

Η χαμηλή τιμή του QPS είναι πλάσματική, καθώς ο μετρούμενος χρόνος συμπεριλαμβάνει το υψηλό υπολογιστικό κόστος της κατασκευής του ευρετηρίου (Index Construction & Training), διαδικασία που σε πραγματικές συνθήκες θα εκτελούταν offline μόνο μία φορά, αφήνοντας τον καθαρό χρόνο αναζήτησης σε επίπεδα millisecond.

```
python3 protein_search.py \  
-d protein_vectors.fvecs \  
-q targets.fasta \  
-s swissprot_50k.fasta \  
-o benchmarks/res_ivfpq.txt \  
-method ivfpq \  
-N 50 \  
--ivf_k 50 \  
--ivf_probe 5
```

Query Protein: A0A009I3Y5

N = 50 (μέγεθος λίστας Top-N για την αξιολόγηση Recall@N)

[1] Συνοπτική σύγκριση μεθόδων

Method	Time/query(s)	QPS	Recall@N vs BLAST
IVF-PQ	6.225	0	0.00
BLAST (Ref)	1.224	0	1.00 (ref)

[2] Top-N γείτονες ανά μέθοδο (εδώ N = 10 για εκτύπωση)

Method: IVF-PQ

Rank	Neighbor ID	L2 Dist	BLAST Identity	In BLAST Top-N?	Bio comment
------	-------------	---------	----------------	-----------------	-------------

1	sp P44033 Y665_HAEIN	0.36	1%	No	Possible false positive
2	sp A3DA85 SELU_SHEB5	0.41	3%	No	Possible false positive
3	sp Q74Q56 SGRR_YERPE	0.41	1%	No	Possible false positive
4	sp Q4QLU9 RLMC_HAEI8	0.41	0%	No	Possible false positive
5	sp A7ZZ64 THIK_ECOHS	0.44	1%	No	Possible false positive
6	sp A7FGL1 MNMC_YERP3	0.44	0%	No	Possible false positive
7	sp C0MEP5 PEPX_STRS7	0.44	1%	No	Possible false positive
8	sp Q1CGA8 RLMC_YERP_N	0.44	0%	No	Possible false positive
9	sp B8ECC0 TPMT_SHEB2	0.44	2%	No	Possible false positive
10	sp Q12R91 TRMN6_SHEDO	0.44	1%	No	Possible false positive

Query Protein: A0A009PCK4

N = 50 (μέγεθος λίστας Top-N για την αξιολόγηση Recall@N)

[1] Συνοπτική σύγκριση μεθόδων

Method	Time/query(s)	QPS	Recall@N vs BLAST
IVF-PQ	6.225	0	0.00
BLAST (Ref)	0.305	3	1.00 (ref)

[2] Top-N γείτονες ανά μέθοδο (εδώ N = 10 για εκτύπωση)

Method: IVF-PQ

Rank	Neighbor ID	L2 Dist	BLAST Identity	In BLAST Top-N?	Bio comment
1	sp P44033 Y665_HAEIN	0.36	1%	No	Possible false positive

2	sp A3DA85 SELU_SHEB5	0.41	3%	No	Possible false positive
3	sp Q74Q56 SGRR_YERPE	0.41	1%	No	Possible false positive
4	sp Q4QLU9 RLMC_HAEI8	0.41	0%	No	Possible false positive
5	sp A7ZZ64 THIK_ECOHS	0.44	1%	No	Possible false positive
6	sp A7FGL1 MNMC_YERP3	0.44	0%	No	Possible false positive
7	sp C0MEP5 PEPX_STRS7	0.44	1%	No	Possible false positive
8	sp B8ECC0 TPMT_SHEB2	0.44	2%	No	Possible false positive
9	sp Q1CGA8 RLMC_YERP3	0.44	0%	No	Possible false positive
10	sp B5F8E4 YCFP_SALA4	0.44	14%	No	Possible false positive

Query Protein: A0A001

N = 50 (μέγεθος λίστας Top-N για την αξιολόγηση Recall@N)

