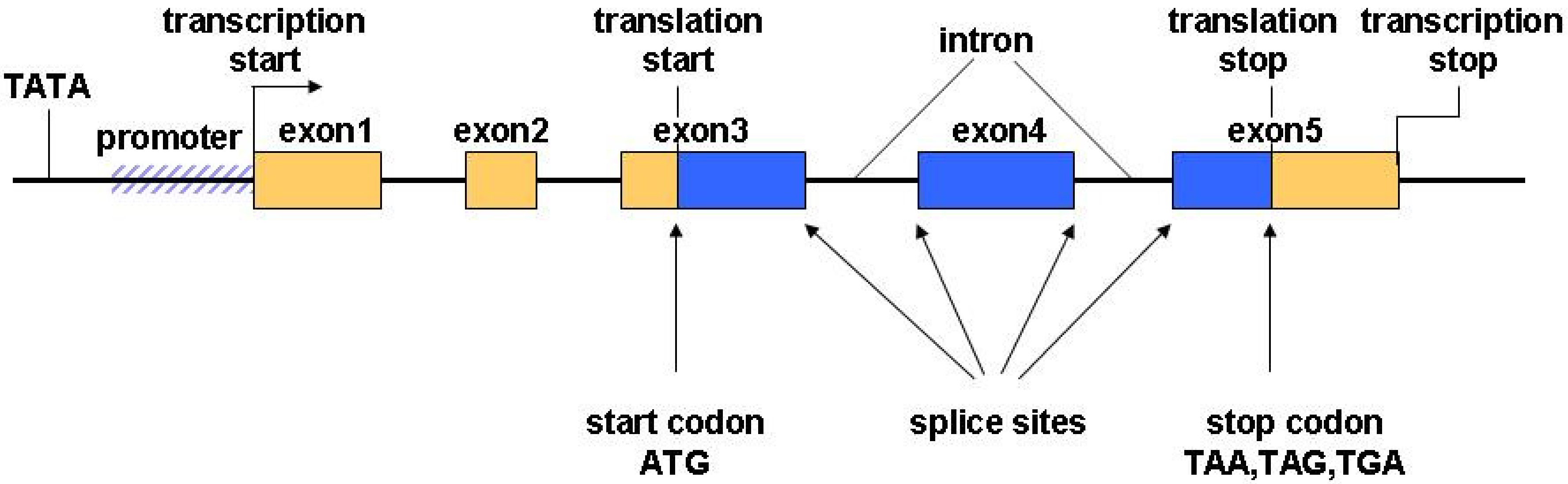
CCUGAGCCAACUAUUGAUGAA

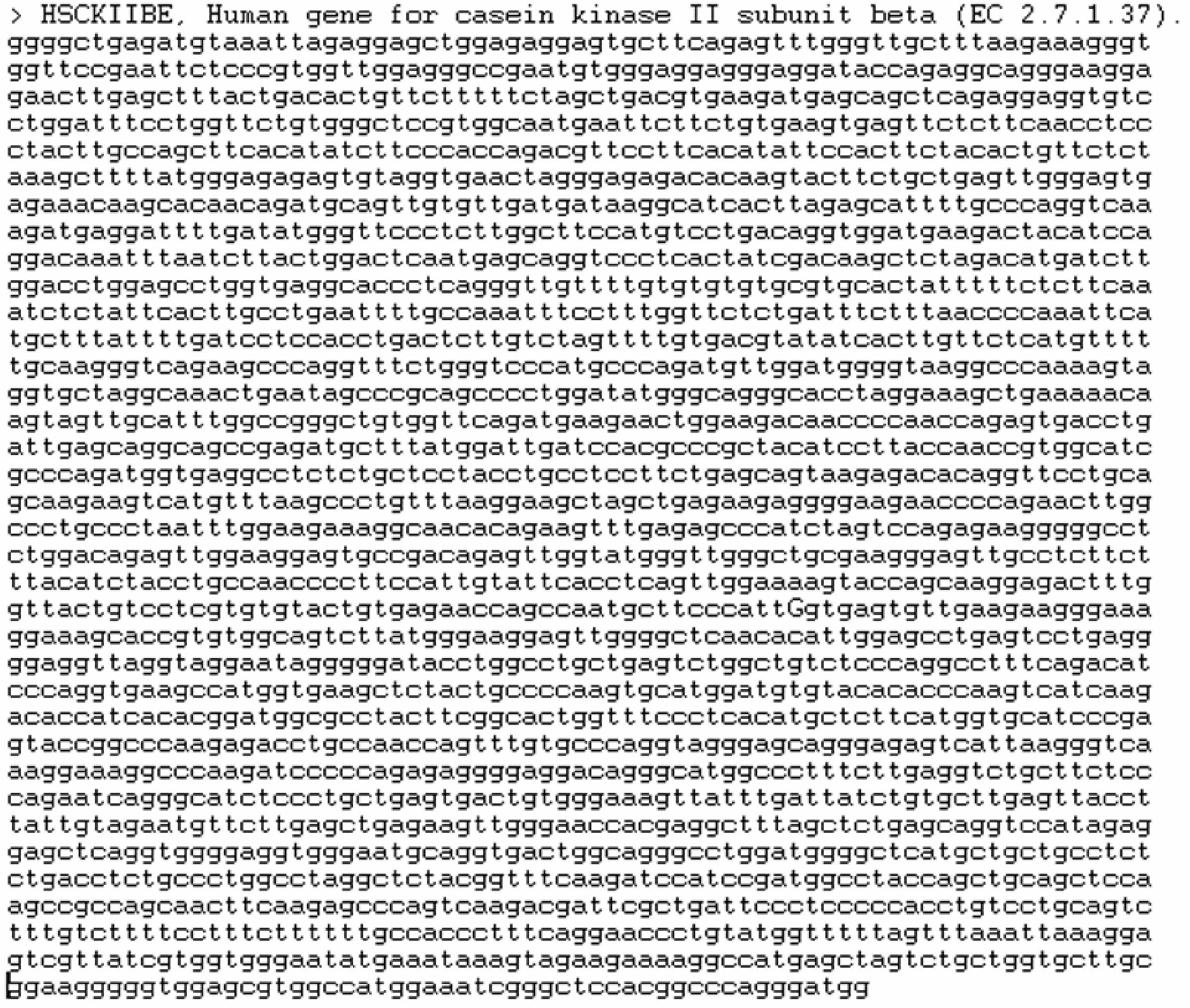
PEPTIDE

Gene structure

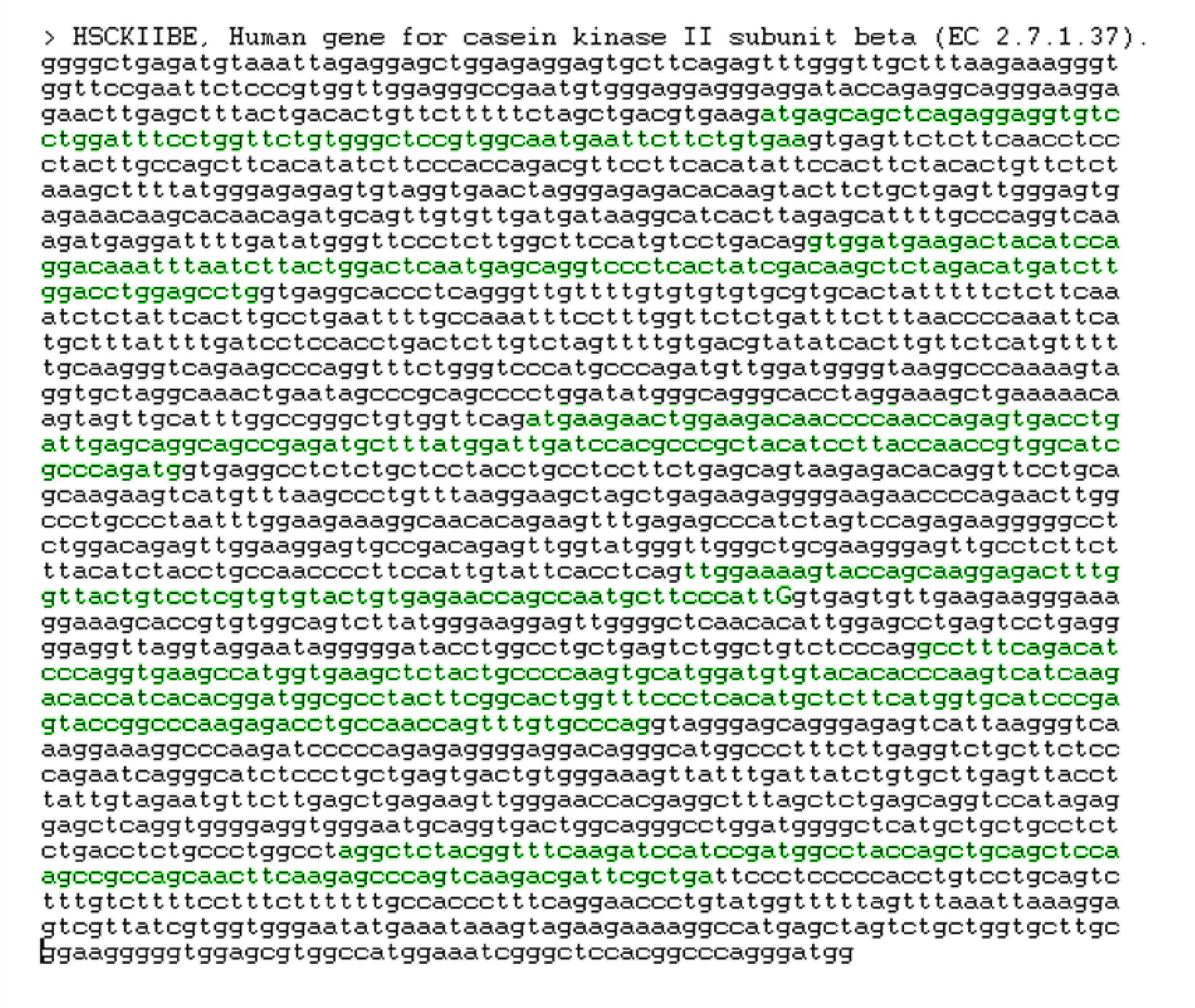


Finding genes





```
> HSCKIIBE, Human gene for casein kinase II subunit beta (EC 2.7.1.37).
ggggctgagatgtaaattagaggagctggagaggagtgcttcagagtttgggttgctttaagaaagggt
ggttccgaattctcccgtggttggagggccgaatgtgggaggagggaggataccagaggcagggaagga
gaacttgagctttactgacactgttctttttctagctgacgtgaagatgagcagctcagaggaggtgtc
ctggatttcctggttctgtgggctccgtggcaatgaattcttctgtgaagtgagttctcttcaacctcc
ctacttgccagcttcacatatcttcccaccagacgttccttcacatattccacttctacactgttctct
aaagcttttatgggagagagtgtaggtgaactagggagagacacaagtacttctgctgagttgggagtg
agaaacaagcacaacagatgcagttgtgttgatgataaggcatcacttagagcatttgcccaggtcaa
agatgaggatttgatatgggttccctcttggcttccatgtcctgacaggtggatgaagactacatcca
ggacaaatttaatcttactggactcaatgagcaggtccctcactatcgacaagctctagacatgatctt
ggacctggagcctggtgaggcaccctcagggttgttttgtgtgtgtgcgtgcactatttttctcttcaa
atctctattcacttgcctgaattttgccaaatttcctttggttctctgatttctttaaccccaaattca
tgctttattttgatcctccacctgactcttgtctagttttgtgacgtatatcacttgttctcatgtttt
tgcaagggtcagaagcccaggtttctgggtcccatgcccagatgttggatggggtaaggcccaaaagta
ggtgctaggcaaactgaatagcccgcagcccctggatatgggcagggcacctaggaaagctgaaaaaca
agtagttgcatttggccgggctgtggttcagatgaagaactggaagacaaccccaaccagagtgacctg
attgagcaggcagccgagatgctttatggattgatccacgcccgctacatccttaccaaccgtggcatc
gcccagatggtgaggcctctctgctcctacctgcctccttctgagcagtaagagacacaggttcctgca
gcaagaagtcatgtttaagccctgtttaaggaagctagctgagaagaggggaagaaccccagaacttgg
ccctgccctaatttggaagaaaggcaacacagaagtttgagagcccatctagtccagagaaggggccct
ctggacagagttggaaggagtgccgacagagttggtatgggttgggctgcgaagggagttgcctcttct
ttacatctacctgccaaccccttccattgtattcacctcagttggaaaagtaccagcaaggagactttg
gttactgtcctcgtgtgtactgtgagaaccagccaatgcttcccattGgtgagtgttgaagaaggaaa
ggaaagcaccgtgtggcagtcttatgggaaggagttggggctcaacacattggagcctgagtcctgagg
ggaggttaggtaggaatagggggatacctggcctgctgagtctggctgtctcccaggcctttcagacat
cccaggtgaagccatggtgaagctctactgccccaagtgcatggatgtgtacacacccaagtcatcaag
acaccatcacacggatggcgcctacttcggcactggtttccctcacatgctcttcatggtgcatcccga
gtaccggcccaagagacctgccaaccagtttgtgcccaggtagggagcagggagagtcattaaggtca
aaggaaaggcccaagatcccccagagaggggaggacagggcatggccctttcttgaggtctgcttctcc
cagaatcagggcatctccctgctgagtgactgtgggaaagttatttgattatctgtgcttgagttacct
tattgtagaatgttcttgagctgagaagttgggaaccacgaggctttagctctgagcaggtccatagag
gagctcaggtggggaggtgggaatgcaggtgactggcagggcctggatggggctcatgctgctgcctct
ctgacctctgccctggcctaggctctacggtttcaagatccatccgatggcctaccagctgcagctcca
agccgccagcaacttcaagagcccagtcaagacgattcgctgattccctcccccacctgtcctgcagtc
tttgtcttttcctttctttttttgccaccctttcaggaaccctgtatggtttttagtttaaattaaagga
gtcgttatcgtggtgggaatatgaaataaagtagaagaaaaggccatgagctagtctgctggtgcttgc
ggaagggggtggagcgtggccatggaaatcgggctccacggcccagggatgg
```



```
> HSCKIIBE, Human gene for casein kinase II subunit beta (EC 2.7.1.37).
ggggctgagatgtaaattagaggagctggagaggagtgcttcagagtttgggttgctttaagaaagggt
ggttccgaattctcccgtggttggagggccgaatgtgggaggagggaggataccagaggcagggaagga
gaacttgagctttactgacactgttctttttctagctgacgtgaagatgagcagctcagaggaggtgtc
ctggatttcctggttctgtgggctccgtggcaatgaattcttctgtgaagtgagttctcttcaacctcc
ctacttgccagcttcacatatcttcccaccagacgttccttcacatattccacttctacactgttctct
