

В прошлый раз

Гены

Предсказание генов

Марковские модели

Сегодня

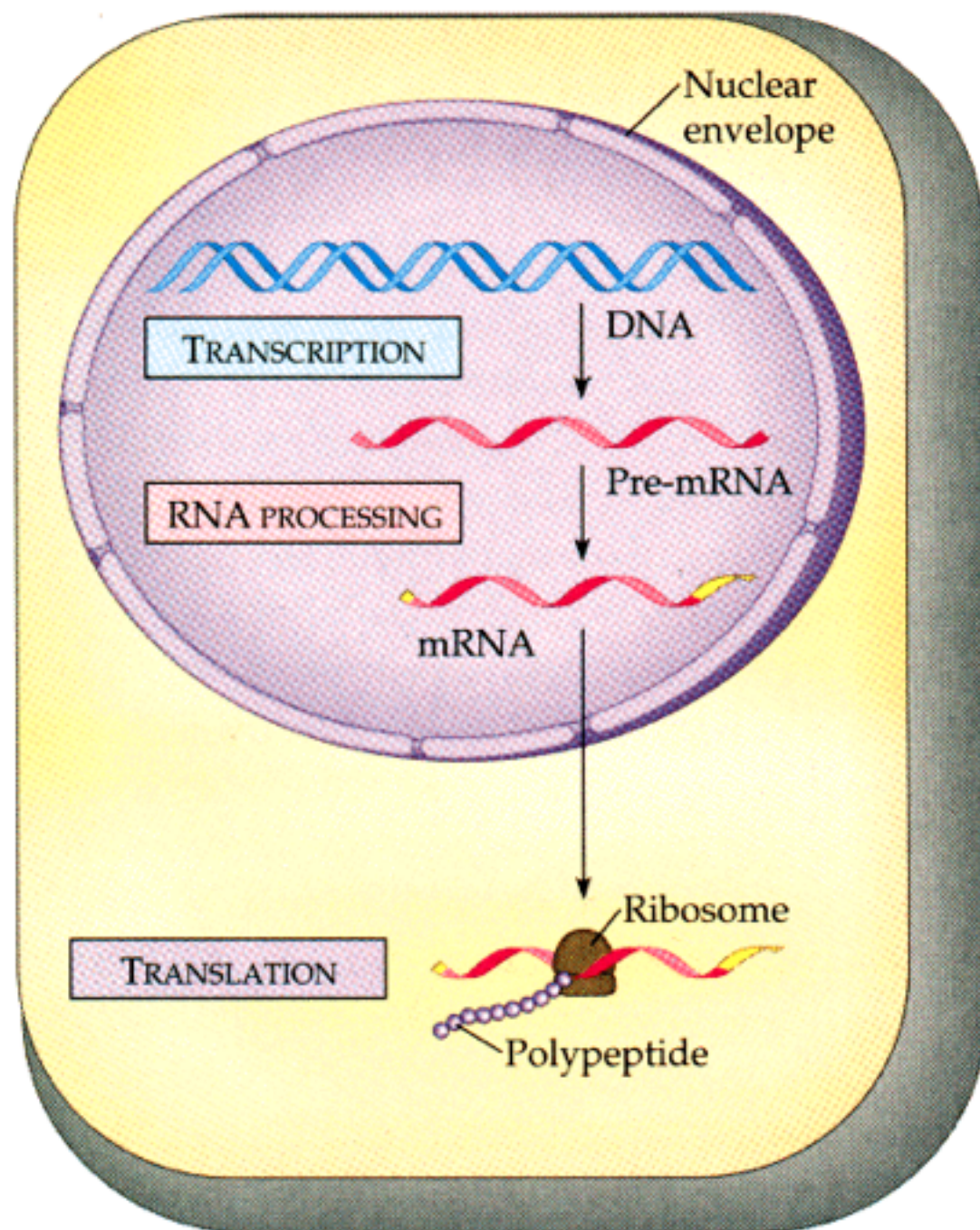
Алгоритм Витерби

Масс-спектрометрия

Эволюционные деревья

Структура РНК

Гены



CCTGAGCCAAC TATTGATGAA



CCUGAGCCAACU AUUGAUGAA

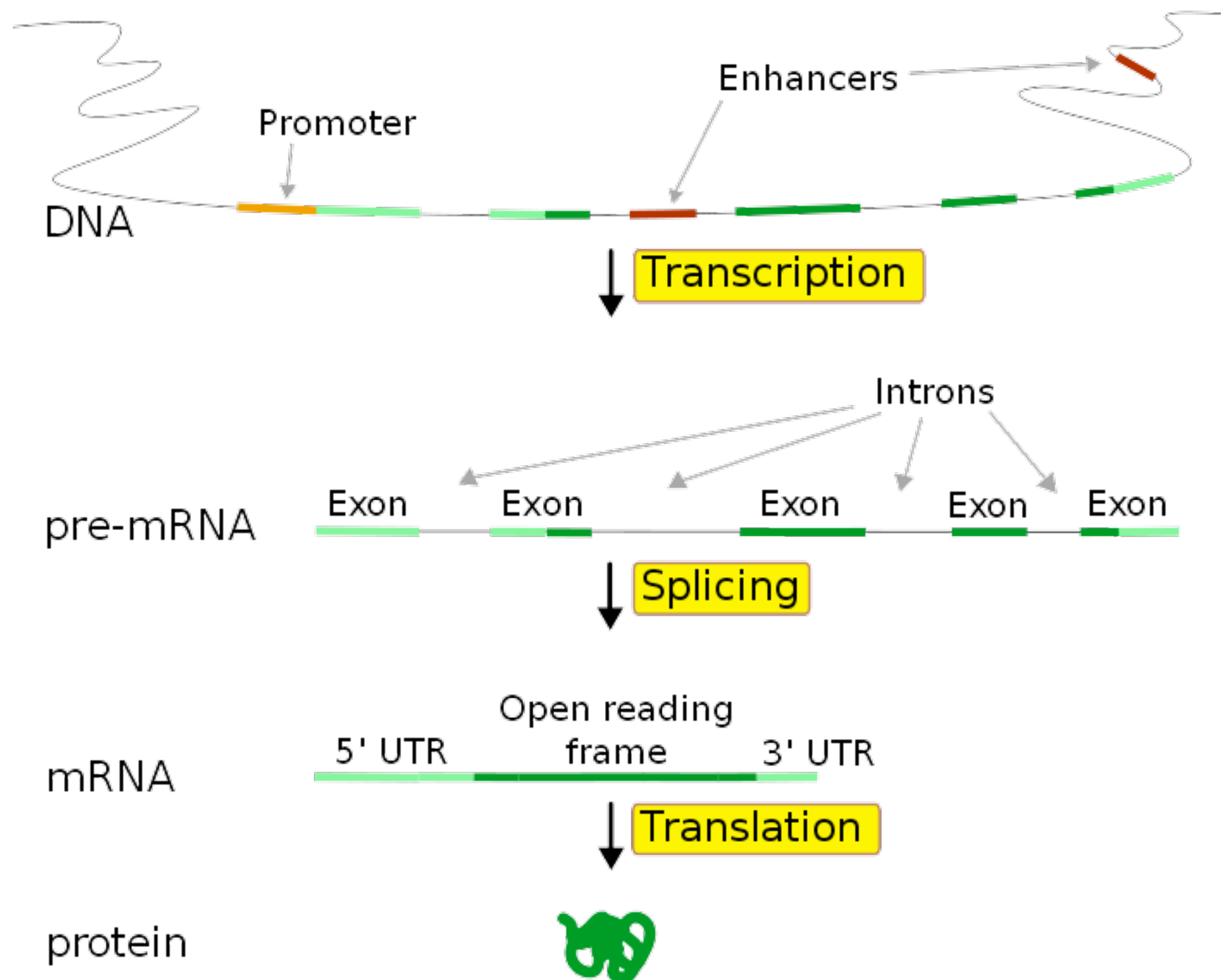


PEPTIDE

ДНК → РНК → Белок

		Second base					
		U	C	A	G		
First base	U	UUU } Phenyl- alanine F UUC } UUA } Leucine L UUG }	UCU } UCC } Serine S UCA } UCG }	UAU } Tyrosine Y UAC } UAA } Stop codon UAG } Stop codon	UGU } Cysteine C UGC } UGA } Stop codon UGG } Tryptophan W	Third base	U C A G
	C	CUU } CUC } Leucine L CUA } CUG }	CCU } CCC } Proline P CCA } CCG }	CAU } Histidine H CAC } CAA } Glutamine Q CAG }	CGU } CGC } Arginine R CGA } CGG }		U C A G
	A	AUU } Isoleucine I AUC } AUA } AUG } Methionine start codon M	ACU } ACC } Threonine T ACA } ACG }	AAU } Asparagine N AAC } AAA } Lysine K AAG }	AGU } Serine S AGC } AGA } Arginine R AGG }		U C A G
	G	GUU } GUC } Valine V GUA } GUG }	GCU } GCC } Alanine A GCA } GCG }	GAU } Aspartic acid D GAC } GAA } Glutamic acid E GAG }	GGU } GGC } Glycine G GGA } GGG }		U C A G

3 нуклеотида (кодон) → 1 аминокислота (aa)



