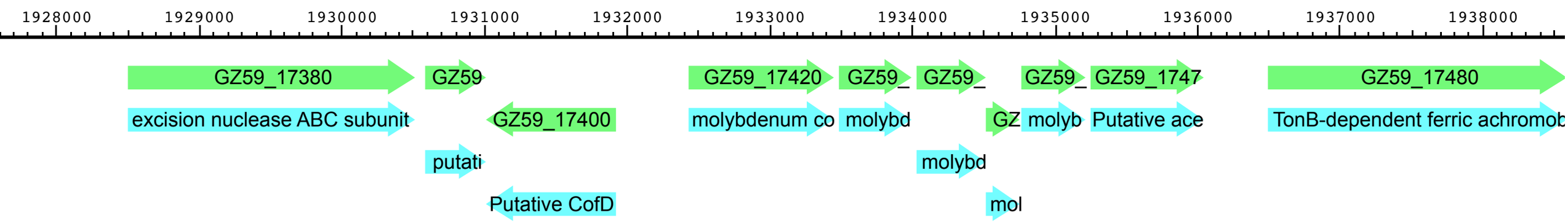
BELARUSIAN STATE UNIVERSITY

Идентификация бактериальных операторов на основе 3D-структур комплексов транскрипционных факторов с ДНК: алгоритм и результаты его применения

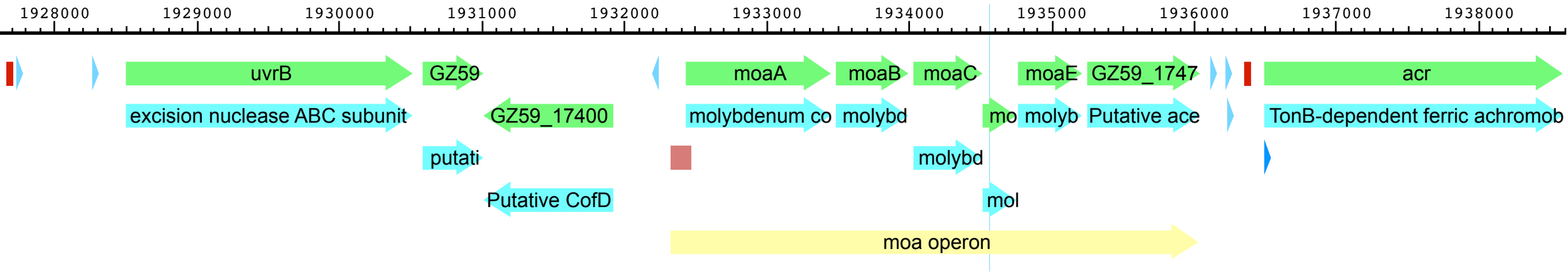
Е. А. Николайчик

Тысячи бактериальных геномов секвенированы, но очень немногие аннотированы качественно



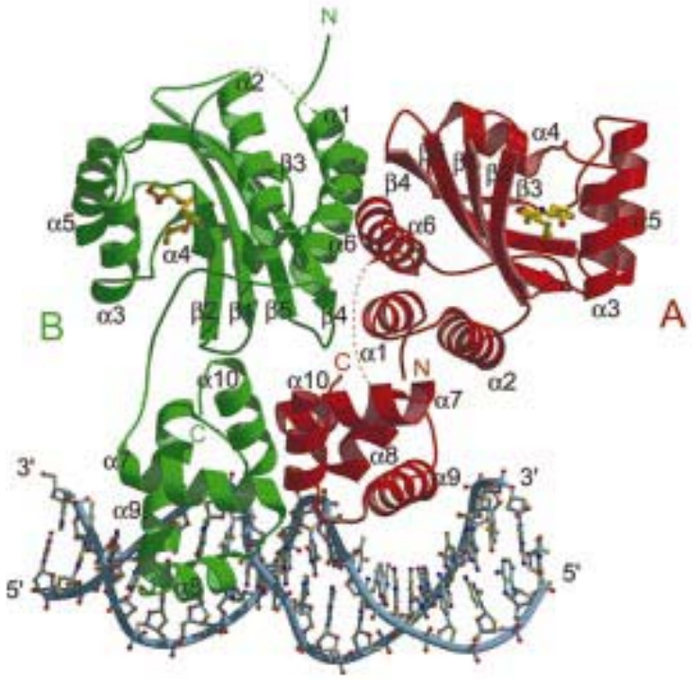
- **Typical annotation: gene, CDS, rRNA, tRNA features only**
- **Transcription control signals (promoters, operators, attenuators, riboswitches, terminators) are rarely annotated**

Without regulatory info one doesn’t know if, when and how strongly a gene is expressed

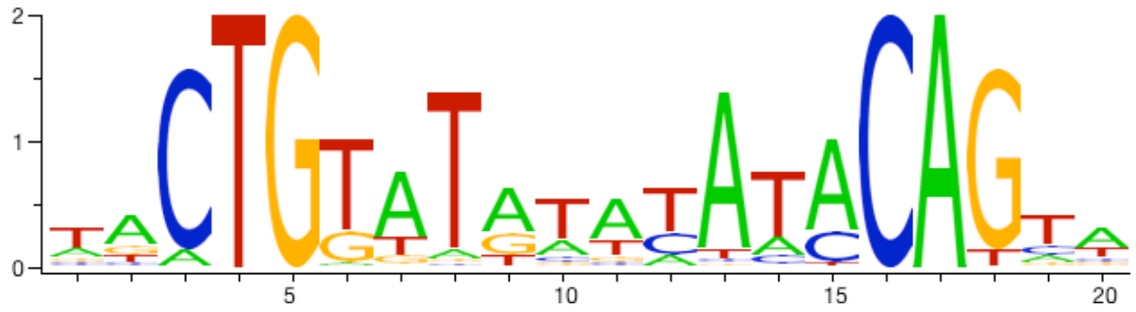


Транскрипционные факторы и сайты их связывания – ключ к пониманию регуляции

TF dimer

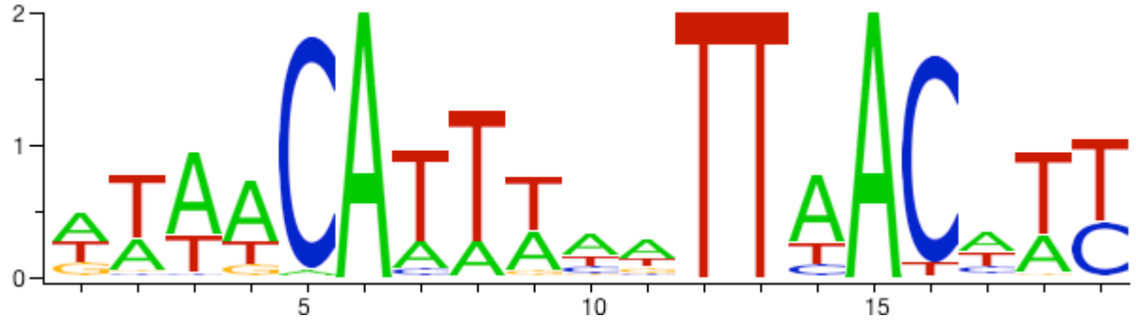


TF binding site
(TFBS or operator)



LexA
(I=21.2 bits)

operator motif
with inverted repeats



ArcA
(I=21.0 бит)

operator motif
with direct repeats

SigmoID версии 1: применение регуляторной информации к неизученным последовательностям геномов

nikolaichik / **SigmoID**

Watch 2 | Star 4 | Fork 0

Code | Issues 0 | Pull requests 0 | Projects 0 | Pulse | Graphs

A Xojo/python tool to ease annotation of sigma-factor and TF binding sites in bacterial genomes

360 commits | **8** branches | **3** releases | **2** contributors | GPL-3.0

Branch: **Version2** | New pull request | Find file | Clone or download

This branch is 6 commits ahead, 1 commit behind master. | Pull request | Compare

nikolaichik Integrated cd-hit for promoter seq clustering — Latest commit `9000d5e` on Sep 8

Build_resources	Integrated cd-hit for promoter seq clustering	2 months ago
Classes	Family TFBSs export to a single meme file	2 months ago

PeerJ

Submitted 26 January 2016
Accepted 29 April 2016
Published 24 May 2016

SigmoID: a user-friendly tool for improving bacterial genome annotation through analysis of transcription control signals

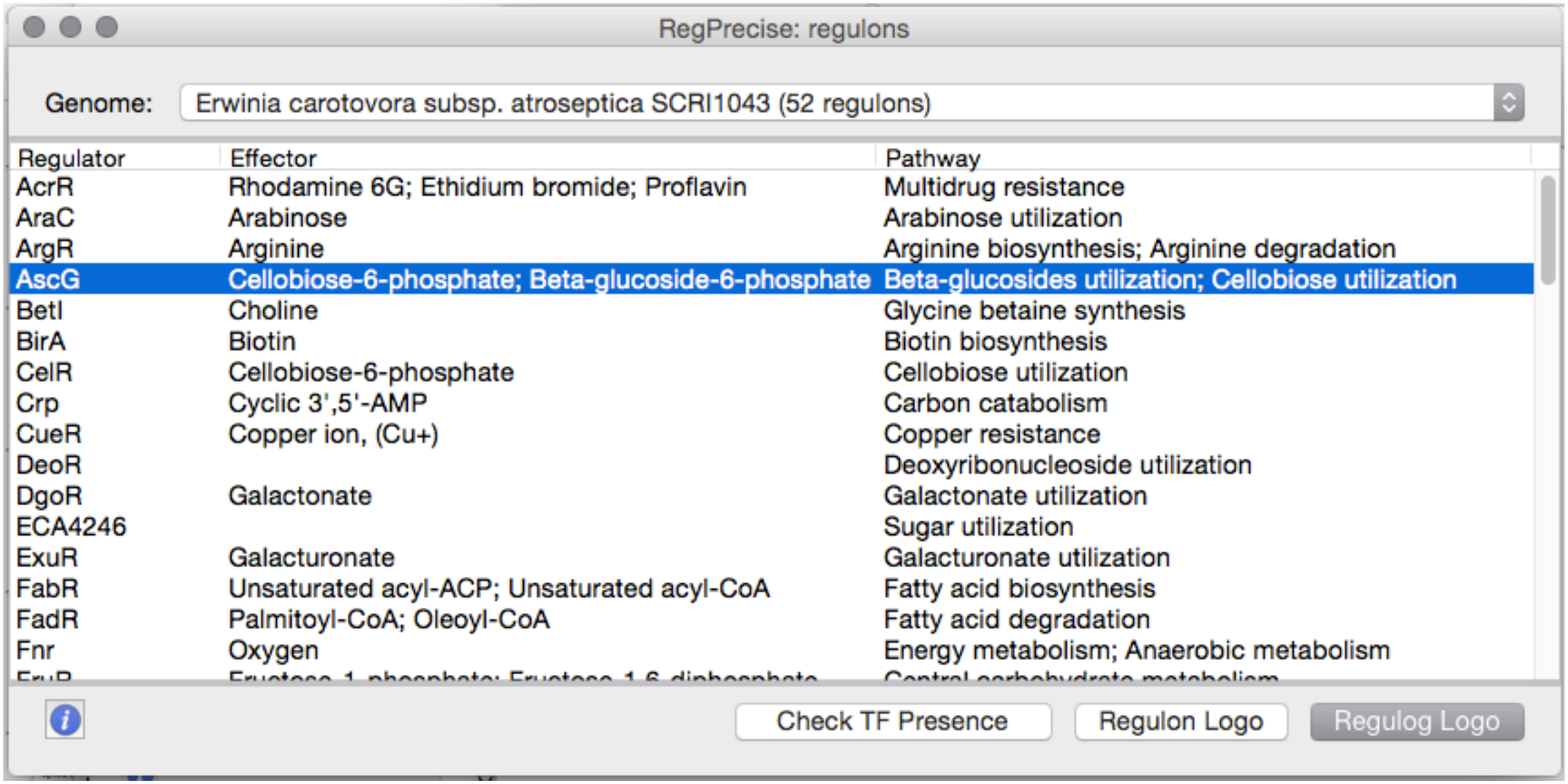
Yevgeny Nikolaichik and Aliaksandr U. Damienikan

