

Meet-U

edition 20²⁰

It is time to fold!!!

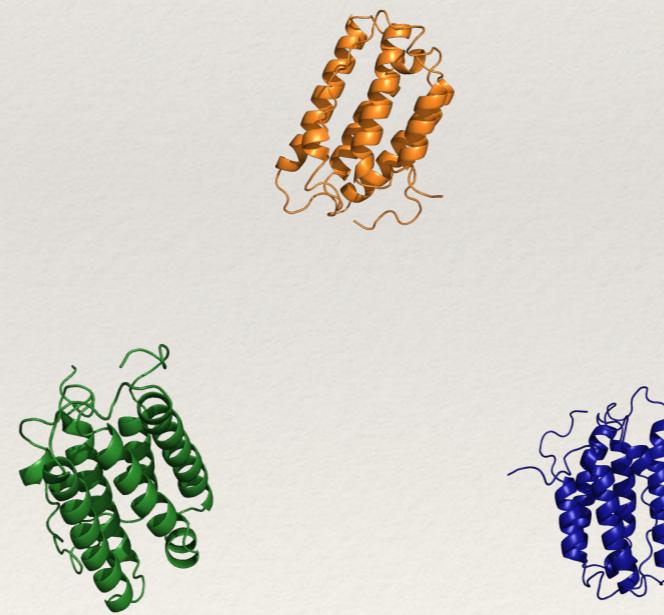
Data presentation

03 / 11 / 2019

HOMSTRAD

HOMSTRAD (HOMologous STructure Alignment Database):

- A curated database of **structure-based alignments** for 1010 homologous protein families.



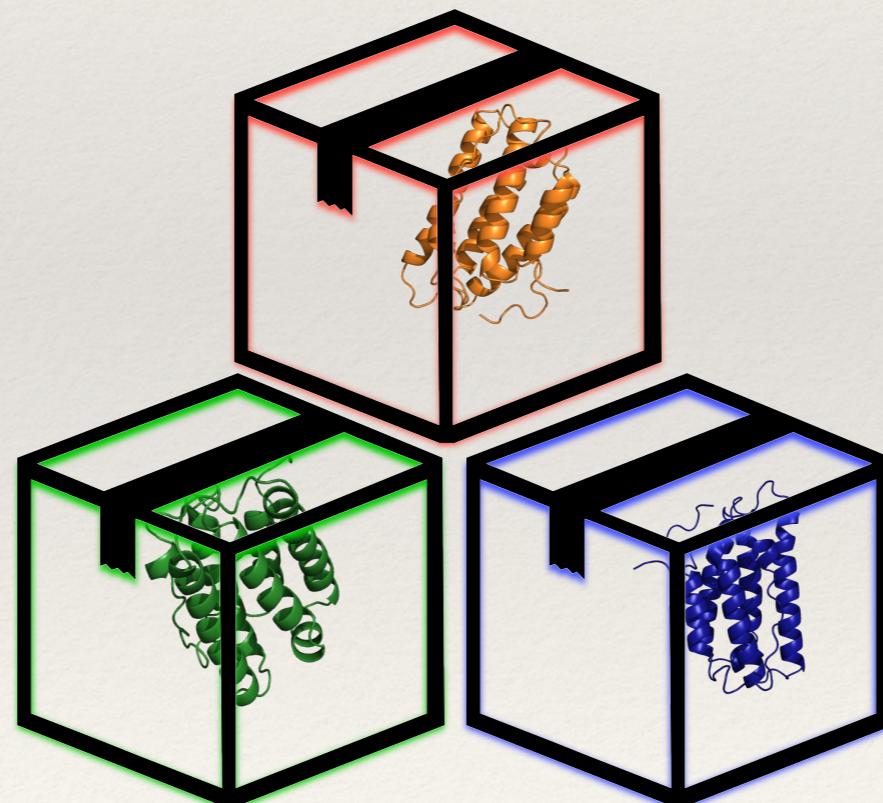
Mizuguchi K, Deane CM, Blundell TL, Overington JP. (1998)

HOMSTRAD: a database of protein structure alignments for homologous families. *Protein Science* 7:2469-2471.

HOMSTRAD

HOMSTRAD (HOMologous STructure Alignment Database):

- A curated database of **structure-based alignments** for 1010 homologous protein families.
- All known **protein structures** are clustered into homologous families (1 representative)



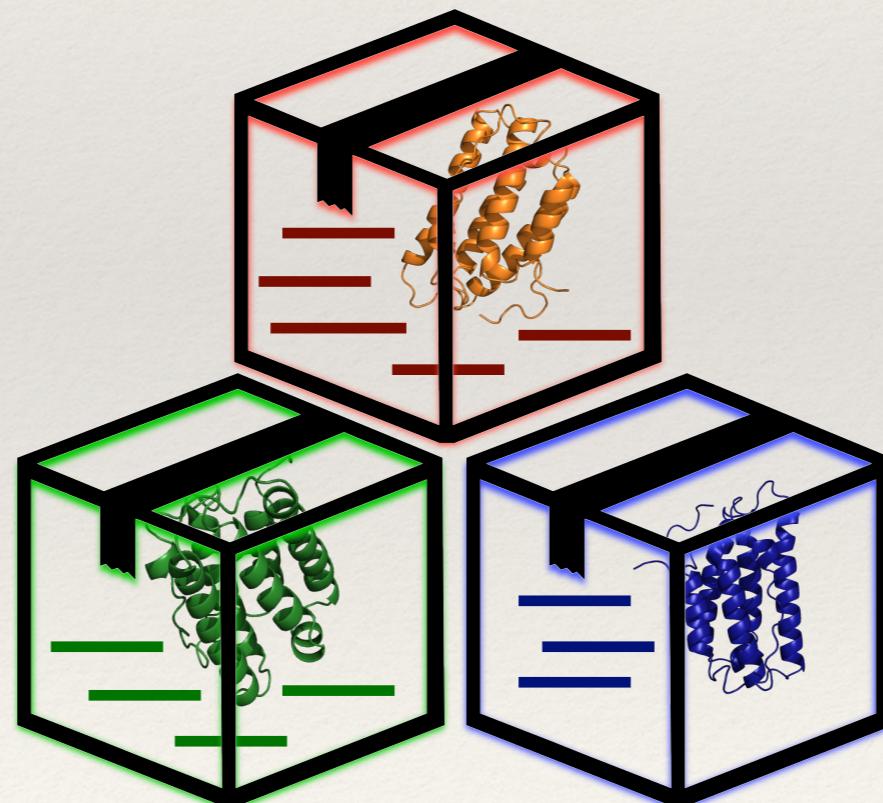
Mizuguchi K, Deane CM, Blundell TL, Overington JP. (1998)

HOMSTRAD: a database of protein structure alignments for homologous families. *Protein Science* 7:2469-2471.

HOMSTRAD

HOMSTRAD (HOMologous STructure Alignment Database):

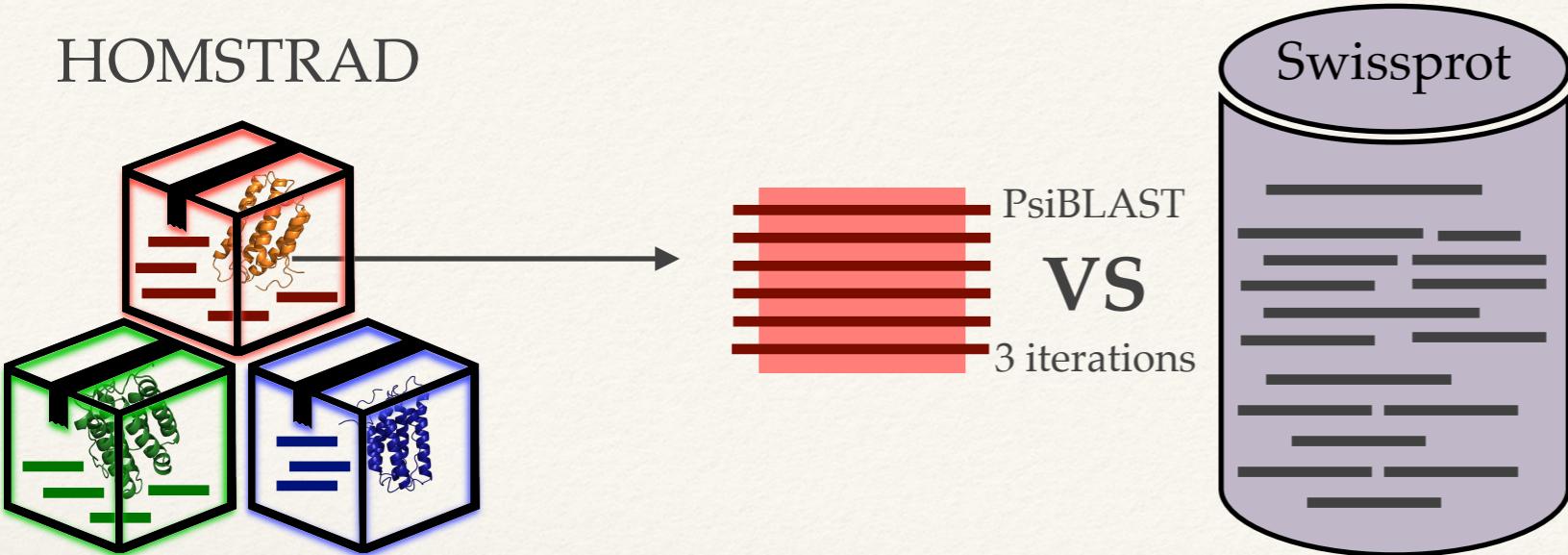
- A curated database of **structure-based alignments** for 1010 homologous protein families.
- All known **protein structures** are clustered into homologous families (1 representative)
- Sequences of representative members of each family are aligned on the basis of their 3D structures



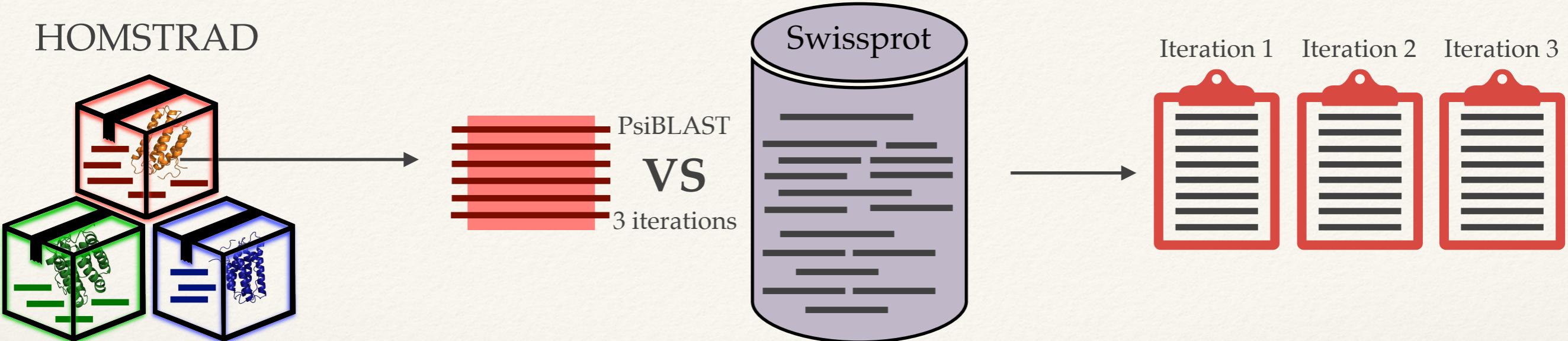
Mizuguchi K, Deane CM, Blundell TL, Overington JP. (1998)

HOMSTRAD: a database of protein structure alignments for homologous families. *Protein Science* 7:2469-2471.