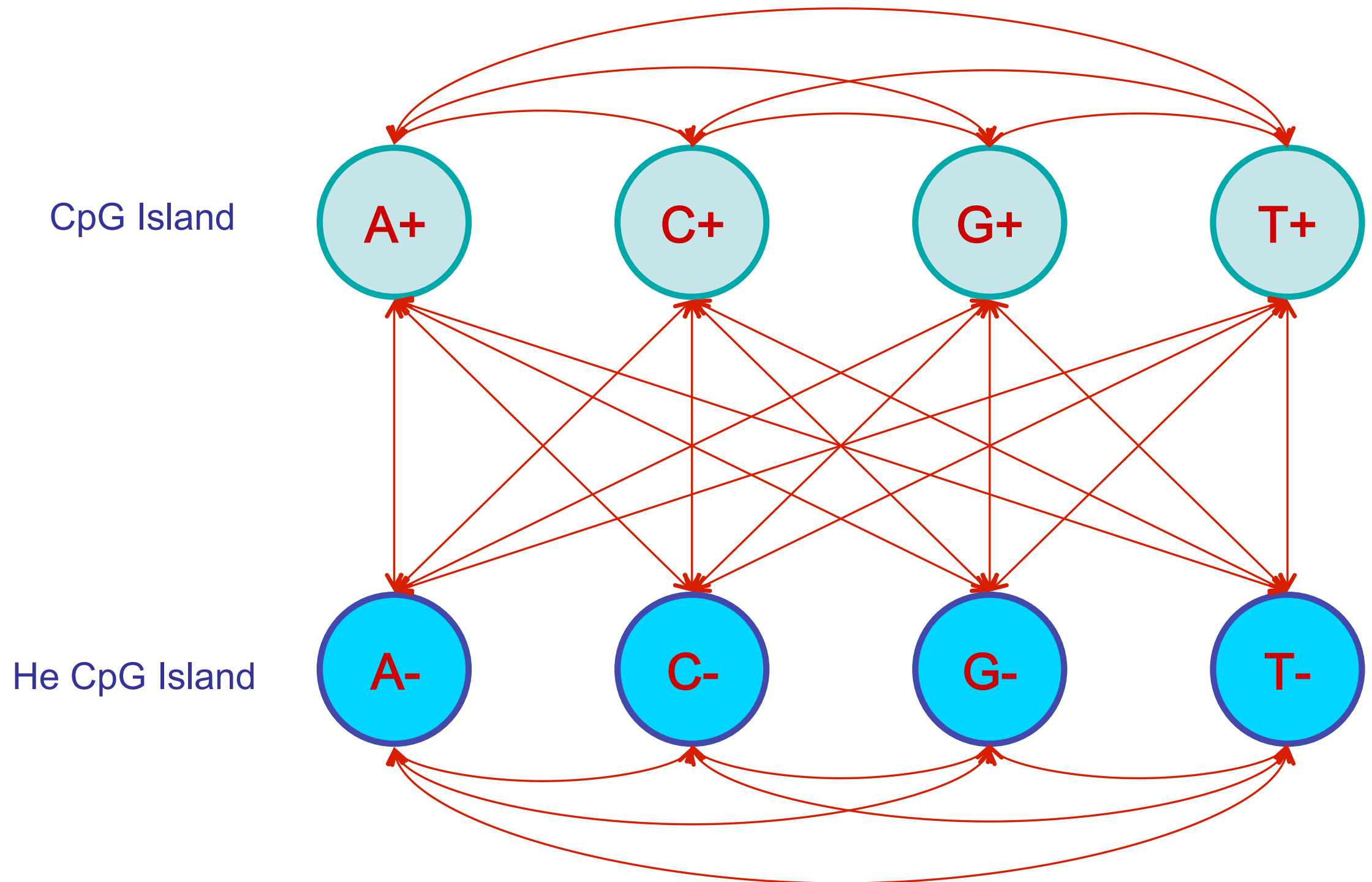
Предсказание генов

aatgcatgцggctatgctaatagcatgцggctatgctaagctgggatccgatgacaatgcat
gцggctatgctaatagcatgцggctatgcaagctgggatccgatgactatgctaagctggga
tccgatgacaatgcatgцggctatgctaatagaatgggtcttgggattttaccttggaatgcta
agctgggatccgatgacaatgcatgцggctatgctaatagaatgggtcttgggattttaccttg
gaatatgctaatagcatgцggctatgctaagctgggatccgatgacaatgcatgцggctatg
ctaatagcatgцggctatgcaagctgggatccgatgactatgctaagctgцggctatgctaa
tgcatgцggctatgctaagctgggatccgatgacaatgcatgцggctatgctaatagcatgц
ggctatgcaagctgggatcctgцggctatgctaatagaatgggtcttgggattttaccttgga
tgctaagctgggatccgatgacaatgcatgцggctatgctaatagaatgggtcttgggattta
ccttggaatatgctaatagcatgцggctatgctaagctgggaatgcatgцggctatgctaag
ctgggatccgatgacaatgcatgцggctatgctaatagcatgцggctatgcaagctgggatc
cgatgactatgctaagctgцggctatgctaatagcatgцggctatgctaagctcatgцggct
atgctaagctgggaatgcatgцggctatgctaagctgggatccgatgacaatgcatgцggc
tatgctaatagcatgцggctatgcaagctgggatccgatgactatgctaagctgцggctatg
ctaatagcatgцggctatgctaagctcggctatgctaatagaatgggtcttgggattttaccttg
gaatgctaagctgggatccgatgacaatgcatgцggctatgctaatagaatgggtcttgggat
ttaccttggaatatgctaatagcatgцggctatgctaagctgggaatgcatgцggctatgct
aagctgggatccgatgacaatgcatgцggctatgctaatagcatgцggctatgcaagctggg
atccgatgactatgctaagctgцggctatgctaatagcatgцggctatgctaagctcatgцg