aaagcttttatgggagagagtgtaggtgaactagggagagacacaagtacttctgctgagttgggagtg
agaaacaagcacaacagatgcagttgtgttgatgataaggcatcacttagagcatttttgcccaggtcaa
agatgaggattttgatatgggttccctcttggcttccatgtcctgacaggtggatgaagactacatcca
ggacaaatttaatcttactggactcaatgagcaggtccctcactatcgacaagctctagacatgatctt
ggacctggagcctggtgaggcaccctcagggttgttttgtgtgtgtgcgtgcactatttttctcttcaa
atctctattcacttgcctgaattttgccaaatttcctttggttctctgatttctttaaccccaaattca
tgctttattttgatcctccacctgactcttgtctagttttgtgacgtatatcacttgttctcatgtttt
tgcaagggtcagaagcccaggtttctgggtcccatgcccagatgttggatggggtaaggcccaaaagta
ggtgctaggcaaactgaatagcccgcagcccctggatatgggcagggcacctaggaaagctgaaaaaca
agtagttgcatttggccgggctgtggttcagatgaagaactggaagacaaccccaaccagagtgacctg
attgagcaggcagccgagatgctttatggattgatccacgcccgctacatccttaccaaccgtggcatc
gcccagatggtgaggcctctctgctcctacctgcctccttctgagcagtaagagacacaggttcctgca
gcaagaagtcatgtttaagccctgtttaaggaagctagctgagaagaggggaagaaccccagaacttgg
ccctgccctaatttggaagaaaggcaacacagaagtttgagagcccatctagtccagagaaggggggcct
ctggacagagttggaaggagtgccgacagagttggtatgggttgggctgcgaagggagttgcctcttct
ttacatctacctgccaacccccttccattgtattcacctcagttggaaaagtaccagcaaggagactttg
gttactgtcctcgtgtgtactgtgagaaccagccaatgcttcccattGgtgagtgttgaagaaggaaa
ggaaagcaccgtgtggcagtcttatgggaaggagttggggctcaacacattggagcctgagtcctgagg
ggaggttaggtaggaataggggatacctggcctgctgagtctggctgtctcccaggcctttcagacat
cccaggtgaagccatggtgaagctctactgccccaagtgcatggatgtgtacacacccaagtcatcaag
acaccatcacacggatggcgcctacttcggcactggtttccctcacatgctcttcatggtgcatcccga
gtaccggcccaagagacctgccaaccagtttgtgcccaggtagggagcagggagagtcattaagggtca
aaggaaaggcccaagatcccccagagaggggaggacagggcatggccctttcttgaggtctgcttctcc
cagaatcagggcatctccctgctgagtgactgtgggaaagttatttgattatctgtgcttgagttacct
tattgtagaatgttcttgagctgagaagttgggaaccacgaggctttagctctgagcaggtccatagag
gagctcaggtggggaggtgggaatgcaggtgactggcagggcctggatggggctcatgctgctgcctct
ctgacctctgccctggcctaggctctacggtttcaagatccatccgatggcctaccagctgcagctcca
agccgccagcaacttcaagagcccagtcaagacgattcgctgattccctcccccacctgtcctgcagtc
tttgtcttttcctttctttttttgccaccctttcaggaaccctgtatggtttttagtttaaattaaagga
gtcgttatcgtggtgggaatatgaaataaagtagaagaaaaggccatgagctagtctgctggtgcttgc
ggaaggggtggagcgtggccatggaaatcgggctccacggcccagggatgg
```

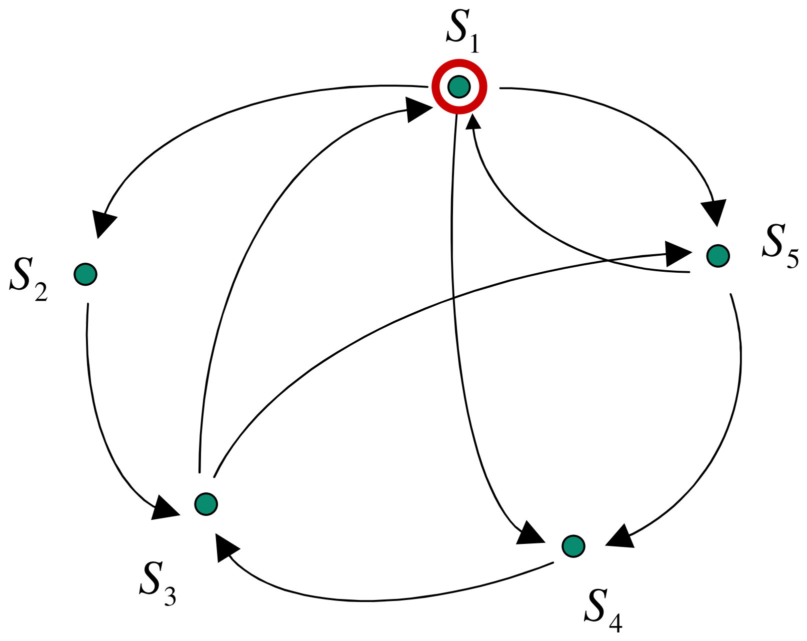
Approaches to genefinding

- Homology
  - BLAST, Procrustes
- Ab initio (or de novo)
  - Genscan, Genie
- Hybrids
  - GenomeScan, GenieEST, Twinscan, SGP, SLAM

Hidden Markov Models

A discrete process

A random process, jumping between a finite number of states



State sequence:

$S_1\ S_2\ S_3\ S_1\ S_4\ S_3\ S_5\ S_1\ S_5\ S_4$

A Markov process

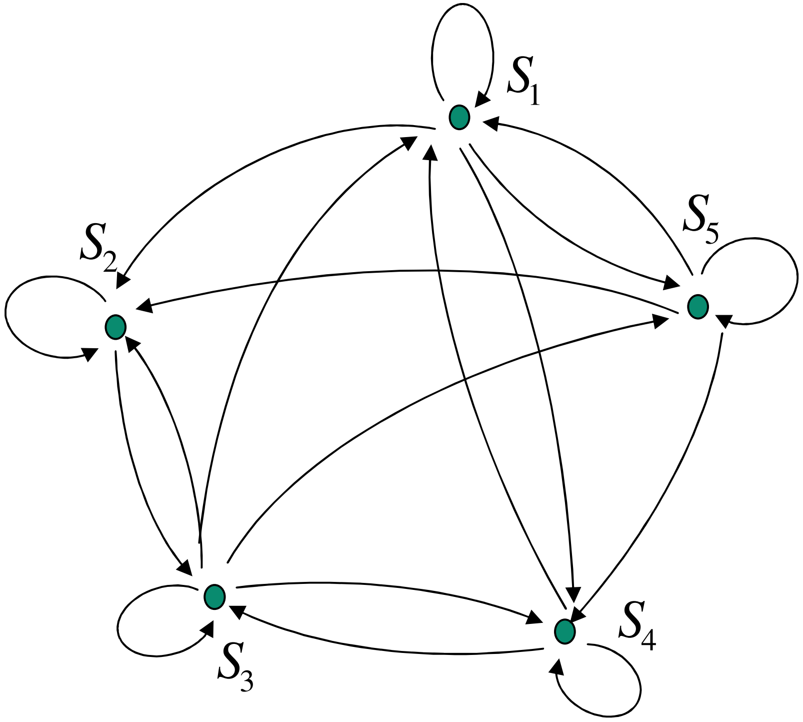
The process is *Markov* if the next state only depends on the current state and not the history.

General description

$$\Pr(X_t = S_j \mid X_{t-1} = S_i, X_{t-2} = S_k, \ldots) =$$

$$\Pr(X_t = S_j \mid X_{t-1} = S_i)$$

Markov description

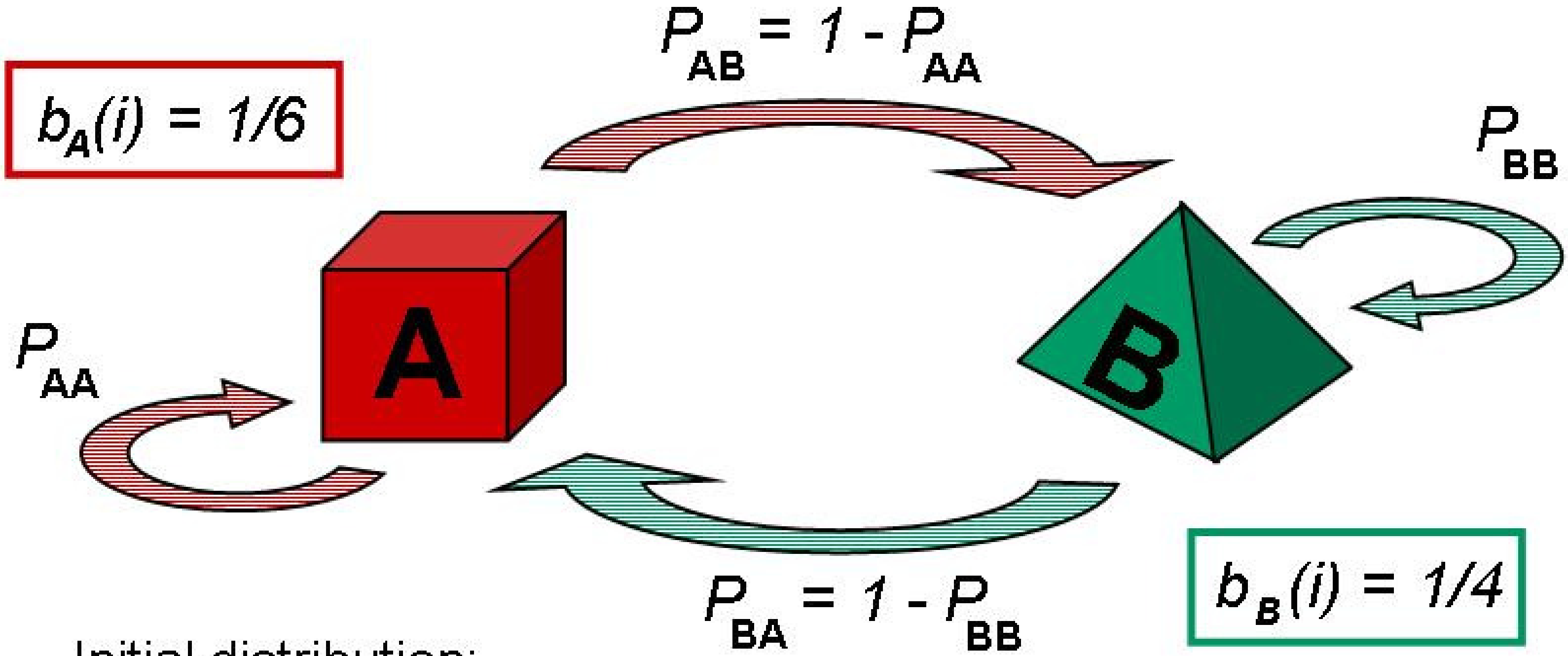


A *hidden* Markov Model

A standard hidden Markov model is comprised of two interrelated processes:

- *Hidden process:* a Markov chain on the state space, generating a state sequence hidden from the observer.
- *Observed process:* generating output through random functions associated with each state.