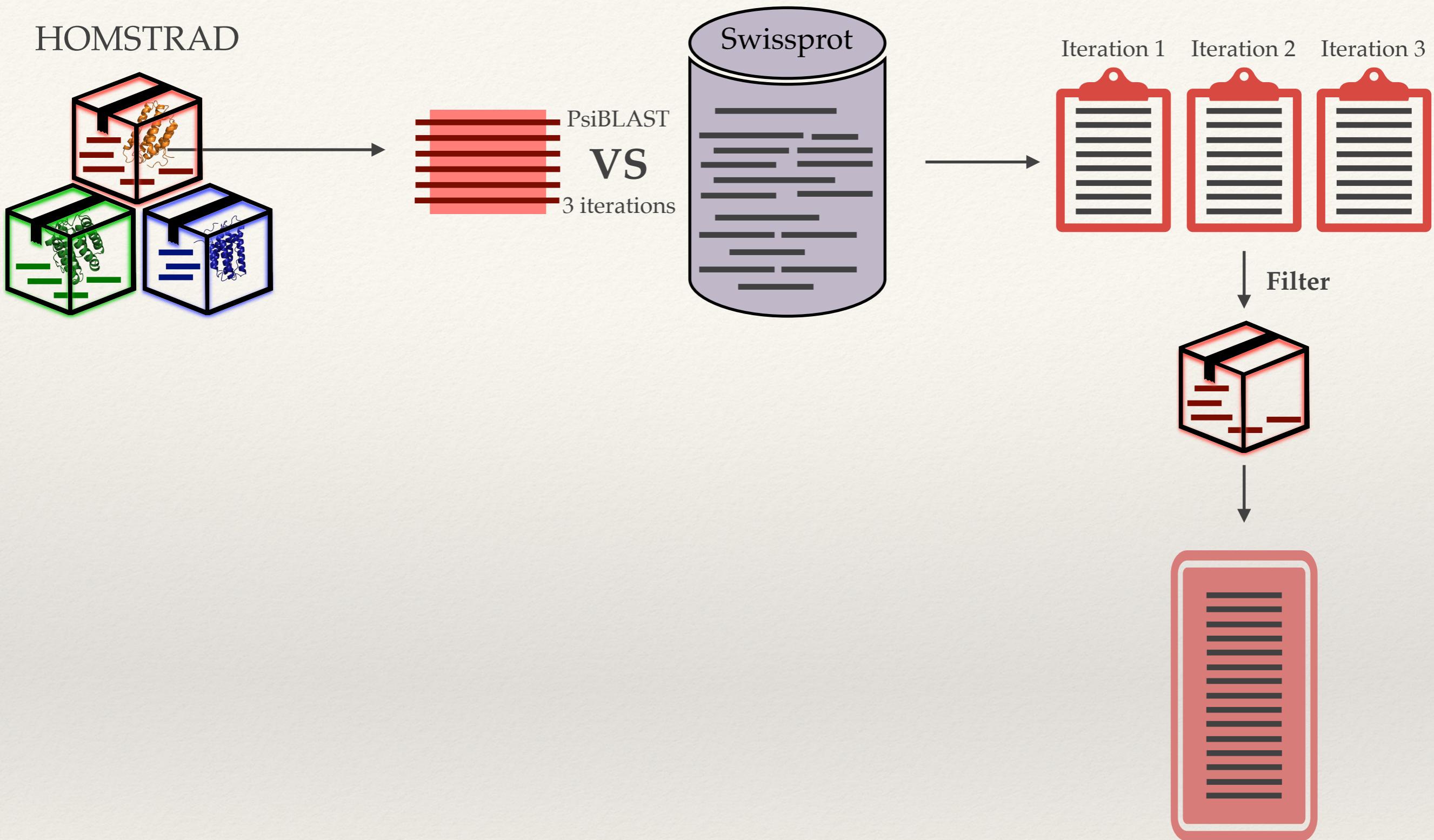
HOMSTRAD



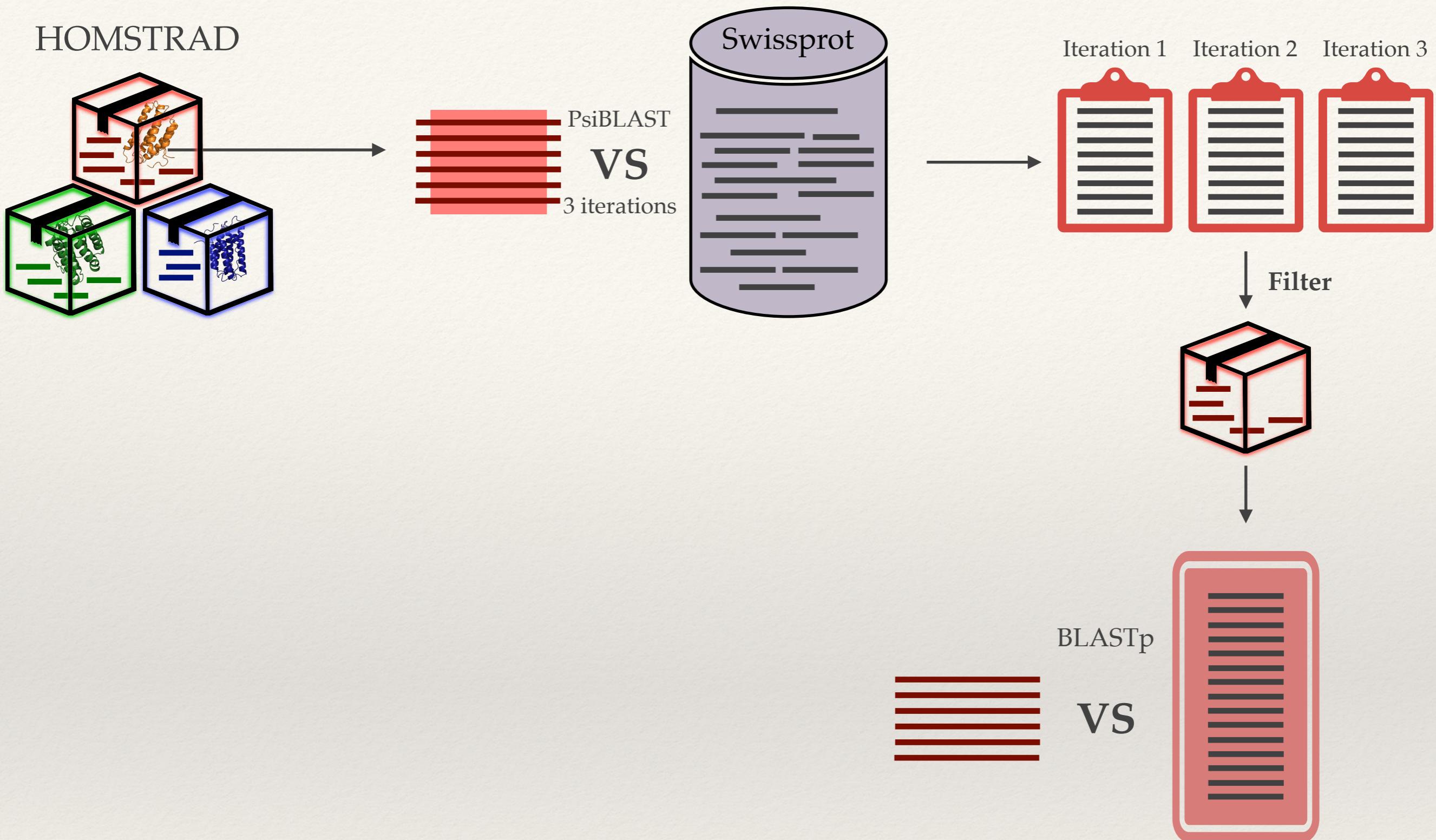
HOMSTRAD



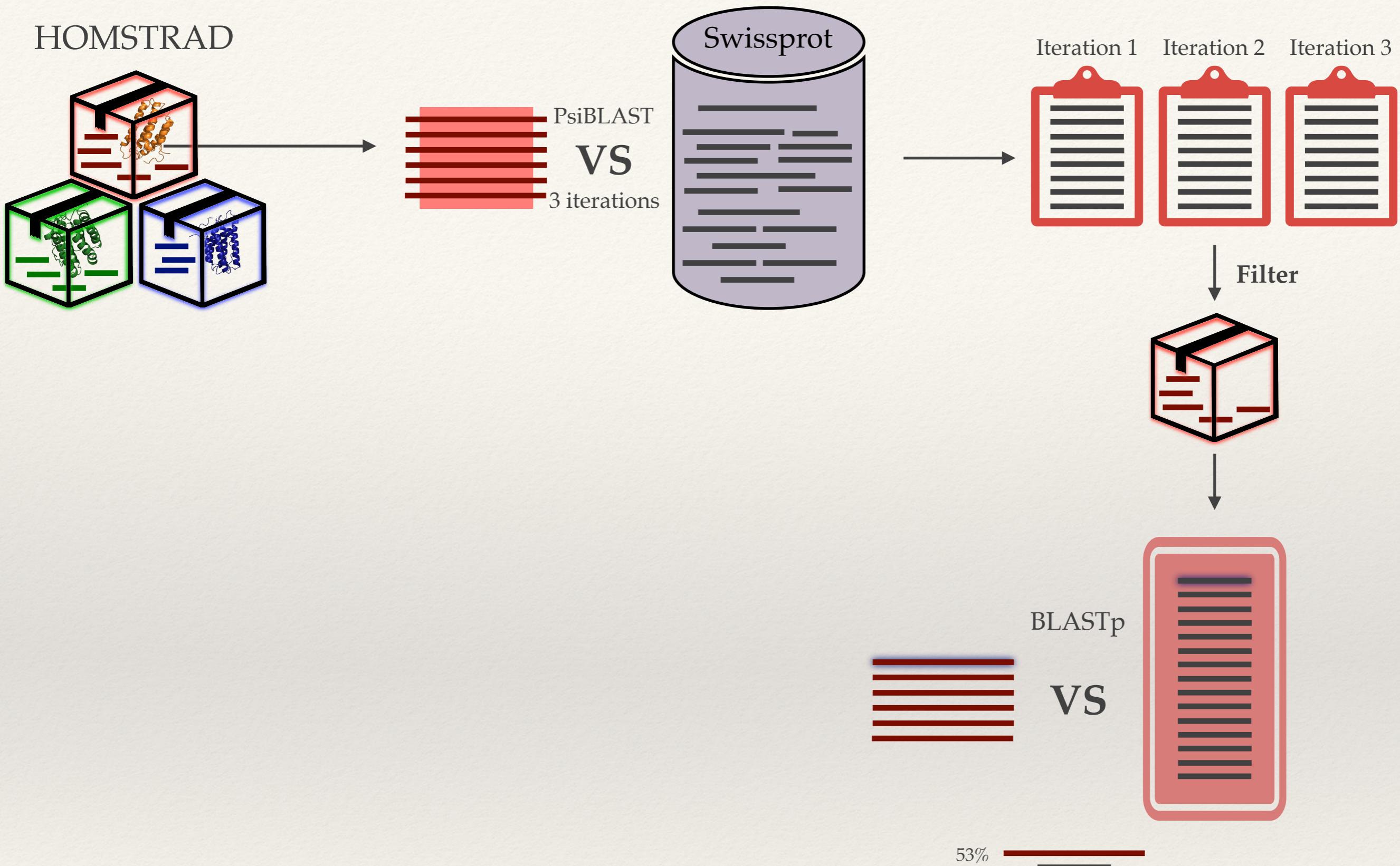
HOMSTRAD



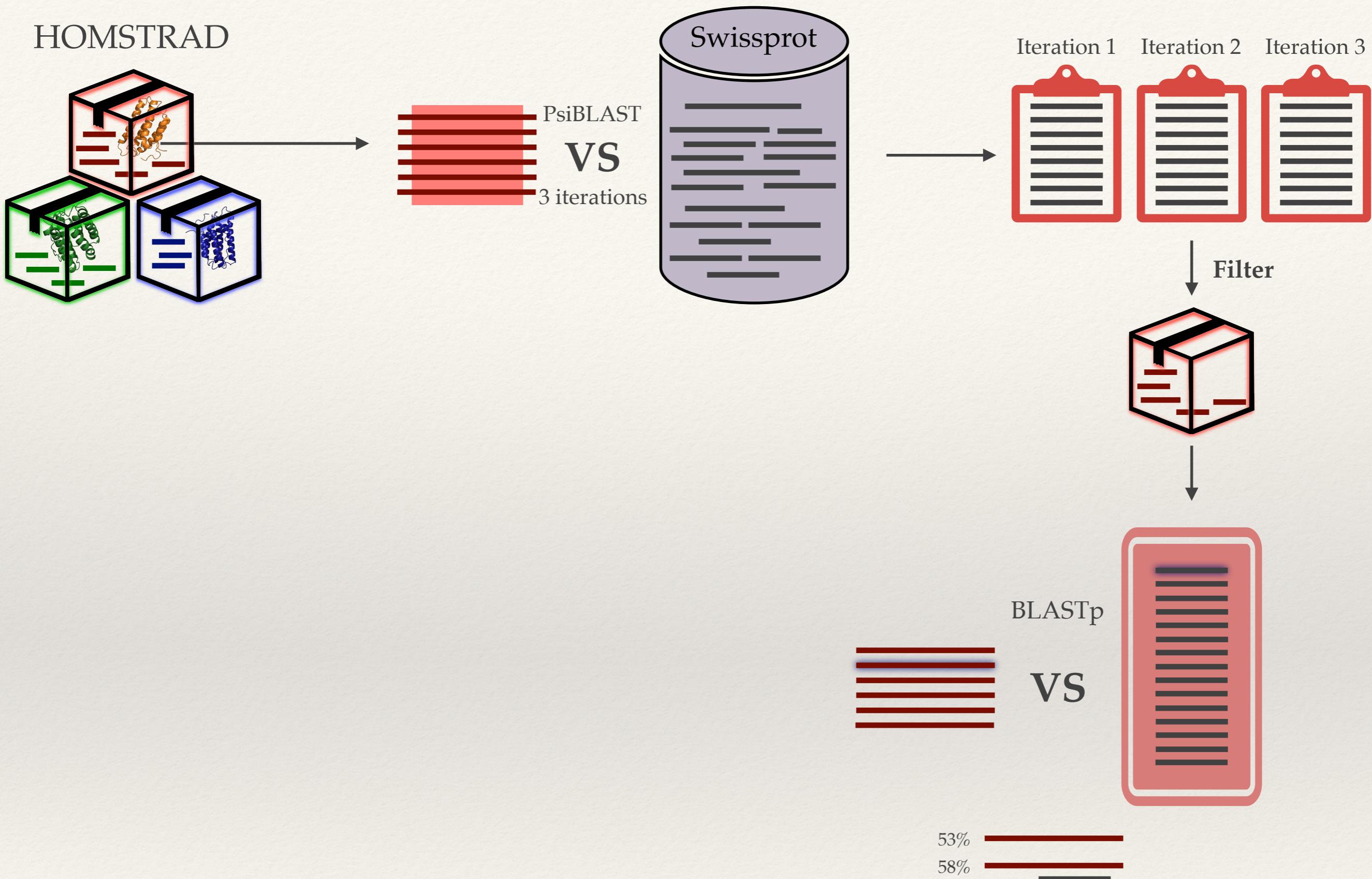
HOMSTRAD



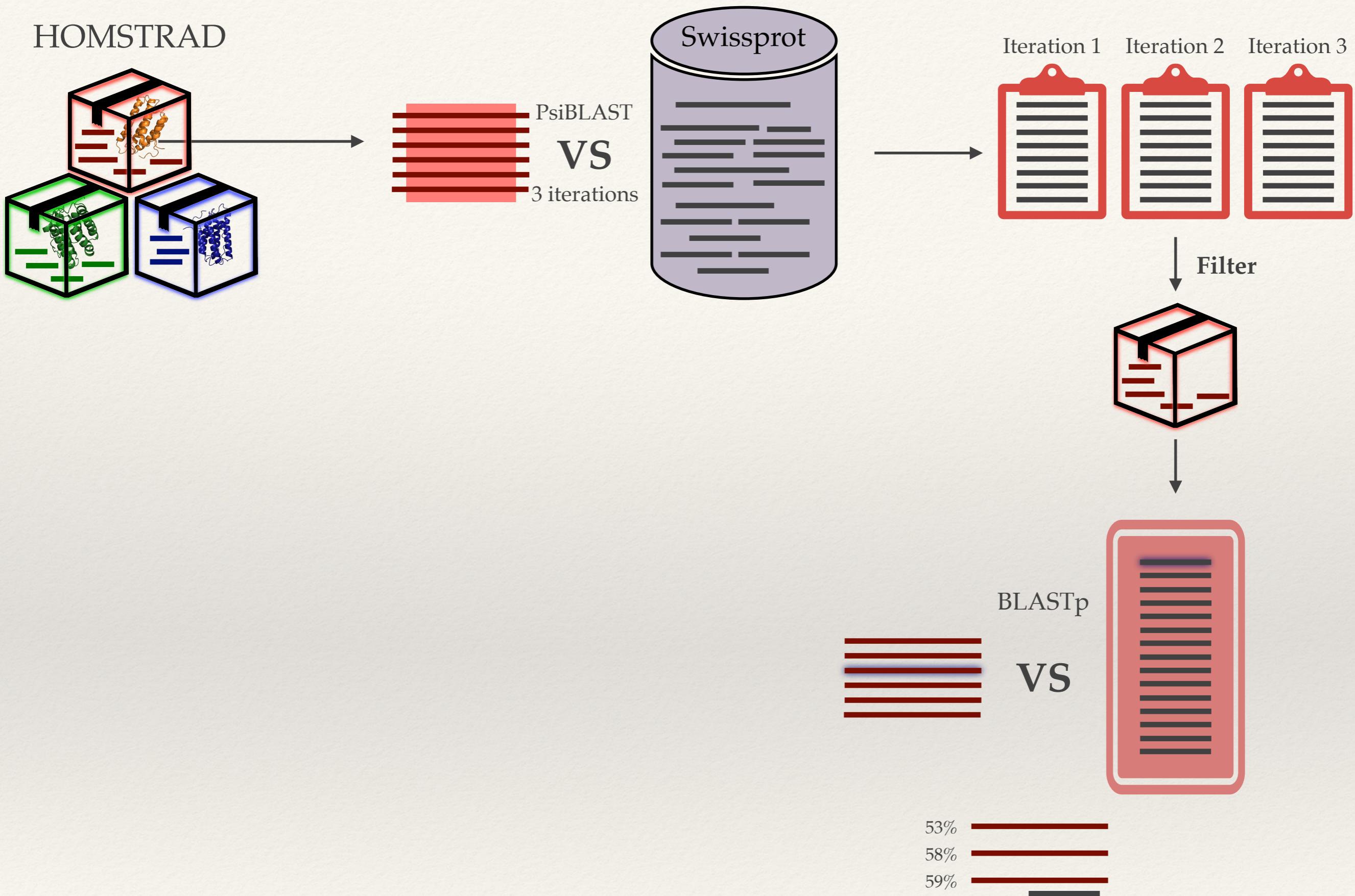
HOMSTRAD