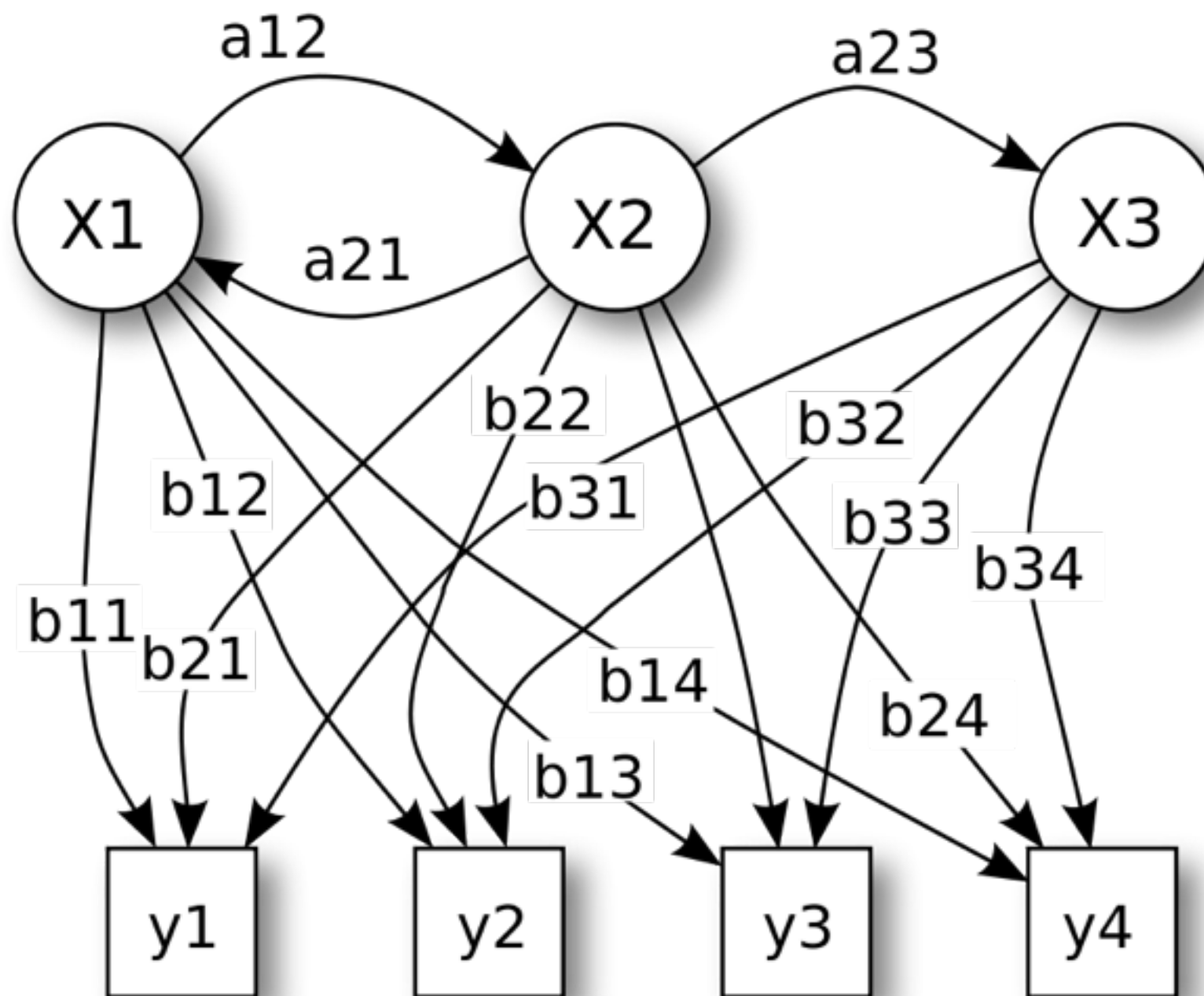
Предсказание генов

aatgcatgцggctatgctaatagcatgцggctatgctaagctgggatccgatgacaatgcat
gцggctatgctaatagcatgцggctatgcaagctgggatccgatgactatgctaagctggga
tccgatgacaatgcatgцggctatgctaatagaatgggtcttgggattttaccttggaatgcta
agctgggatccgatgacaatgcatgцggctatgctaatagaatgggtcttgggattttaccttg
gaatatgctaatagcatgцggctatgctaagctgggatccgatgacaatgcatgцggctatg
ctaatagcatgцggctatgcaagctgggatccgatgactatgctaagctgцggctatgctaa
tgcatgцggctatgctaagctgggatccgatgacaatgcatgцggctatgctaatagcatgc
ggctatgcaagctgggatcctgцggctatgctaatagaatgggtcttgggattttaccttgga
tgctaagctgggatccgatgacaatgcatgцggctatgctaatagaatgggtcttgggattta
ccttggaatatgctaatagcatgцggctatgctaagctgggaatgcatgцggctatgctaag
ctgggatccgatgaca**atgcatgцggctatgctaatagcatgцggctatgcaagctgggatc**
cgatgactatgctaagctgцggctatgctaatagcatgцggctatgctaagctcatgцggct
atgctaagctgggaatgcatgцggctatgctaagctgggatccgatgacaatgcatgцggc
tatgctaatagcatgцggctatgcaagctgggatccgatgactatgctaagctgцggctatg
ctaatagcatgцggctatgctaagctcggctatgctaatagaatgggtcttgggattttaccttg
gaatgctaagctgggatccgatgacaatgcatgцggctatgctaatagaatgggtcttgggat
ttaccttggaatatgctaatagcatgцggctatgctaagctgggaatgcatgцggctatgct
aagctgggatccgatgacaatgcatgцggctatgctaatagcatgцggctatgcaagctggg
atccgatgactatgctaagctgцggctatgctaatagcatgцggctatgctaagctcatgцg

