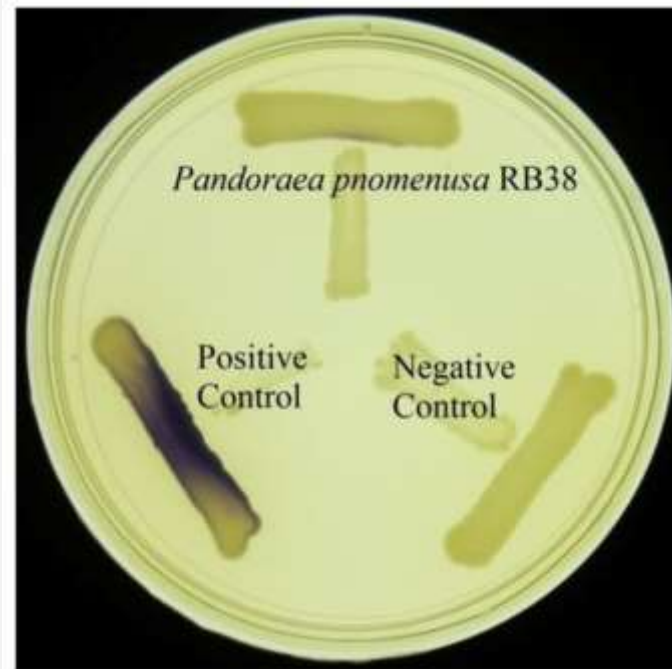
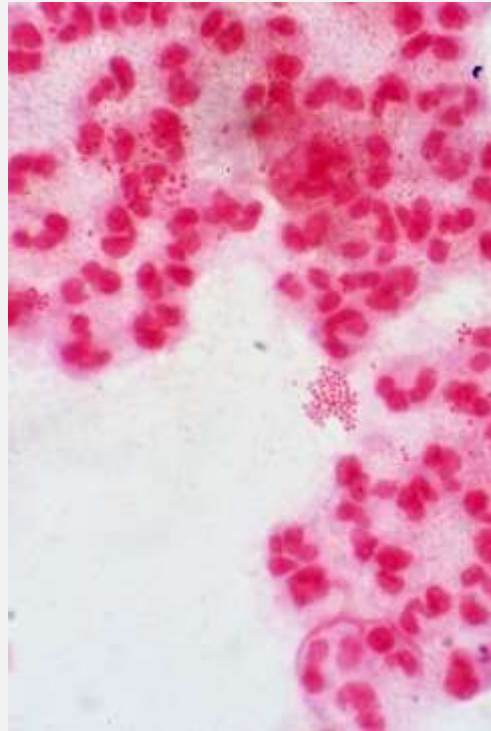
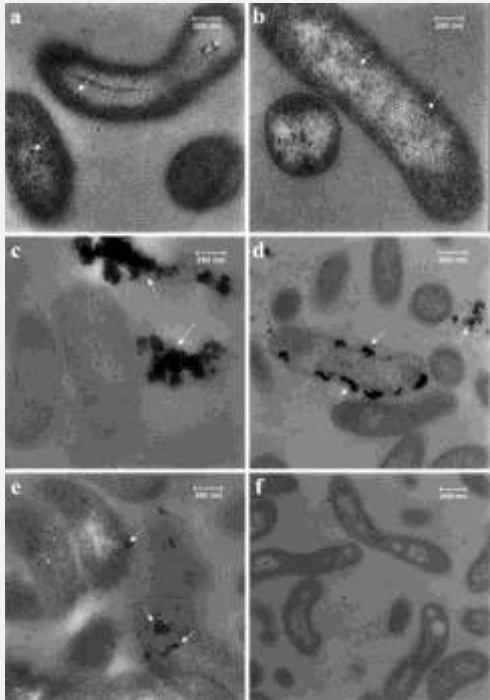


# ПРЕДСКАЗАНИЕ Z-DNA И ЕЁ КОНСЕРВАТИВНОСТИ СРЕДИ ТАКСОНА BETAPROTEOBACTERIA

Подготовлено студентами 3 курса майнора  
«Биоинформатика»