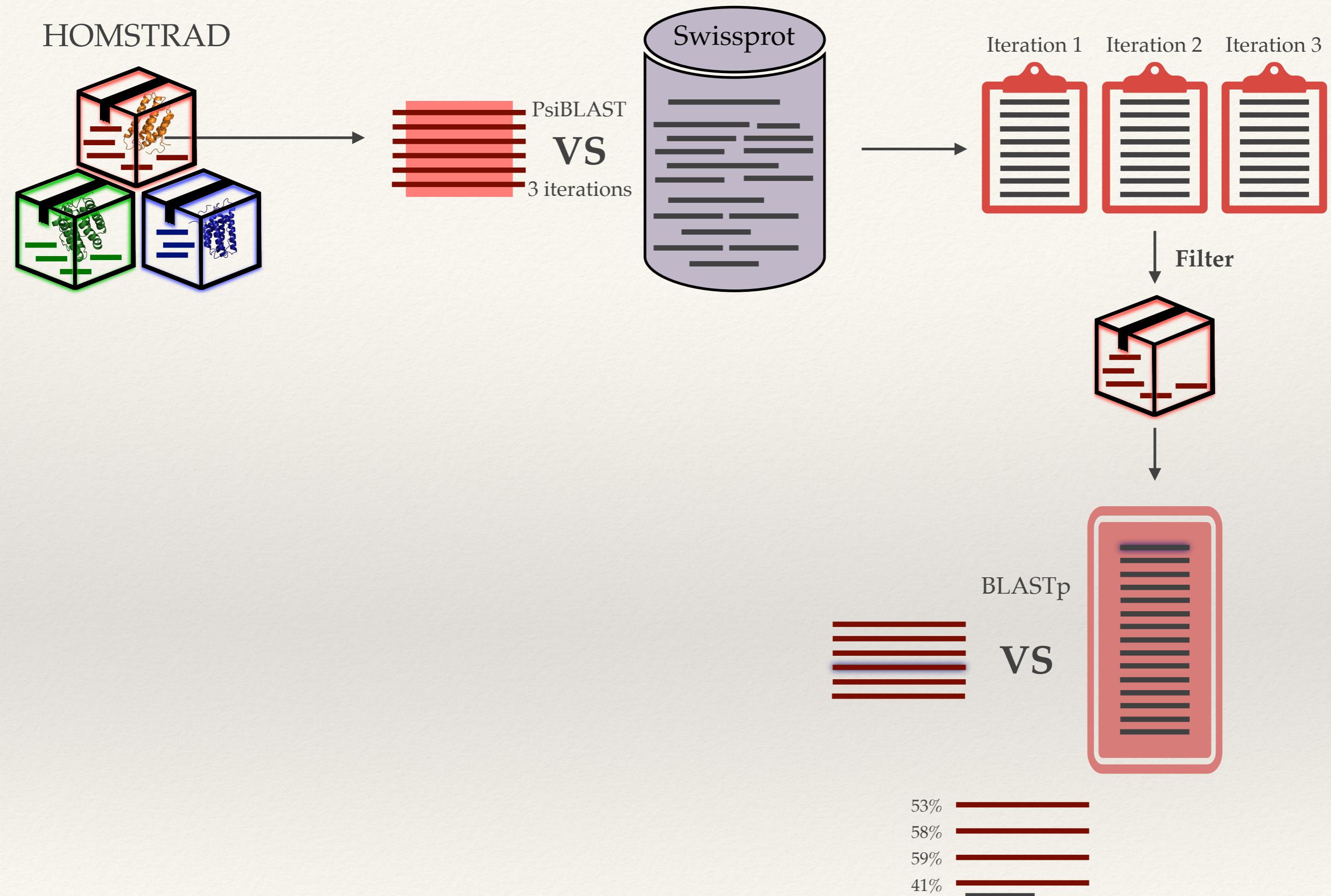
HOMSTRAD



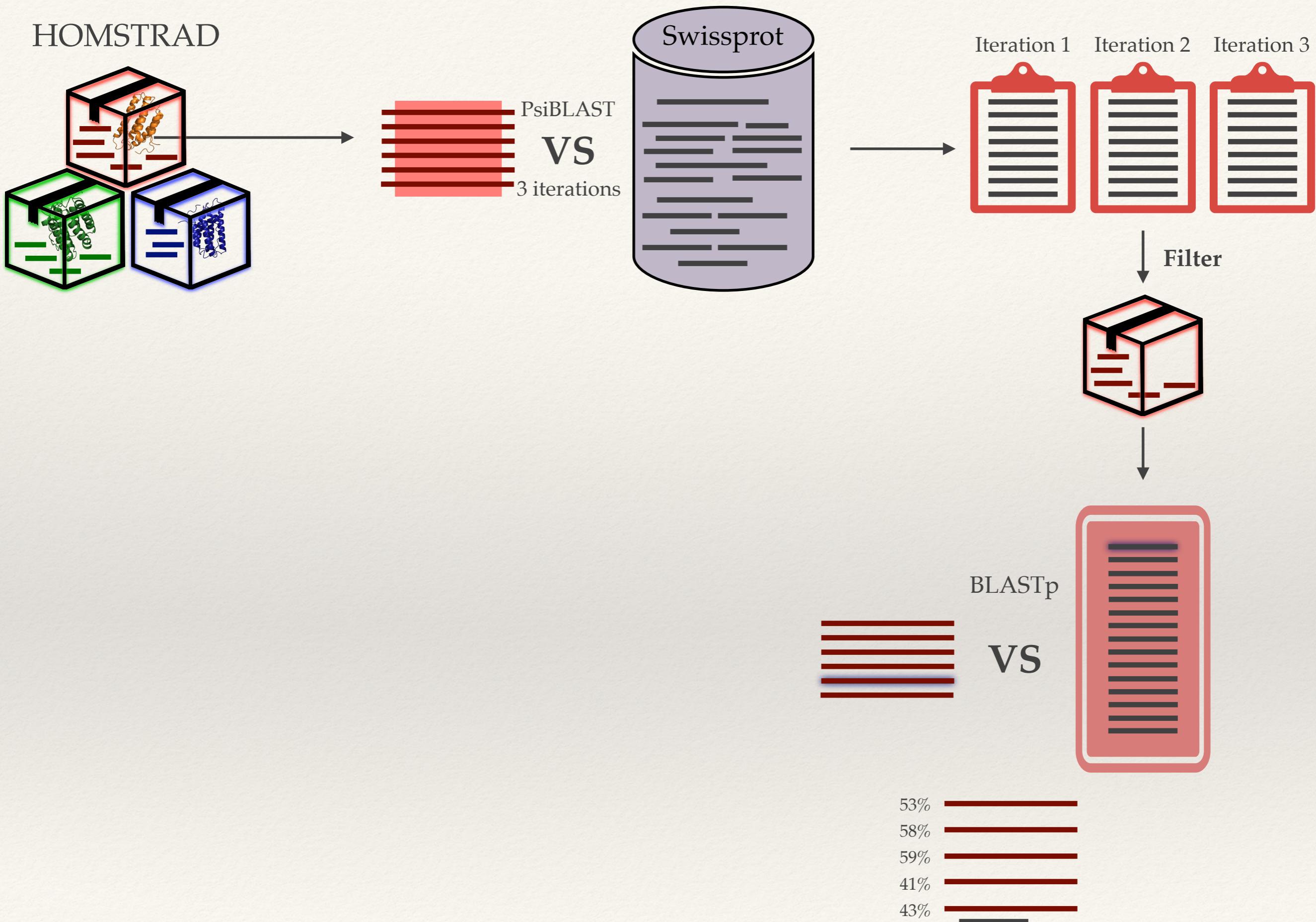
HOMSTRAD



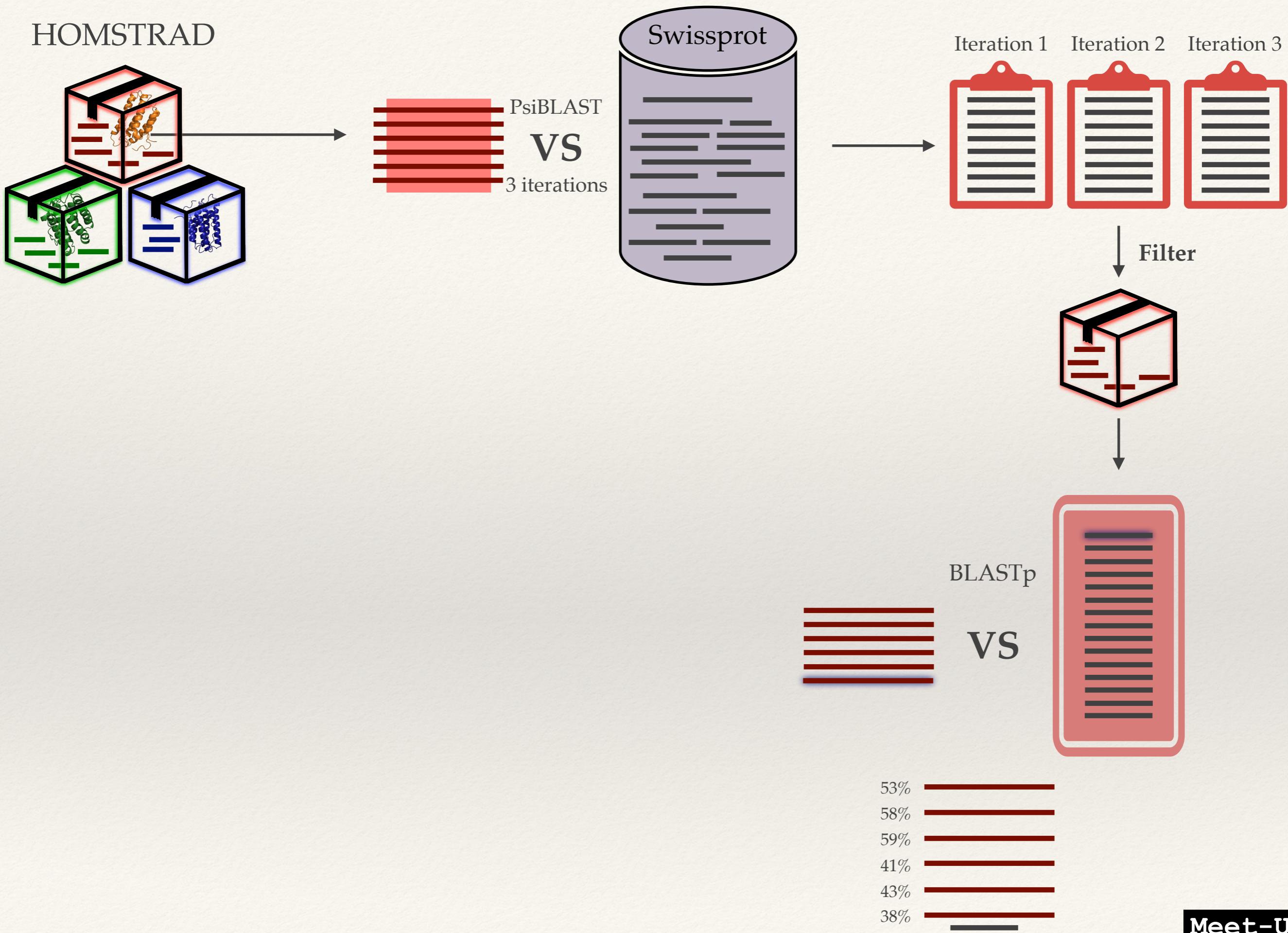
HOMSTRAD



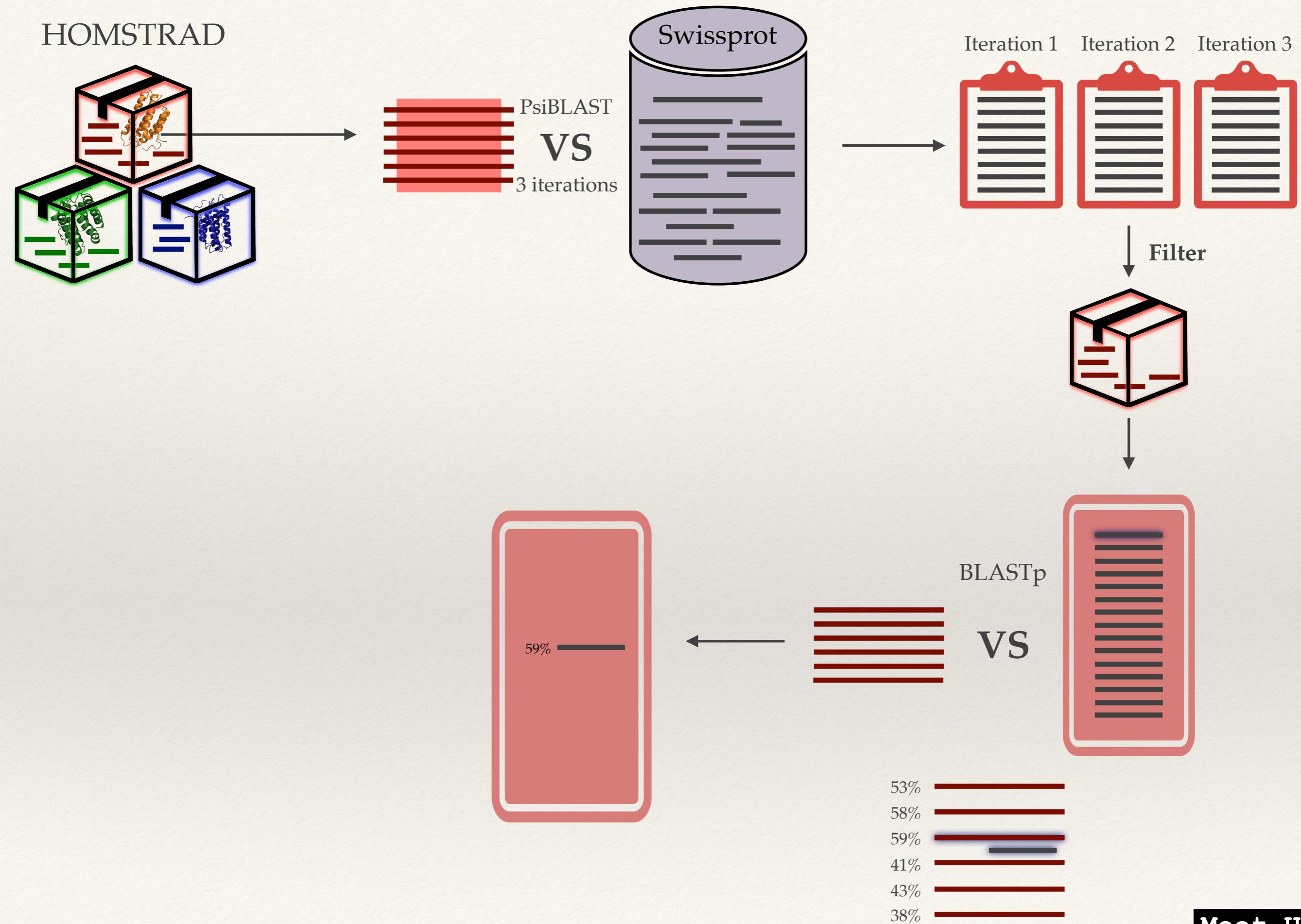
HOMSTRAD



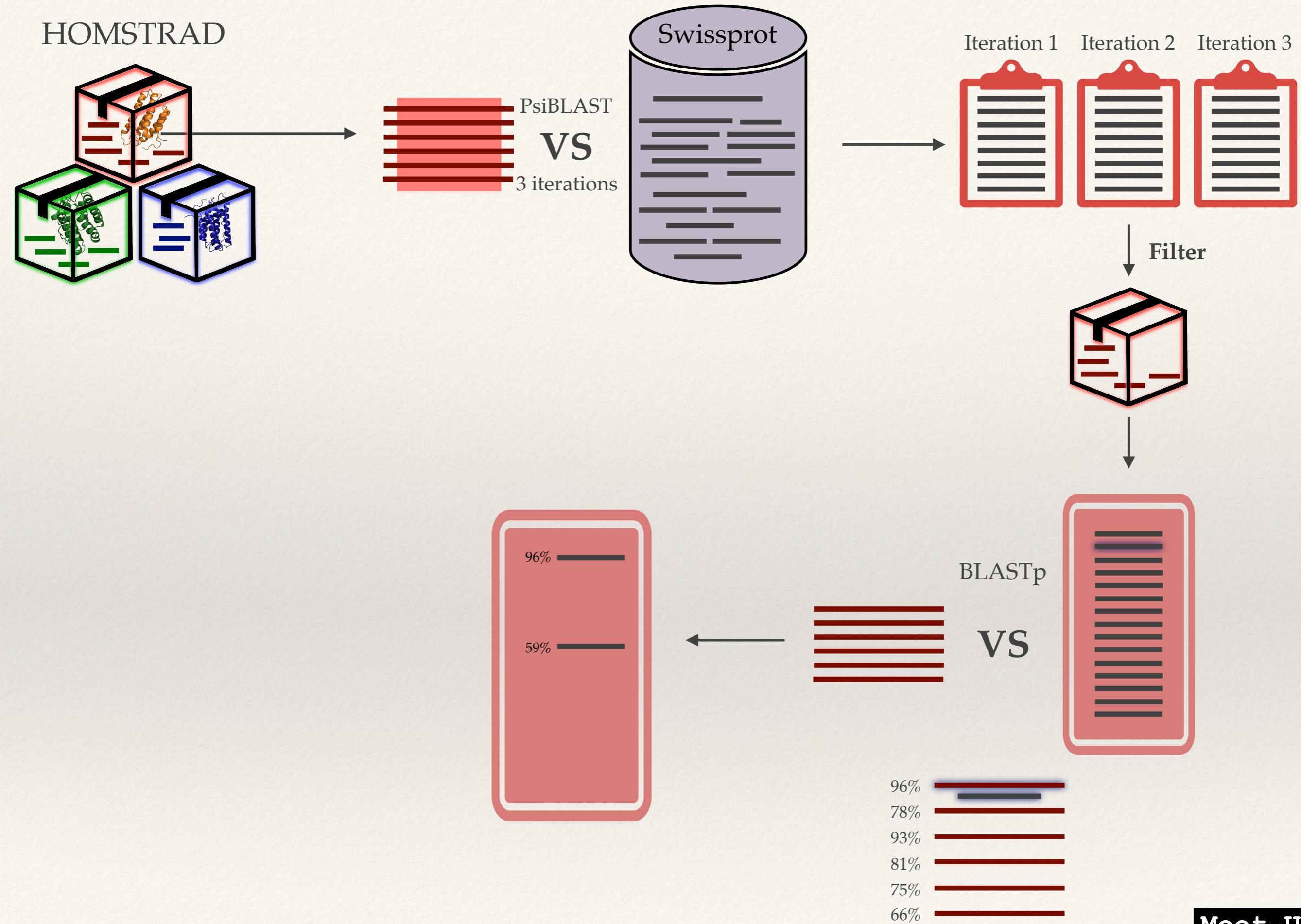
HOMSTRAD