HMM

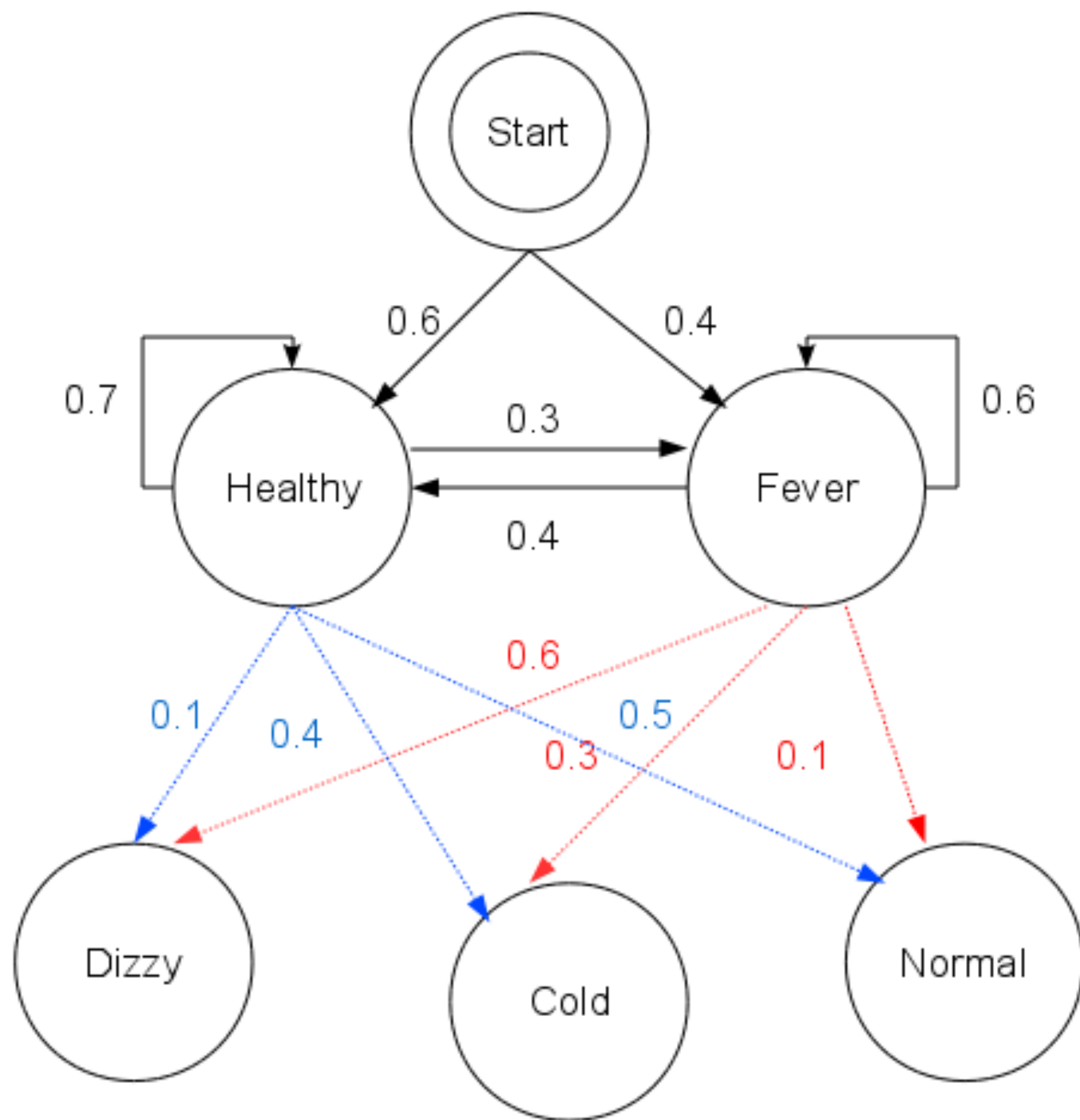


HMM

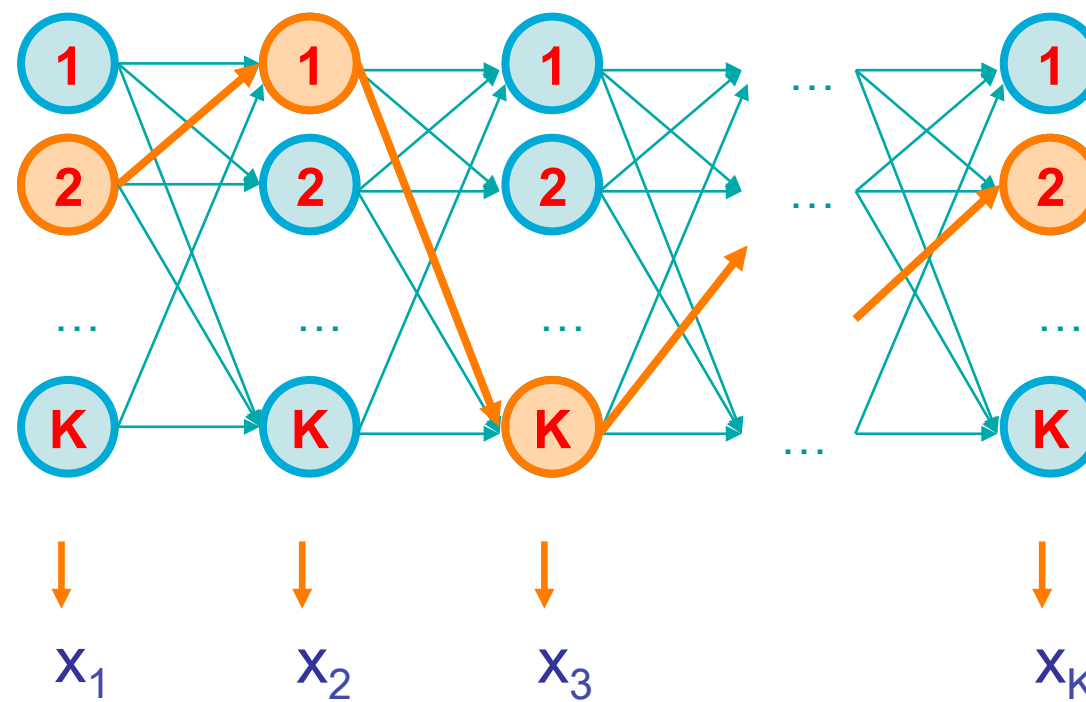


Алгоритм Витерби

Поиск наиболее подходящего списка состояний, который получает наиболее вероятную последовательность произошедших событий.



Алгоритм Витерби



Multiple sequence alignment