# BETAPROTEOBACTERIA

- Класс грамм-отрицательных бактерий; насчитывает более 75 родов и 400 видов



Слева направо: *Comamonas aquatica*, *Neisseria gonorrhoeae*,  
*Pandoraea pnomenusa*

## Исследуемые роды:

*Neisseria*

*Herbaspirillum*

*Massilia*

*Comamonas*

*Paraburkholderia*

*Rhodoferax*

*Pandoraea*

*Bordetella*

*Acidovorax*

## СПИСОК ГРУППЫ

Имя	Род
Белова Олеся	Neisseria
Княжевский Владимир	Herbaspirillum
Кулешова Полина	Massilia
Овчинникова Анастасия	Comamonas
Савинов Максим	Paraburkholderia
Ершов Кирилл	Rhodoferax
Поморцев Леонид	Pandorea
Попов Игорь	Bordetella
Космачёв Алексей	Acidovorax
Антонников Григорий	Cupriavidus

# ИСХОДНЫЕ ГЕНОМЫ

Род	Средняя длина генома	Среднее кол-во генов	Среднее кол-во участков <b>Z-DNA</b>
Neisseria	3133146	2271	17808
Herbaspirillum	5412755	4871	85574
Massilia	5133376	6101	156244
Commamonas	4832368	4432	367859
Paraburkholderi a	8173433	7384	110335
Rhodoferax	4714642	4490	71594
Pandoraea	9632013	4902	120566
Bordetella	5132278	4697	111961
Acidovorax	4098593	4463	120033
Cupriavidus	6150640	6378	90420

# ИССЛЕДОВАННЫЕ БЕЛКИ

Genus	ZH_score mean	ZH_score max	ZH count	Function
45	31849,91	302785,5	49	30S ribosomal protein S21
50	4567,26	48804,94	48	30S ribosomal protein S11
49	3762,23	10894,72	48	glutamate-1-semialdehyde 2,1-aminomutase
50	16659,39	94590,41	49	glutamine hydrolyzing GMP synthase
50	23703,9	138924,1	46	transcription antitermination factor NusB
50	5162,9	28780,5	46	acetyl-CoA carboxylase biotin carboxylase subunit
50	11458,56	138924,1	46	cell division protein FtsL

# ТЕПЛОВАЯ КАРТА КЛАСТЕРОВ



## НАИБОЛЕЕ ИНТЕРЕСНЫЕ КЛАСТЕРЫ

Genus	ZH mean	ZH max	ZH count	Functions
45	31849,91	302785,5	49	30S ribosomal protein S2I
50	16659,39	94590,41	49	glutamine-hydrolyzing GMP synthase
50	23703,9	138924,1	46	transcription antitermination factor NusB

- 30S ribosomal protein S2I – структурный компонент малой субъединицы рибосомы; связывается с rRNA
- glutamine-hydrolyzing GMP synthase – участвует в синтезе GMP из IMP (синтез нуклеотидов), аминокислотном метаболизме
- transcription antitermination factor NusB – фактор транскрипции; связывается с rRNA, позволяет при транскрипции игнорировать определённые стоп-кодоны