A simple Hidden Markov Model (HMM)



$b_A(i) = 1/6$

$P_{AB} = 1 - P_{AA}$

P_{BB}

P_{AA}

$P_{BA} = 1 - P_{BB}$

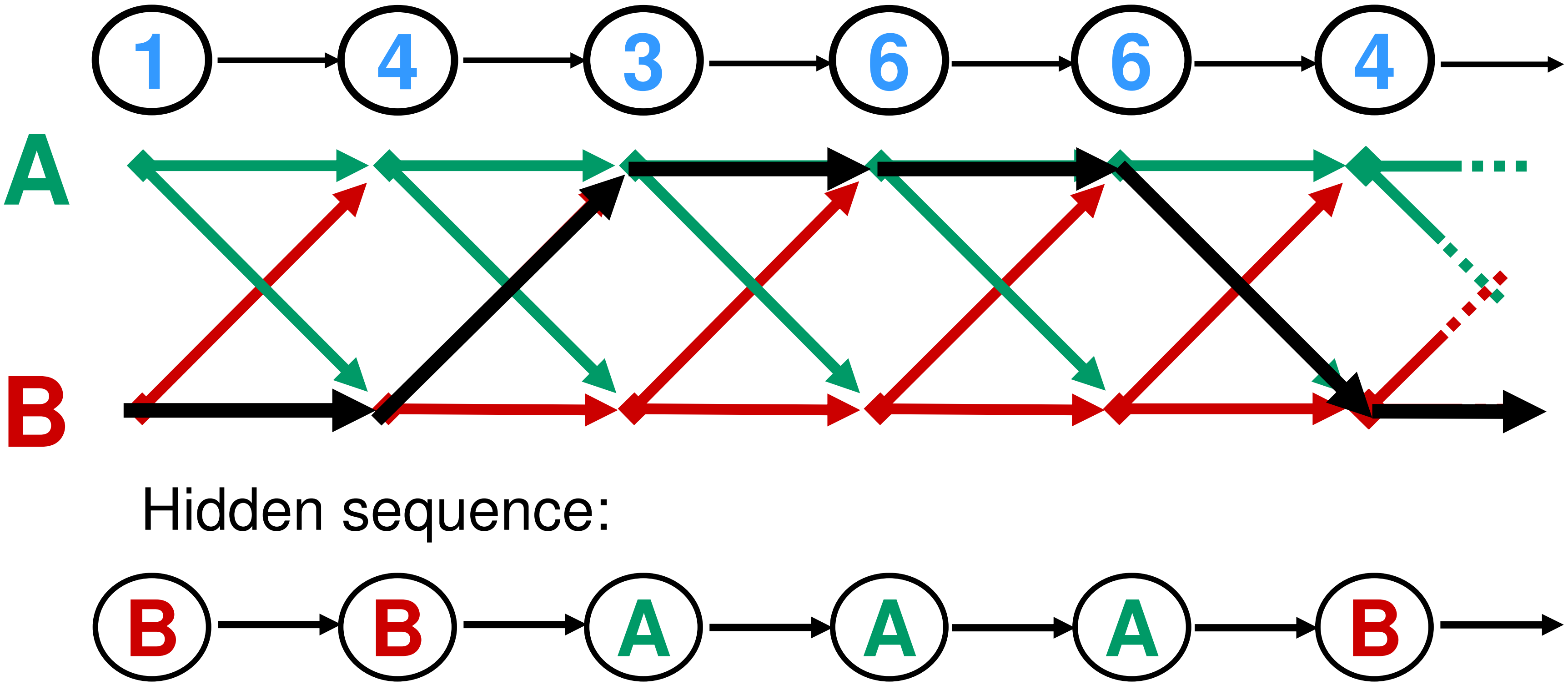
$b_B(i) = 1/4$

Initial distribution:

$\pi = (\pi_A, \pi_B)$

A lattice view

Observed sequence:



Hidden sequence:

Two fundamental problems

- The probability of the observed data given the model.

The forward algorithm

- The best hidden state sequence given the data.

The Viterbi algorithm

The HMM algorithms

The forward variables:

$$\alpha_t(i) = \Pr(\text{obs. up to } t\text{, ending in state } i \text{ at time } t)$$

The backward variables:

$$\beta_t(i) = \Pr(\text{obs. after } t \mid \text{ending in state } i \text{ at time } t)$$

The Viterbi variables:

$$\delta_t(i) = \max_{X_1,\ldots,X_t} \Pr(\text{obs. up to } t\text{, ending in state } i \text{ at time } t)$$

The hidden process:

Let $\{X_t\}_{t=1}^{T}$ denote the Markov process assuming values in the state space $S_1,\ldots,S_N$.

The process begins in a state determined by the initial distribution $\{\pi_i\}_{i=1}^{N}$ and evolves through the state space according to transition probabilities

$$a_{ij} = \Pr(X_{t+1} = j \mid X_t = i)$$

The observed process:

Let $\{Y_t\}_{t=1}^{T}$ be the observed process generating output as random functions of the state X_t according to some output function

$$b_j(Y_t \mid Y_1,\dots,Y_{t-1}) = \Pr(Y_t \mid X_t = j, Y_1,\dots,Y_{t-1})$$

The joint probability

The joint probability of the two interrelated processes becomes

$$\Pr(X_{t+1}=j, Y_{t+1} \mid X_t = i, X_1,\ldots, X_{t-1}, Y_1,\ldots, Y_{t-1}) =$$

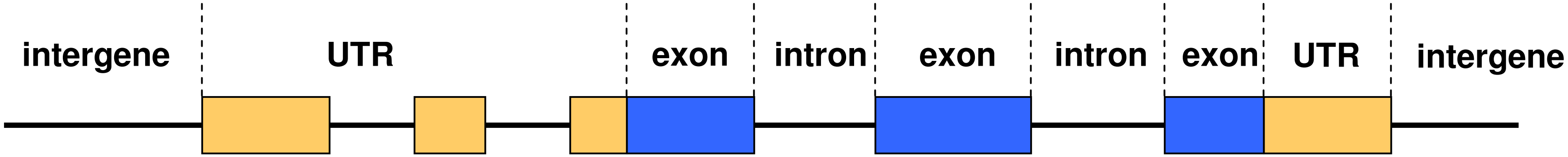
$$= \Pr(X_{t+1}=j \mid X_t = i)\Pr(Y_{t+1} \mid X_{t+1}, Y_1,\ldots Y_t) =$$

$$= \underbrace{a_{ij}}_{\textit{transition probability}} \underbrace{b_j(Y_{t+1} \mid Y_1,\ldots, Y_t)}_{\textit{output distribution}}$$

Gene finding

- Generalized HMMs

HMMs for gene finding

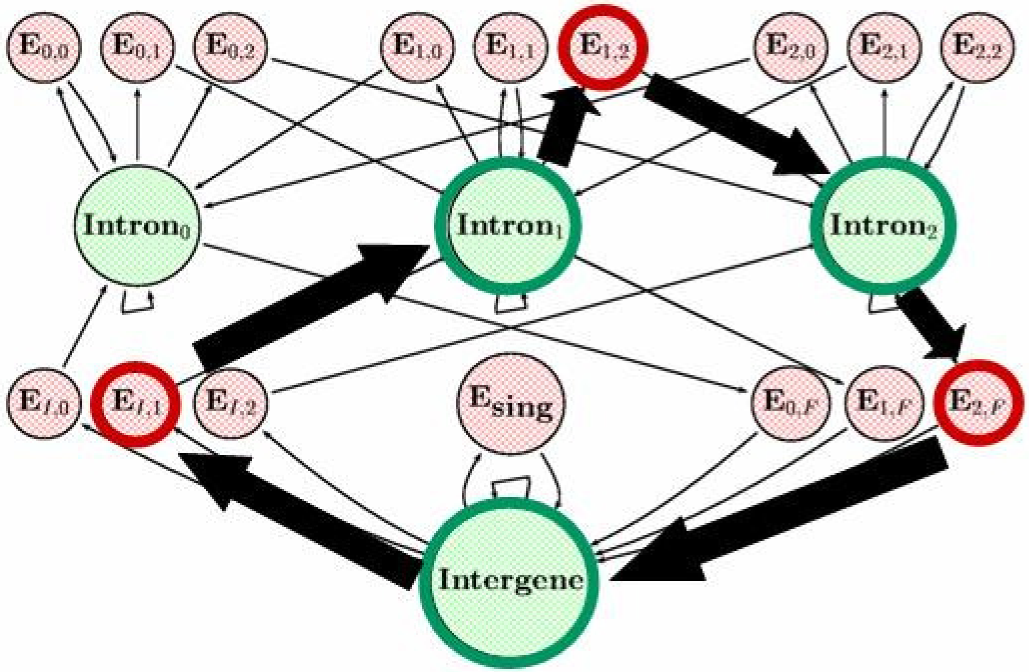


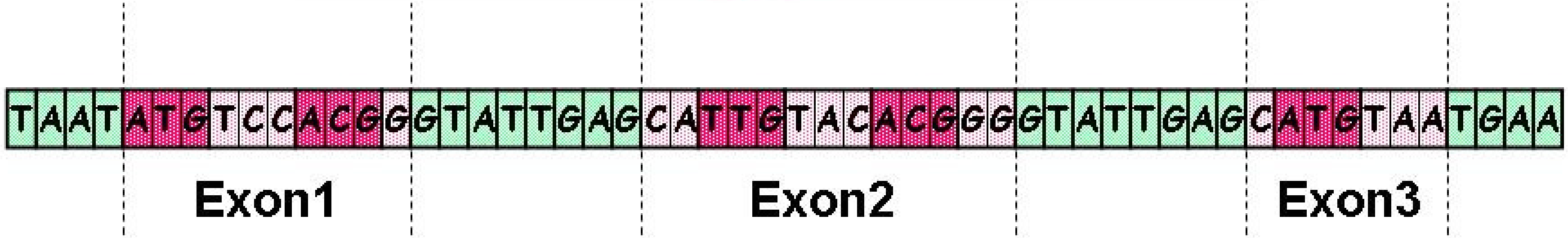
Observed:

CTTGATGCTGGCACGTTCTGCTTCATCGGAGACAAATTACGGCTTTCCGGAGCA

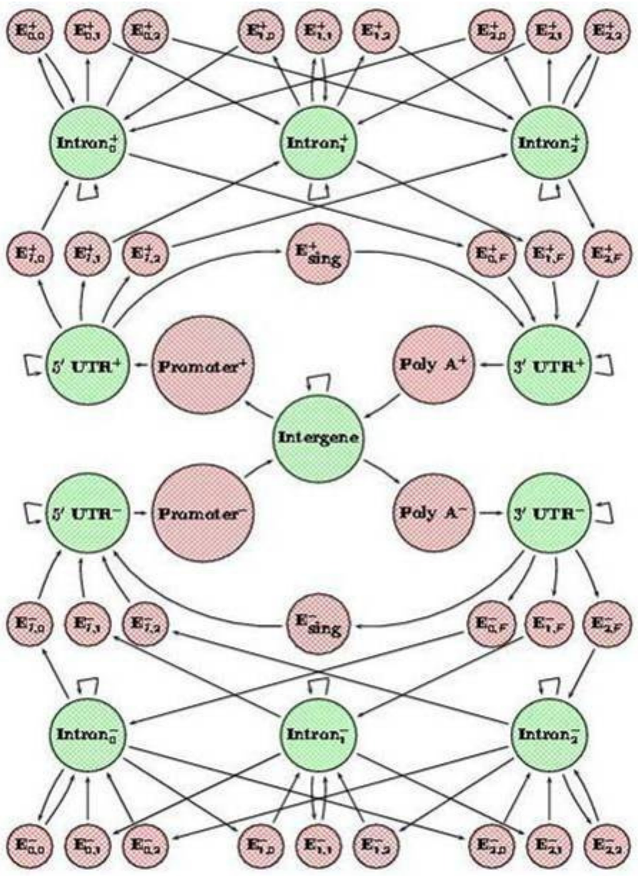
Hidden:

CTTGATGCTGGCACGTTCTGCTTCATCGGAGACAAATTACGGCTTTCCGGAGCA

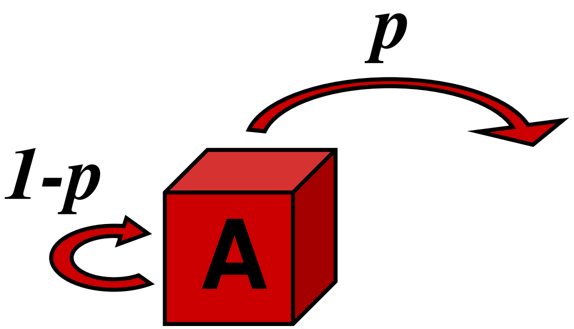

$E_{0,0}$
$E_{0,1}$
$E_{0,2}$
$E_{1,0}$
$E_{1,1}$
$E_{1,2}$
$E_{2,0}$
$E_{2,1}$
$E_{2,2}$
$Intron_0$
$Intron_1$
$Intron_2$
$E_{I,0}$
$E_{I,1}$
$E_{I,2}$
E_{sing}
$E_{0,F}$
$E_{1,F}$
$E_{2,F}$
Intergene


T A A T A T G T C C A C G G G T A T T G A G C A T T G T A C A C G G G G T A T T G A G C A T G T A A T G A A
Exon1
Exon2
Exon3

The Genscan state space



State duration times

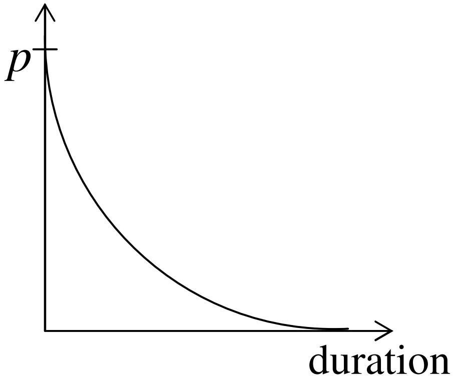


$$\Pr(\text{leaving state}) = p$$

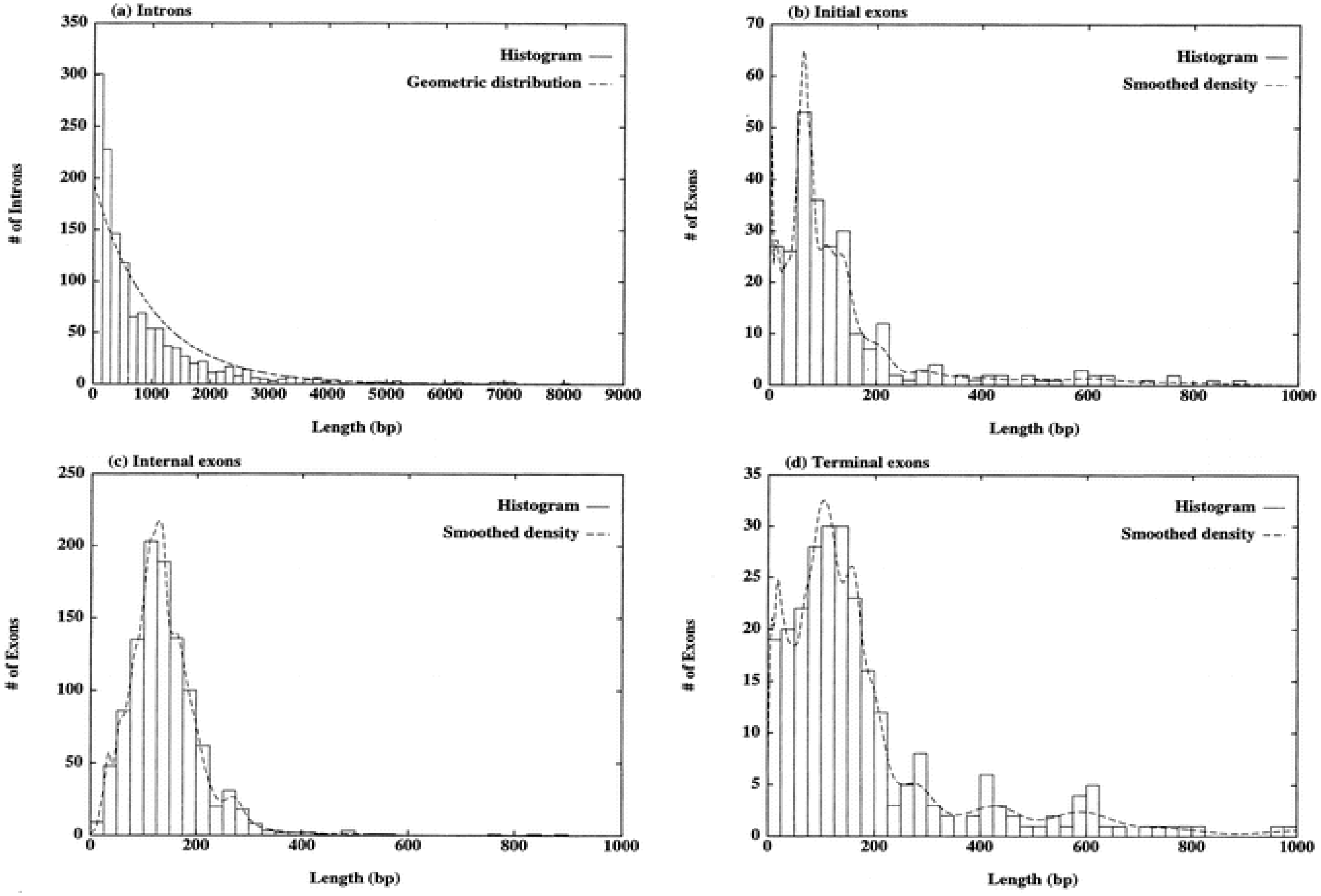
$$\Pr(\text{staying in state}) = 1 - p$$

$$\Pr(\text{output of exactly } r \text{ in state}) = (1\text{-}\mathrm{p})^{\mathrm{r-1}}\, p$$

The geometric distribution



Observed duration times



Generalized HMMs

When in state X_l the duration d_l is chosen from a generalized length distribution

$$f_{X_l}(d_l) = \Pr(\text{state duration} = d_l \mid X_l)$$

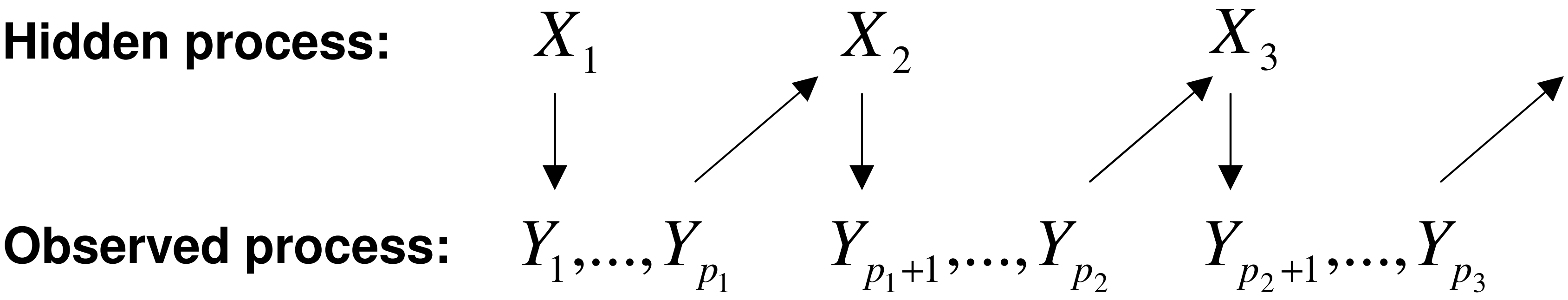
Now the indices for the observed and the hidden process may differ, and we introduce partial sums

$$p_l = \sum_{k=1}^{l} d_k, \quad p_0 = 0$$

Generalized HMMs, cont.

- In state X_l choose state duration d_l.
- Generate output $Y_{p_{l-1}+1},\dots,Y_{p_l}$ according to

$$b_{X_l}(Y_{p_{l-1}+1},\dots,Y_{p_l} \mid Y_1,\dots,Y_{p_{l-1}})$$



Hidden process: X_1 X_2 X_3

Observed process: $Y_1,\dots,Y_{p_1}$ $Y_{p_1+1},\dots,Y_{p_2}$ $Y_{p_2+1},\dots,Y_{p_3}$

Generalized HMMs, cont.

Assume that we observe a sequence of outputs $Y_1,\ldots,Y_T$ from a sequence of hidden states $X_1,\ldots,X_L$ with durations $d_1,\ldots,d_L$ (assume $p_L = T$). The joint probability of hidden and observed data becomes

$$\Pr(Y_1^T, X_1^L, d_1^L) =$$

$$= \pi_{X_1} f_{X_1}(d_1) b_{X_1}(Y_1^{p_1}) \prod_{l=2}^{L} \underbrace{a_{X_{l-1},X_l}}_{\text{transition probability}} \overbrace{f_{X_l}(d_l)}^{\text{duration distribution}} \underbrace{b_{X_l}(Y_{p_{l-1}+1}^{p_l} \mid Y_1^{p_{l-1}})}_{\text{output distribution}}$$

Sequence alignment

- Pair HMMs

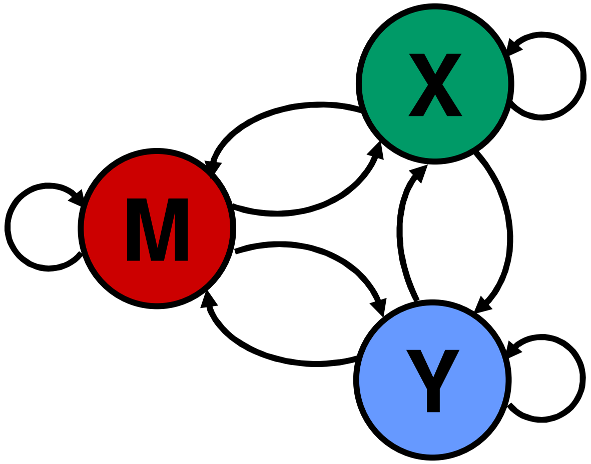
Sequence alignment

```
 50     .    :    .    :    .    :    .    :    .    :
247 GGTGAGGTCGAGGACCCTGCA  CGGAGCTGTATGGAGGGCA   AGAGC
    |:   ||  ||||:  |||| --:||  ||| |::|   |||---||||
368 GAGTCGGGGGAGGGGGCTGCTGTTGGCTCTGGACAGCTTGCATTGAGAGG

100     .    :    .    :    .    :    .    :    .    :
292 TTC          CTACAGAAAAGTCCCAGCAAGGAGCCACACTTCACTG
    |||----------|| |   |::| |: ||||::|:||:-||  ||:| |
418 TTCTGGCTACGCTCTCCCTTAGGGACTGAGCAGAGGGCT CAGGTCGCGG

150     .    :    .    :    .    :    .    :    .    :
332                ATGTCGAGGGGAAGACATCATTCGGGATGTCAGTG
    ---------------||||||||||||||||||||||:||||||||||||
467 TGGGAGATGAGGCCAATGTCGAGGGGAAGACATCATTTGGGATGTCAGTG

200     .    :    .    :    .    :    .    :    .    :
367 TTCAACCTCAGCAATGCCATCATGGGCAGCGGCATCCTGGGACTCGCCTA
    |||||:||||||||:||||||||||||||:|| ||:|||||:||||||||
517 TTCAATCTCAGCAACGCCATCATGGGCAGTGGAATTCTGGGGCTCGCCTA
```

Sequence comparisons. Why?