Department of Molecular Biology, Belarusian State University, Minsk, Belarus

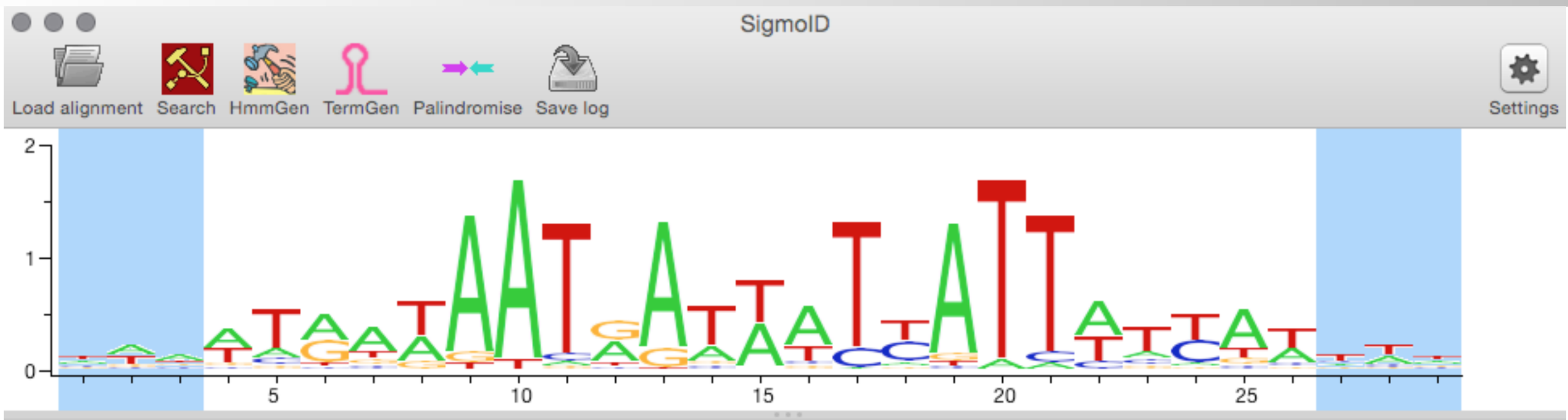
ABSTRACT

The majority of bacterial genome annotations are currently automated and based

С помощью SigmoID можно извлечь информацию о сайтах связывания из RegPrecise…



… оптимизировать операторный мотив для конкретного вида и штамма…



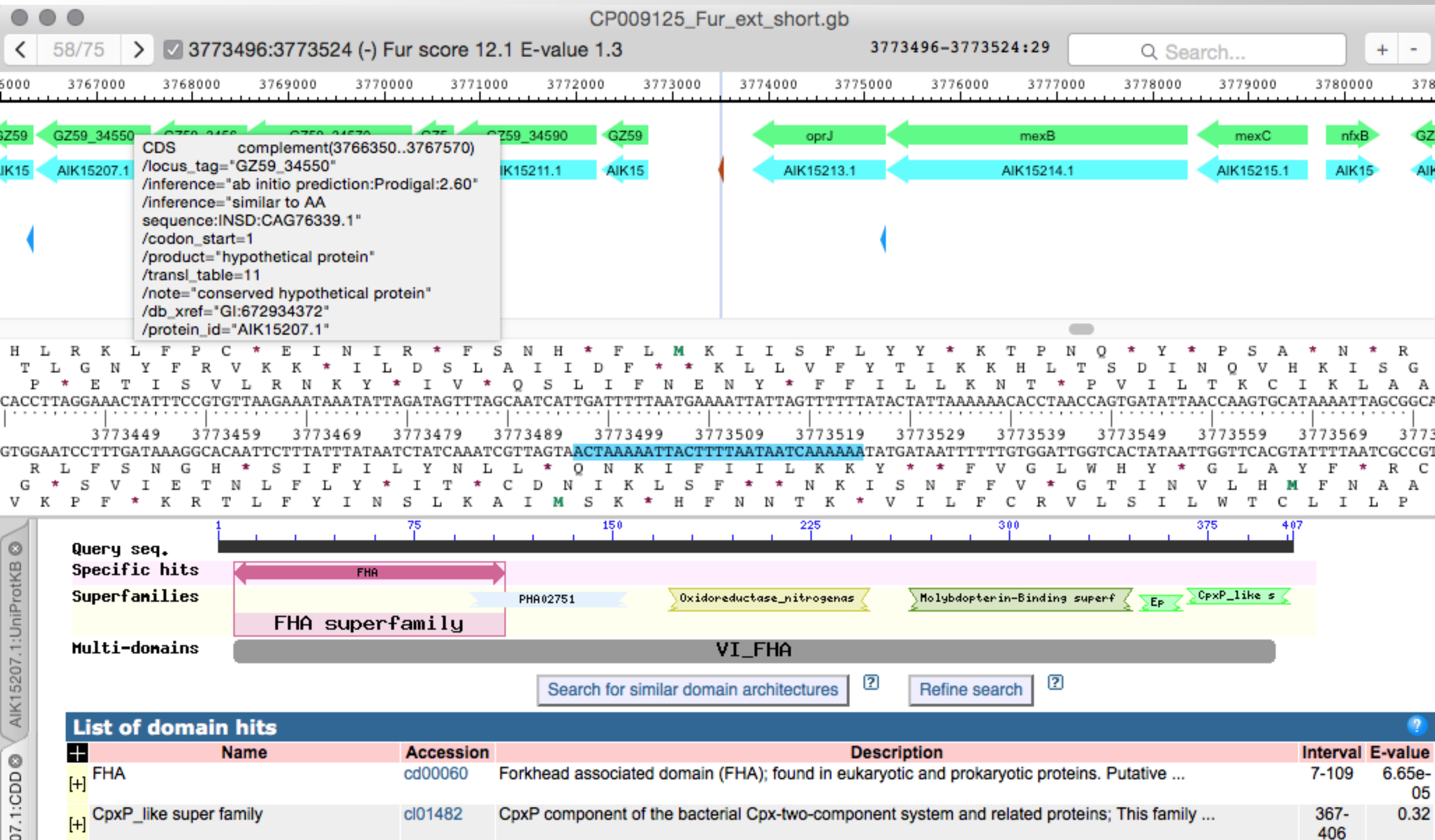

```
    score  bias    Evalue   hmmfrom    hmm to     alifrom    ali to      envfrom    env to     sq len      acc
    ------ ----- --------- ------- ------- ..    --------- ---------    --------- --------- ---------     ----
?    10.1   0.0       8.9         3        24 ..    1981275   1981296 ..    1981273   1981299 ..    4991806    0.88

  Alignment:
  score: 10.1 bits
  alimaskTmp       3 tataataAtaattattaTtatt 24
                       ata t Ataa tat+aTta t
    CP009125 1981275 CATATTGATAAATATCATTACT 1981296
                     689999**************965 PP



Internal pipeline statistics summary:
-------------------------------------
Query model(s):                            1  (29 nodes)
Target sequences:                          1  (9983612 residues searched)
Residues passing SSV filter:         3274939  (0.328); expected (0.3)
Residues passing bias filter:        3274939  (0.328); expected (0.3)
Residues passing Vit filter:         3274939  (0.328); expected (1)
```

…найти сайты связывания регулятора в геноме и при необходимости скорректировать аннотацию.



Есть коллекции операторных последовательностей, но они ограничены по объему, видовой специфике, чатоте обновлений

Organism specific:
RegulonDB (E. coli) – 141 TF, ~50 usable
DBTBS (B. subtilis) – 40 usable (Outdated)
RhizoRegNet - 46 usable (defunct)

Multi-species:
CollecTF – 84 of 252
Prodoric2 – 261 (few usable)
RegTransBase – 141 (defunct)
RegPrecise – 1184 usable

Что можно сделать для энтеробактерии с известными операторными мотивами?

Пример – *Pectobacterium atrosepticum* 21A:

- ✦ найдено почти 100 ортологов ТФ с известными операторными мотивами
- ✦ Идентифицировано около 1000 операторов
- ✦ Найденные операторы контролируют около 40% всех транскрипционных единиц.