# НАИБОЛЕЕ ИНТЕРЕСНЫЕ КЛАСТЕРЫ - 30S RIBOSOMAL PROTEIN S21

1. Z-DNA_for_protein_WP_003680926.1	CTGCGCGC
2. Z-DNA_for_protein_WP_003680926.1(2)	CACGCAG
3. Z-DNA_for_protein_WP_040658767.1	CGCCCA
4. Z-DNA_for_protein_WP_040658767.1(2)	CTGCGCGC
5. Z-DNA_for_protein_WP_003680926.1(3)	CACGCAG
6. Z-DNA_for_protein_WP_085415572.1	CGCCCA
7. Z-DNA_for_protein_WP_005665410.1	AGTTGCGGC
8. Z-DNA_for_protein_WP_006462949.1	ACTGCGCGC
9. Z-DNA_for_protein_WP_006462949.1(2)	CGCCCA
10. Z-DNA_for_protein_WP_006462949.1(3)	ACTGCGCGC
11. Z-DNA_for_protein_WP_006462949.1(4)	ACTGCGCGC
12. Z-DNA_for_protein_WP_005665410.1(2)	ACTGCGCGC
13. Z-DNA_for_protein_WP_005665410.1(3)	ACTGCGCGC
14. Z-DNA_for_protein_WP_005665410.1(4)	CGCCCA
15. Z-DNA_for_protein_WP_005665410.1(5)	CACGCAG
16. Z-DNA_for_protein_WP_005665410.1(6)	CGCCCA
17. Z-DNA_for_protein_WP_003052999.1	CGCCCA
18. Z-DNA_for_protein_WP_003052999.1(2)	CGCGCAGGT
19. Z-DNA_for_protein_WP_007833691.1	CGCCCA
20. Z-DNA_for_protein_WP_007833691.1(2)	CGCCCA
21. Z-DNA_for_protein_WP_087279623.1	CTGCGCGC
22. Z-DNA_for_protein_WP_013434742.1	CGCTCACGCGCT
23. Z-DNA_for_protein_WP_013434742.1(2)	CGCCCA
24. Z-DNA_for_protein_WP_158952615.1(2)	ACGCGCGGCG
25. Z-DNA_for_protein_WP_158952615.1(3)	TCCACGCGCCAGTTGCGCGCGCAA
26. Z-DNA_for_protein_WP_158952615.1(4)	CGCGCCCCCGCGA
27. Z-DNA_for_protein_WP_028196180.1	GACGCGGCGCGGC
28. Z-DNA_for_protein_WP_028196180.1(2)	TCGCGCGCCG
29. Z-DNA_for_protein_WP_006050883.1	CGTGCACGTTCCGA
30. Z-DNA_for_protein_WP_006050883.1(2)	TCGCGCATDA
31. Z-DNA_for_protein_WP_115778899.1	CGCGA
32. Z-DNA_for_protein_WP_029586606.1	CACGCAA
33. Z-DNA_for_protein_WP_011465162.1	CGCACGTCTGCA
34. Z-DNA_for_protein_WP_011465162.1(2)	TTGCGCGC
35. Z-DNA_for_protein_WP_068628364.1	CGCGCAGGT
36. Z-DNA_for_protein_WP_029709788.1	CTTGCAGCGC
37. Z-DNA_for_protein_WP_077561342.1	AACTGCGCGC
38. Z-DNA_for_protein_WP_006218592.1	GCGTTC
39. Z-DNA_for_protein_WP_006218592.1(4)	CGCGACAA