Sequence 1:

Sequence 2:

Sequence 3:

Sequence 4:

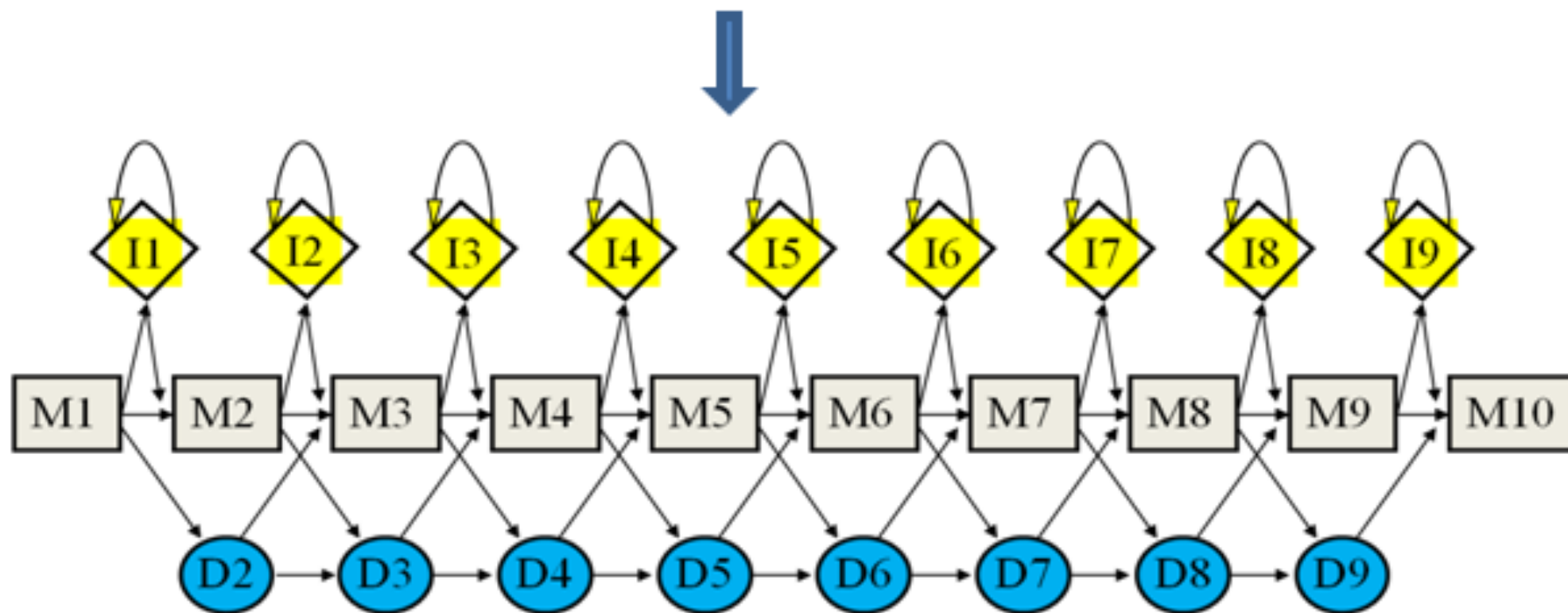
Sequence 5:

Sequence 6:

Sequence 7:

F	K	L	L	S	H	C	L	L	V
F	K	A	F	G	Q	T	M	F	Q
Y	P	I	V	G	Q	E	L	L	G
F	P	V	V	K	E	A	I	L	K
F	K	V	L	A	A	V	I	A	D
L	E	F	I	S	E	C	I	I	Q
F	K	L	L	G	N	V	L	V	C