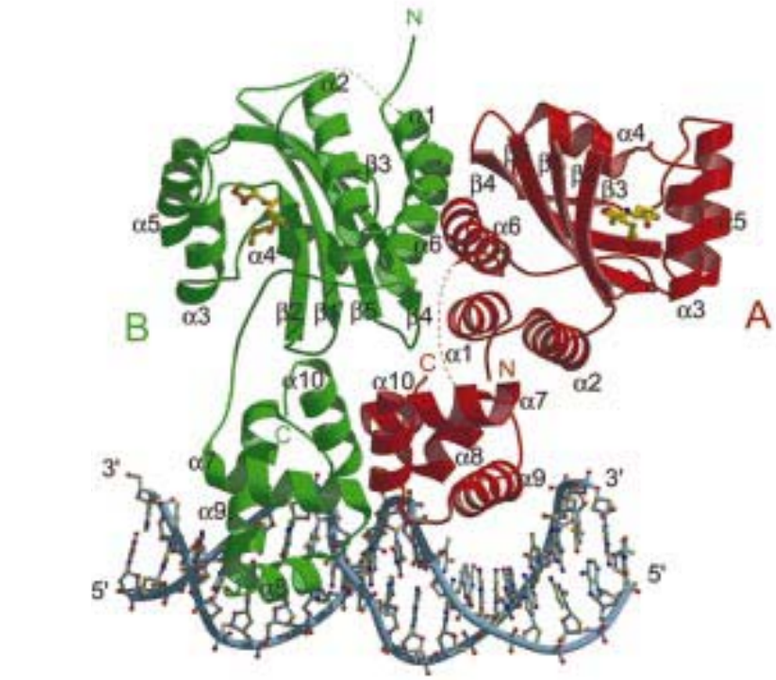
Проблемы:

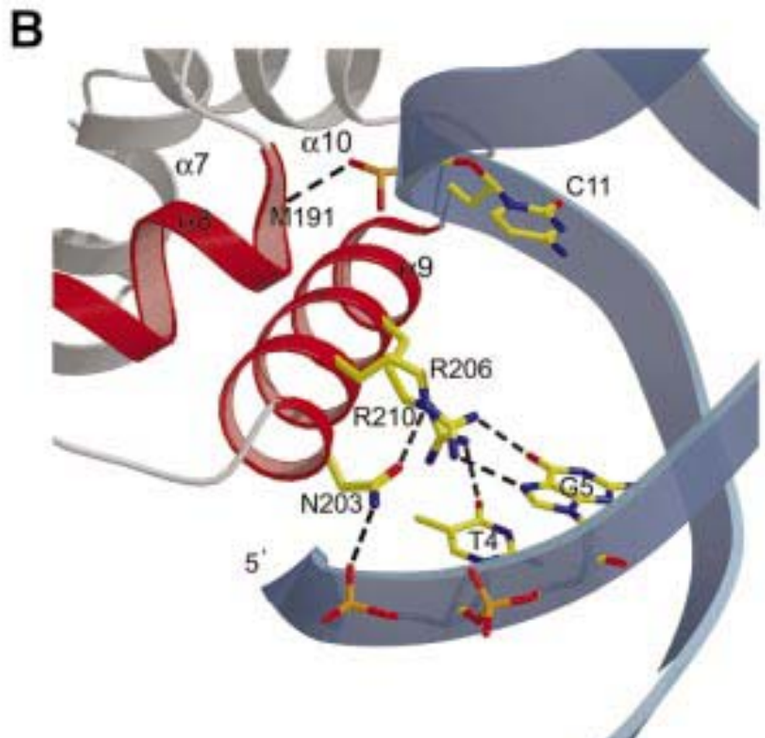
- только для 6 ТФ есть экспериментальное подтверждение их операторов
- процедура опирается на предположение (которое никогда не проверяется!) о консервативности операторов у родственных организмов

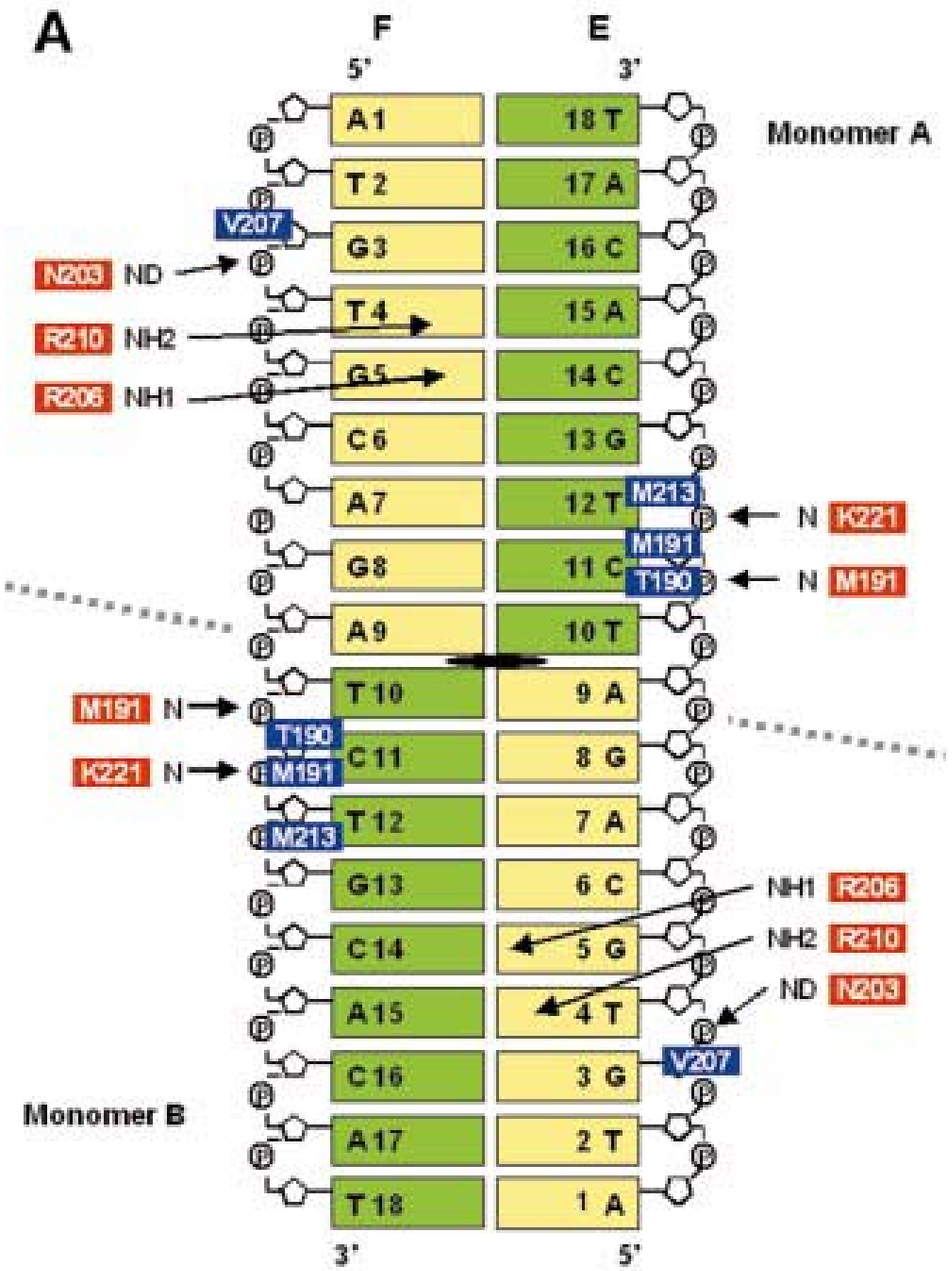
Есть ли лучший способ?

Последовательности операторов считываются несколькими **критичными остатками** ДНК-связывающих доменов

Many TF-TFBS structures are solved, CR-nucleotide contacts are known, but this info is never used!



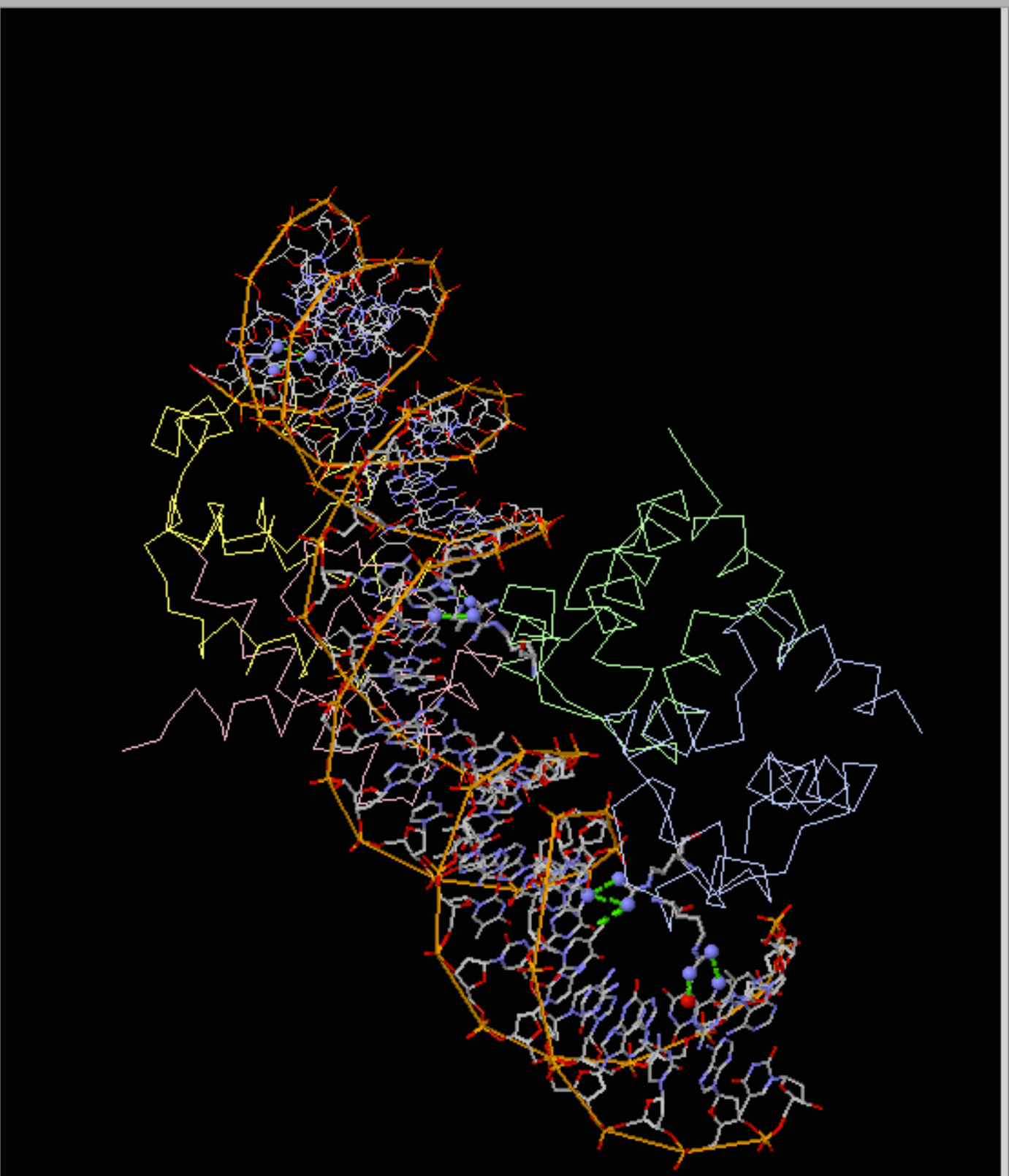




A. Vannini et al. The crystal structure of the quorum sensing protein TraR bound to its autoinducer and target DNA. EMBO J. 21:4393-4401 (2002)

NPIDB (npidb.belozersky.msu.ru) – удобная база данных для поиска критичных аминокислотных остатков