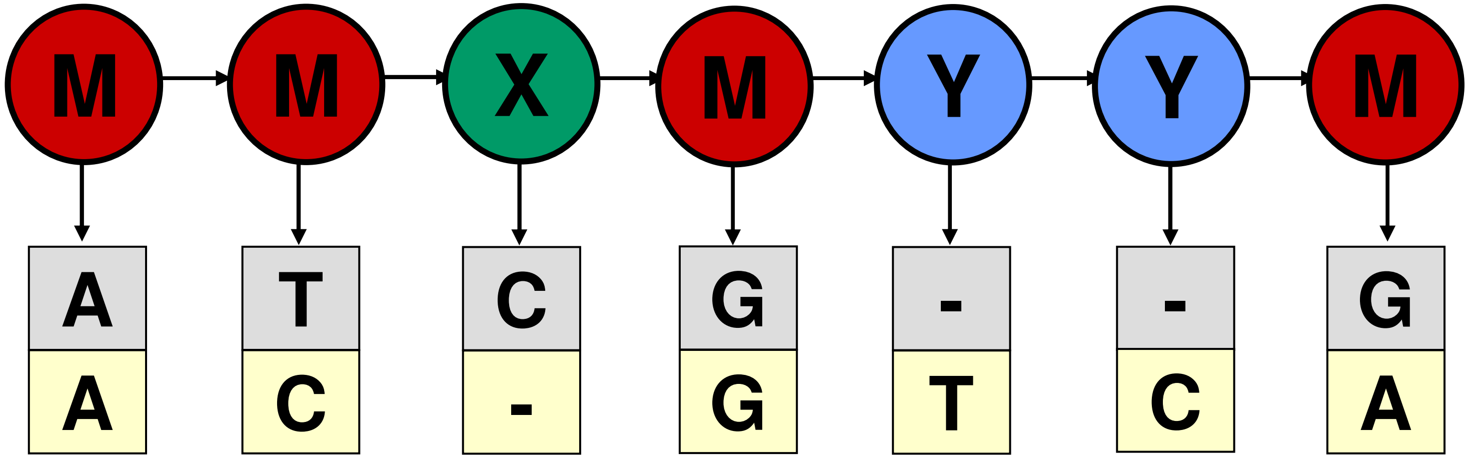
- Are the sequences related?
- What regions are related?
- How evolutionary distant are they?
- Info about the evolutionary process.

Pair HMMs for alignment



M = (mis)match
X = insert seq1
Y = insert seq2

Pair HMMs



Output sequence:

ATCG- - G
AC- GTCA

Observed sequences:

ATCGG
ACGTCA

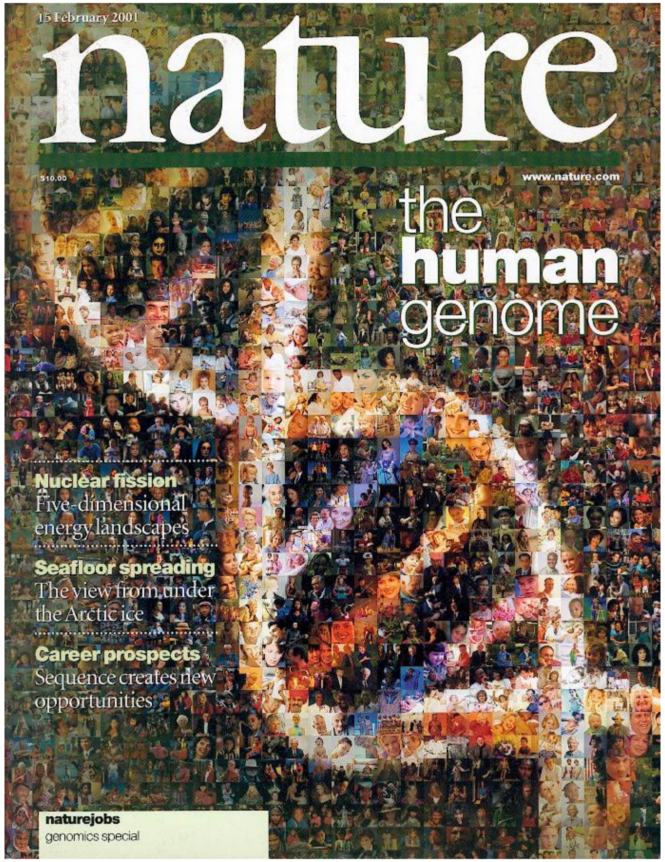
Comparative gene finding

- Generalized Pair HMMs

Fraunhofer ITWM

CHALMERS
Research Centre
Industrial Mathematics

Comparing human and mouse



15 February 2001

nature

$10.00

www.nature.com

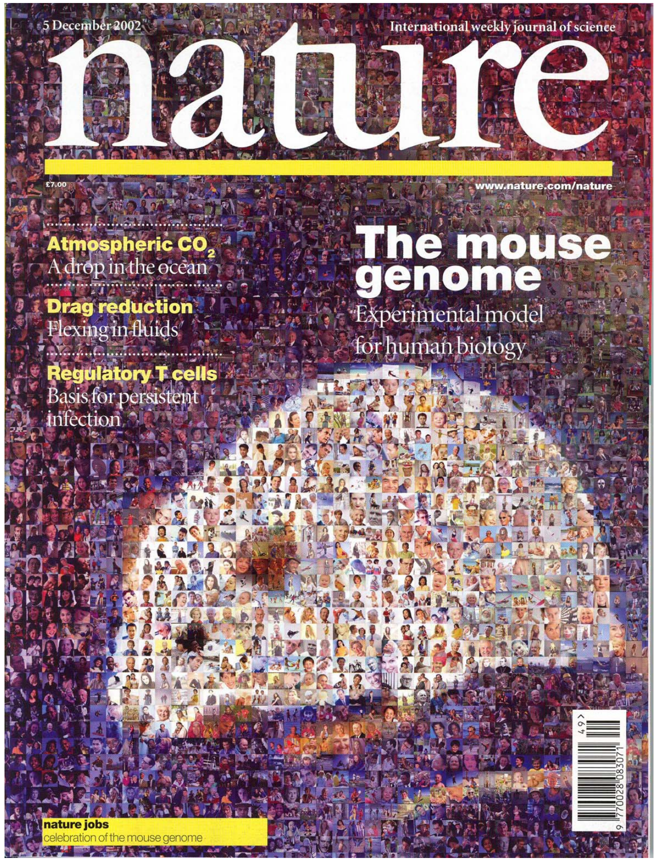
the **human** genome

Nuclear fission
Five-dimensional energy landscapes

Seafloor spreading
The view from under the Arctic ice

Career prospects
Sequence creates new opportunities

naturejobs
genomics special



5 December 2002

International weekly journal of science

nature

£7.00

www.nature.com/nature

Atmospheric CO_2
A drop in the ocean

Drag reduction
Flexing in fluids

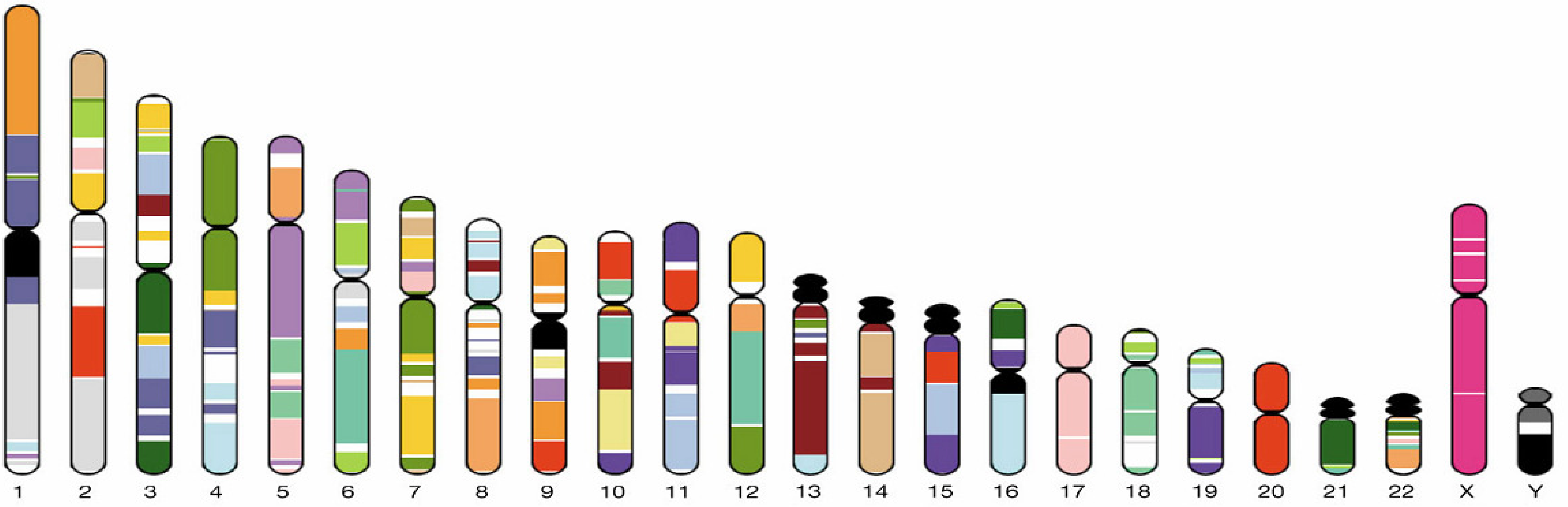
Regulatory T cells
Basis for persistent infection

The mouse genome
Experimental model for human biology

nature jobs
celebration of the mouse genome

Why mouse?

