

```
python3 protein_search.py \
-d protein_vectors.fvecs \
-q targets.fasta \
-s swissprot_50k.fasta \
-o benchmarks/res_hypercube.txt \
-method hypercube \
-N 50 \
--hc_k 12 \
--hc_M 5000 \
--hc_probes 2 \
--hc_w 1.5
```

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
Hypercube	0.250	4	0.00
BLAST (Ref)	0.679	1	1.00 (ref)

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

Method: Hypercube

Rank	Neighbor ID	L2 Dist	BLAST Identity	In BLAST Top-N?	Bio comment
1	sp Q8X8V8 MTFA_ECO57	0.35	1%	No	Possible false positive
2	sp B3SRR8 NSP6_ROTH7	0.36	19%	No	Possible false positive

3	sp Q06464 GPMI_ANTSP	0.36	1%	No	Possible false positive
4	sp B2K2M0 YTFE_YERP	0.39	1%	No	Possible false positive
5	sp Q8ZB88 YTFE_YERPE	0.39	1%	No	Possible false positive
6	sp Q46QJ0 SLYX_CUPPJ	0.40	18%	No	Possible false positive
7	sp Q9B6D6 NU4M_YARLI	0.41	1%	No	Possible false positive
8	sp P60957 LGT_ECOL6	0.41	2%	No	Possible false positive
9	sp Q29005 5HT2B_PIG	0.41	17%	No	Possible false positive
10	sp P43142 GP182_MOUSE	0.41	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
--------	---------------	-----	-------------------

Hypercube	0.250	4	0.00
-----------	-------	---	------

BLAST (Ref)	0.387	2	1.00 (ref)
-------------	-------	---	------------

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

Method: Hypercube

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

1	sp O74738 SET10_SCHPO	0.34	2%	No	Possible false positive
---	-----------------------	------	----	----	-------------------------

2	sp A5ISZ2 Y1521_STAA9	0.36	15%	No	Possible false positive
---	-----------------------	------	-----	----	-------------------------

3	sp Q9CM18 SPRT_PASMU	0.36	20%	No	Possible false positive
---	----------------------	------	-----	----	-------------------------

4	sp Q2G1M8 SBNF_STAA8	0.39	0%	No	Possible false positive
5	sp Q7MAX6 TRMA_PHOLL	0.40	1%	No	Possible false positive
6	sp Q74Q56 SGRR_YERPE	0.40	1%	No	Possible false positive
7	sp Q1CMQ0 SGRR_YERPN	0.41	1%	No	Possible false positive
8	sp Q66EM6 SGRR_YERPS	0.41	1%	No	Possible false positive
9	sp Q8X8V8 MTFA_ECO57	0.41	1%	No	Possible false positive
10	sp P0C2V9 SYCT_YEREN	0.41	23%	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
--------	---------------	-----	-------------------

Hypercube	0.250	4	0.10
-----------	-------	---	------

BLAST (Ref)	0.887	1	1.00 (ref)
-------------	-------	---	------------

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

Method: Hypercube

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

1	sp P9WQJ3 FATRP_MYCTU	0.30	1%	Yes	-
2	sp Q13BH6 NDVA_RHOPS	0.32	1%	Yes	-
3	sp Q9ZEP3 CSEC_STRCO	0.33	1%	No	Possible false positive
4	sp A0K739 SSUB1_BURCH	0.34	2%	No	Possible false positive

5	sp Q51719 YCOBA_PROFR	0.39	1%	No	Possible false positive
6	sp Q3JSR6 SSUB_BURP1	0.40	2%	No	Possible false positive
7	sp P23886 CYDC_ECOLI	0.40	1%	Yes	-
8	sp Q21BF6 MODC_RHOPB	0.40	1%	No	Possible false positive
9	sp Q82MV1 SSUB1_STRAW	0.41	0%	No	Possible false positive
10	sp A6W3V7 RECF_KINRD	0.41	1%	No	Possible false positive

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
--------	---------------	-----	-------------------

Hypercube	0.250	4	0.30
-----------	-------	---	------

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

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

Method: Hypercube

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

1	sp P9WQJ3 FATRP_MYCTU	0.31	1%	Yes	-
2	sp Q13BH6 NDVA_RHOPS	0.33	2%	Yes	-
3	sp Q03024 APRD_PSEAE	0.35	0%	Yes	-
4	sp Q20Z38 NDVA_RHOPB	0.36	2%	Yes	-
5	sp Q8P8W4 MSBA_XANCP	0.41	1%	Yes	-