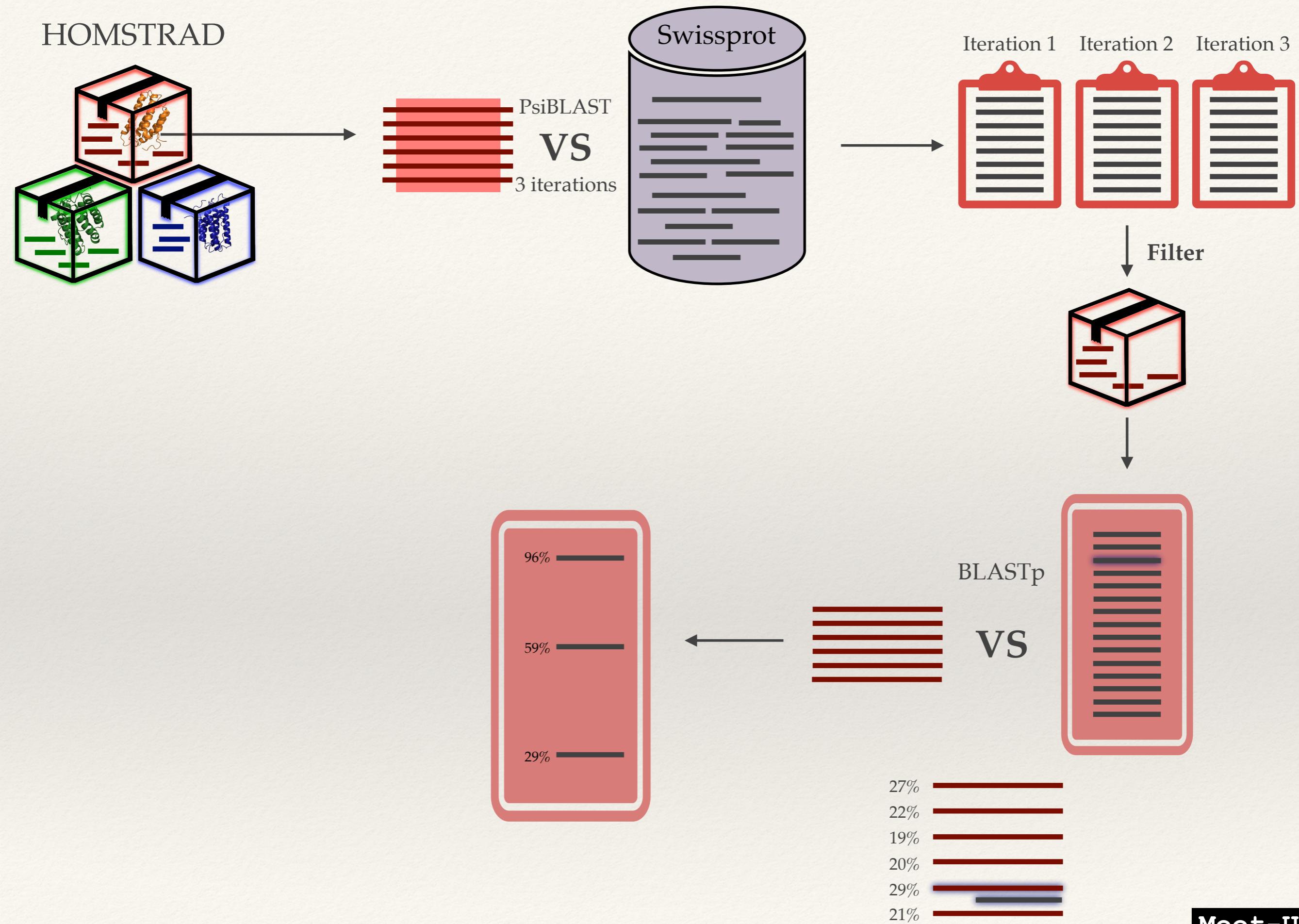
HOMSTRAD



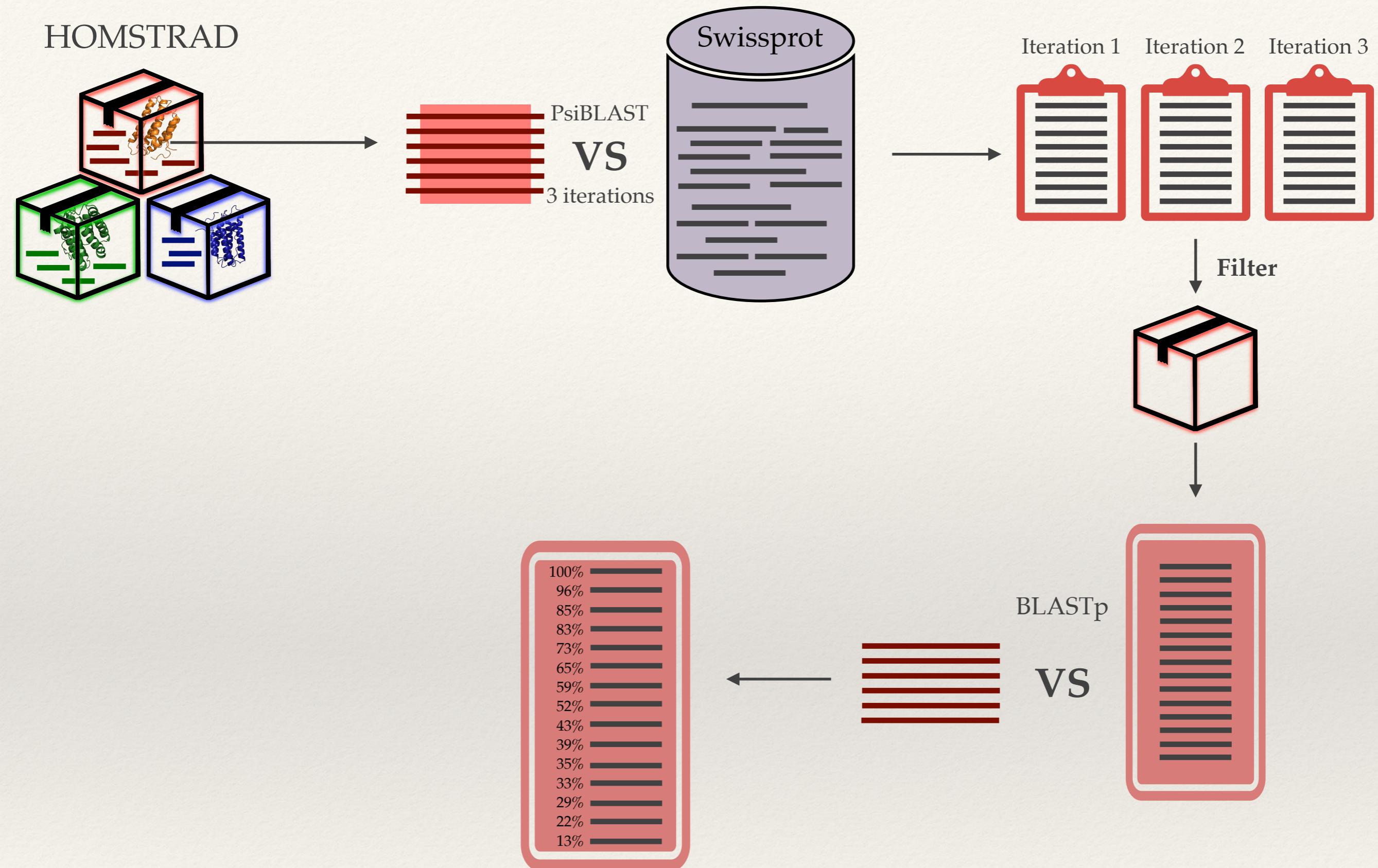
HOMSTRAD



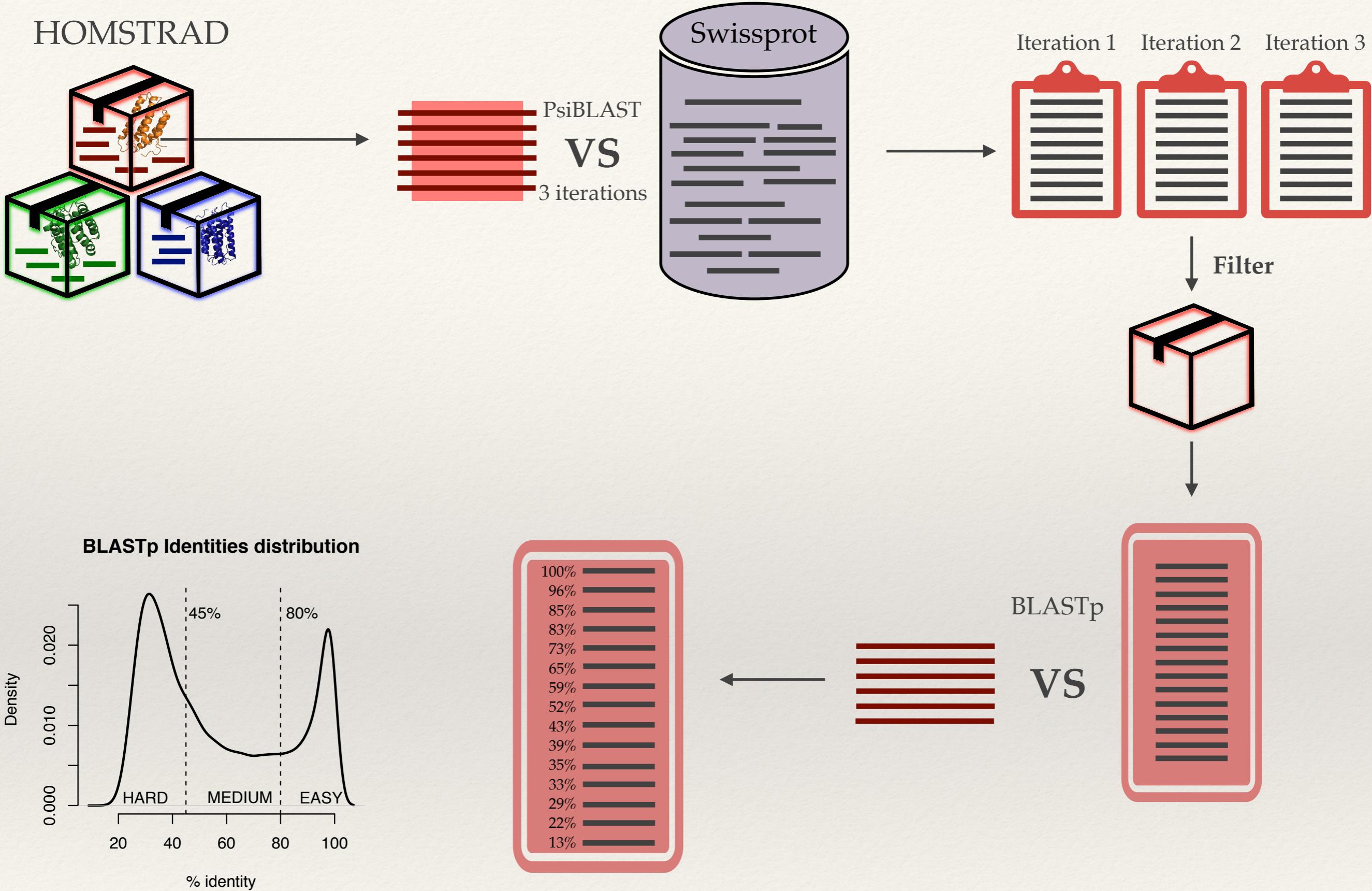
HOMSTRAD



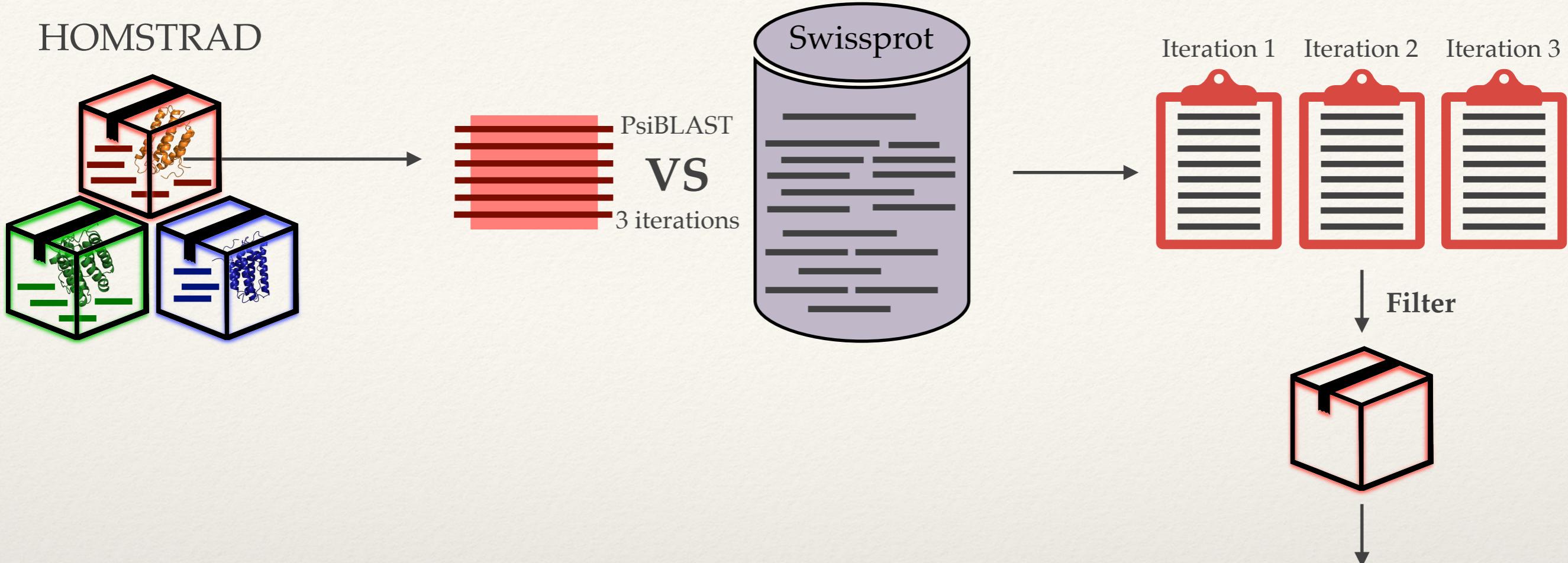
HOMSTRAD



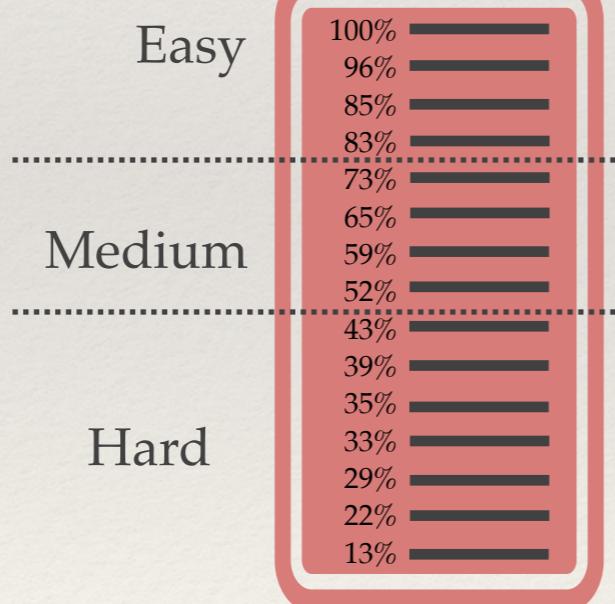