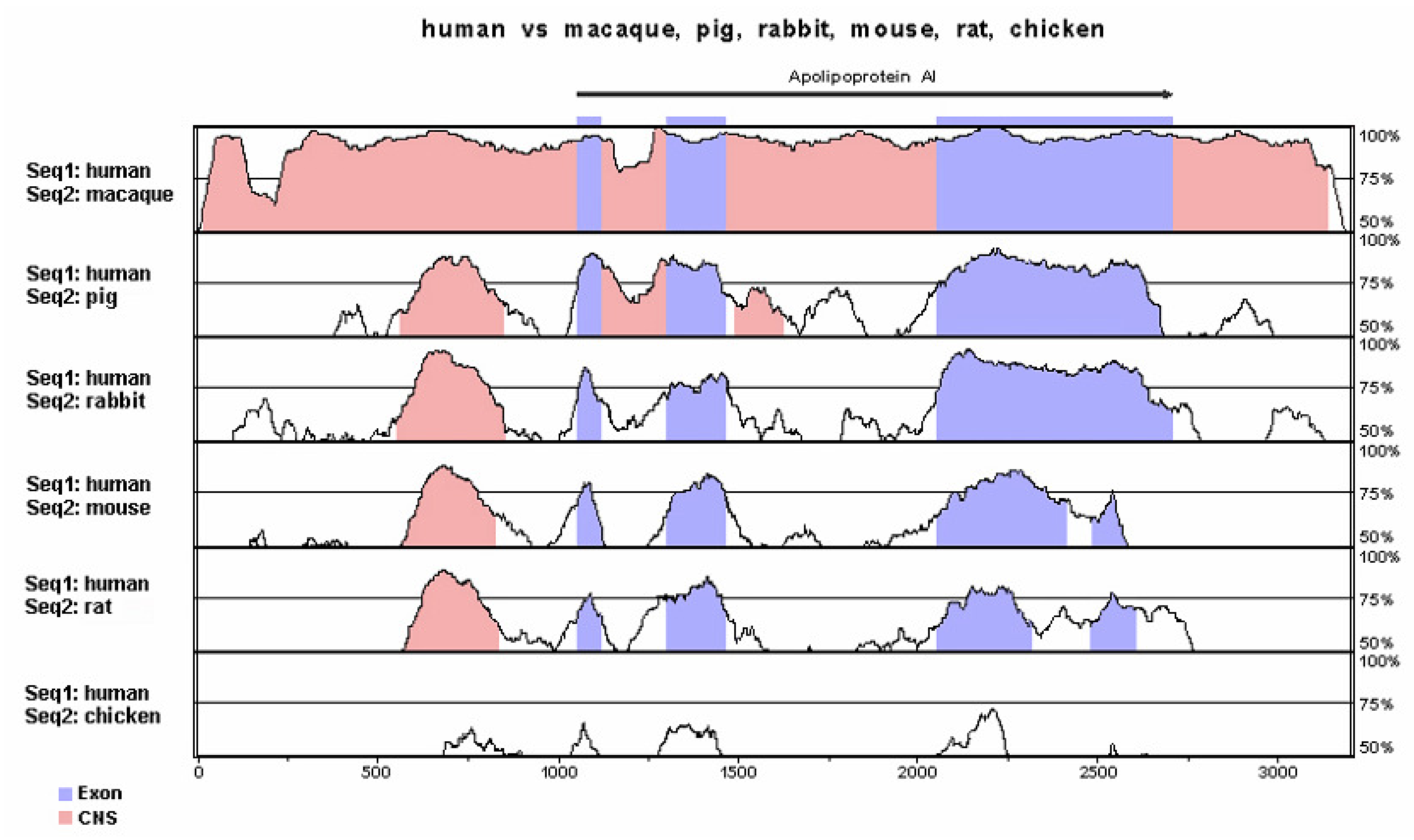
Human



Mouse

■ 1 ■ 2 ■ 3 ■ 4 ■ 5 ■ 6 ■ 7 ■ 8 ■ 9 ■ 10

■ 11 ■ 12 ■ 13 ■ 14 ■ 15 ■ 16 ■ 17 ■ 18 ■ 19 ■ X ■ Y



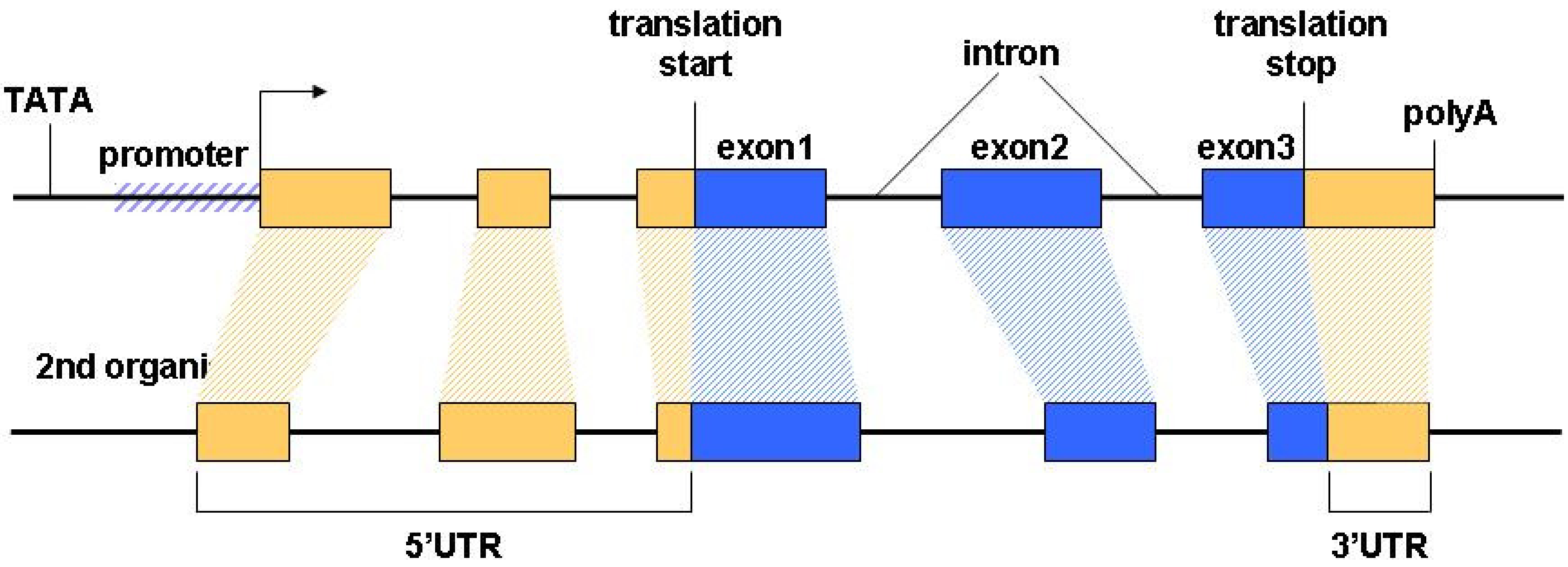
Fraunhofer ITWM

CHALMERS
Research Centre
Industrial Mathematics

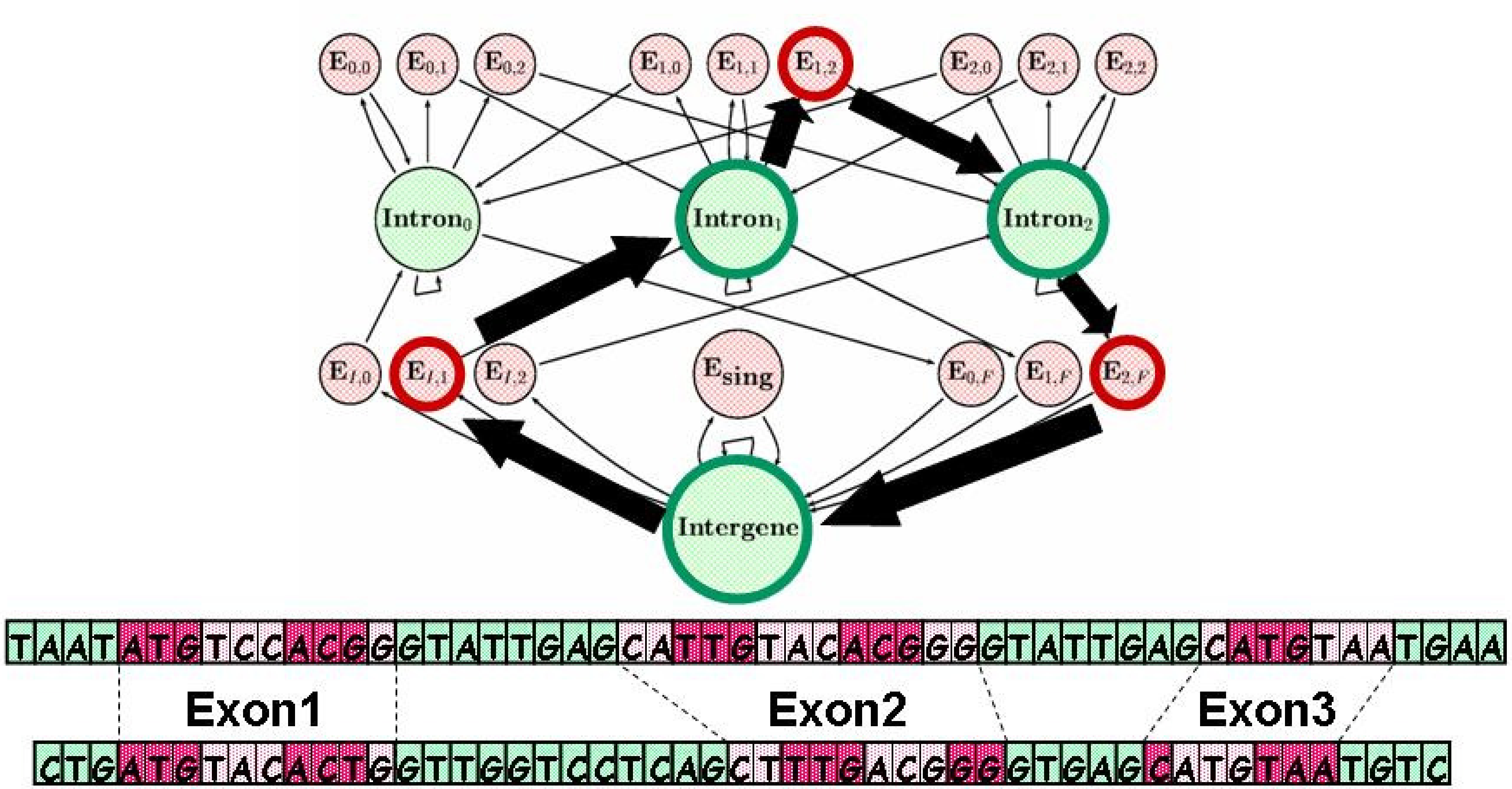
Comparison of 1196 orthologous gene pairs
(Makalowski et al., 1996)

- Sequence identity
  - exons: 84.6%
  - protein: 85.4%
  - introns: 36%
  - 5’ UTRs: 67%
  - 3’ UTRs: 69%

Comparative gene finding



Generalized Pair HMMs



TAATATGTCCACGGGTATTGAGCATTGTACACGGGGTATTGAGCATGTAATGAA

CTGATGTACACTGGTTGGTCCTCAGCTTTGACGGGGTGAGCATGTAATGTC

Generalized Pair HMMs

- Same hidden process $X_1,\ldots,X_L$ on state space $S_1,\ldots,S_N$
- Two output sequences: $Y_1,\ldots,Y_T$ and $Z_1,\ldots,Z_U$
- Two sets of durations: $d_1,\ldots,d_L$ and $e_1,\ldots,e_L$
- Two sets of partial sums:

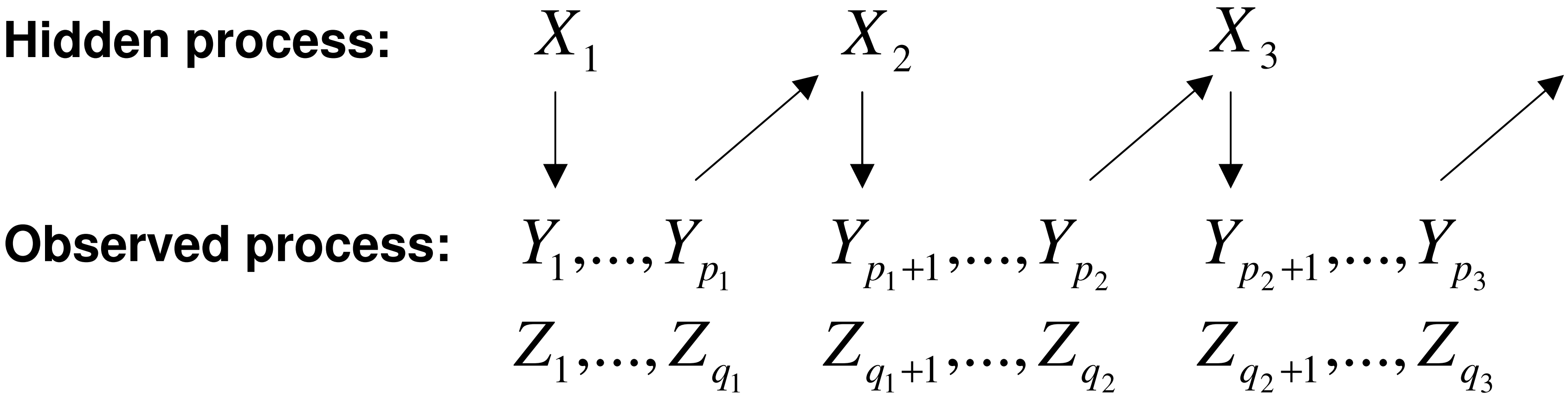
$$p_l = \sum_{k=1}^{l} d_k, \quad p_0 = 0, \quad p_L = T \quad \text{and}$$

$$q_l = \sum_{k=1}^{l} e_k, \quad q_0 = 0, \quad q_L = U$$

Generalized Pair HMMs, cont.

- In state X_l choose state durations (d_l, e_l) from some joint distribution $f_{X_l}(d_l, e_l)$.
- Output $Y_{p_{l-1}+1}, \ldots, Y_{p_l}$ and $Z_{q_{l-1}+1}, \ldots, Z_{q_l}$ jointly generated from

$$b_{X_l}(Y_{p_{l-1}+1}^{p_l}, Z_{q_{l-1}+1}^{q_l} \mid Y_1^{p_{l-1}}, Z_1^{q_{l-1}})$$



Hidden process: X_1 X_2 X_3

Observed process: $Y_1, \ldots, Y_{p_1}$ $Y_{p_1+1}, \ldots, Y_{p_2}$ $Y_{p_2+1}, \ldots, Y_{p_3}$

$Z_1, \ldots, Z_{q_1}$ $Z_{q_1+1}, \ldots, Z_{q_2}$ $Z_{q_2+1}, \ldots, Z_{q_3}$

Generalized Pair HMMs, cont.