6	sp Q51719 YCOBA_PROFR	0.41	1%	No	Possible false positive
7	sp Q82MV1 SSUB1_STRAW	0.42	1%	No	Possible false positive
8	sp Q3JSR6 SSUB_BURP1	0.42	2%	No	Possible false positive
9	sp P9WQJ4 OPPD_MYCTO	0.42	1%	No	Possible false positive
10	sp Q7VZE5 CYSA_BORPE	0.43	2%	No	Possible false positive

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
--------	---------------	-----	-------------------

Hypercube	0.250	4	0.36
-----------	-------	---	------

BLAST (Ref)	0.272	3	1.00 (ref)
-------------	-------	---	------------

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

Method: Hypercube

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

1	sp Q5HI09 GRAR_STAAC	0.18	1%	Yes	-
2	sp P33112 SPAR_BACIU	0.18	2%	Yes	-
3	sp Q99VW2 GRAR_STAAN	0.19	1%	Yes	-
4	sp Q8CQ37 GRAR_STAES	0.19	2%	Yes	-
5	sp Q99RR6 HSSR_STAAM	0.21	1%	Yes	-
6	sp P28257 YCF27_GALSU	0.22	2%	Yes	-

7	sp Q9AE24 RPRY_BACFR	0.22	2%	Yes	-
8	sp P0CAW8 CTRA_CAUVC	0.22	1%	Yes	-
9	sp A0A0H2UQ68 WALR_...	0.24	1%	Yes	-
10	sp P0A4H8 CIAR_STRR6	0.25	1%	Yes	-

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
--------	---------------	-----	-------------------

-----

Hypercube	0.250	4	0.00
-----------	-------	---	------

BLAST (Ref)	0.460	2	1.00 (ref)
-------------	-------	---	------------

-----

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

Method: Hypercube

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

-----

1	sp Q8EEF3 RUVB_SHEON	0.14	1%	No	Possible false positive
2	sp Q03R36 RUVB_LEVBA	0.22	0%	No	Possible false positive
3	sp Q03YJ6 RUVB_LEUMM	0.22	0%	No	Possible false positive
4	sp Q15RN6 RUVB_PSEA6	0.22	0%	No	Possible false positive
5	sp Q8DEV2 CLPB_VIBVU	0.22	1%	No	Possible false positive
6	sp Q9KU18 CLPB_VIBCH	0.23	1%	No	Possible false positive
7	sp Q88V03 RUVB_LACPL	0.24	0%	No	Possible false positive

8	sp Q889C2 CLPB_PSESM	0.24	1%	No	Possible false positive
9	sp Q5HRM8 CLPC_STAEQ	0.24	0%	No	Possible false positive
10	sp Q7VYV6 CLPB_BORPE	0.24	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
Hypercube	0.250	4	0.03
BLAST (Ref)	0.660	1	1.00 (ref)

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

Method: Hypercube

Rank	Neighbor ID	L2 Dist	BLAST Identity	In BLAST Top-N?	Bio comment
1	sp A7MSB2 RAPA_VIBC1	0.14	1%	Yes	-
2	sp Q1MQP3 SECA_LAWIP	0.14	1%	No	Possible false positive
3	sp A8GSU6 SECA_RICRS	0.15	1%	No	Possible false positive
4	sp Q03ZQ8 SECA_LEUMM	0.15	1%	No	Possible false positive
5	sp Q6MEX9 SECA_PARUW	0.15	2%	No	Possible false positive
6	sp Q15Q25 SECA_PSEA6	0.16	2%	No	Possible false positive
7	sp Q21MG1 SECA_SACD2	0.16	1%	No	Possible false positive
8	sp B7I8J5 SECA_ACIB5	0.16	2%	No	Possible false positive

9	sp Q5HA05 SECA_EHRRW	0.16	1%	No	Possible false positive
10	sp A0KPW5 SECA_AERHH	0.16	2%	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
--------	---------------	-----	-------------------

Hypercube	0.250	4	0.13
-----------	-------	---	------

BLAST (Ref)	0.675	1	1.00 (ref)
-------------	-------	---	------------

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

Method: Hypercube

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

1	sp Q0HR19 RAPA_SHESR	0.11	1%	Yes	-
2	sp Q0HMR2 RAPA_SHESM	0.11	1%	Yes	-
3	sp A0KGL5 RAPA_AERHH	0.11	1%	Yes	-
4	sp Q48LP5 RAPA_PSE14	0.12	2%	Yes	-
5	sp Q3KGV8 RAPA_PSEPF	0.12	2%	Yes	-
6	sp Q9HYT6 RAPA_PSEAE	0.12	1%	Yes	-
7	sp A7FKM4 UVRB_YERP3	0.12	0%	No	Possible false positive
8	sp C5BEJ2 UVRB_EDWI9	0.13	0%	No	Possible false positive
9	sp B2UCW7 SECA_RALPJ	0.13	1%	No	Possible false positive