NPIDB: 4X4D structure



NPIDB: Interaction

H-bonds | Hydrophobic Interactions | Help

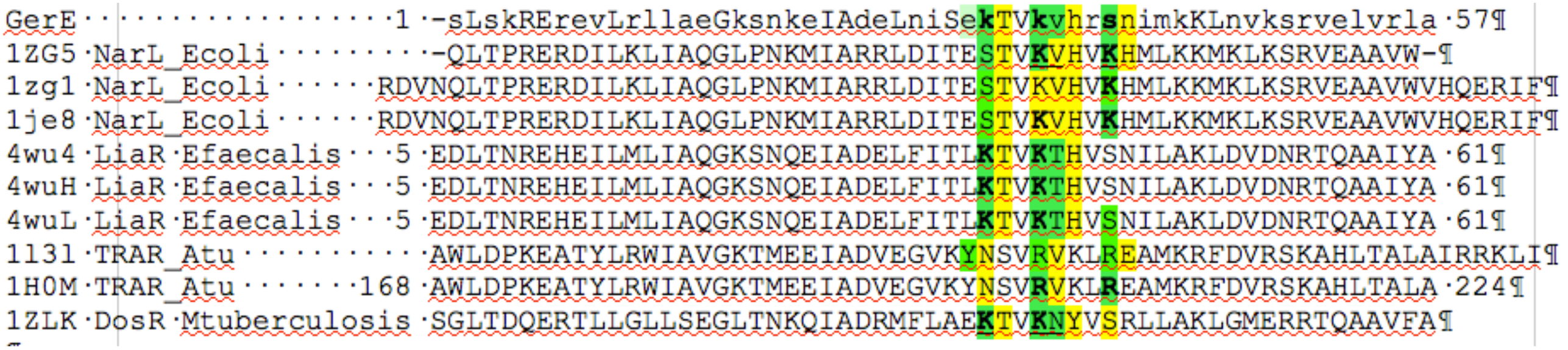
- Show specific contacts
- Show all contacts
- Hide all contacts

[Download]

Nucleic atom	dist	Protein atom
DG3:F.OP2/1	2.98	GLN24:D.NE2/1
DG3:F.OP2/1	3.04	SER39:D.OG/1
DG3:F.N7/1	3.49	ARG35:D.NH1/1
DG3:F.N7/1	3.34	ARG35:D.NH2/1
DG3:F.O6/1	2.71	ARG35:D.NH2/1
DT4:F.OP2/1	2.68	ARG43:D.NE/1
DT4:F.OP2/1	2.4	ARG43:D.NH2/1
DT4:F.O5'/1	2.89	ARG43:D.NH2/1
DA12:F.OP1/1	3.19	ASN47:C.ND2/1
DA13:F.OP1/1	3.41	THR49:C.N/1
DA13:F.OP1/1	2.56	SER52:C.OG/1
DA13:F.OP2/1	2.11	TYR37:C.OH/1
DT14:F.OP2/1	3.4	ASP34:C.N/1
DC17:F.OP1/1	3.03	ARG17:B.NH1/1
DC17:F.OP2/1	2.93	GLN24:B.N/1
DC17:F.N4/1	2.66	ARG35:C.NH2/1
DA18:F.OP2/1	2.83	GLN24:B.NE2/1
DA18:F.N6/1	3.01	ARG35:B.NH1/1
DC19:F.OP2/1	3.01	ARG43:B.NE/1
DA27:F.OP1/1	2.83	ASN47:A.ND2/1
DA28:F.OP1/1	2.99	THR49:A.N/1
DA28:F.OP1/1	2.79	SER52:A.OG/1
DA28:F.OP2/1	3.1	TYR37:A.OH/1
DG29:F.N7/1	3.37	ARG46:A.NH1/1

Для основных семейств транскрипционных факторов определены структуры их ДНК-связывающих доменов в комплексе с ДНК

LuxR-семейство: структуры сильно варьируют, но контактирующие с ДНК (критичные) аминокислотные остатки **всегда в трех соседних витках распознающей α-спирали**



Идентифицировать и классифицировать транскрипционные факторы легко с помощью моделей PFAM

The clan contains the following 256 members:

AbiEi_3_N	AbiEi_4	ANAPC2	AphA_like	Arg_repressor	ARID	B-block_TFIIIC
Bac_DnaA_C	BetR	Bot1p	BrkDBD	C_LFY_FLO	Cdc6_C	CENP-B_N
Cro	Crp	CSN8_PSD8_EIF3K	Cullin_Nedd8	CUT	DDRGK	DEP
Dimerisation	Dimerisation2	DsrD	DUF1133	DUF1153	DUF1323	DUF134
DUF1441	DUF1492	DUF1495	DUF1670	DUF1804	DUF1836	DUF1870
DUF2089	DUF2250	DUF2316	DUF2582	DUF3116	DUF3253	DUF3853
DUF3860	DUF3908	DUF433	DUF4364	DUF4447	DUF480	DUF722
DUF739	DUF742	DUF977	E2F_TDP	EAP30	ELL	ESCRT-II
Ets	Exc	F-112	FaeA	Fe_dep_repr_C	Fe_dep_repress	FeoC
FokI_C	FokI_N	Forkhead	Ftsk_gamma	FUR	GcrA	GerE
GntR	HARE-HTH	HemN_C	HNF-1_N	Homeobox	Homeobox_KN	Homez
HPD	HrcA_DNA-bdg	HSF_DNA-bind	HTH_1	HTH_10	HTH_11	HTH_12
HTH_13	HTH_15	HTH_16	HTH_17	HTH_18	HTH_19	HTH_20
HTH_21	HTH_22	HTH_23	HTH_24	HTH_25	HTH_26	HTH_27
HTH_28	HTH_29	HTH_3	HTH_30	HTH_31	HTH_32	HTH_33
HTH_34	HTH_35	HTH_36	HTH_37	HTH_38	HTH_39	HTH_40
HTH_41	HTH_42	HTH_43	HTH_45	HTH_46	HTH_47	HTH_5
HTH_6	HTH_7	HTH_8	HTH_9	HTH_AraC	HTH_AsnC-type	HTH_CodY
HTH_Crp_2	HTH_DeoR	HTH_IclR	HTH_Mga	HTH_micro	HTH_OrfB_IS605	HTH_psq
HTH_Tnp_1	HTH_Tnp_1_2	HTH_Tnp_4	HTH_Tnp_IS1	HTH_Tnp_IS630	HTH_Tnp_ISL3	HTH_Tnp_Mu_1
HTH_Tnp_Mu_2	HTH_Tnp_Tc3_1	HTH_Tnp_Tc3_2	HTH_Tnp_Tc5	HTH_WhiA	HxlR	IBD
IF2_N	IRF	KicB	KORA	KorB	La	LacI
LexA_DNA_bind	Linker_histone	LZ_Tnp_IS481	MADF_DNA_bdg	MarR	MarR_2	MerR
MerR-DNA-bind	MerR_1	MerR_2	Mga	Mnd1	Mor	MotA_activ
MqsA_antitoxin	MRP-L20	Myb_DNA-bind_2	Myb_DNA-bind_3	Myb_DNA-bind_4	Myb_DNA-bind_5	Myb_DNA-bind_6
Myb_DNA-bind_7	Myb_DNA-binding	Neugrin	NUMOD1	OST-HTH	P22_Cro	PaaX
PadR	PAX	PCI	Penicillinase_R	Phage_AlpA	Phage_antitermQ	Phage_CI_repr
Phage_CII	Phage_rep_org_N	Phage_terminase	Pou	Pox_D5	PuR_N	Put_DNA-bind_N
Rap1-DNA-bind	Rep_3	RepA_C	RepA_N	RepC	RepL	Replic_Relax
RFX_DNA_binding	Ribosomal_S19e	Ribosomal_S25	Rio2_N	RNA_pol_Rpc34	RP-C	RPA
RPA_C	RQC	Rrf2	RTP	RuvB_C	SAC3_GANP	SANT_DAMP1_like
SatD	SelB-wing_1	SelB-wing_2	SelB-wing_3	SgrR_N	Sigma54_CBD	Sigma54_DBD
Sigma70_ECF	Sigma70_ner	Sigma70_r2	Sigma70_r3	Sigma70_r4	Sigma70_r4_2	SLIDE
SMC_ScpB	SpoIIID	STN1_2	Sulfolobus_pRN	SWIRM	TBPIP	Terminase_5
TetR_N	TFIIE_alpha	TFIIE_beta	TFIIF_alpha	TFIIF_beta	Tn7_Tnp_TnsA_C	Tn916-Xis
TraI_2_C	Trans_reg_C	TrfA	TrmB	Trp_repressor	UPF0122	Vir_act_alpha_C
YdaS_antitoxin	YjcQ	YokU	z-alpha			

ТФ *Pectobacterium* spp. принадлежат более чем к 30 семействам, но 80% всех ТФ – к 12 основным

Семейство	Pfam #	Число ТФ
LysR	PF00126	59
AraC	PF00165	24
GntR	PF00392	22
LacI	PF00356	20
XRE	PF01381	19
TetR_N	PF00440	17
Trans_reg_C	PF00486	14
LuxR	PF00196	12
HxlR	PF01638	8
Fis	PF02954	8
DeoRC	PF00455	6
MarR	SM00347	6

Как можно использовать структурную информацию?

BIOINFORMATICS **ORIGINAL PAPER**

Vol. 26 no. 21 2010, pages 2672–2677
doi:10.1093/bioinformatics/btq501

Sequence analysis

Advance Access publication August 31, 2010

Novel sequence-based method for identifying transcription factor binding sites in prokaryotic genomes

Gurmukh Sahota and Gary D. Stormo*

Department of Genetics, Washington University School of Medicine, Saint Louis, MO 63108, USA