The joint probability of hidden and observed data now $Y_1, \ldots, Y_T$ becomes

$$\Pr(Y_1^T, Z_1^U, X_1^L, d_1^L, e_1^L) =$$

$$= \pi_{X_1} f_{X_1}(d_1, e_1) b_{X_1}(Y_1^{p_1}, Z_1^{q_1}) \prod_{l=2}^{L} \underbrace{a_{X_{l-1},X_l}}_{\text{transition probability}} \overbrace{f_{X_l}(d_l, e_l)}^{\text{duration distribution}} \underbrace{b_{X_l}(Y_{p_{l-1}+1}^{p_l}, Z_{q_{l-1}+1}^{q_l} \mid Y_1^{p_{l-1}}, Z_1^{q_{l-1}})}_{\text{output distribution}}$$

Reducing computational complexity

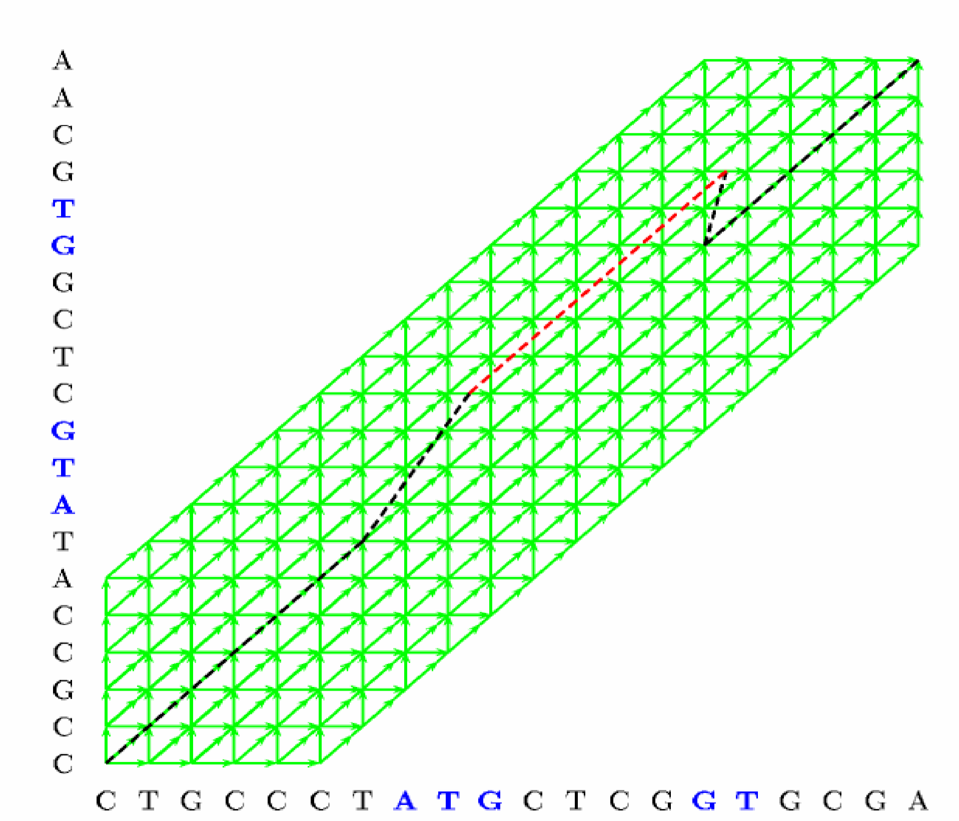
Computational complexity

Model	Time	Space
HMM	N^2T	NT
PHMM	N^2TU	NTU
GHMM	D^2N^2T	NT
GPHMM	D^4N^2TU	NTU

$N =$ no. states $\quad T =$ length seq1

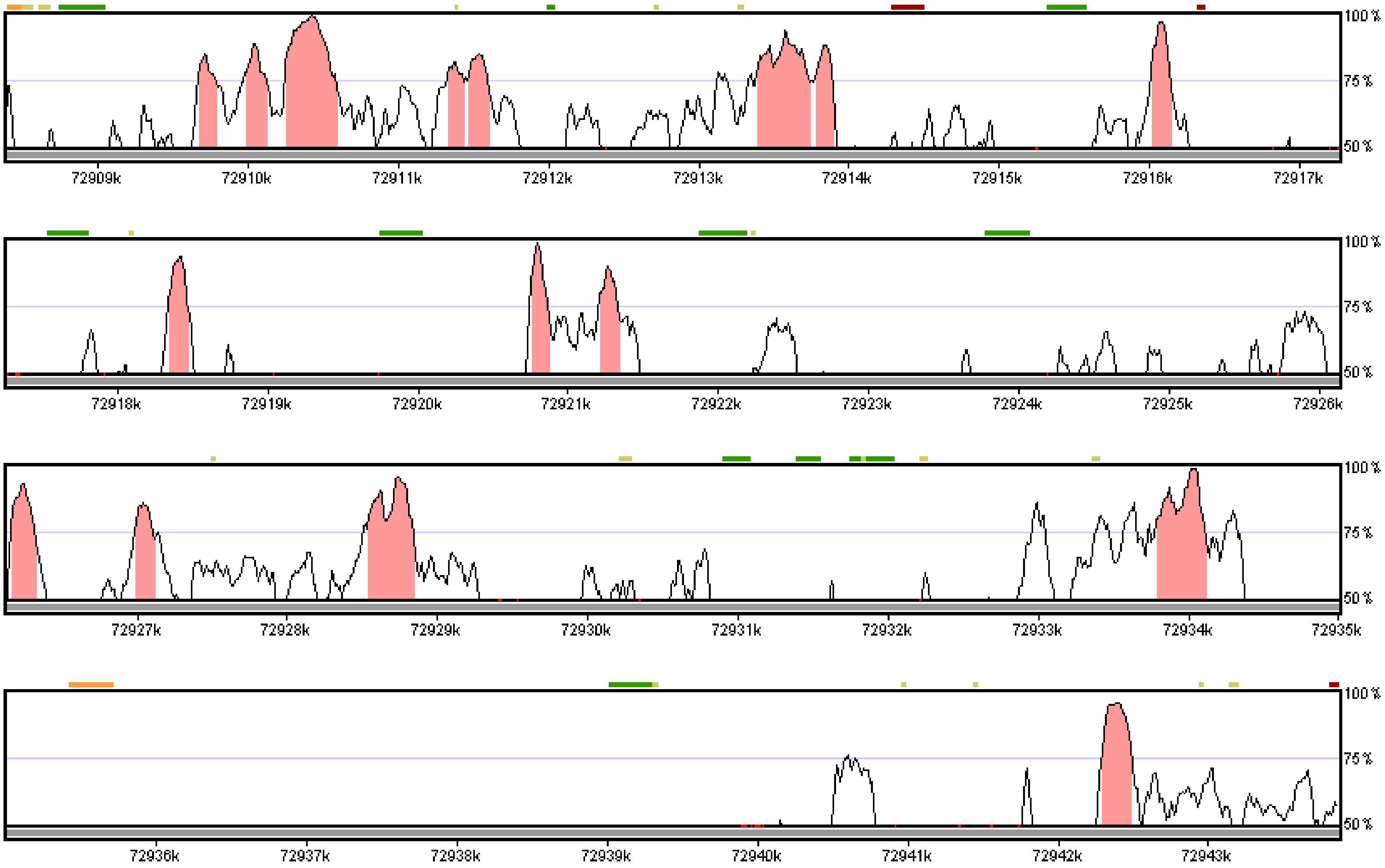
$D =$ max duration $\quad U =$ length seq2