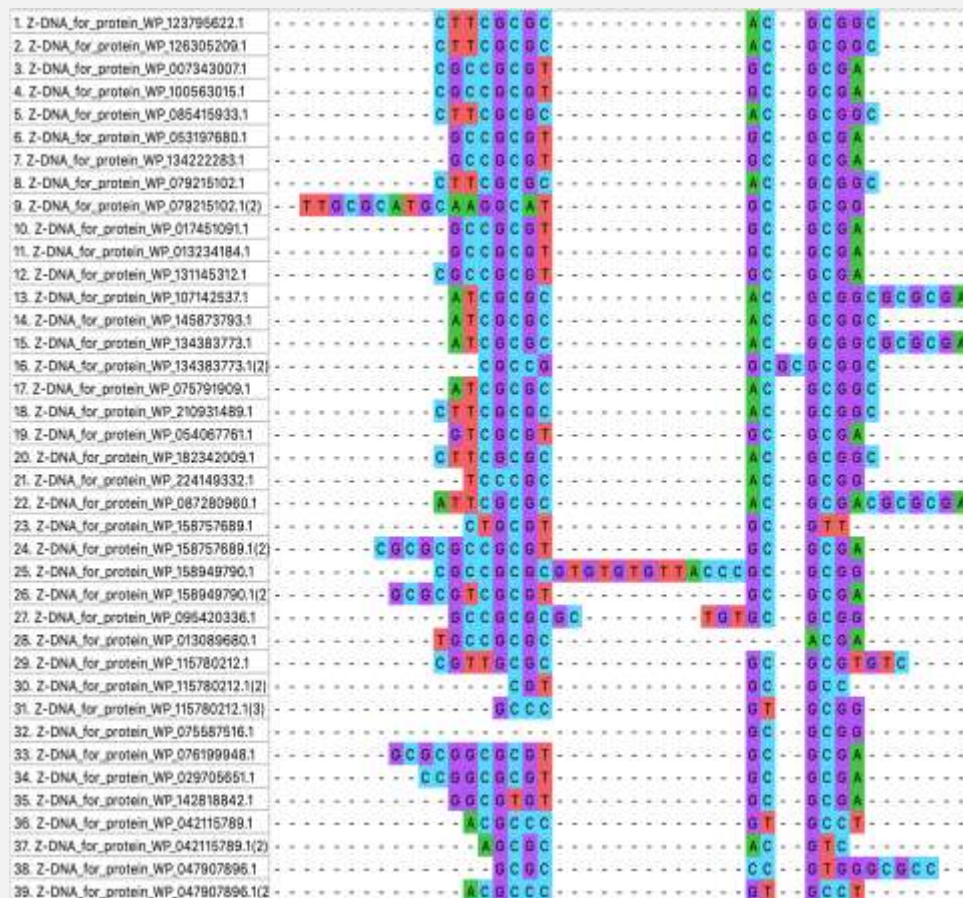
Нуклеотидное выравнивание Z-DNA

1. WP_003052999.1_MULTISPECIES_30S_ribosomal_protein_S21_Commonadaceae	MTTIVKKNKFFVVALRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
2. WP_003052999.1_MULTISPECIES_30S_ribosomal_protein_S21_Commonadaceae(2)	MTTIVKKNKFFVVALRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
3. WP_003680926.1_MULTISPECIES_30S_ribosomal_protein_S21_Betaproteobacteria	MPAIRVKNKFFEVAMRRFKRAVEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
4. WP_003680926.1_MULTISPECIES_30S_ribosomal_protein_S21_Neisseriaceae	MPAIRVKNKFFEVAMRRFKRAVEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
5. WP_003680926.1_MULTISPECIES_30S_ribosomal_protein_S21_Neisseriaceae(2)	MPAIRVKNKFFEVAMRRFKRAVEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
6. WP_005665410.1_MULTISPECIES_30S_ribosomal_protein_S21_Burkholderiales	MTTIVKKNKFFEVAMRRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
7. WP_005665410.1_MULTISPECIES_30S_ribosomal_protein_S21_Burkholderiales(2)	MTTIVKKNKFFEVAMRRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
8. WP_005665410.1_MULTISPECIES_30S_ribosomal_protein_S21_Burkholderiales(3)	MTTIVKKNKFFEVAMRRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
9. WP_005665410.1_MULTISPECIES_30S_ribosomal_protein_S21_Burkholderiales(4)	MTTIVKKNKFFEVAMRRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
10. WP_005665410.1_MULTISPECIES_30S_ribosomal_protein_S21_Burkholderiales(5)	MTTIVKKNKFFEVAMRRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
11. WP_005665410.1_MULTISPECIES_30S_ribosomal_protein_S21_Burkholderiales(6)	MTTIVKKNKFFEVAMRRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
12. WP_005799770.1_MULTISPECIES_30S_ribosomal_protein_S21_Commonadaceae	MTTIVKKNKFFDVALLRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
13. WP_006050883.1_MULTISPECIES_30S_ribosomal_protein_S21_Burkholderiaceae	MTTIVKKNKFFDVALLRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
14. WP_006050883.1_MULTISPECIES_30S_ribosomal_protein_S21_Burkholderiaceae(2)	MTTIVKKNKFFDVALLRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
15. WP_006050883.1_MULTISPECIES_30S_ribosomal_protein_S21_Burkholderiaceae(3)	MTTIVKKNKFFDVALLRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
16. WP_006050883.1_MULTISPECIES_30S_ribosomal_protein_S21_Burkholderiaceae(4)	MTTIVKKNKFFDVALLRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
17. WP_006218592.1_MULTISPECIES_30S_ribosomal_protein_S21_Alcaligenaceae	MPFIVLKNKFFFAALLRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
18. WP_006218592.1_MULTISPECIES_30S_ribosomal_protein_S21_Alcaligenaceae(2)	MPFIVLKNKFFFAALLRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
19. WP_006218592.1_MULTISPECIES_30S_ribosomal_protein_S21_Alcaligenaceae(3)	MPFIVLKNKFFFAALLRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
20. WP_006218592.1_MULTISPECIES_30S_ribosomal_protein_S21_Alcaligenaceae(4)	MPFIVLKNKFFFAALLRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
21. WP_006462949.1_MULTISPECIES_30S_ribosomal_protein_S21_Herbaspirillum	MTTIVKKNKFFEVAMRRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
22. WP_006462949.1_MULTISPECIES_30S_ribosomal_protein_S21_Herbaspirillum(2)	MTTIVKKNKFFEVAMRRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
23. WP_006462949.1_MULTISPECIES_30S_ribosomal_protein_S21_Herbaspirillum(3)	MTTIVKKNKFFEVAMRRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
24. WP_007833691.1_MULTISPECIES_30S_ribosomal_protein_S21_Betaproteobacteria	MTTIVKKNKFFDVALLRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