Associate Editor: Dmitrij Frishman

ABSTRACT

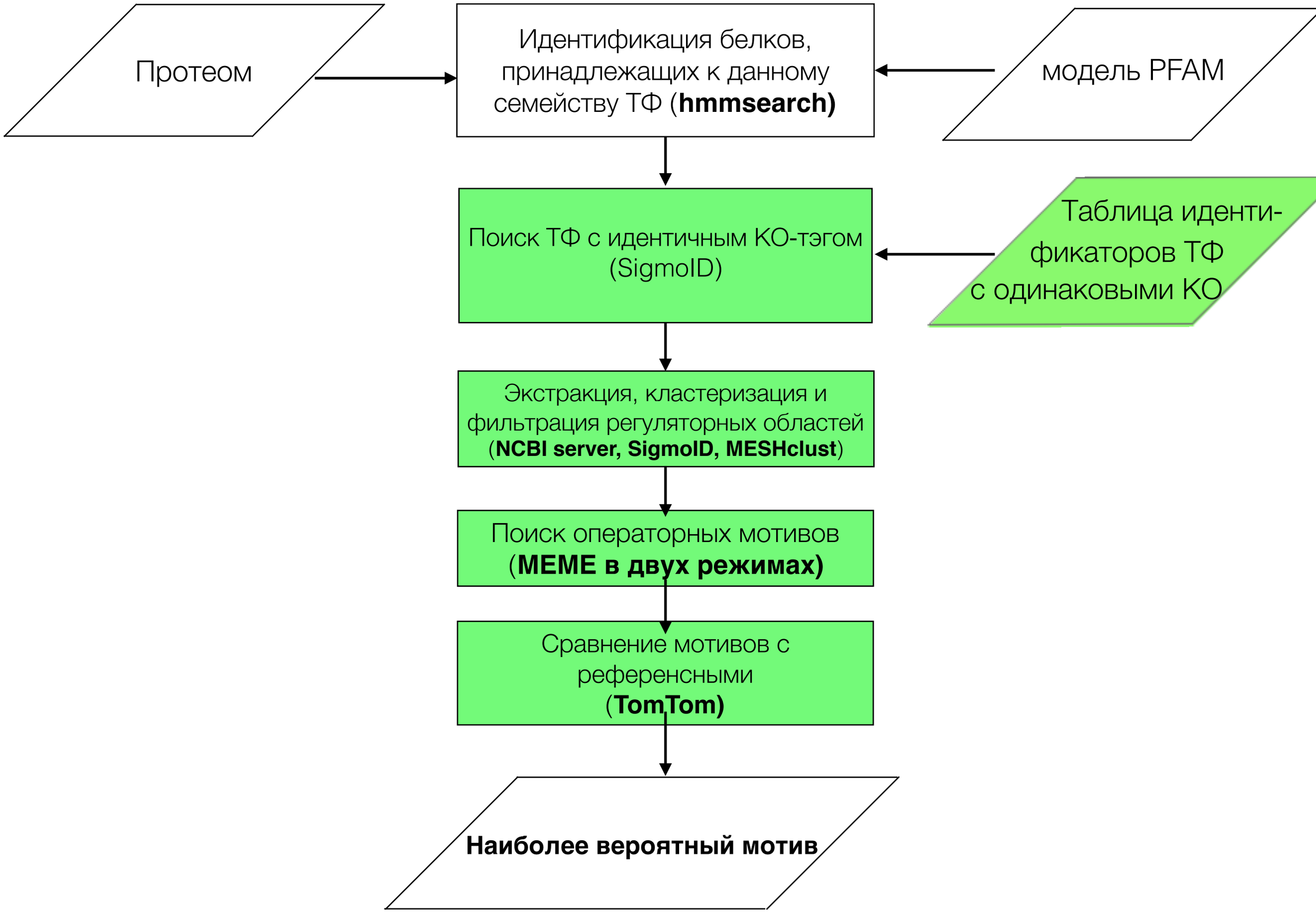
Motivation: Computational techniques for microbial genomic sequence analysis are becoming increasingly important. With next-generation sequencing technology and the human microbiome project underway, current sequencing capacity is significantly greater than the speed at which organisms of interest can be studied experimentally. Most related computational work has been focused on sequence assembly, gene annotation and metabolic network reconstruction. We have developed a method that will primarily use available sequence data in order to determine prokaryotic transcription factor (TF) binding specificities.

Results: Specificity determining residues (critical residues) were identified from crystal structures of DNA–protein complexes and TFs with the same critical residues were grouped into specificity classes. The putative binding regions for each class were defined as the set of promoters for each TF itself (autoregulatory) and the immediately upstream and downstream operons. MEME was used to find putative motifs within each separate class. Tests on the LacI and TetR TF families, using RegulonDB annotated sites, showed the sensitivity of prediction 86% and 80%, respectively.

classification of gene (Qin *et al.*, 2010; Selengut *et al.*, 2010) and metabolic network reconstruction (Ye and Doak, 2009). Many of these analyses are accomplished through the identification of homologous proteins with known function and the inference of functional conservation in the newly sequenced species. As of yet, there has been little computational work focused on transcriptional regulation in these prokaryotic systems. In this article, we present a novel sequence-based method to infer the specificities of prokaryotic transcription factors (TFs) through the comparisons of their DNA-binding domains and applying a motif-finding algorithm to likely binding regions.

Most prokaryotic TFs contain a helix-turn-helix (HTH) fold, where the second helix, also known as the recognition helix, primarily contacts DNA (Harrison, 1991; Perez-Rueda and Collado-Vides, 2000; Santos *et al.*, 2009). Using crystal structures of protein–DNA complexes, we can determine a set of residues that is important for defining the specificity of the protein, the ‘critical residues’. Commonly, these HTH TFs bind as homodimers with palindromic DNA specificities. Previous studies have utilized those features to identify regulatory motifs in related bacterial species

Улучшенный алгоритм поиска ССТФ в SigmoID v.2



Текущая имплементация в программе SigmoID

de novo TFBS inference

Name	Accession	CR tag residues	Description
GerE	PF00196.16	31,35,36,39,40	Bacterial regulatory proteins, luxR family
GntR	PF00392.19	28,37,38,39,58,59	Bacterial regulatory proteins, gntR family
GntR	PF00392.19	27,28,37,38,39,40,43,58,59,62	Bacterial regulatory proteins, gntR family
HTH_1	PF00126.25	27, 28, 29, 31, 32, 49, 51	Bacterial regulatory helix-turn-helix protein, lysR family
HTH_1	PF00126.25	27, 28, 29, 31, 32, 49, 50, 51	Bacterial regulatory helix-turn-helix protein, lysR family
HTH_1	PF00126.25	26, 27, 28, 29, 31, 32, 49, 51	Bacterial regulatory helix-turn-helix protein, lysR family
HTH_27	PF13463.4	32, 33, 34, 35, 37, 38, 53, 57, 59	Winged helix DNA-binding domain
HTH_3	PF01381.20	23, 24, 27, 28, 33, 35	cro/C1-type helix-turn-helix domain
HTH_3	PF01381.20	12, 13, 22, 23, 24, 25, 27, 28, 33, 34	cro/C1-type helix-turn-helix domain
HTH_MARR	SM00347	37, 38, 39, 40, 42, 43, 58, 62, 64	helix_turn_helix multiple antibiotic resistance protein
HTH_MARR	SM00347	37, 38, 39, 40, 42, 43, 64	helix_turn_helix multiple antibiotic resistance protein
HxlR	PF01638.15	30, 32, 34	HxlR-like helix-turn-helix
LacI	PF00356.18	2,3,12,13,14,17,18,24,25,26	Bacterial regulatory proteins, lacI family
TetR_N	PF00440.20	20,29,30,31,32,34,35	Bacterial regulatory proteins, tetR family
XRE_superfamily	SM00530	13, 14, 23, 24, 25, 26, 28, 29, 34, 35	cro/C1-type HTH domain

Select All | Deselect All

Protein database to search

- ○ UniProt (full)
- ◉ UniProt Reference Proteins
  - ☑ Look in full UniProt if less than 10 hits

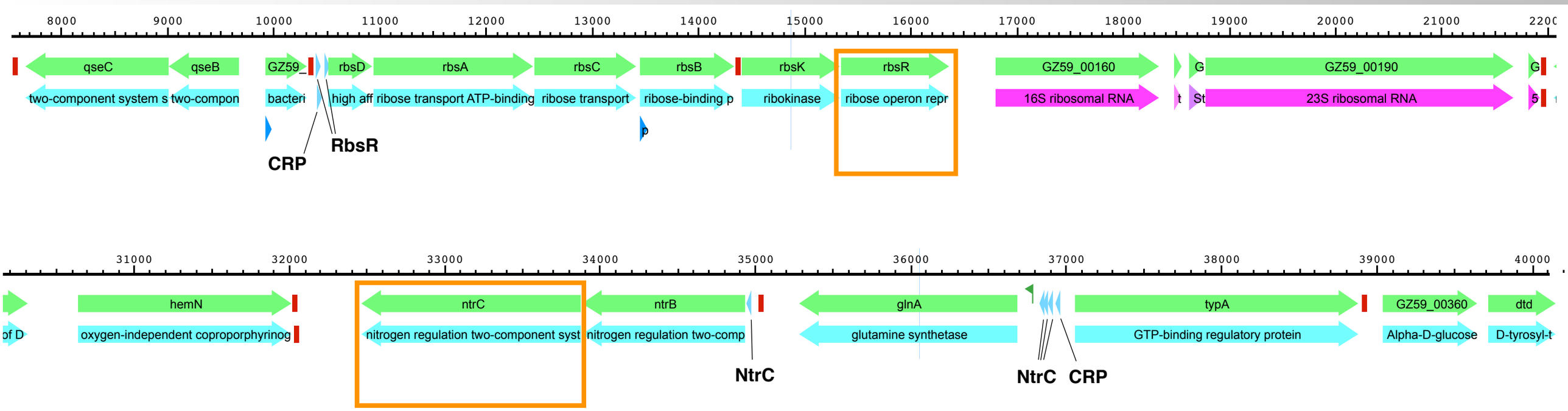
MEME settings

- ☑ Palindromic
- Minimal motif width: 16
- Maximal motif width: 23

Output Folder: [ ] Choose

Cancel | Run!

Ключевой этап конвейера: экстракция регуляторных последовательностей

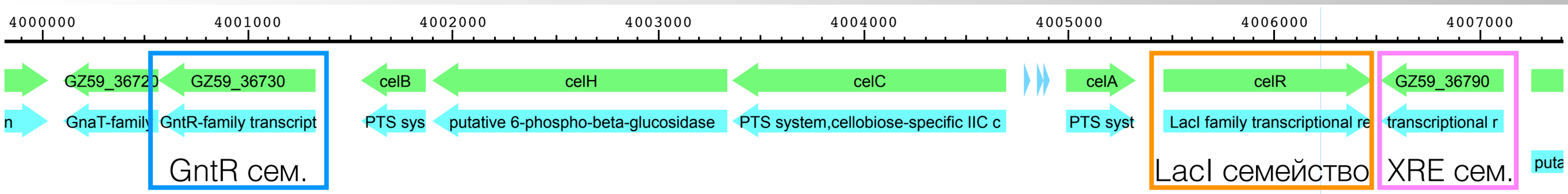


Принцип: экстракция трех промоторных областей (-400..+50): своего оперона и двух соседних с объединением регуляторных областей для дивергонов

Основные проблемы:

1. Ошибки аннотации: пропуск генов (или лишние гены) не позволяют корректно находить регуляторные области.
2. Пропущенные поля в аннотации (могут не дать найти нужную последовательность вообще)
3. Короткие фрагменты геномов

MEME почти всегда находит несколько мотивов. Почему их много и какой "правильный"?