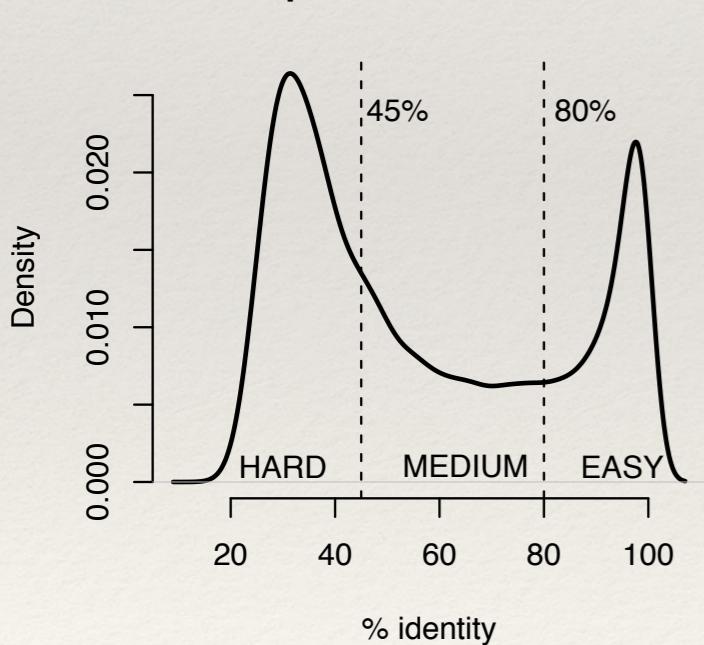
HOMSTRAD



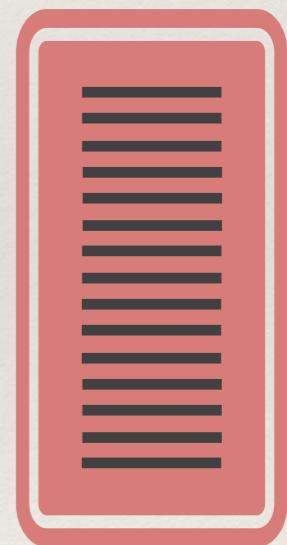
HOMSTRAD



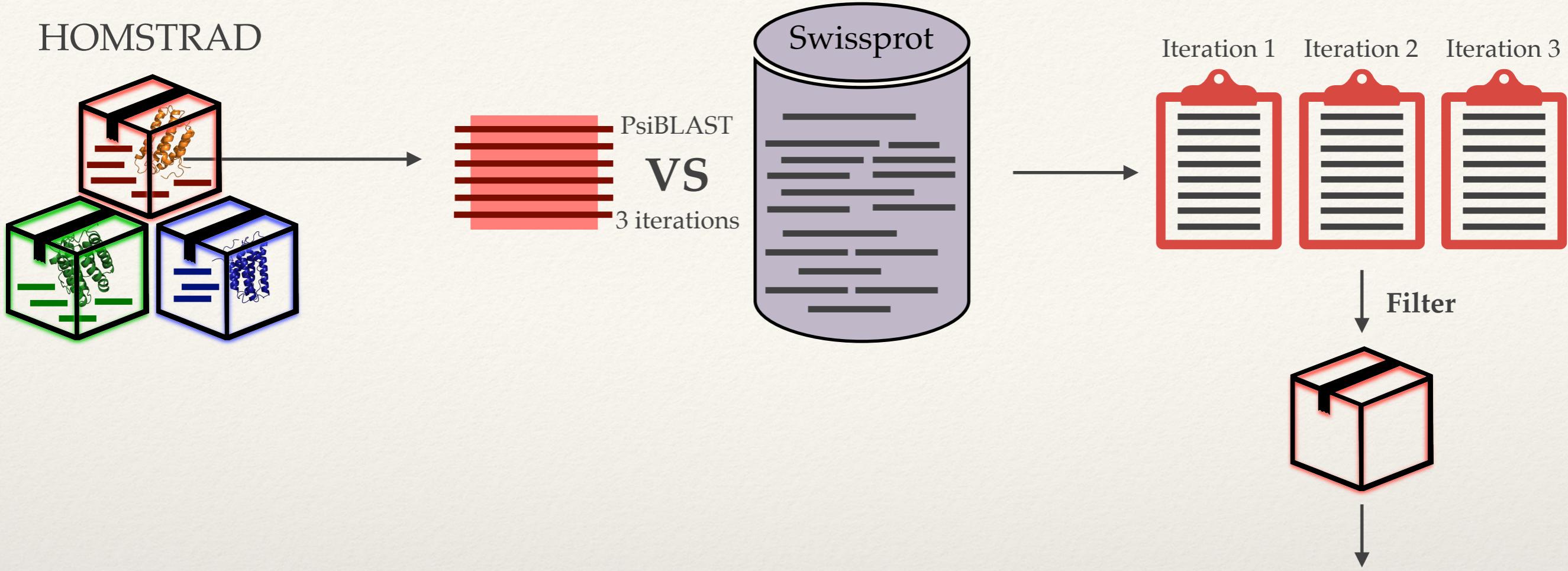
BLASTp Identities distribution



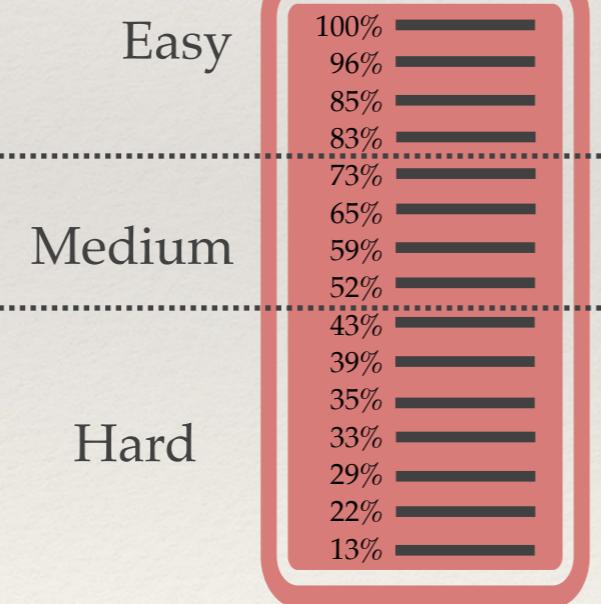
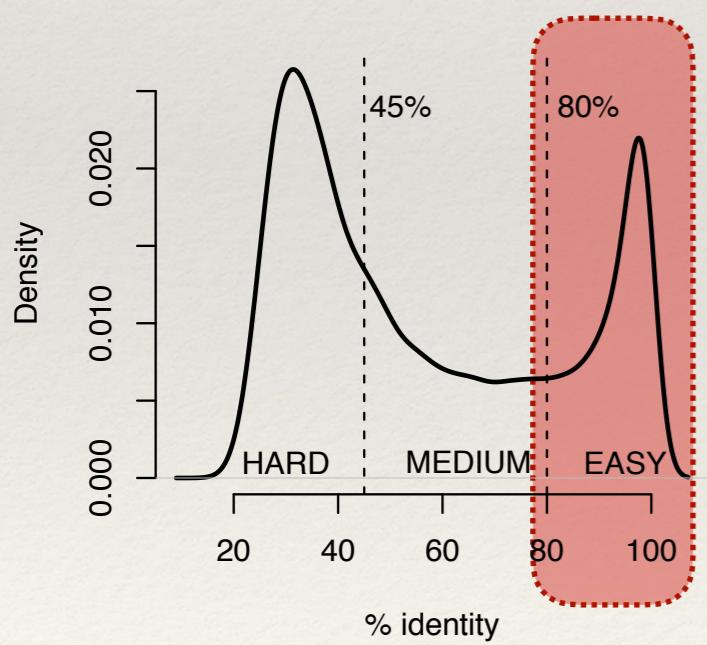
BLASTp
VS