[1] Συνοπτική σύγκριση μεθόδων

Method	Time/query(s)	QPS	Recall@N vs BLAST
IVF-PQ	6.225	0	0.08
BLAST (Ref)	0.602	1	1.00 (ref)

[2] Top-N γείτονες ανά μέθοδο (εδώ N = 10 για εκτύπωση)

Method: IVF-PQ

Rank	Neighbor ID	L2 Dist	BLAST Identity	In BLAST Top-N?	Bio comment
1	sp Q51719 YCOBA_PROFR	0.34	1%	No	Possible false positive
2	sp Q20Z38 NDVA_RHOPB	0.40	1%	Yes	-

3	sp Q21BF6 MODC_RHOPB	0.40	1%	No	Possible false positive
4	sp Q8P8W4 MSBA_XANCP	0.42	1%	Yes	-
5	sp P33973 PKN1_MYXXA	0.45	1%	No	-
6	sp Q9HNV0 COFC_HALSA	0.46	1%	No	-
7	sp P9WMX0 MFTF_MYCTO	0.46	2%	No	-
8	sp P9WIL0 PANE_MYCTO	0.46	1%	No	-
9	sp P0DMK6 IRLS_BURPS	0.47	1%	No	-
10	sp Q67LC0 FTSH1_SYMTH	0.47	2%	No	-

Query Protein: A0A002

N = 50 (μέγεθος λίστας Top-N για την αξιολόγηση Recall@N)

[1] Συνοπτική σύγκριση μεθόδων

Method	Time/query(s)	QPS	Recall@N vs BLAST
--------	---------------	-----	-------------------

IVF-PQ	6.225	0	0.08
--------	-------	---	------

BLAST (Ref)	0.549	1	1.00 (ref)
-------------	-------	---	------------

[2] Top-N γείτονες ανά μέθοδο (εδώ N = 10 για εκτύπωση)

Method: IVF-PQ

Rank	Neighbor ID	L2 Dist	BLAST Identity	In BLAST Top-N?	Bio comment
------	-------------	---------	----------------	-----------------	-------------

1	sp Q51719 YCOBA_PROFR	0.36	1%	No	Possible false positive
2	sp Q20Z38 NDVA_RHOPB	0.41	2%	Yes	-
3	sp Q21BF6 MODC_RHOPB	0.42	1%	No	Possible false positive

4	sp Q8P8W4 MSBA_XANCP	0.45	1%	Yes	-
5	sp P0DMK6 IRLS_BURPS	0.47	1%	No	-
6	sp Q67JX3 ECFA1_SYMTH	0.48	1%	No	-
7	sp Q5P6D5 MACB_AROAE	0.49	1%	No	-
8	sp Q3KJQ7 TAUB_PSEPF	0.50	1%	No	-
9	sp Q2SB47 HMUV_HAHCH	0.50	1%	No	-
10	sp D1BZ82 MSHA_XYLCX	0.50	2%	No	-

Query Protein: A0A009HL96

N = 50 (μέγεθος λίστας Top-N για την αξιολόγηση Recall@N)

[1] Συνοπτική σύγκριση μεθόδων

Method	Time/query(s)	QPS	Recall@N vs BLAST
--------	---------------	-----	-------------------

IVF-PQ	6.225	0	0.06
--------	-------	---	------

BLAST (Ref)	0.289	3	1.00 (ref)
-------------	-------	---	------------

[2] Top-N γείτονες ανά μέθοδο (εδώ N = 10 για εκτύπωση)

Method: IVF-PQ

Rank	Neighbor ID	L2 Dist	BLAST Identity	In BLAST Top-N?	Bio comment
------	-------------	---------	----------------	-----------------	-------------

1	sp Q5HI09 GRAR_STAAC	0.18	1%	Yes	-
2	sp A5VUU3 DIVK_BRUO2	0.42	25%	Yes	Twilight zone
3	sp O07586 CUER_BACSU	0.48	14%	No	-
4	sp B7LLW0 NSRR_ESCF3	0.51	15%	No	-