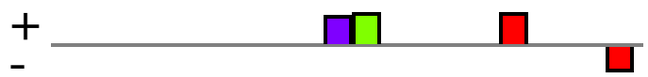
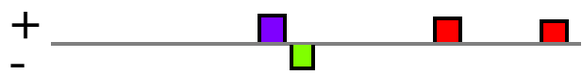


zoops mode

Logo	E-value	Sites
	9.3e-016	11
	6.1e-010	11
	4.9e-009	13
	2.2e-005	10
	1.0e-009	14

anr mode

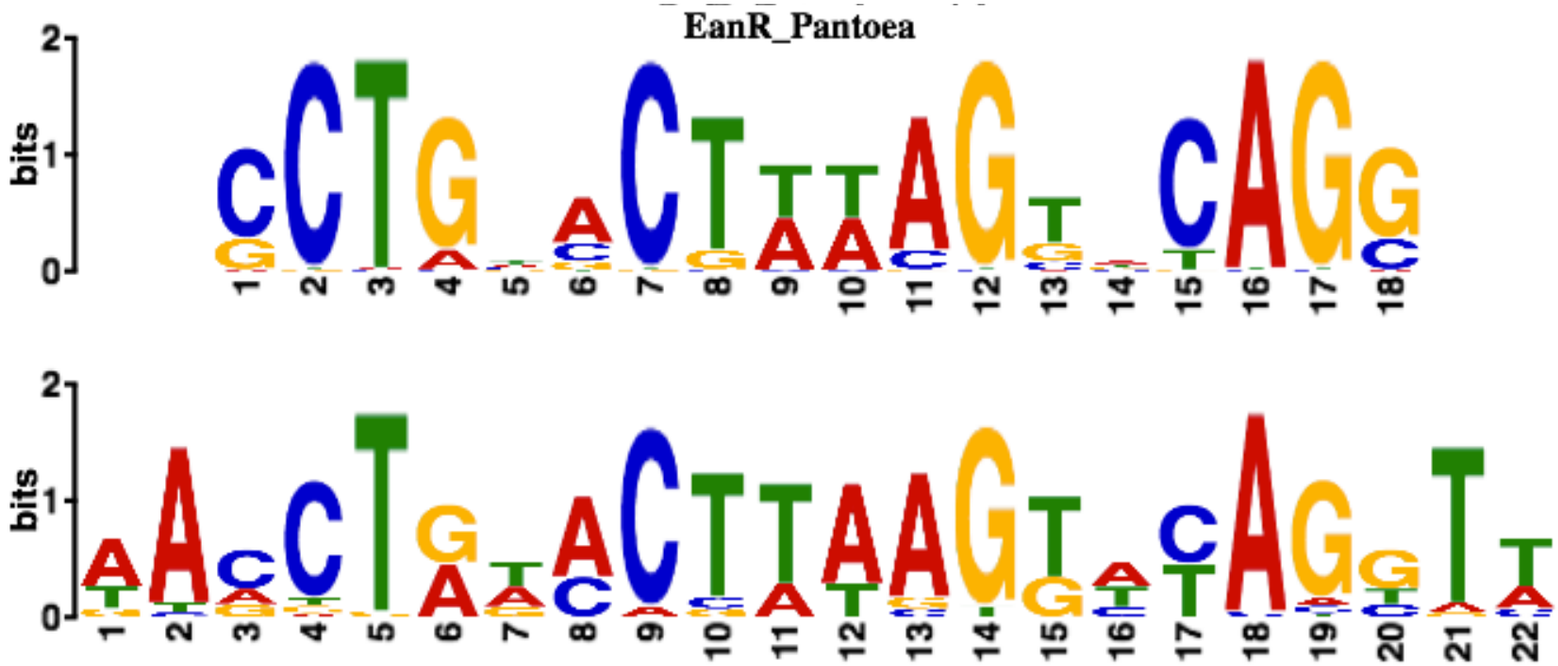
Logo	E-value	Sites
	3.0e-020	24
	3.0e-016	15
	1.3e-006	13
	4.2e-005	13
	6.6e-003	5

Name	*p*-value	Motif Location
1. ECA3649|CAG76547|Pectobacterium	6.17e-10	
2. ascG|Dda3937_02578|ADM99913|Dickeya	7.40e-9	
3. WN53_01995|AKG68009|Serratia	5.08e-11	
4. Ent638_3283|ABP61947|Enterobacter	2.77e-14	
5. Z042_06230|AHG19260|Chania	5.15e-9	

Выбрать корректный мотив позволяет TomTom

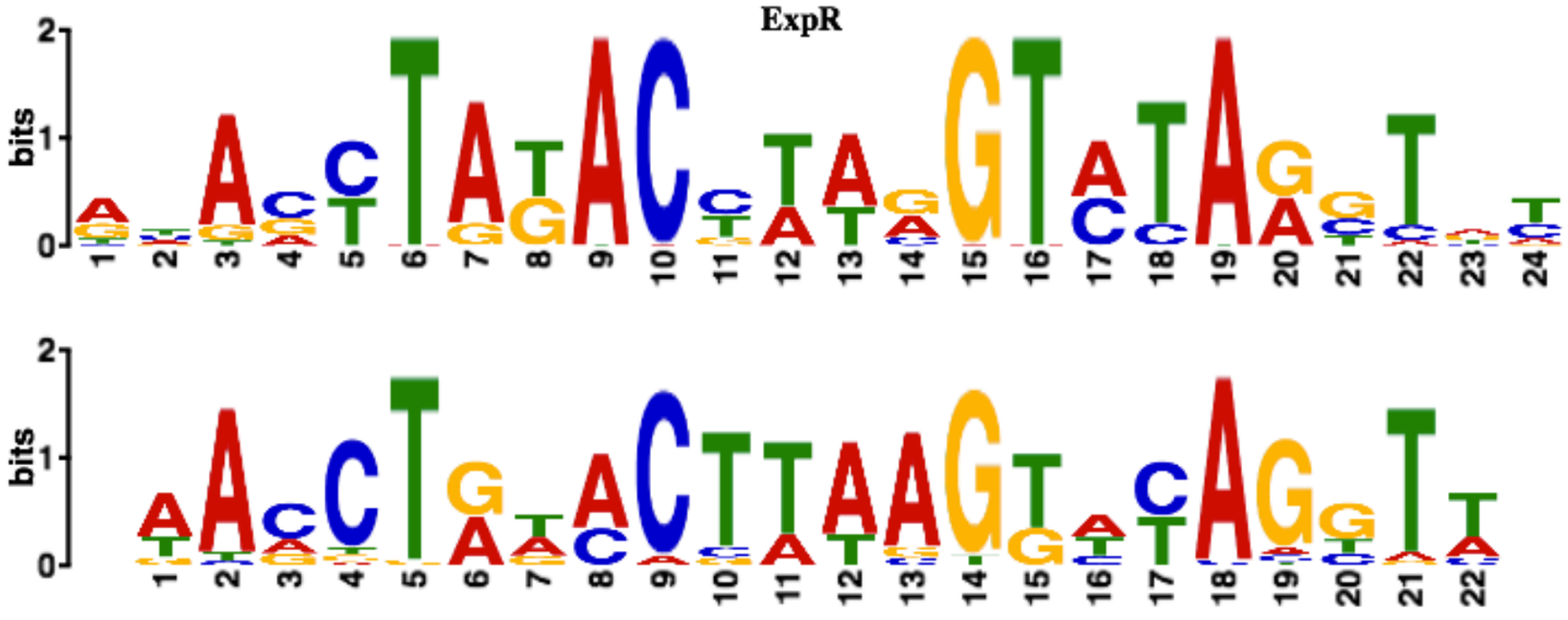
Name	EanR_Pantoea (LuxR_family)
Database	LuxR anchor
***p*-value**	6.32e-14
***E*-value**	1.13e-10
***q*-value**	5.55e-11
Overlap	18
Offset	-2
Orientation	Normal

Show logo download options



Name	ExpR (LuxR_family)
Database	LuxR anchor
***p*-value**	1.22e-09
***E*-value**	2.19e-06
***q*-value**	7.18e-07
Overlap	22
Offset	1
Orientation	Normal

Show logo download options



Верификация сайта связывания белка VirR *P. atrosepticum*

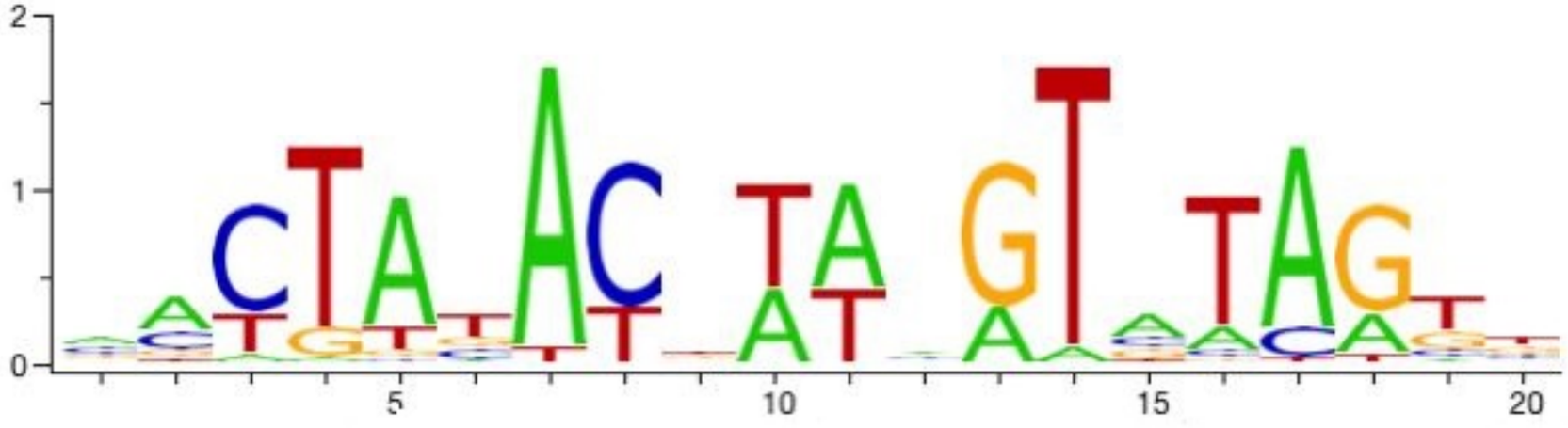
В геноме *P. atrosepticum* на сегодня проанализировано 156 регуляторов из 10 семейств

- Правильно определены 39 из 42 известных мотивов (93%)
- Для неохарактеризованных белков типичный для семейства мотив найден в 74% случаев

Семейство ТФ	Модель PFAM	Число ТФ с известными 3D-структурами	число ТФ в геноме *Pat* 21A		
			всего (известных)	доступны для анализа	идентифицировано корректных мотивов
LacI	PF00356	7	21 (14)	18	17 (94%)
GntR	PF00392	3	22 (9)	18	14 (78%)
LuxR	PF00196	4	12 (4)	12	8 (67%)
TetR	PF00440	7	17 (3)	16	13 (81%)
LysR	PF00126	1	59 (2)	50	32 (64%)
XRE	SM00347	11	19 (1)	15	12 (80%)
HxlR	PF01638	1	8 (1)	8	6 (75%)
MarR	SM00530	5	6 (0)	6	5 (83%)
OmpR	PF00486	5	13 (4)	13	9 (69%)
bEBP	PF02954	4	8 (5)	8	8 (100%)