25. WP_007833691.1_MULTISPECIES_30S_ribosomal_protein_S21_Proteobacteria	MTTIVKKNKFFDVALLRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
26. WP_007833691.1_MULTISPECIES_30S_ribosomal_protein_S21_Proteobacteria(2)	MTTIVKKNKFFDVALLRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
27. WP_007833691.1_MULTISPECIES_30S_ribosomal_protein_S21_Proteobacteria(3)	MTTIVKKNKFFDVALLRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
28. WP_007833691.1_MULTISPECIES_30S_ribosomal_protein_S21_Proteobacteria(4)	MTTIVKKNKFFDVALLRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
29. WP_008919901.1_MULTISPECIES_30S_ribosomal_protein_S21_Burkholderiaceae	MTTIVKKNKFFEVAMRRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
30. WP_011465162.1_30S_ribosomal_protein_S21_Rhodospirillumrubrum	MTTIVKKNKFFDVALLRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
31. WP_011903537.1_MULTISPECIES_30S_ribosomal_protein_S21_Polynucleobacter	MTTIVLKNKFFEVALLRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
32. WP_011903537.1_MULTISPECIES_30S_ribosomal_protein_S21_Polynucleobacter(2)	MTTIVLKNKFFEVALLRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
33. WP_011903537.1_MULTISPECIES_30S_ribosomal_protein_S21_Polynucleobacter(3)	MTTIVLKNKFFEVALLRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
34. WP_011903537.1_MULTISPECIES_30S_ribosomal_protein_S21_Polynucleobacter(4)	MTTIVLKNKFFEVALLRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
35. WP_012417952.1_MULTISPECIES_30S_ribosomal_protein_S21_Alcaligenaceae	MPFIVLKNKFFFAALLRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
36. WP_013434742.1_MULTISPECIES_30S_ribosomal_protein_S21_Burkholderiaceae	MTTIVKKNKFFEVAMRRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
37. WP_028196180.1_MULTISPECIES_30S_ribosomal_protein_S21_Burkholderiaceae	MTTIVKKNKFFEVAMRRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
38. WP_029709788.1_30S_ribosomal_protein_S21_Rhodospirillumrubrum	MTTIVKKNKFFDVALLRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
39. WP_040658767.1_30S_ribosomal_protein_S21_Neisseria_bacilliformis	MPAIRVKNKFFEVAMRRFKRAVEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
40. WP_08628364.1_MULTISPECIES_30S_ribosomal_protein_S21_Commonadaceae	MTTIVKKNKFFDVALLRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
41. WP_079586606.1_MULTISPECIES_30S_ribosomal_protein_S21_Rhodospirillum	MTTIVKKNKFFDVALLRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
42. WP_077561342.1_30S_ribosomal_protein_S21_Rhodospirillumrubrum	MTTIVKKNKFFDVALLRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
43. WP_085415572.1_30S_ribosomal_protein_S21_Neisseria_cavia	MPAIRVKNKFFEVAMRRFKRAVEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
44. WP_087279623.1_30S_ribosomal_protein_S21_Commonadaceae	MTTIVKKNKFFDVALLRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
45. WP_114659583.1_30S_ribosomal_protein_S21_Polynucleobacter_necessarius	MTTIVLKNKFFEVALLRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
46. WP_115778899.1_MULTISPECIES_30S_ribosomal_protein_S21_Paraburkholderia	MTTIVKKNKFFEVAMRRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
47. WP_158952615.1_30S_ribosomal_protein_S21_Paraburkholderia_adiposa	MTTIVKKNKFFEVAMRRFKRTIEKTLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY
48. WP_159817519.1_MULTISPECIES_30S_ribosomal_protein_S21_Proteobacteria	MTTIVKKNKFFDVALLRRFKRTIEKLLLTDLARAEFYKPTAEKRRKAAAVRHHYKRVSMQLPKKLY