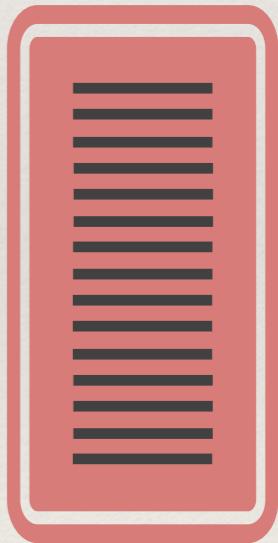
HOMSTRAD



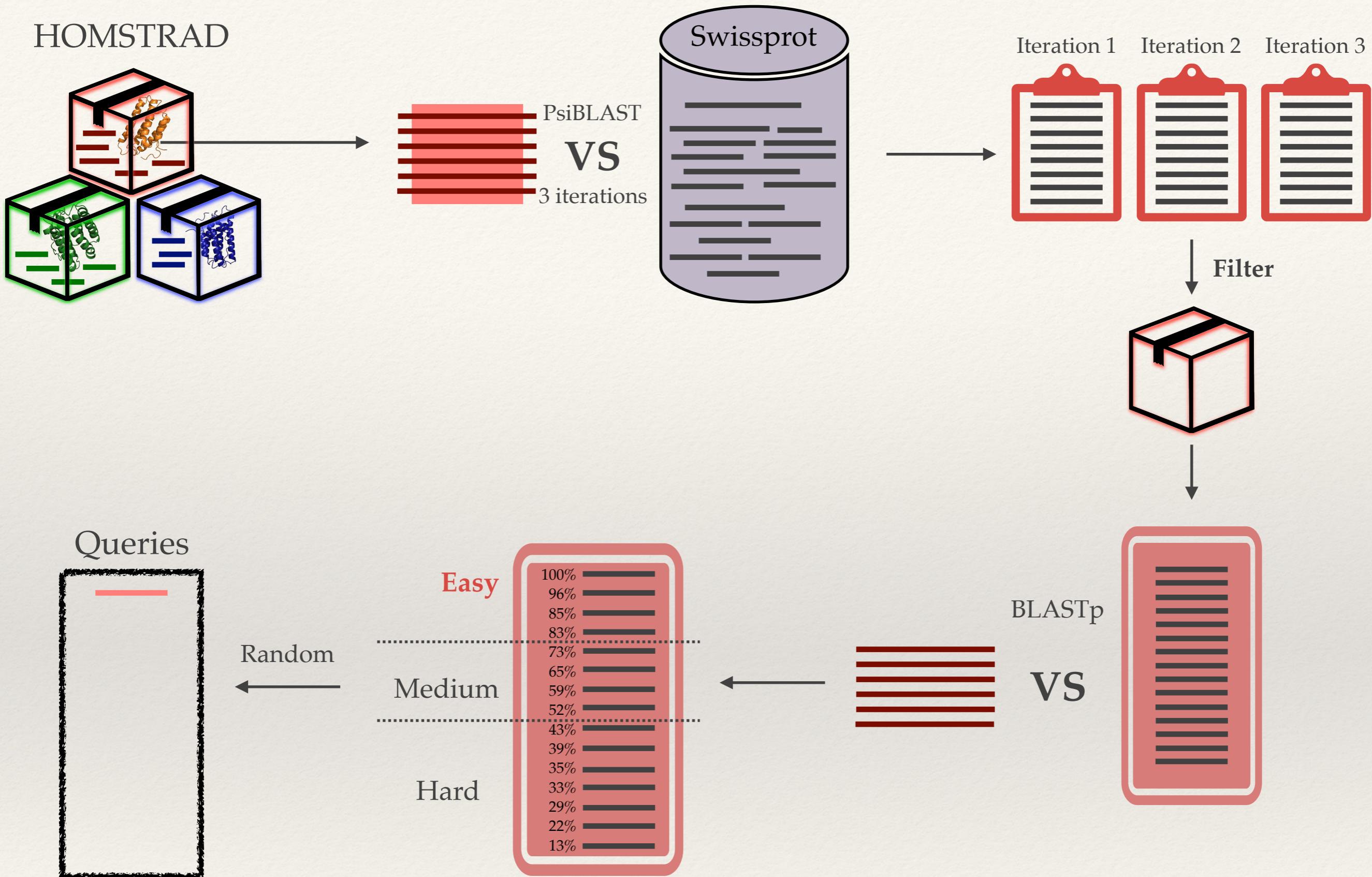
BLASTp Identities distribution



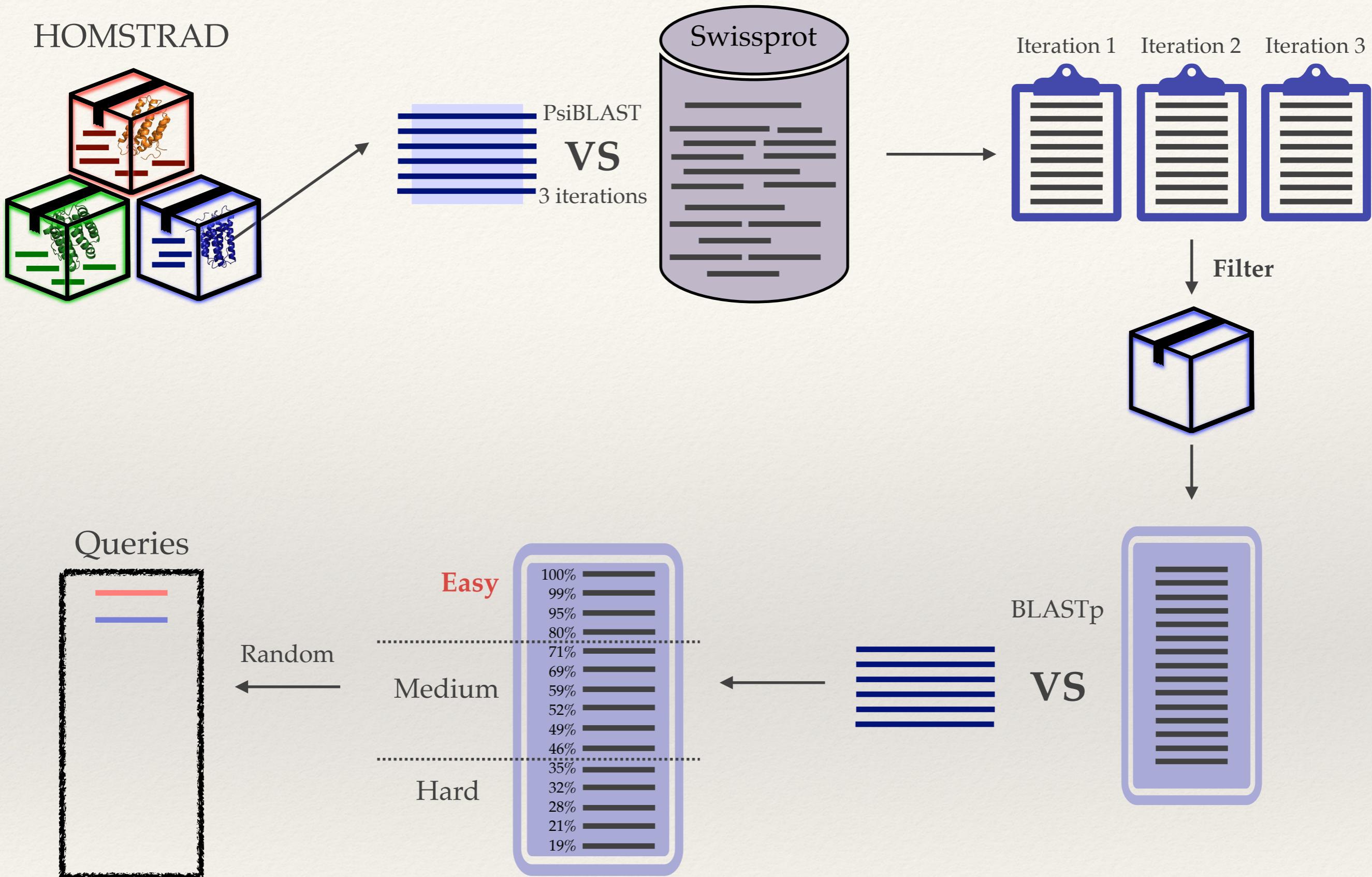
BLASTp
VS



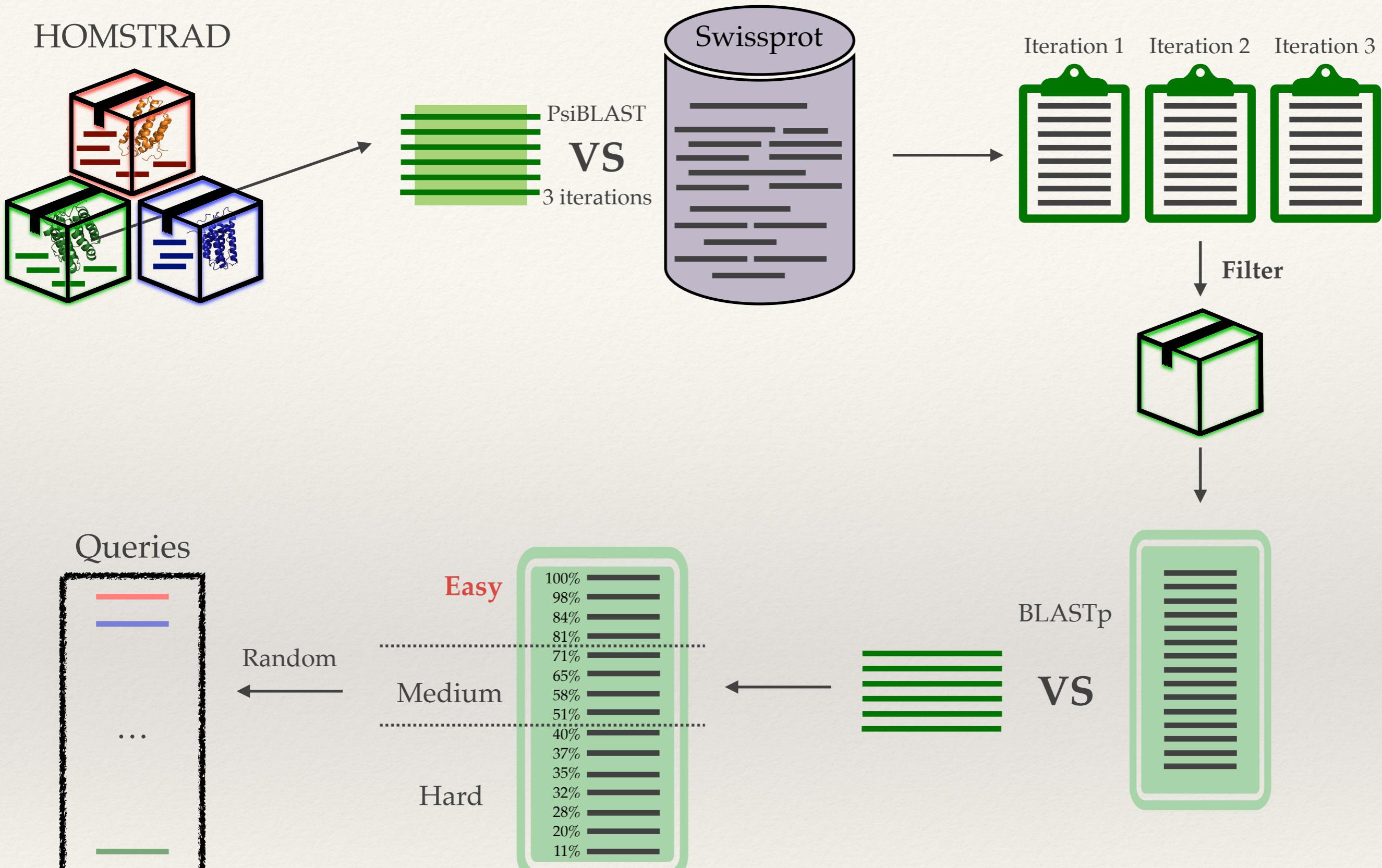
HOMSTRAD



HOMSTRAD



HOMSTRAD



Upstream Input

Queries

398 queries sequences

...

More or less easy
Homologous with the family

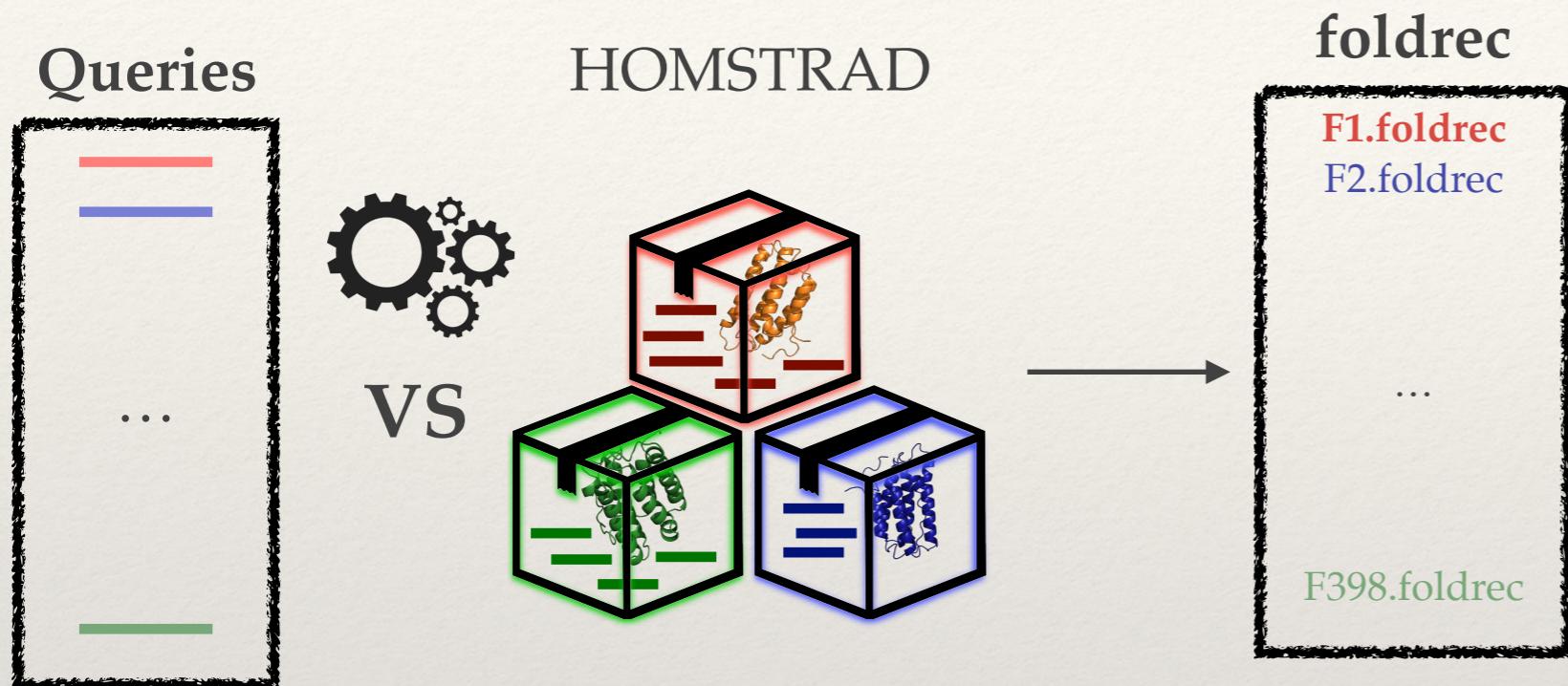
Do not participate in the
HOMSTRAD-family sequences