5	sp P44617 Y293_HAEIN	0.51	9%	No	-
6	sp O34994 CCPN_BACSU	0.52	1%	No	-
7	sp Q6GK52 LYTS_STAAR	0.52	1%	No	-
8	sp P62182 RP28_BACAN	0.52	2%	No	-
9	sp Q6G6T5 SARZ_STAAS	0.52	20%	No	Twilight zone
10	sp Q71YJ7 RSGA1_LISMF	0.53	1%	No	-

Query Protein: A0A009HLV9

N = 50 (μέγεθος λίστας Top-N για την αξιολόγηση Recall@N)

[1] Συνοπτική σύγκριση μεθόδων

Method	Time/query(s)	QPS	Recall@N vs BLAST
--------	---------------	-----	-------------------

IVF-PQ	6.225	0	0.08
--------	-------	---	------

BLAST (Ref)	0.404	2	1.00 (ref)
-------------	-------	---	------------

[2] Top-N γείτονες ανά μέθοδο (εδώ N = 10 για εκτύπωση)

Method: IVF-PQ

Rank	Neighbor ID	L2 Dist	BLAST Identity	In BLAST Top-N?	Bio comment
------	-------------	---------	----------------	-----------------	-------------

1	sp B5F3J7 RUVB_SALA4	0.23	0%	No	Possible false positive
2	sp P0A812 RUVB_ECOLI	0.24	0%	No	Possible false positive
3	sp Q3K999 SFNR_PSEPF	0.25	2%	Yes	-
4	sp B4SDZ7 RUVB_PELPB	0.26	1%	No	Possible false positive
5	sp Q15RN6 RUVB_PSEA6	0.26	0%	No	Possible false positive

6	sp P0DF50 RUVB_STRP3	0.26	1%	No	Possible false positive
7	sp A8A3I6 NORR_ECOHS	0.27	1%	Yes	-
8	sp A4STH5 NORR_AERS4	0.28	1%	Yes	-
9	sp Q7WGB3 RUVB_BORBR	0.28	0%	No	Possible false positive
10	sp Q7U9W7 RUVB_PARMW	0.28	1%	No	Possible false positive

Query Protein: A0A009HN45

N = 50 (μέγεθος λίστας Top-N για την αξιολόγηση Recall@N)

[1] Συνοπτική σύγκριση μεθόδων

Method	Time/query(s)	QPS	Recall@N vs BLAST
--------	---------------	-----	-------------------

IVF-PQ	6.225	0	0.03
--------	-------	---	------

BLAST (Ref)	0.606	1	1.00 (ref)
-------------	-------	---	------------

[2] Top-N γείτονες ανά μέθοδο (εδώ N = 10 για εκτύπωση)

Method: IVF-PQ

Rank	Neighbor ID	L2 Dist	BLAST Identity	In BLAST Top-N?	Bio comment
------	-------------	---------	----------------	-----------------	-------------

1	sp A6M2Z0 UVRB_CLOB8	0.17	1%	No	Possible false positive
2	sp A8GSU6 SECA_RICRS	0.19	1%	No	Possible false positive
3	sp B7NA78 UVRB_ECOLU	0.20	0%	No	Possible false positive
4	sp Q5HHR0 UVRB_STAAC	0.20	0%	No	Possible false positive
5	sp Q8NXM0 UVRB_STAAW	0.21	0%	No	Possible false positive
6	sp Q47EI2 UVRB_DECAR	0.21	0%	No	Possible false positive

7	sp C0PWY7 UVRB_SALPC	0.21	0%	No	Possible false positive
8	sp Q6MRS1 DNAA_BDEBA	0.21	0%	No	Possible false positive
9	sp Q6MEX9 SECA_PARUW	0.21	2%	No	Possible false positive
10	sp P24125 V51K_BPL79	0.21	0%	No	Possible false positive

Query Protein: A0A009HQC9

N = 50 (μέγεθος λίστας Top-N για την αξιολόγηση Recall@N)