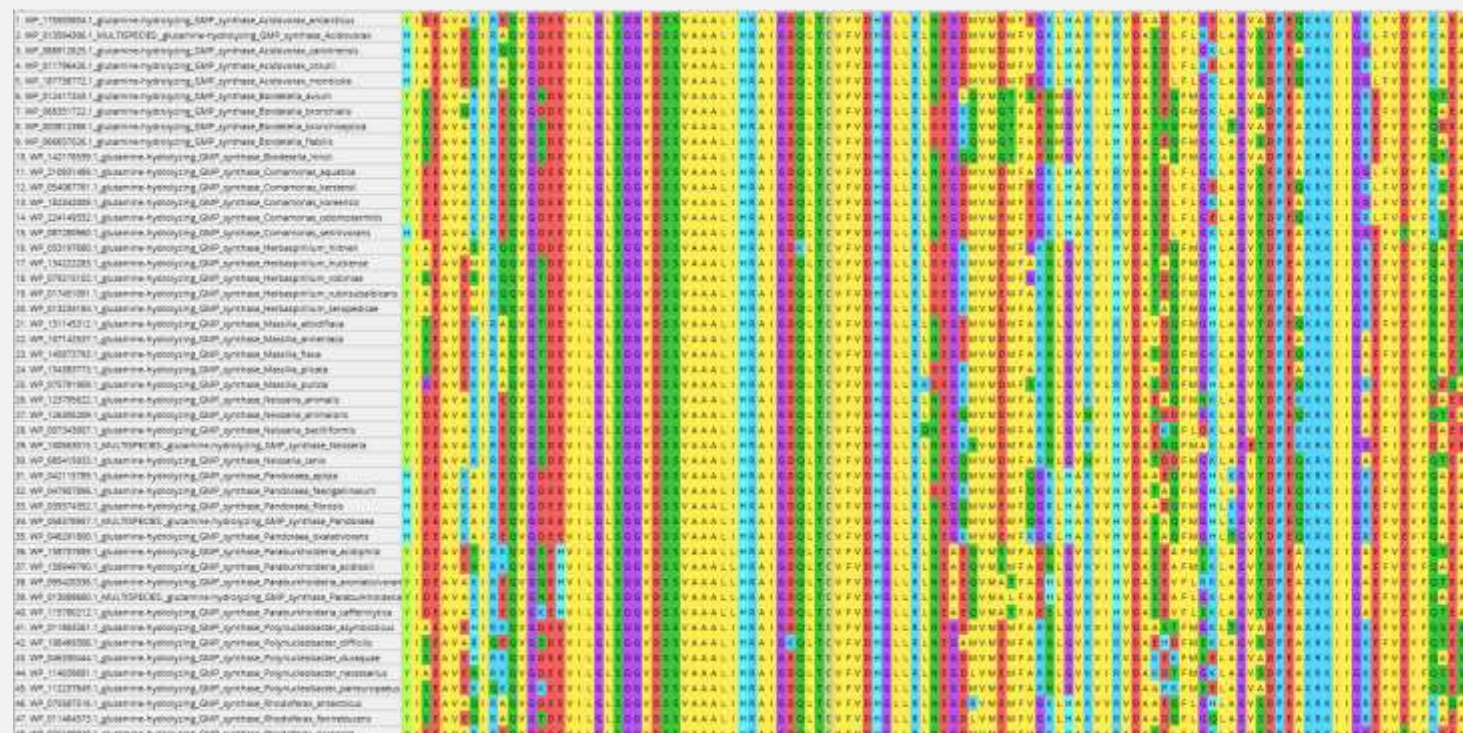
Аминокислотное выравнивание белков



# НАИБОЛЕЕ ИНТЕРЕСНЫЕ КЛАСТЕРЫ – GLUTAMINE-HYDROLYZING GMP SYNTHASE



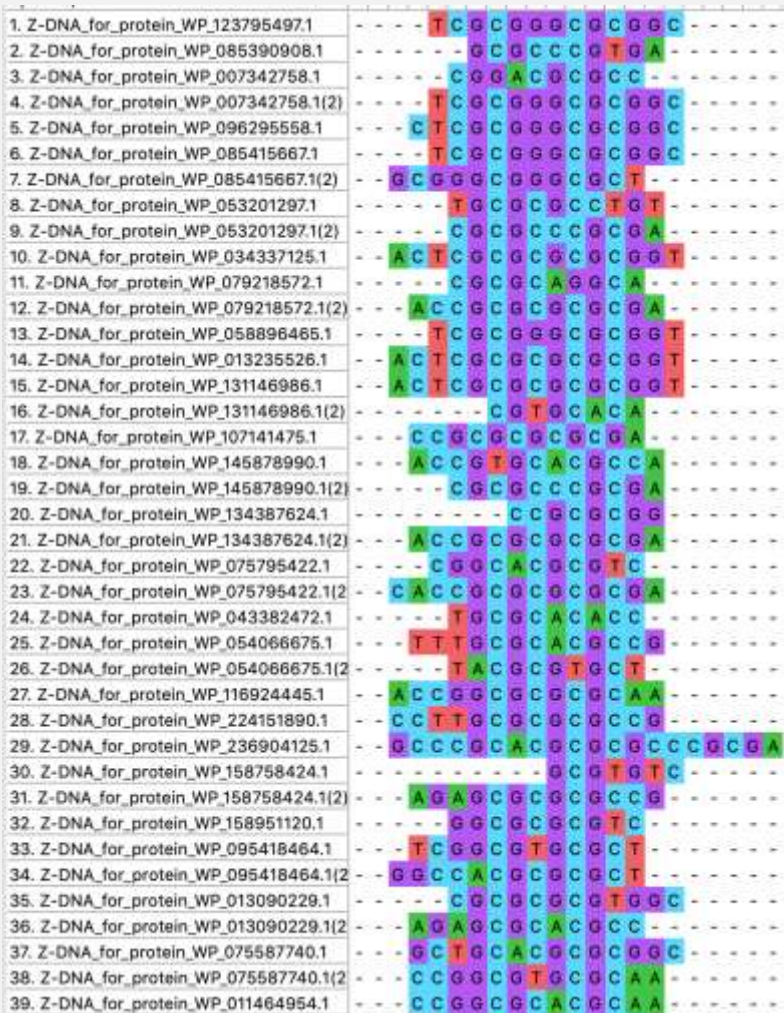