INPUT

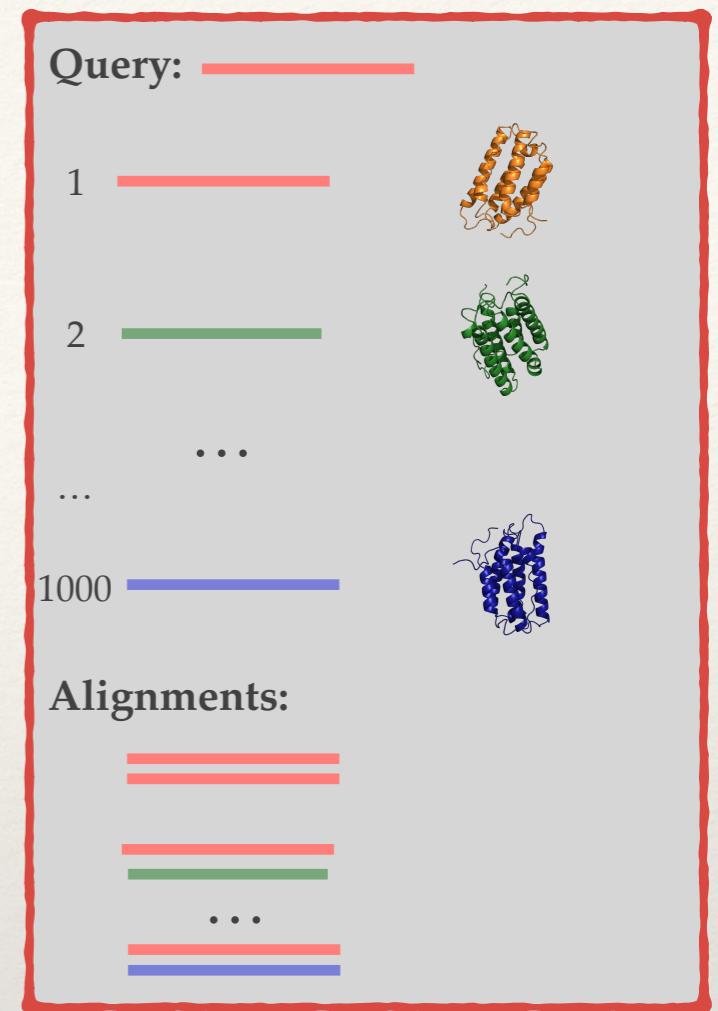
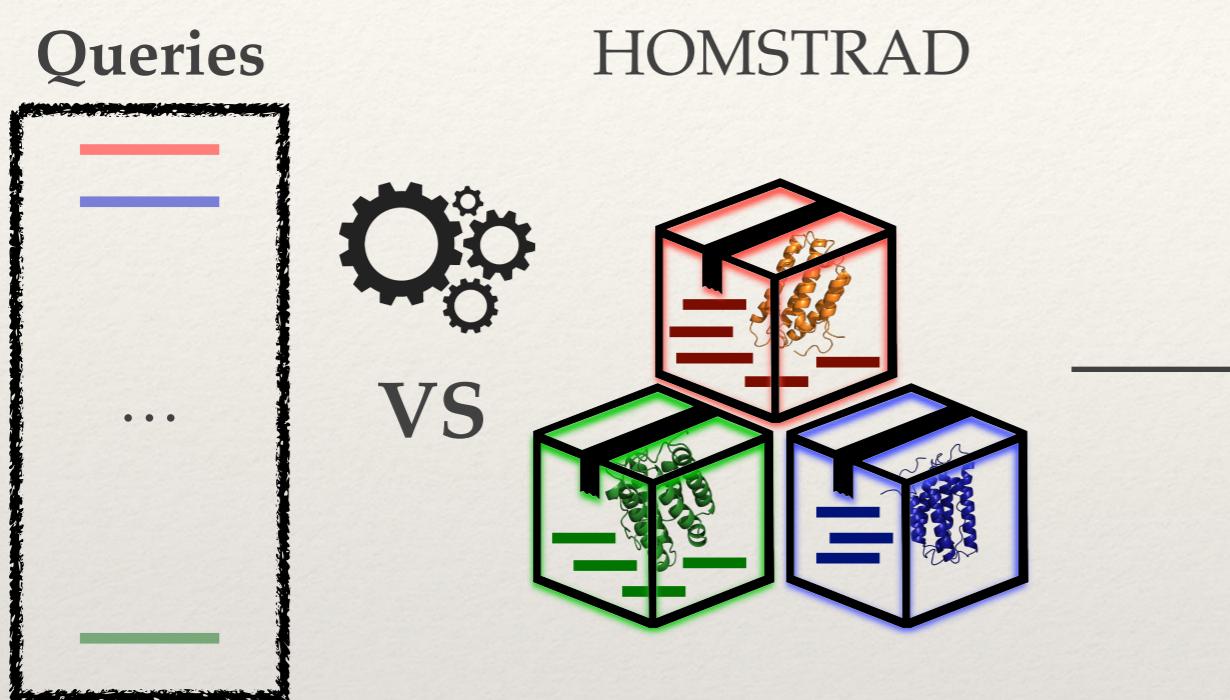
Homstrad Family	UniProt Accession	NCBI Accession	% identity
--------------------	----------------------	-------------------	---------------

```
>mdm2 | P31015.4 | WP_011194573.1 | 24.24%
MPKGEPKIKMVEPIRLIPREDREAAIKAAYNPFLRLRSSDVYIDLTDGTGAMSQFQWSAMML
GDESYAGASSYYRLKEAVTDITGYEYVLPTHQGRGAEKSAFAQLITRPGMYVLSNMFFDTTRGHV
QLAGGRPIDLLLDPVTEEYHPFKGNMDTARLEAFIQEHGAENIACIVMTVTNNSSAGGQPVS MANI
RETSRIARKYGILLLFDVARYAENCHFIRMREEGYADKAPIDIAREMF SYGDGLMMMSAKKD ALVN
IGGLLAFKDEELYTRVGGTVVPFEGFLTYGGLAGRDLEAMAVGLREALDPDYLAYRVGQVEY LGN
LLRSAGIPIQWPVGHHAVFIDA AKFLPHIPWDQFP GHALTVALYQEGGVRTVEVGSLVMGRDPET
GENVRSPFETRLAIPRRVYTNLHEDVAETVINA FQKREQIRGVKF TREPKVLRHTAHFDLV
>TNFR_c6 | P68636.1 | NP_570600.1 | 98.56%
MKSVLYSYILFLSCI INGRDVTPYAPSNGKCDNEYKRHNLCCLSCPPGTYASRLCDSKTNTQC
TPCGSGTFTSRNNHLPACLSNCNRCDSNQVETRSCNTTHNRICECSPGYYCILKGSSGCKACVSQ
TKCGIGYGVSGHTSAGDVICSPCGL GTYSRTVSSADKCEPVPNSNTNYIDVEINLYPVNDTSCTR
TTTGISESISTSELTITMNHKDCDPVFREEYFSVLNKVATSGFFTGENRYQNISKVCTLNFEIK
CNNKGSSSKQLTAKNNDGIMPHSETVTLVGCLSSVDIYI LYNSNTNTQDYETDTISYHAGNVL
VD SHMPGSCDIHKLITNSKPTHFL
>FN1 | Q5R8J0.1 | NP_001126112.1 | 98.0%
MNAMKRLCCVLLCGAVFALPSQEIHARVRRGARS YQVICRDEKTQMIYQQHQSWLRPVLRSNR
VEYCWCNSGRAQCHSVPRCSEPRCFNGGT CQQALYFSDFVCQCPEGFAGKCCEIDTRATCYED
QGISYRGTwSTAESGAECTNWNS SALAQKP YSGRRPDAIRLGLGNHNYCRNPDRDSKPWCYVFKA
GKYSSEFCSTPACSEGNSDCYFGNGLAYRGTHSLTESGASCLLWNSMILIGKVYTAQNPNAQALG
LGKHNYCRNP DGDAPWCHVLKNRRLTWEYCDVPSCTCGLRQYSQPQFRIKGGFADIA SHPWQ
AAIFARHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHTVILGRTYRVVPGE EEQKFEVE
KYIVHKEFDDDTYDNDI ALLQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTCELSGYKHEAL
SPFYSERLKEAHVRLYPSSRCTSQHLLNRTVADNMLCAGDTRSGGPQANLHDACQGDGGPLVCL
NDGRMTLVGII SWGLCGEKDVPGVYTKVNYLDWIHDNMRP
>Me-amine-deh_L | P22641.2 | WP_036750427.1 | 98.4%
MLGNFRFDDMVEKLSRRVAGRTSRRGAIGRLGT VLAGAALVPLLVD RGRVSRANAAGPAEGVD
PRAK WQPQDN DIQACDYWRHCSIDGNICDCSGGS LTNCPPGTKL ATASWVASCYNPTDGQS YLIA
YRDCCGYNVSGRCPCLNTEGELPVYRPEFANDIIWCFAEDDAMTYHCTISPIVGKAS
>DISIN | Q10743.2 | NP_062127.1 | 43.42%
MVLPTVLILL SWAAGLGGQYGNPLNKYIRHYEGLSYNVDSLHQKHQR ARAVSHEDQFL LDFH
AHGRQFNLRMKRDTSLFSDEFKVETSNKVLDYDTSHIYTGHYGEEGFSFSGHSVVDGRFEGFIQT
RG GTFYIEPAERYIKDRILPFHSVIYHEDDINY PHKYGPQGGCADHSVFERMRKYQMTGVEEGTR
```

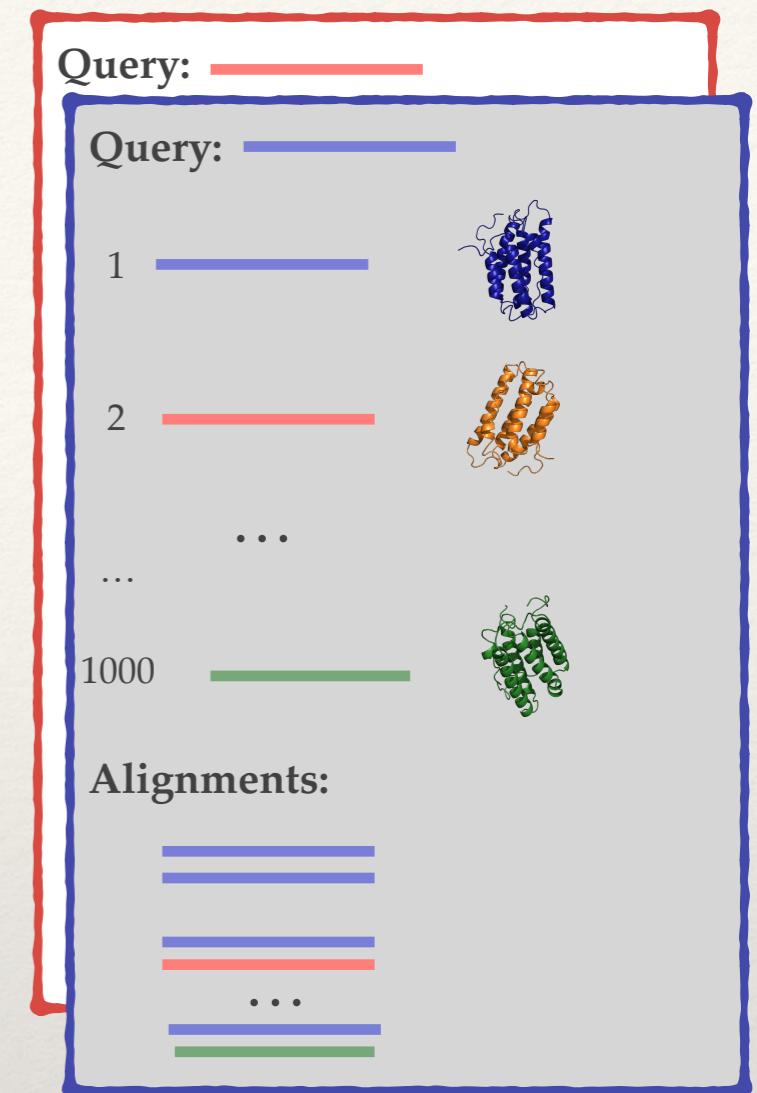
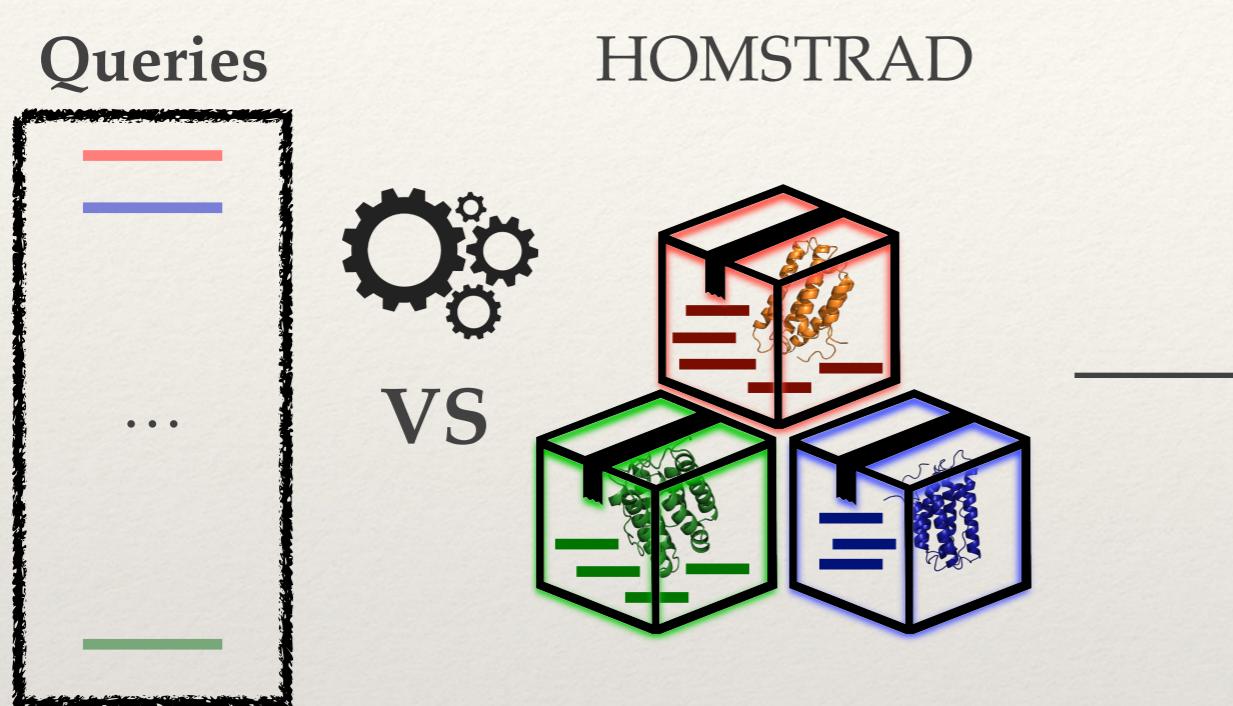
Upstream Output - Downstream Input



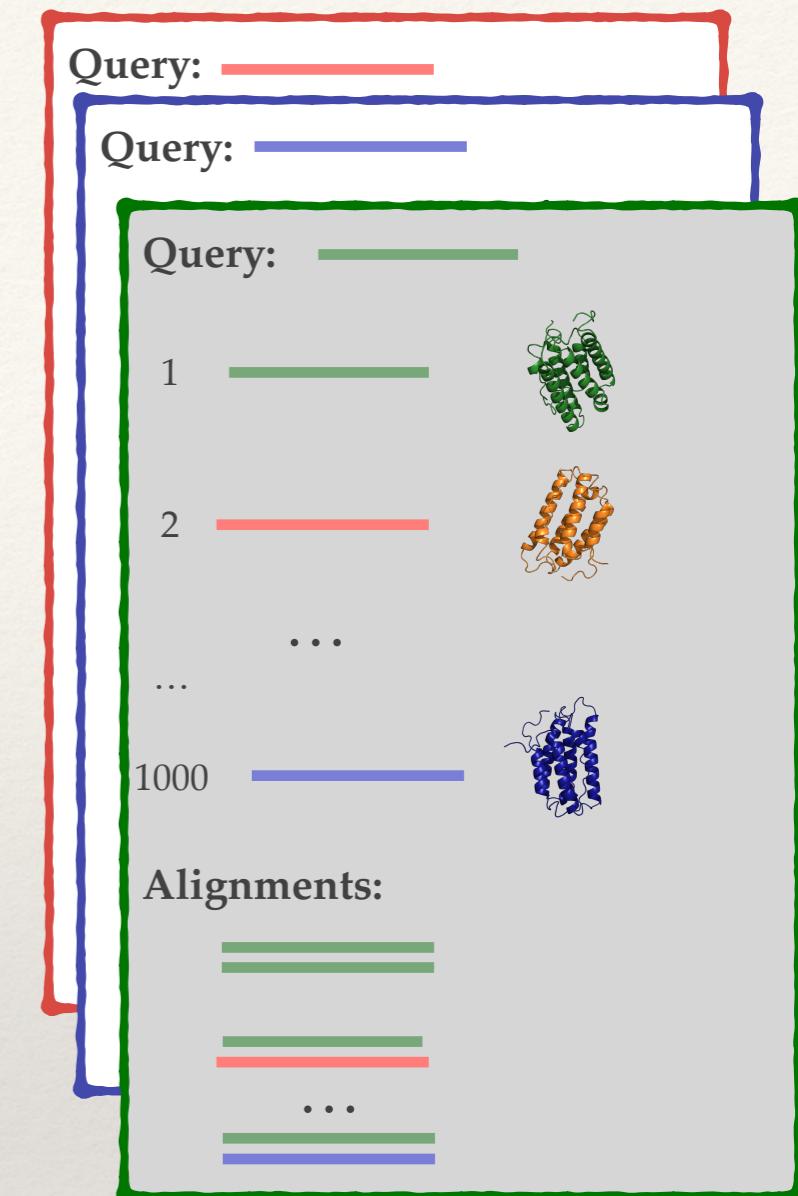
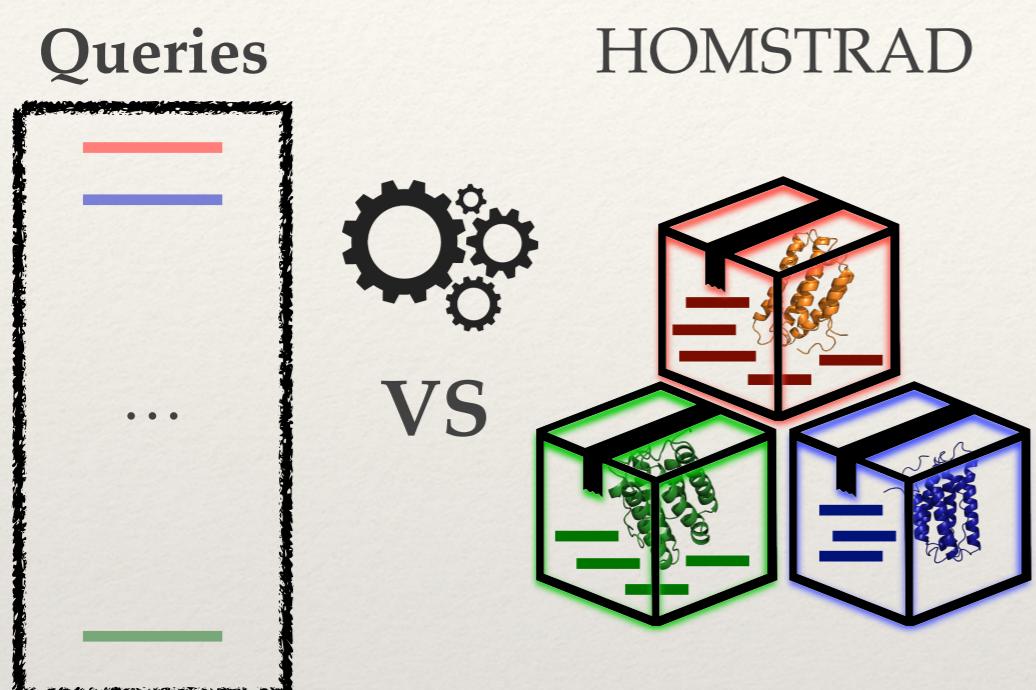
Upstream Output - Downstream Input