Практический пример 1: два паралогичных ТФ с одинаковыми КО распознают одинаковый мотив

```
1zg5 NarL_Ecoli                QLTPRERDILKLIAQGLPNKMIARRLDITESTVKVHVKHMLKKMKLKSRVEAAV
1zg1 NarL_Ecoli                QLTPRERDILKLIAQGLPNKMIARRLDITESTVKVHVKHMLKKMKLKSRVEAAV
1je8 NarL_Ecoli                QLTPRERDILKLIAQGLPNKMIARRLDITESTVKVHVKHMLKKMKLKSRVEAAV
4wu4 LiaR_Efaecalis            DLTNREHEILMLIAQGKSNQEIADELFITLKTVKTHVSNILAKLDVDNRTQAAI
4wuH LiaR_Efaecalis            DLTNREHEILMLIAQGKSNQEIADELFITLKTVKTHVSNILAKLDVDNRTQAAI
4wuL LiaR_Efaecalis            DLTNREHEILMLIAQGKSNQEIADELFITLKTVKTHVSNILAKLDVDNRTQAAI
1H0M TRAR_Atu                  WLDPKEATYLRWIAVGKTMEEIADVEGVKYNSVRVKLREAMKRFDVRSKAHLTA
1ZLK DosR_Mtuberculosis        GLTDQERTLLGLLSEGLTNKQIADRMFLAEKTVKNYVSRLLAKLGMERRTQAAV
YP_049663 VirR_Patrosepticum   IFSQRENEILYWASMGKTYPEIALILDIKISTVKFHIGNVVKKLGVLNAKHAIR
AAX77678  ExpR_Pcarotovorum    IFSQRENEILYWASMGKTYLEVAIILGIKTSTVKFHIGNVVKKLGVLNAKHAIR
```



Практический пример 2: расшифровка регуляции факторов вирулентности у *Pectobacterium atrosepticum*

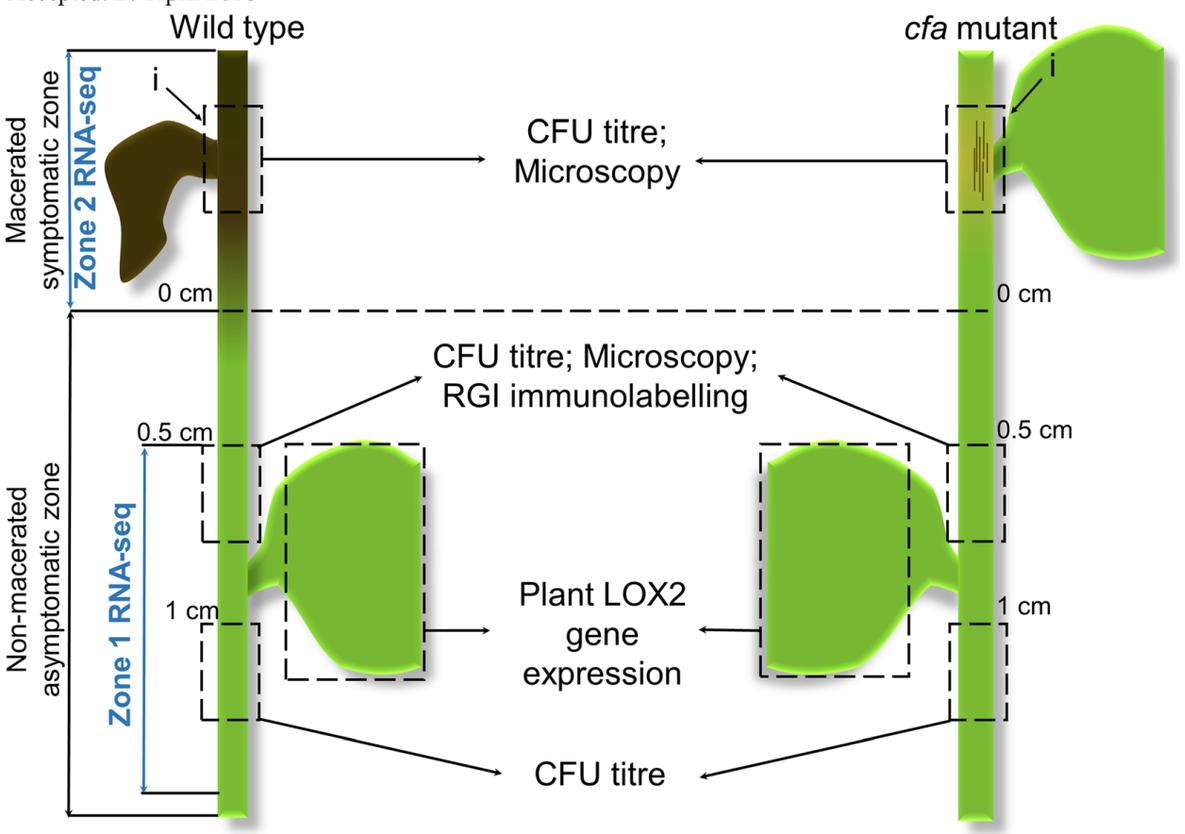
Eur J Plant Pathol
https://doi.org/10.1007/s10658-018-1496-6

SI: PLANT PATHOLOGY FOR INNOVATIVE AGROECOLOGY

Transcriptome profiling helps to identify potential and true molecular switches of stealth to brute force behavior in *Pectobacterium atrosepticum* during systemic colonization of tobacco plants

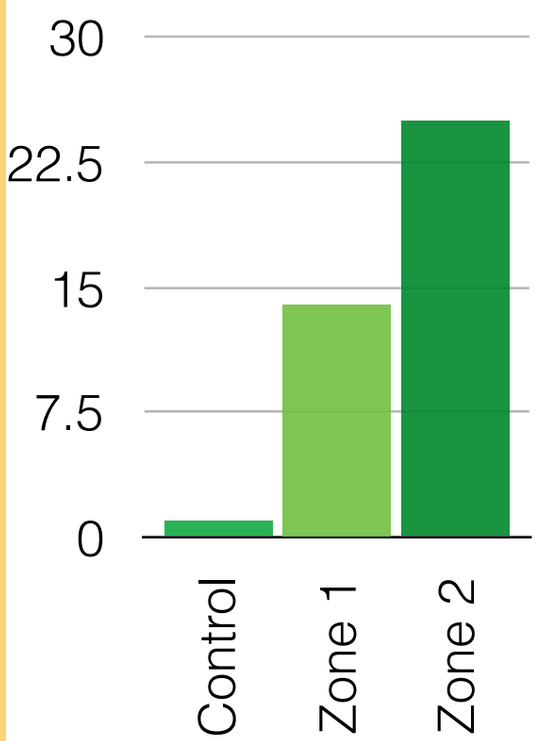
Vladimir Gorshkov · Rim Gubaev · Olga Petrova · Amina Daminova · Natalia Gogoleva · Marina Ageeva · Olga Parfirova · Maxim Prokchorchik · Yevgeny Nikolaichik · Yuri Gogolev

Accepted: 27 April 2018

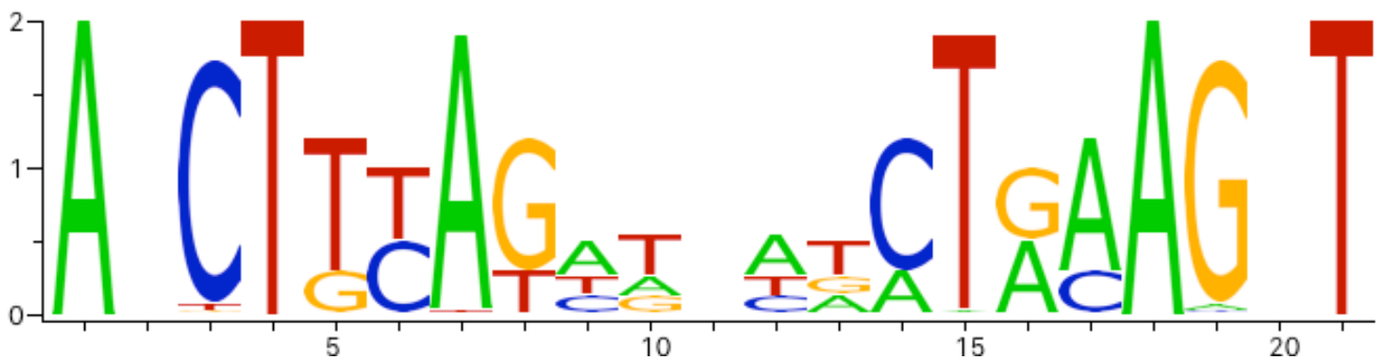


Locus tag	Gene	UniProt annotation	Family	Expression level (Log2FC) Zone 1 vs. Zone 2	Zone 1 vs. in vitro	Zone 2 vs. in vitro
ECA0131		LysR-family transcriptional regulator	LysR	nonDEG	-5.32	-3.67
ECA0336		LysR-family transcriptional regulator	LysR	nonDEG	-1.07	nonDEG
ECA0427		LysR-family transcriptional regulator	LysR	nonDEG	3.81	4.66
ECA0905		LysR-family transcriptional regulator	LysR	nonDEG	1.13	1.67
ECA0916		LysR-family transcriptional regulator	LysR	nonDEG	-2.46	-1.58
ECA1085		LysR-family transcriptional regulator	LysR	nonDEG	-1.80	nonDEG
ECA1482		LysR-family transcriptional regulator	LysR	nonDEG	-1.17	-1.36
ECA1965		LysR-family transcriptional regulator	LysR	nonDEG	1.37	nonDEG
ECA2227		LysR-family transcriptional regulator	LysR	nonDEG	1.35	1.44
ECA2284	cysB	cys regulon transcriptional activator	LysR	nonDEG	-1.85	-1.73
ECA2724	rscR	LysR-family transciptional regulator	LysR	nonDEG	-1.52	-2.04
ECA2973		LysR-family transcriptional regulator	LysR	nonDEG	1.19	nonDEG
ECA3556		LysR-family transcriptional regulator	LysR	nonDEG	-1.08	nonDEG
ECA3879	nhaR	transcriptional activator protein	LysR	nonDEG	-1.13	-1.03
ECA4305	sftR	LysR-family transcriptional regulator	LysR	nonDEG	1.65	2.28
ECA0610		LysR-family transcriptional regulator	LysR	3.69	4.05	nonDEG
ECA4483	nac	nitrogen assimilation regulatory protein	LysR	nonDEG	-1.07	nonDEG
ECA2036	pecS	regulatory protein	MarR	nonDEG	1.95	nonDEG
ECA1911		MarR-family transcriptional regulator	MarR	nonDEG	1.36	nonDEG
ECA1954		putative transcriptional regulator	MarR	nonDEG	1.44	1.24
ECA1644		putative DNA-binding protein	MarR	nonDEG	3.23	2.81
ECA2910		putative plasmid replication protein	MarR	nonDEG	3.49	2.88
ECA3511	emrR	negative regulator of multidrug resistan	MarR	nonDEG	1.86	1.90
ECA1194	cueR	copper efflux regulator	MerR	nonDEG	-1.07	-2.08
ECA4253	metJ	repressor of the methionine regulon	MetJ	nonDEG	1.01	nonDEG
ECA2437	rdgB	regulator of pectin lyase production	Mor	nonDEG	nonDEG	2.19
ECA3695		phage regulatory protein protein	Mor	nonDEG	-1.43	-1.82
ECA0089	mtlR	mannitol operon repressor	MtlR	nonDEG	nonDEG	1.38
ECA1203	rcsB	two-component response regulator	NarL	nonDEG	nonDEG	-1.25
ECA1901	narP	nitrate/nitrite response regulator	NarL	nonDEG	-2.49	-2.40
ECA2089	hrpY	two-component response regulator	NarL	nonDEG	-1.30	-1.11
ECA0027	glnG	nitrogen regulation two-component syst	NtrC	nonDEG	1.32	1.73
ECA0785		two component system response regula	OmpR	nonDEG	5.65	7.22