Нуклеотидное выравнивание Z-DNA



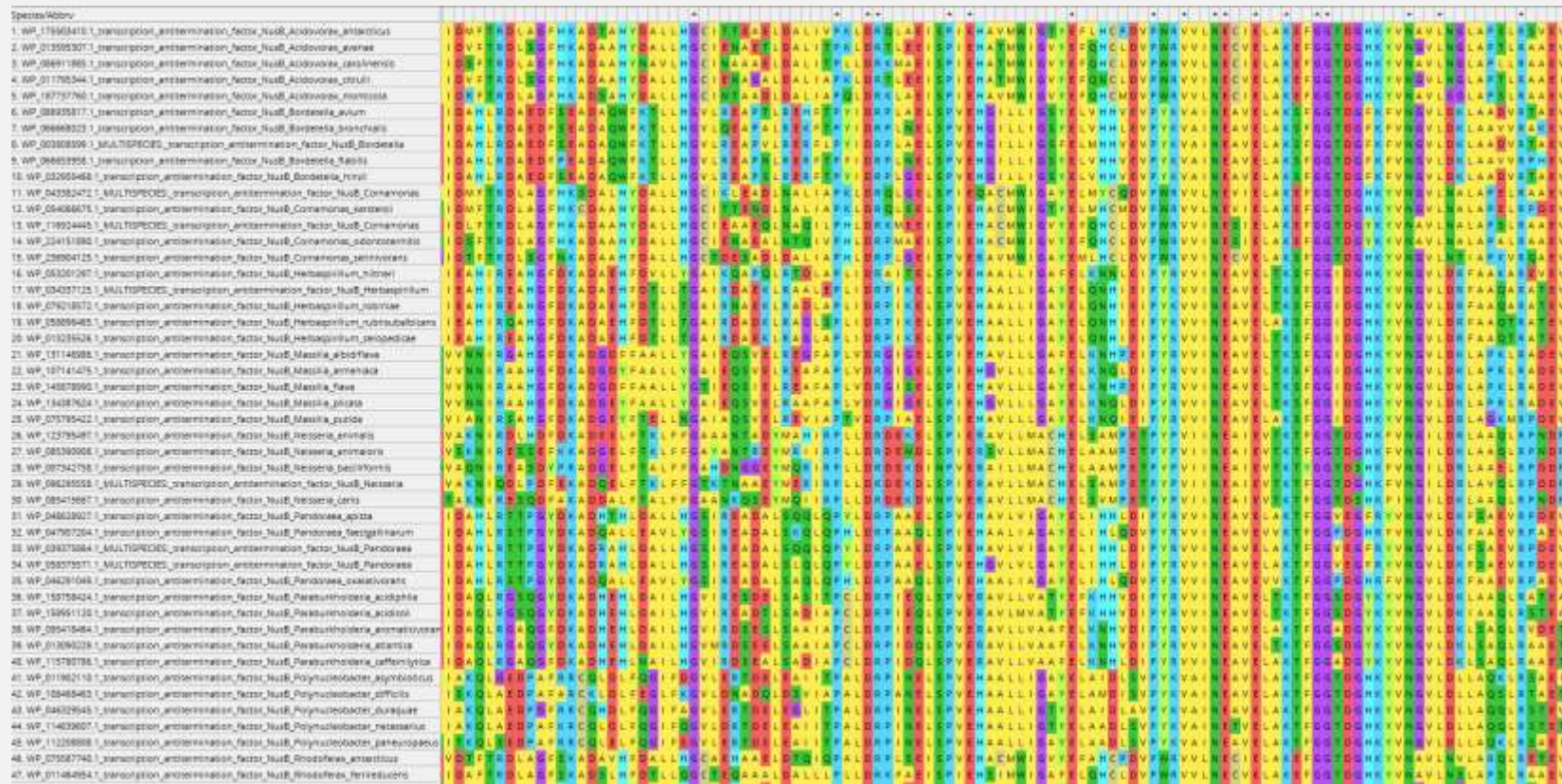
Аминокислотное выравнивание белков



НАИБОЛЕЕ ИНТЕРЕСНЫЕ КЛАСТЕРЫ –  
 TRANSCRIPTION ANTITERMINATION  
 FACTOR NUSB



# Нуклеотидное выравнивание Z-DNA



## Аминокислотное выравнивание белков

# **СПАСИБО ЗА ВНИМАНИЕ!**

Мы готовы ответить на ваши вопросы