PFAM



I = insert state

M = match state

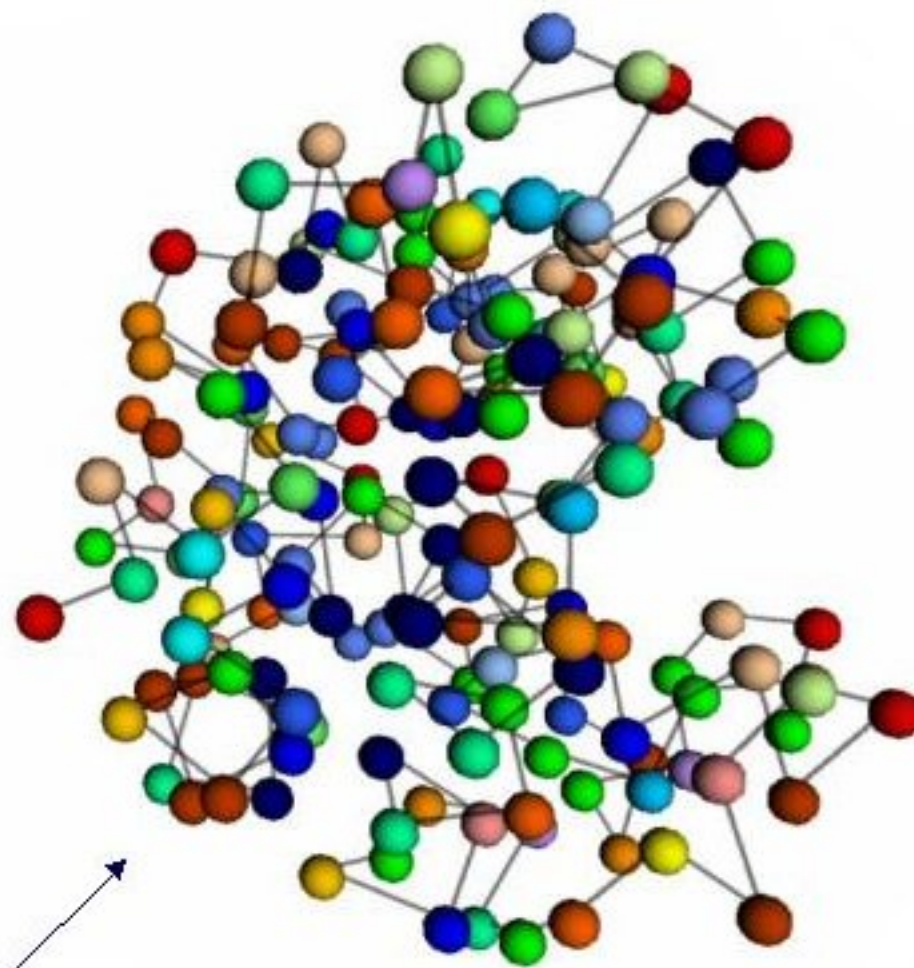
D = delete state

Белки

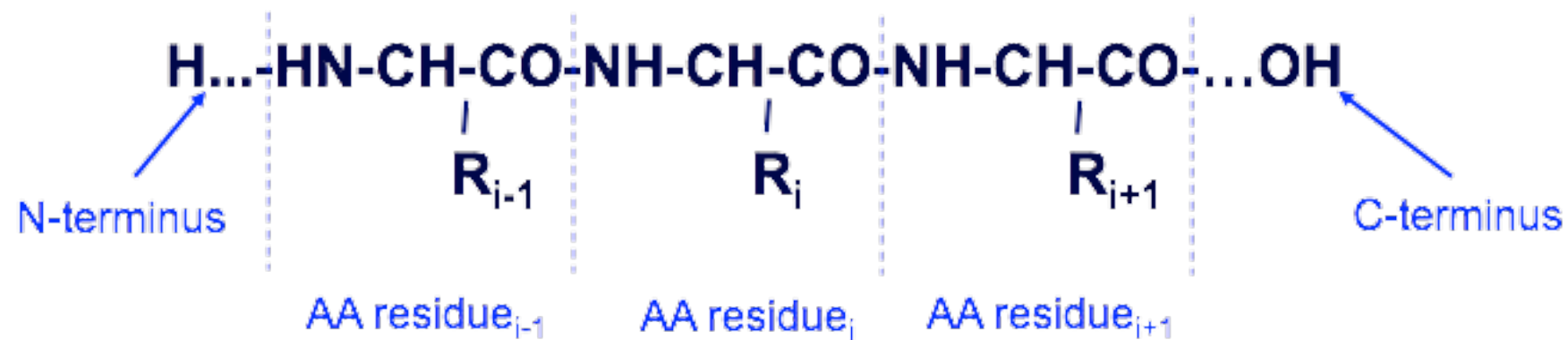
R T D C Y G N V N R I D T T G
A S C K T A K P E G L S Y C G
V S A S K K I A E R D L Q A M
D R Y K T I I K K V G E K L C
V E P A V I A G I I S R E S H
A G K V L K N G W G D R G N G
F G L M Q V D K R S H K P Q G
T W N G E V H I T Q G T T I L
I N F I K T I Q K K F P S W T
K D Q Q L K G G I S A Y N A G
A G N V R S Y A R M D I G T T
H D D Y A N D V V A R A Q Y Y
K Q H G Y