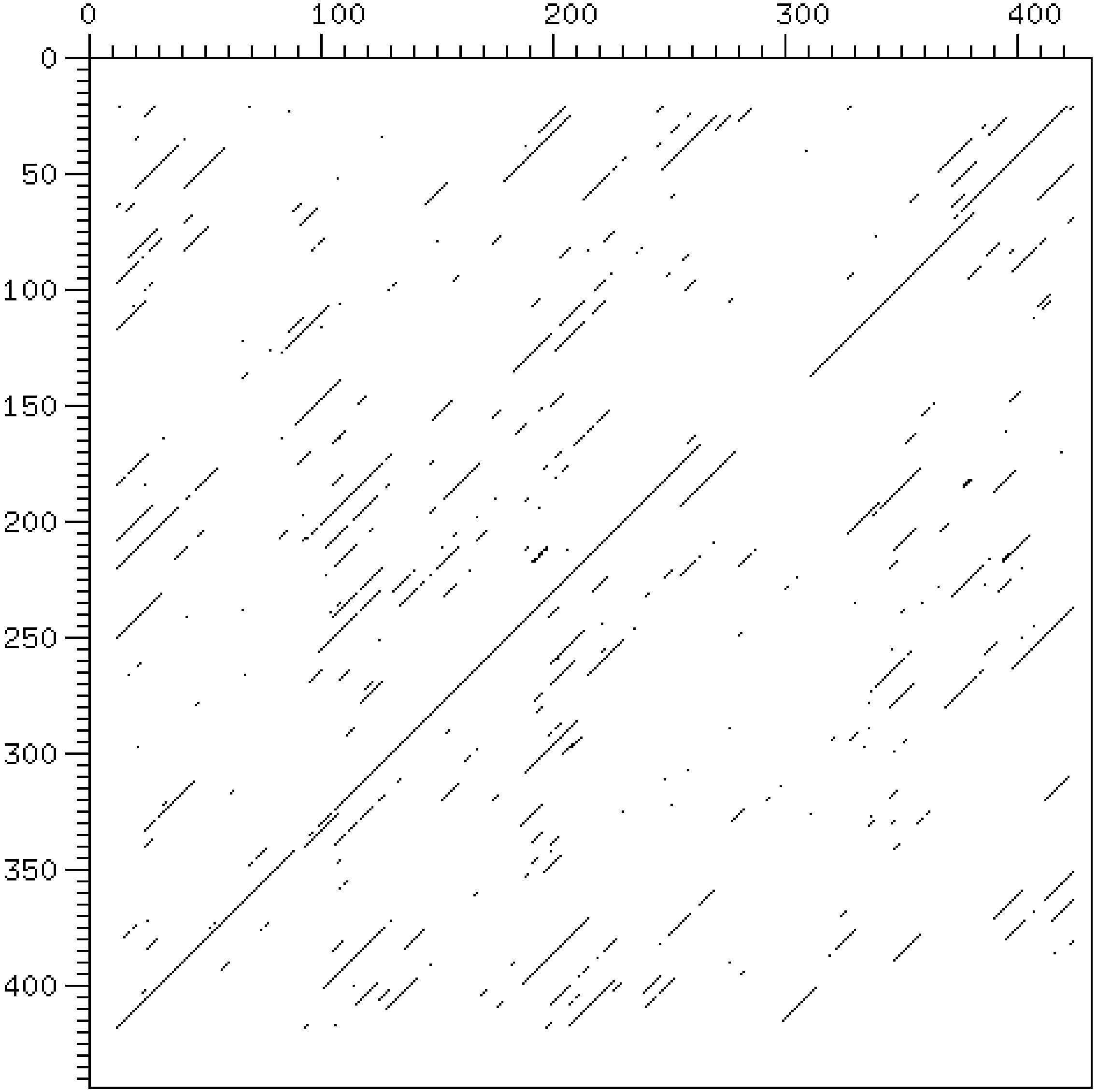
Approximate alignment

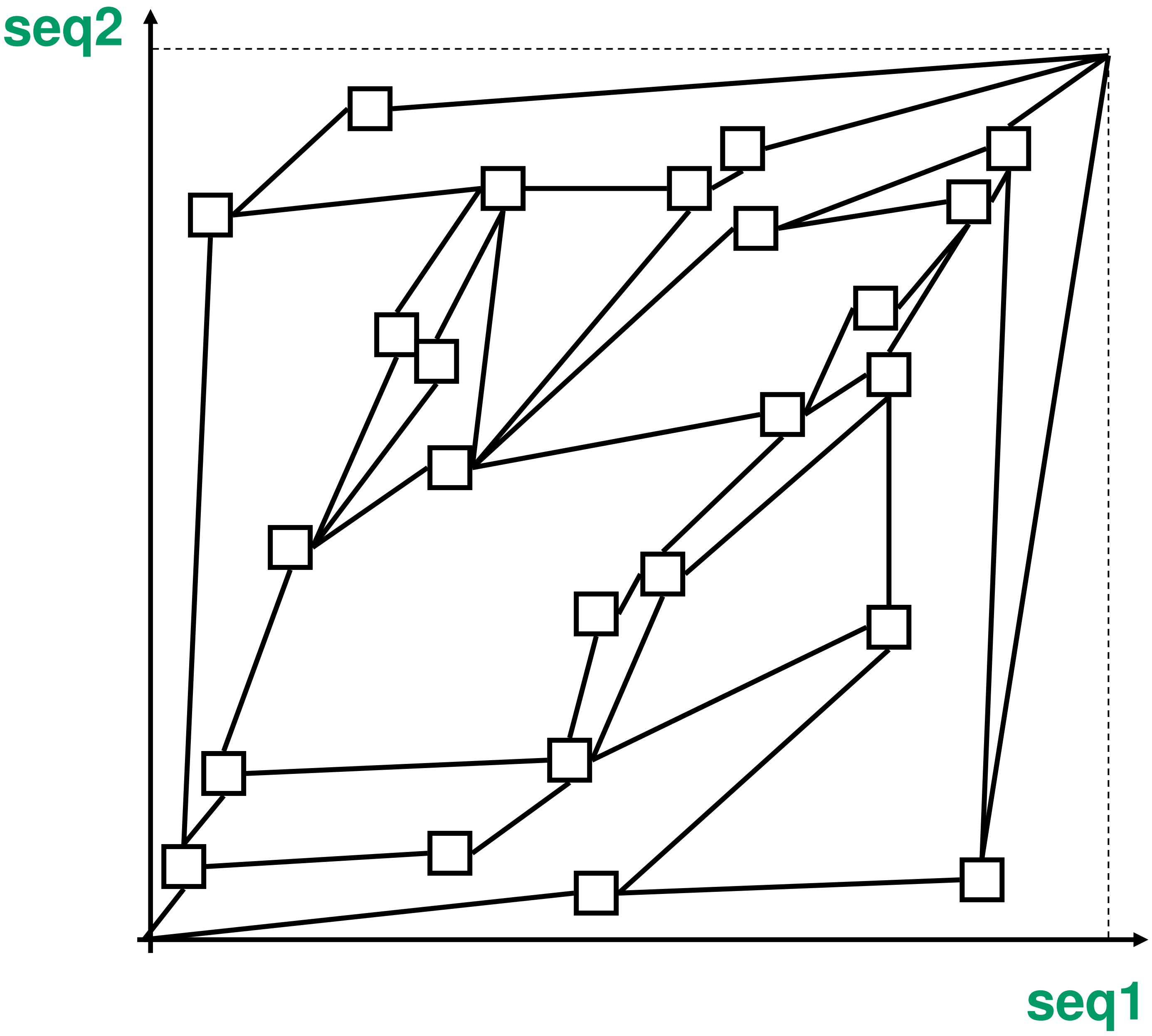

A A C G T G G C T C G T A T A C C G C C
C T G C C C T A T G C T C G G T G C G A

Reduces

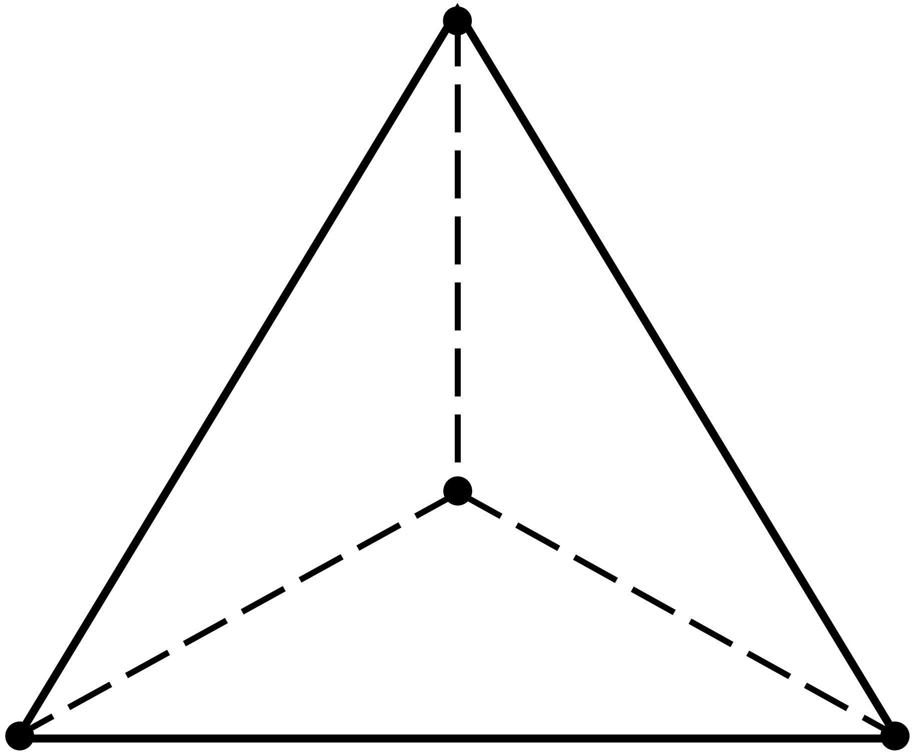
TU **-factor to *hT***


100 %
75 %
50 %
72909k
72910k
72911k
72912k
72913k
72914k
72915k
72916k
72917k
100 %
75 %
50 %
72918k
72919k
72920k
72921k
72922k
72923k
72924k
72925k
72926k
100 %
75 %
50 %
72927k
72928k
72929k
72930k
72931k
72932k
72933k
72934k
72935k
100 %
75 %
50 %
72936k
72937k
72938k
72939k
72940k
72941k
72942k
72943k

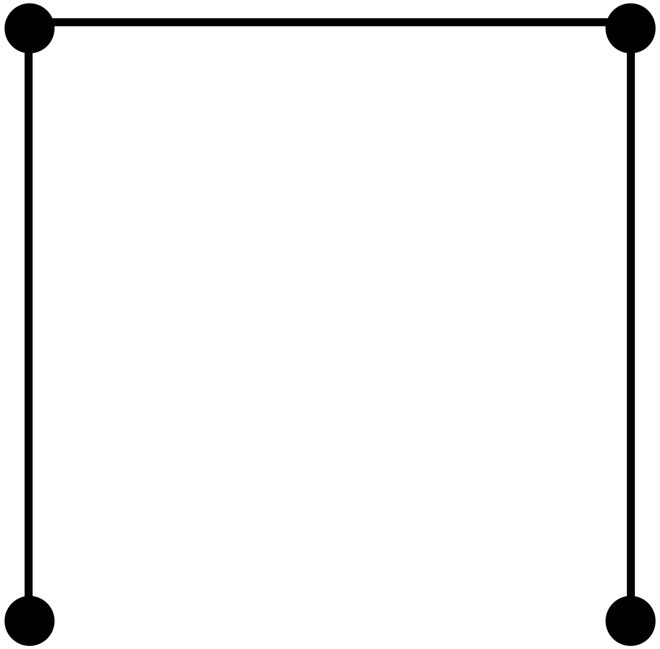

0
100
200
300
400
0
50
100
150
200
250
300
350
400


seq2
seq1

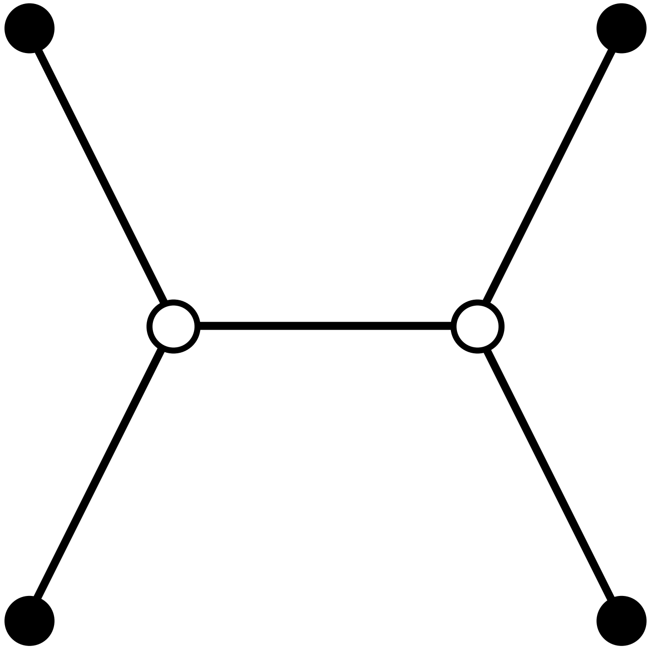
Steiner trees



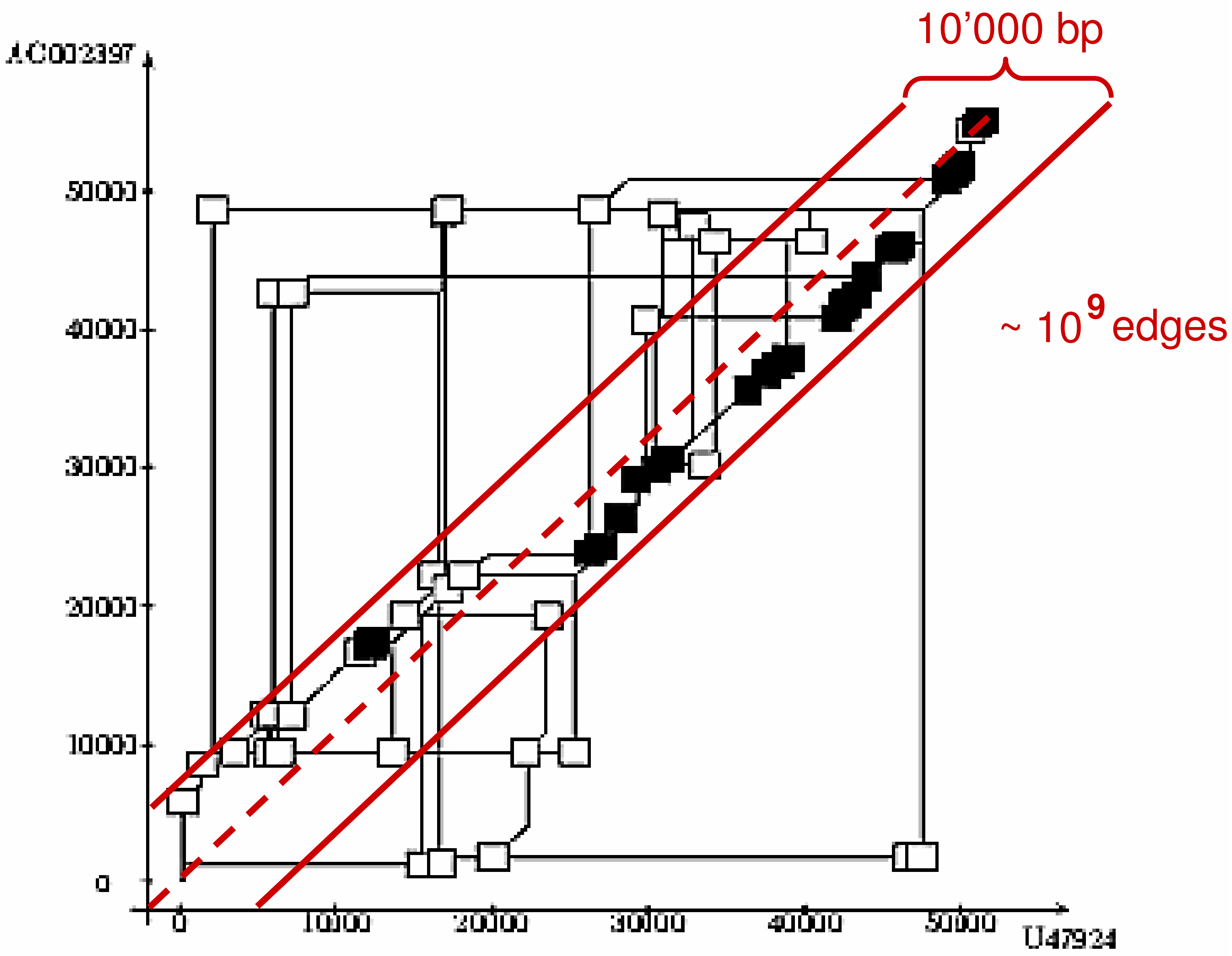
Steiner trees



Minimum spanning tree



Steiner tree



2-approximation in O(n³)