[1] Συνοπτική σύγκριση μεθόδων

Method	Time/query(s)	QPS	Recall@N vs BLAST
--------	---------------	-----	-------------------

IVF-PQ	6.225	0	0.02
--------	-------	---	------

BLAST (Ref)	0.635	1	1.00 (ref)
-------------	-------	---	------------

[2] Top-N γείτονες ανά μέθοδο (εδώ N = 10 για εκτύπωση)

Method: IVF-PQ

Rank	Neighbor ID	L2 Dist	BLAST Identity	In BLAST Top-N?	Bio comment
------	-------------	---------	----------------	-----------------	-------------

1	sp Q48LP5 RAPA_PSE14	0.13	2%	Yes	-
2	sp B7NA78 UVRB_ECOLU	0.21	0%	No	Possible false positive
3	sp C0PWY7 UVRB_SALPC	0.21	0%	No	Possible false positive
4	sp Q47EI2 UVRB_DECAR	0.22	1%	No	Possible false positive
5	sp B7I8J5 SECA_ACIB5	0.22	1%	No	Possible false positive
6	sp O84707 SECA_CHLTR	0.22	1%	No	Possible false positive
7	sp A1W465 SECA_ACISJ	0.23	1%	No	Possible false positive



8	sp Q6MEX9 SECA_PARUW	0.24	2%	No	Possible false positive
9	sp Q72XS9 SECA_BACC1	0.24	1%	No	Possible false positive
10	sp Q21MG1 SECA_SACD2	0.24	1%	No	Possible false positive

Query Protein: A0A009HPM0

N = 50 (μέγεθος λίστας Top-N για την αξιολόγηση Recall@N)

[1] Συνοπτική σύγκριση μεθόδων

Method	Time/query(s)	QPS	Recall@N vs BLAST
IVF-PQ	6.225	0	0.00
BLAST (Ref)	0.494	2	1.00 (ref)

[2] Top-N γείτονες ανά μέθοδο (εδώ N = 10 για εκτύπωση)

Method: IVF-PQ

Rank	Neighbor ID	L2 Dist	BLAST Identity	In BLAST Top-N?	Bio comment
1	sp Q1CGA8 RLMC_YERPN	0.31	0%	No	Possible false positive
2	sp Q3Z4Q0 SELU_SHISS	0.32	0%	No	Possible false positive
3	sp Q0T790 SELU_SHIF8	0.32	0%	No	Possible false positive
4	sp A3DA85 SELU_SHEB5	0.32	1%	No	Possible false positive
5	sp B0VKL3 RLMD_ACIBS	0.32	1%	No	Possible false positive
6	sp A1AJ76 MUTL_ECOK1	0.33	2%	No	Possible false positive
7	sp Q4QLU9 RLMC_HAEI8	0.33	0%	No	Possible false positive
8	sp A3D8E4 RSMC_SHEB5	0.33	1%	No	Possible false positive

9	sp B7M5S7 GLPB_ECO8A	0.33	1%	No	Possible false positive
10	sp B5EYX0 GLPB_SALA4	0.33	3%	No	Possible false positive

Query Protein: A0A009IB02

N = 50 (μέγεθος λίστας Top-N για την αξιολόγηση Recall@N)

[1] Συνοπτική σύγκριση μεθόδων

Method	Time/query(s)	QPS	Recall@N vs BLAST
--------	---------------	-----	-------------------

IVF-PQ	6.225	0	0.18
--------	-------	---	------

BLAST (Ref)	0.731	1	1.00 (ref)
-------------	-------	---	------------

[2] Top-N γείτονες ανά μέθοδο (εδώ N = 10 για εκτύπωση)

Method: IVF-PQ

Rank	Neighbor ID	L2 Dist	BLAST Identity	In BLAST Top-N?	Bio comment
------	-------------	---------	----------------	-----------------	-------------