Primary Sequence

3D Structure



Масс-спектрометрия



Масс-спектрометрия

PRTEIN

Масс-спектрометрия

PRTEIN

P RTEIN

Масс-спектрометрия

PRTEIN

P RTEIN

PR TEIN

Масс-спектрометрия

PRTEIN

P RTEIN

PR TEIN

PRT EIN

PRTE IN

PRTEI N

Масс-спектрометрия

$$m(P) = 97$$

PRTEIN

$$m(R) = 156$$

P RTEIN

$$m(T) = 101$$

PR TEIN

$$m(E) = 129$$

PRT EIN

$$m(I) = 113$$

PRTE IN

$$m(N) = 114$$

PRTEI N

Масс-спектрометрия

$$m(P) = 97$$

PRTEIN

710

$$m(R) = 156$$

P RTEIN

97

613

$$m(T) = 101$$

PR TEIN

253

457

$$m(E) = 129$$

PRT EIN

354

356

$$m(I) = 113$$

PRTE IN

483

227

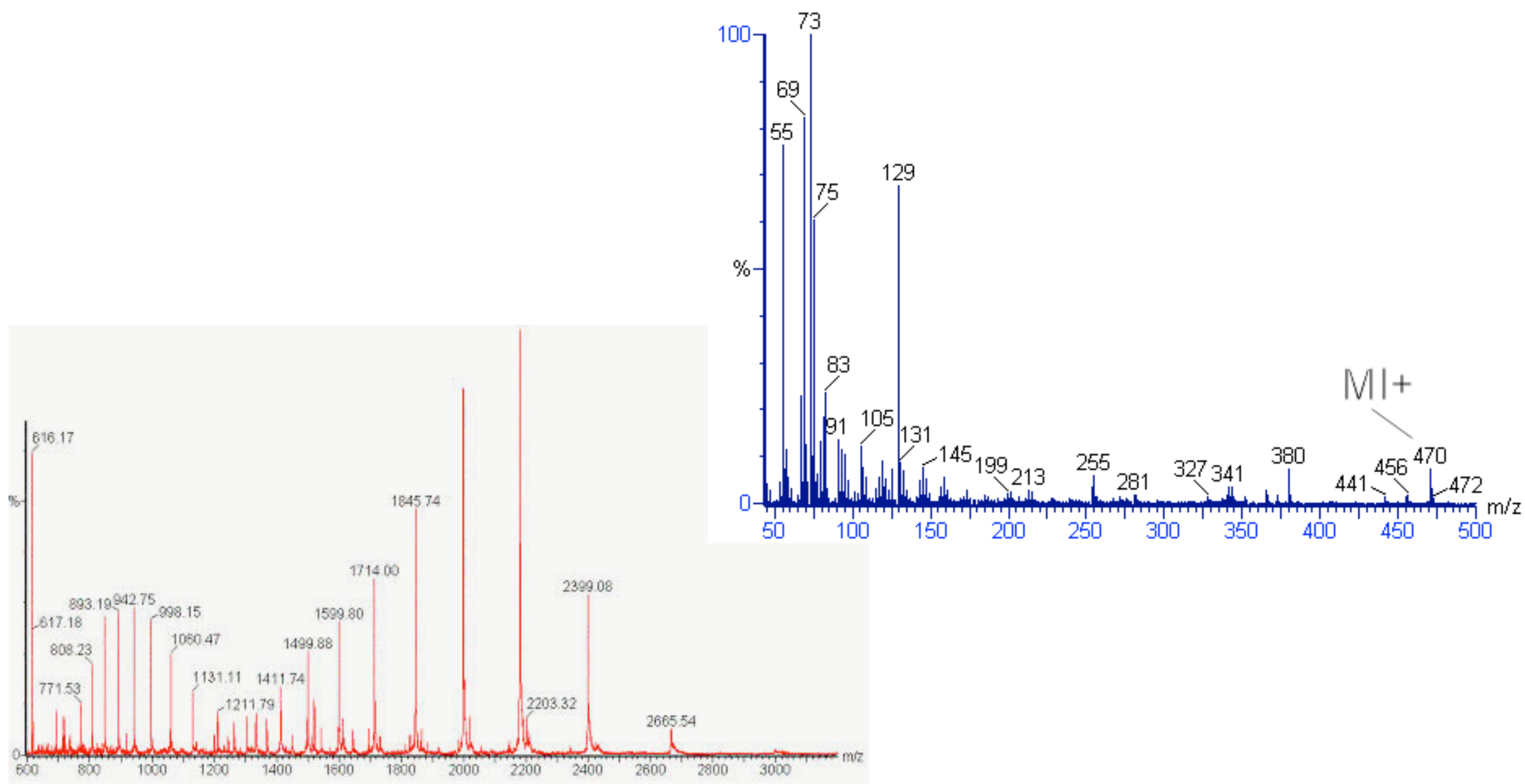
$$m(N) = 114$$

PRTEI N

596

114

Спектры



Сборка белка

Де ново

По базе данных