10	sp Q3K063 SECA2_STRA1	0.13	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
--------	---------------	-----	-------------------

Hypercube	0.250	4	0.02
-----------	-------	---	------

BLAST (Ref)	0.706	1	1.00 (ref)
-------------	-------	---	------------

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

Method: Hypercube

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

1	sp Q9CMJ8 RIMK_PASMU	0.27	1%	No	Possible false positive
2	sp A3QHZ8 RSMC_SHELP	0.28	0%	No	Possible false positive
3	sp B1IRN3 RLMG_ECOLC	0.29	0%	No	Possible false positive
4	sp Q5WSS4 MNMG_LEGPL	0.29	1%	No	Possible false positive
5	sp B0TUZ2 RLMC_SHEHH	0.30	1%	No	Possible false positive
6	sp B5BG31 RLMG_SALPK	0.31	0%	No	Possible false positive
7	sp Q033L1 MNMG_LACP3	0.31	1%	No	Possible false positive
8	sp Q7N126 LEUD_PHOLL	0.31	1%	No	Possible false positive
9	sp Q9XBS9 Y1353_ZYMMO	0.31	16%	No	Possible false positive
10	sp Q7MIT5 MNMC_VIBVY	0.31	0%	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
Hypercube	0.250	4	0.32
BLAST (Ref)	0.811	1	1.00 (ref)

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

Method: Hypercube

Rank	Neighbor ID	L2 Dist	BLAST Identity	In BLAST Top-N?	Bio comment
1	sp P10802 ODP2_AZOVI	0.28	1%	Yes	-
2	sp P49786 BCCP_BACSU	0.46	15%	No	-
3	sp P22439 ODPX_BOVIN	0.46	1%	Yes	-
4	sp P43874 BCCP_HAEIN	0.46	16%	Yes	-
5	sp P57003 TONB_NEIMA	0.50	1%	No	-
6	sp A0QJ48 RS16_MYCA1	0.51	11%	No	-
7	sp P45523 FKBA_ECOLI	0.53	1%	No	-
8	sp A0QV37 RS16_MYCS2	0.55	11%	No	-
9	sp B4RCD1 RS16_PHEZH	0.56	12%	No	-
10	sp A8ZV83 RL17_DESOH	0.56	1%	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
Hypercube	0.250	4	0.05
BLAST (Ref)	0.513	1	1.00 (ref)

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

Method: Hypercube

Rank	Neighbor ID	L2 Dist	BLAST Identity	In BLAST Top-N?	Bio comment
1	sp A2QI22 YTM1_ASPNC	0.24	1%	No	Possible false positive
2	sp A2RAG5 ATG18_ASPNC	0.25	1%	No	Possible false positive
3	sp O35828 CORO7_RAT	0.25	1%	No	Possible false positive
4	sp Q9P7C0 YKU5_SCHPO	0.26	0%	No	Possible false positive
5	sp Q4R4I8 COPB2_MACFA	0.27	0%	No	Possible false positive
6	sp Q4P553 HAT2_MYCMD	0.27	1%	No	Possible false positive
7	sp A3LNM0 IPI3_PICST	0.28	2%	No	Possible false positive
8	sp Q6KAU8 A16L2_MOUSE	0.28	1%	No	Possible false positive
9	sp Q4WTQ4 FYV10_ASPFU	0.28	1%	No	Possible false positive
10	sp A6ZRQ4 IPI3_YEAS7	0.29	1%	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
--------	---------------	-----	-------------------

Hypercube	0.250	4	0.03
-----------	-------	---	------

BLAST (Ref)	0.719	1	1.00 (ref)
-------------	-------	---	------------

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

Method: Hypercube

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

1	sp Q6CXX3 HIR1_KLULA	0.25	0%	Yes	-
2	sp P0CP30 NPL4_CRYNJ	0.25	0%	No	Possible false positive
3	sp P38419 LOXC1_ORYSJ	0.25	1%	No	Possible false positive
4	sp Q4WE58 TRM6_ASPFU	0.26	0%	No	Possible false positive
5	sp Q9SKN1 MON1_ARATH	0.27	1%	No	Possible false positive
6	sp Q5BFX0 EXO84_EMENI	0.28	1%	No	Possible false positive
7	sp Q10QA2 CCD53_ORYSJ	0.28	0%	No	Possible false positive
8	sp Q9NV92 NFIP2_HUMAN	0.28	1%	No	Possible false positive
9	sp Q6FLA4 ERB1_CANGA	0.28	1%	No	Possible false positive
10	sp P54283 CACB1_RAT	0.28	1%	No	Possible false positive