1	sp P10802 ODP2_AZOVI	0.28	1%	Yes	-
2	sp A5VXN9 RL7_PSEP1	0.53	15%	No	-
3	sp Q60A07 RL7_METCA	0.60	14%	No	-
4	sp B1IGG3 RL7_CLOBK	0.61	10%	No	-
5	sp Q0RRT2 RL7_FRAAA	0.61	13%	No	-
6	sp Q7NQE5 RL7_CHRVO	0.61	15%	No	-
7	sp B0VLZ7 RL7_ACIBS	0.61	13%	No	-
8	sp Q1CN79 RL7_YERPN	0.62	10%	No	-
9	sp C4LBV3 RL7_TOLAT	0.62	16%	No	-

10 | sp|C0Q9Y0|RL7\_DESAH | 0.62 | 10% | No | -

Query Protein: A0A010Q3W2

N = 50 (μέγεθος λίστας Top-N για την αξιολόγηση Recall@N)

[1] Συνοπτική σύγκριση μεθόδων

Method | Time/query(s) | QPS | Recall@N vs BLAST

IVF-PQ | 6.225 | 0 | 0.05

BLAST (Ref) | 0.491 | 2 | 1.00 (ref)

[2] Top-N γείτονες ανά μέθοδο (εδώ N = 10 για εκτύπωση)

Method: IVF-PQ

Rank | Neighbor ID | L2 Dist | BLAST Identity | In BLAST Top-N? | Bio comment

1	sp Q8TEB1 DCA11_HUMAN	0.28	0%	No	Possible false positive
2	sp A2RAG5 ATG18_ASPNC	0.29	1%	No	Possible false positive
3	sp Q9S7I8 CDC22_ARATH	0.30	0%	No	Possible false positive
4	sp O82266 SWA1_ARATH	0.31	2%	Yes	-
5	sp P93107 PF20_CHLRE	0.31	1%	No	Possible false positive
6	sp Q05B30 WDR91_XENTR	0.31	0%	No	Possible false positive
7	sp O74340 DCA13_SCHPO	0.32	1%	No	Possible false positive
8	sp P0CS37 HAT2_CRYNB	0.32	2%	No	Possible false positive
9	sp Q8BX09 RBBP5_MOUSE	0.33	1%	No	Possible false positive
10	sp Q5BIM8 ERCC8_BOVIN	0.33	0%	No	Possible false positive

Query Protein: A0A010Q8R4

N = 50 (μέγεθος λίστας Top-N για την αξιολόγηση Recall@N)

[1] Συνοπτική σύγκριση μεθόδων

Method	Time/query(s)	QPS	Recall@N vs BLAST
--------	---------------	-----	-------------------

IVF-PQ	6.225	0	0.05
--------	-------	---	------

BLAST (Ref)	-1.200	0	1.00 (ref)
-------------	--------	---	------------

[2] Top-N γείτονες ανά μέθοδο (εδώ N = 10 για εκτύπωση)

Method: IVF-PQ

Rank	Neighbor ID	L2 Dist	BLAST Identity	In BLAST Top-N?	Bio comment
------	-------------	---------	----------------	-----------------	-------------

1	sp Q4WWM8 SEC2_ASPFU	0.28	1%	No	Possible false positive
2	sp B8MJJ8 RTC5_TALSN	0.29	0%	No	Possible false positive
3	sp E4UPA0 RTC5_ARTGP	0.29	1%	No	Possible false positive
4	sp P0CO69 MDM12_CRYNB	0.29	1%	No	Possible false positive
5	sp Q4WD42 FSQA_ASPFU	0.29	1%	No	Possible false positive
6	sp P0CP30 NPL4_CRYNJ	0.30	0%	No	Possible false positive
7	sp Q5BFX0 EXO84_EMENI	0.30	1%	No	Possible false positive
8	sp A2QNNQ8 GRC3_ASPNC	0.30	1%	No	Possible false positive
9	sp Q24246 DYIN_DROME	0.30	1%	No	Possible false positive
10	sp Q5AZC9 SNX41_EMENI	0.31	1%	No	Possible false positive