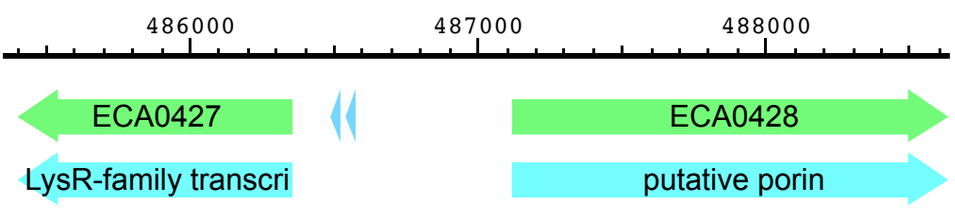
ECA0427 скорее всего контролирует RsmA – глобальный регулятор вторичного метаболизма



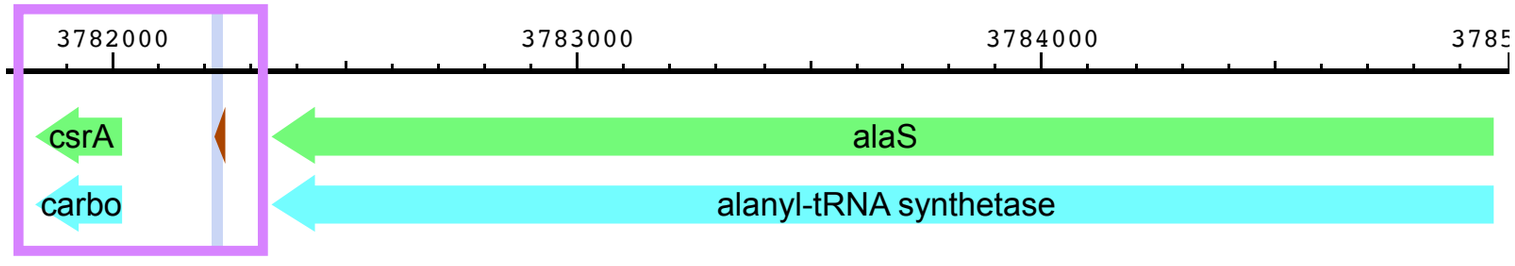
RNA-Seq data for ECA0427



ECA0427 binding site logo

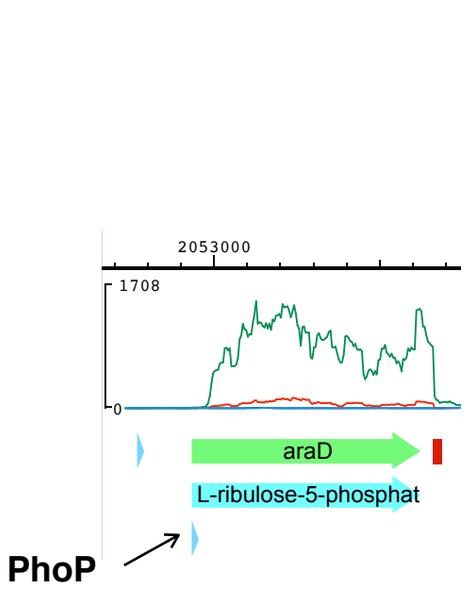


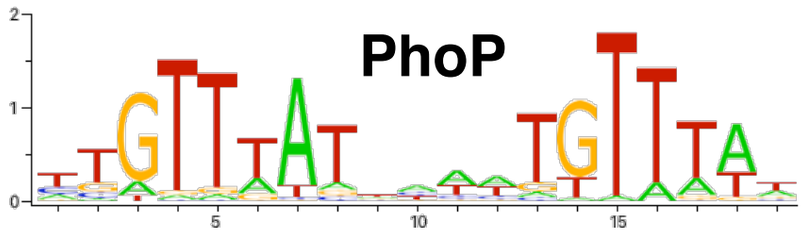
ECA0427 genome locus

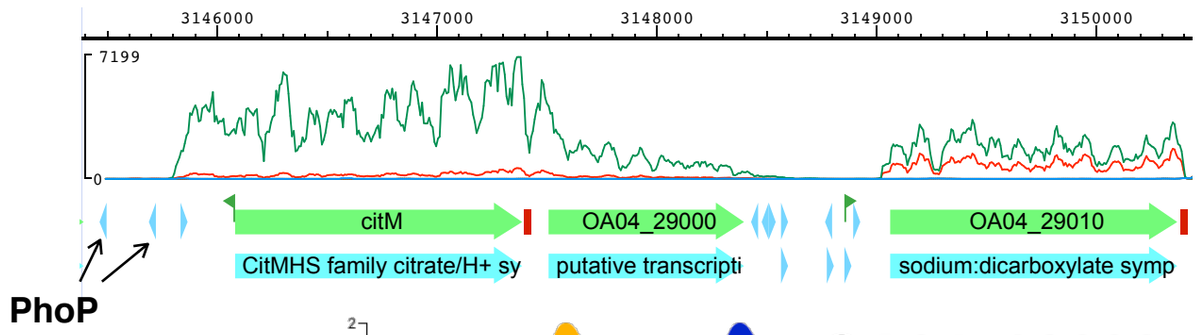


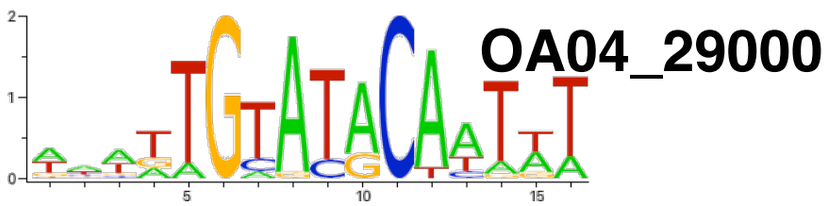
Target locus

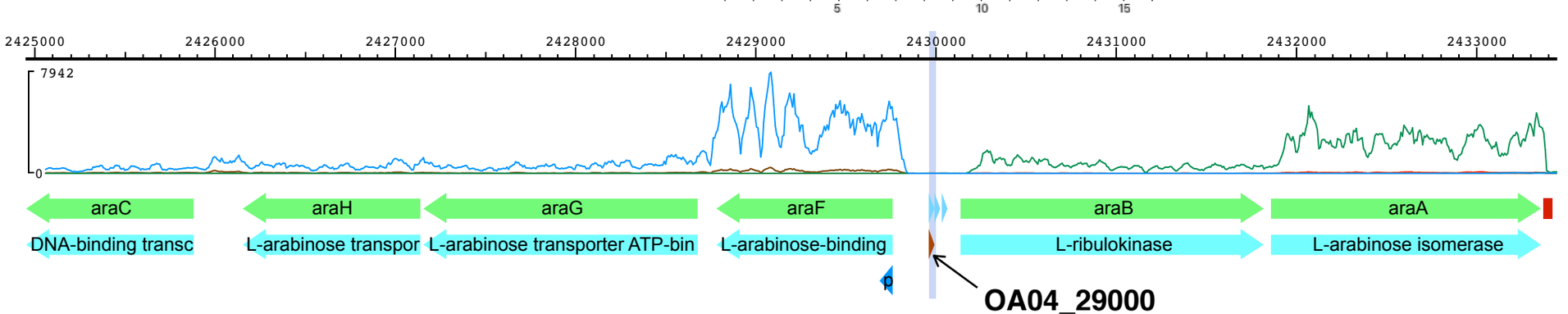
Практический пример 3: анализ данных RNA-seq позволяет различать прямую и непрямую регуляцию

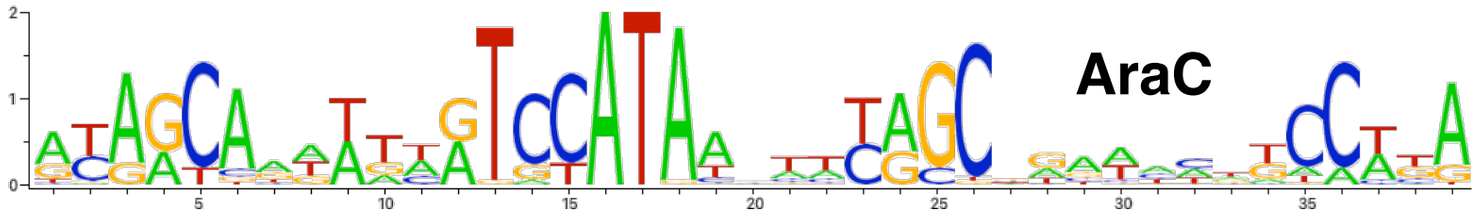


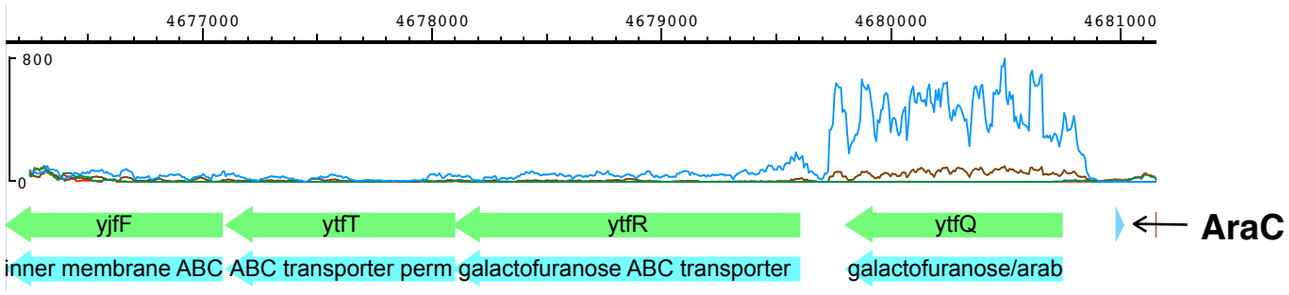












Спасибо за внимание!

The team:

Aliaxandr Damienikan

Pavel Vychik

Yevgeny Nikolaichik

The tools:

XOJO®