Upstream Output - Downstream Input

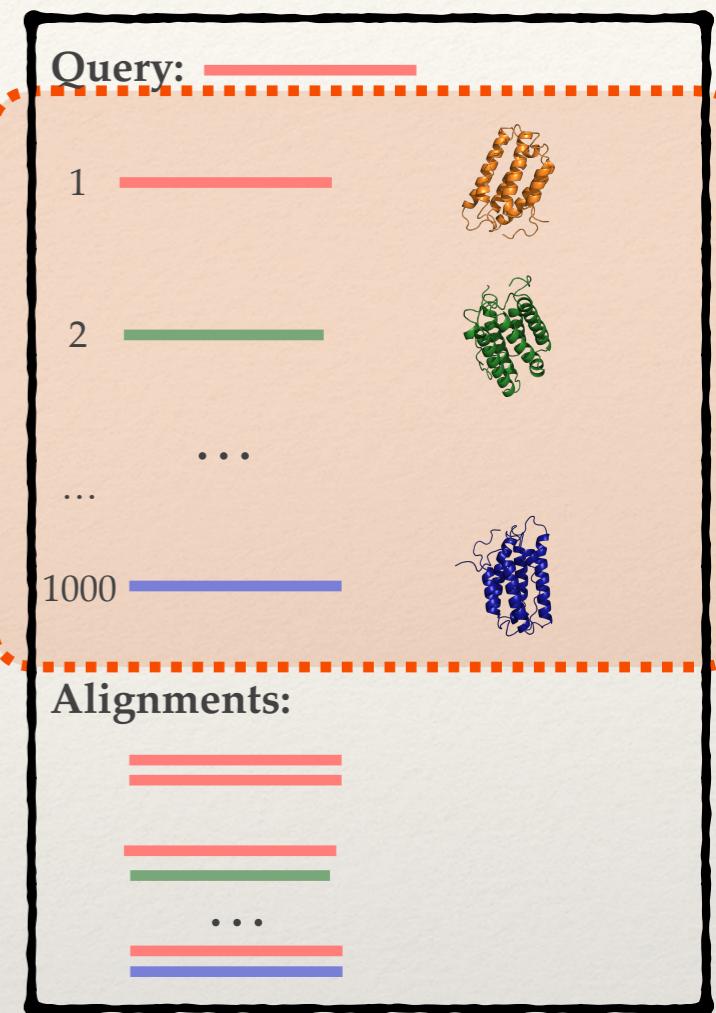


Upstream Output - Downstream Input



Upstream Output - Downstream Input

*** HITS RANKED ***													
	SEQUENCE QUERY FILE : aakinase P37306.2 WP_000855379.1 98.71, 297 aa.												
#	Score	Ungaped_score	Pvalue_Q	Pscore	PQTscore	P-Value_T	Q. Length	T. Length	Q. begin-end	T. begin-end	HITS		
1	567.713	206.331	1.72E-03	12.320	0.998	0.488	0.00E+00	297	307	1-237	1-307	aakinase :	c.73.1.1
2	512.939	246.690	4.40E-05	12.940	1.000	0.409	0.00E+00	297	467	1-297	1-467	ADP_PFK_GK :	c.72.1.3
3	483.730	282.551	3.22E-04	12.631	1.000	0.458	2.26E-202	297	555	1-297	1-555	PGI :	c.80.1.2
4	459.763	242.312	1.28E-03	12.379	0.999	0.658	0.00E+00	297	276	1-237	1-276	MCR_alpha :	a.89.1.1
5	452.120	208.195	1.91E-03	12.298	0.998	0.366	4.83E-284	297	472	1-232	1-472	ghf15 :	a.102.1.1
6	450.328	238.503	2.09E-03	12.280	0.998	0.661	0.00E+00	297	261	1-296	1-261	COX3 :	f.25.1.1
7	443.333	262.667	2.95E-03	12.206	0.997	0.627	0.00E+00	297	380	35-297	1-380	actin :	c.55.1.1
8	434.427	255.442	4.46E-03	12.113	0.996	0.698	0.00E+00	297	295	23-297	1-295	SRP54 :	a.24.13.1
9	434.344	108.558	4.48E-03	12.112	0.996	0.153	9.24E-62	297	854	1-237	1-854	prc :	b.41.1.1
10	434.020	250.129	4.55E-03	12.109	0.995	0.485	3.08E-289	297	456	1-297	1-456	lyase_1 :	a.127.1.1
11	433.437	182.914	4.67E-03	12.103	0.995	0.289	5.75E-65	297	510	1-237	1-510	ATP-sulfurylase :	b.122.1.3
12	430.231	219.194	5.38E-03	12.069	0.995	0.521	0.00E+00	297	343	1-237	1-343	peroxidase :	a.93.1.1
13	420.599	209.144	8.09E-03	11.969	0.992	0.332	2.85E-174	297	512	1-297	1-512	chorismate_bind :	d.161.1.1
14	419.007	240.283	8.64E-03	11.952	0.991	0.558	0.00E+00	297	331	1-297	1-331	S_T_dehydratase :	c.79.1.1
15	417.379	204.316	9.23E-03	11.935	0.991	0.607	0.00E+00	297	236	1-237	1-236	Isochorismatase :	c.33.1.3
16	416.743	180.297	9.46E-03	11.928	0.991	0.428	0.00E+00	297	357	1-237	1-357	BChl_A :	b.75.1.1
17	415.287	240.253	1.00E-02	11.913	0.990	0.446	7.96E-136	297	491	21-297	1-491	PK :	c.1.12.1



Upstream Output - Downstream Input

*** HITS RANKED ***													
SEQUENCE QUERY FILE : aakinase P37306.2 WP_000855379.1 98.71, 297 aa.													
#	Score	Ungaped_score	Pvalue_Q	Pscore	PQTscore	P-Value_T	Q. Length	T. Length	Q. begin-end	T. begin-end	HITS		
1	567.713	206.331	1.72E-03	12.320	0.998	0.488	0.00E+00	297	307	1-237	1-307	aakinase :	c.73.1.1
2	512.939	246.690	4.40E-05	12.940	1.000	0.409	0.00E+00	297	467	1-297	1-467	ADP_PFK_GK :	c.72.1.3
3	483.730	282.551	3.22E-04	12.631	1.000	0.458	2.26E-202	297	555	1-297	1-555	PGI :	c.80.1.2
4	459.763	242.312	1.28E-03	12.379	0.999	0.658	0.00E+00	297	276	1-237	1-276	MCR_alpha :	a.89.1.1
5	452.120	208.195	1.91E-03	12.298	0.998	0.366	4.83E-284	297	472	1-232	1-472	ghf15 :	a.102.1.1
6	450.328	238.503	2.09E-03	12.280	0.998	0.661	0.00E+00	297	261	1-296	1-261	COX3 :	f.25.1.1
7	443.333	262.667	2.95E-03	12.206	0.997	0.627	0.00E+00	297	380	35-297	1-380	actin :	c.55.1.1
8	434.427	255.442	4.46E-03	12.113	0.996	0.698	0.00E+00	297	295	23-297	1-295	SRP54 :	a.24.13.1
9	434.344	108.558	4.48E-03	12.112	0.996	0.153	9.24E-62	297	854	1-237	1-854	prc :	b.41.1.1
10	434.020	250.129	4.55E-03	12.109	0.995	0.485	3.08E-289	297	456	1-297	1-456	lyase_1 :	a.127.1.1
11	433.437	182.914	4.67E-03	12.103	0.995	0.289	5.75E-65	297	510	1-237	1-510	ATP-sulfurylase :	b.122.1.3
12	430.231	219.194	5.38E-03	12.069	0.995	0.521	0.00E+00	297	343	1-237	1-343	peroxidase :	a.93.1.1
13	420.599	209.144	8.09E-03	11.969	0.992	0.332	2.85E-174	297	512	1-297	1-512	chorismate_bind :	d.161.1.1
14	419.007	240.283	8.64E-03	11.952	0.991	0.558	0.00E+00	297	331	1-297	1-331	S_T_dehydratase :	c.79.1.1
15	417.379	204.316	9.23E-03	11.935	0.991	0.607	0.00E+00	297	236	1-237	1-236	Isochorismatase :	c.33.1.3
16	416.743	180.297	9.46E-03	11.928	0.991	0.428	0.00E+00	297	357	1-237	1-357	BChl_A :	b.75.1.1
17	415.287	240.253	1.00E-02	11.913	0.990	0.446	7.96E-136	297	491	21-297	1-491	PK :	c.1.12.1

Ranking & Scoring

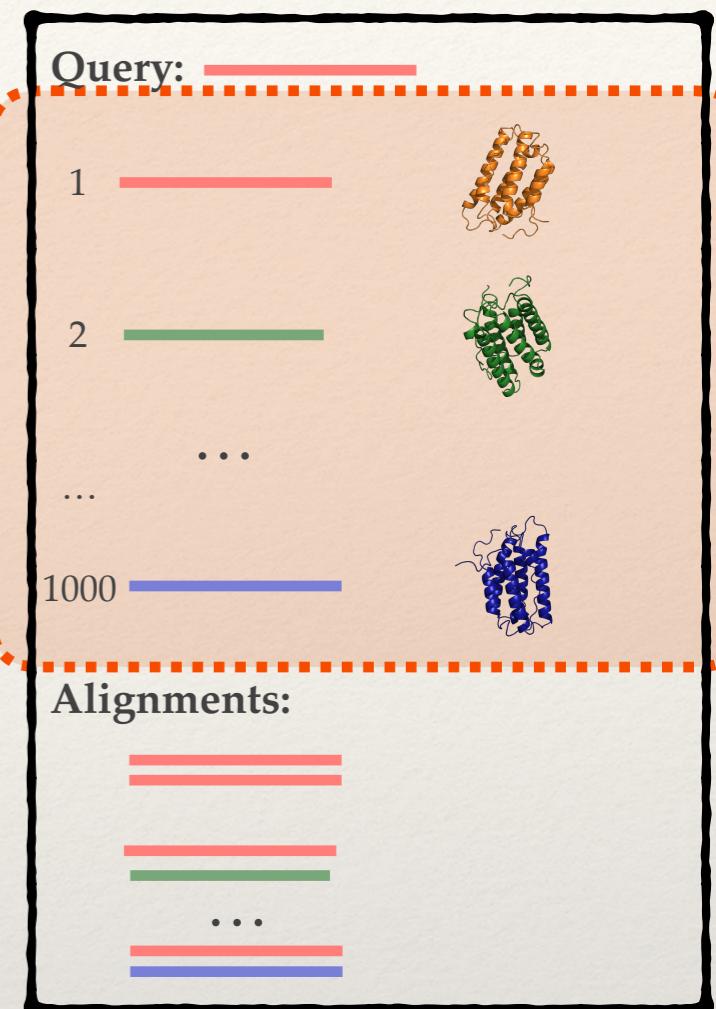
Query & Template
Alignment positions

Template family &
SCOP ID

To respect

Can be filled by XXXXs
or other information

Title



Upstream Output - Downstream Input

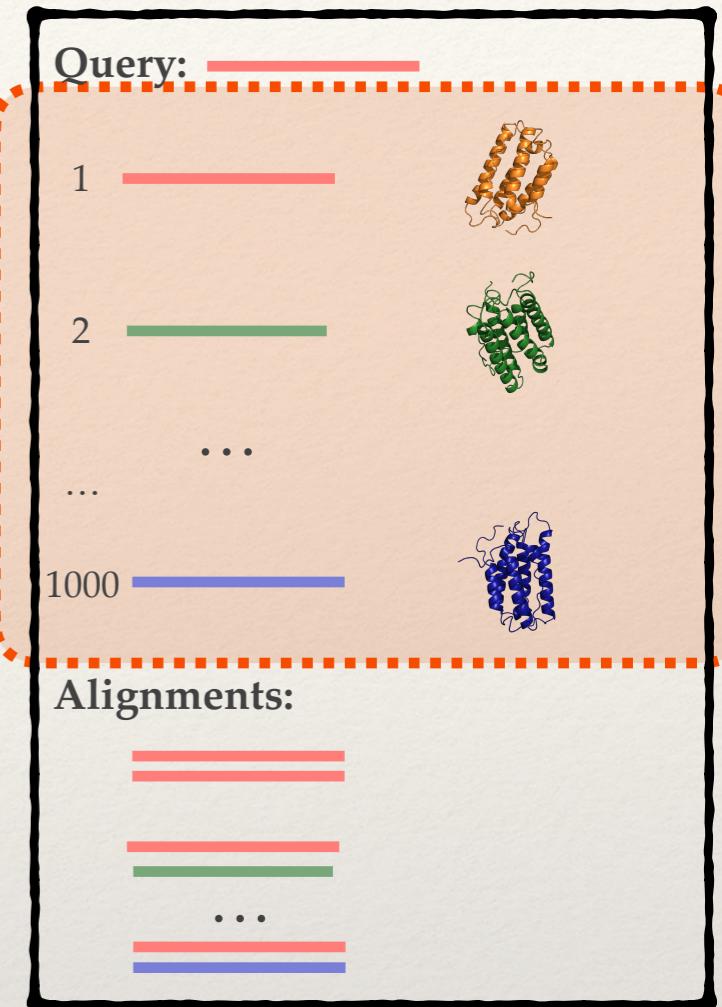
*** HITS RANKED ***											
#	Score	Ungaped_score	Pvalue_Q	Pscore	PQTscore	P-Value_T	Q. Length	T. Length	Q. begin-end	T. begin-end	HITS
1	567.713	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	1-237	1-307	aakinase : c.73.1.1		
2	512.939	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	1-297	1-467	ADP_PFK_GK : c.72.1.3		
3	483.730	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	1-297	1-555	PGI : c.80.1.2		
4	459.763	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	1-237	1-276	MCR_alpha : a.89.1.1		
5	452.120	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	1-232	1-472	ghf15 : a.102.1.1		
6	450.328	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	1-296	1-261	COX3 : f.25.1.1		
7	443.333	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	35-297	1-380	actin : c.55.1.1		
8	434.427	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	23-297	1-295	SRP54 : a.24.13.1		
9	434.344	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	1-237	1-854	prc : b.41.1.1		
10	434.020	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	1-297	1-456	lyase_1 : a.127.1.1		
11	433.437	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	1-237	1-510	ATP-sulfurylase : b.122.1.3		
12	430.231	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	1-237	1-343	peroxidase : a.93.1.1		
13	420.599	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	1-297	1-512	chorismate_bind : d.161.1.1		
14	419.007	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	1-297	1-331	S_T_dehydratase : c.79.1.1		
15	417.379	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	1-237	1-236	Isochorismatase : c.33.1.3		
16	416.743	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	1-237	1-357	BChl_A : b.75.1.1		
17	415.287	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	XXXX XXXX XXXX XXXX XXXX XXXX XXXX XXXX	21-297	1-491	PK : c.1.12.1		

Ranking & Scoring

Query & Template
Alignment positions

Template family &
SCOP ID

Title

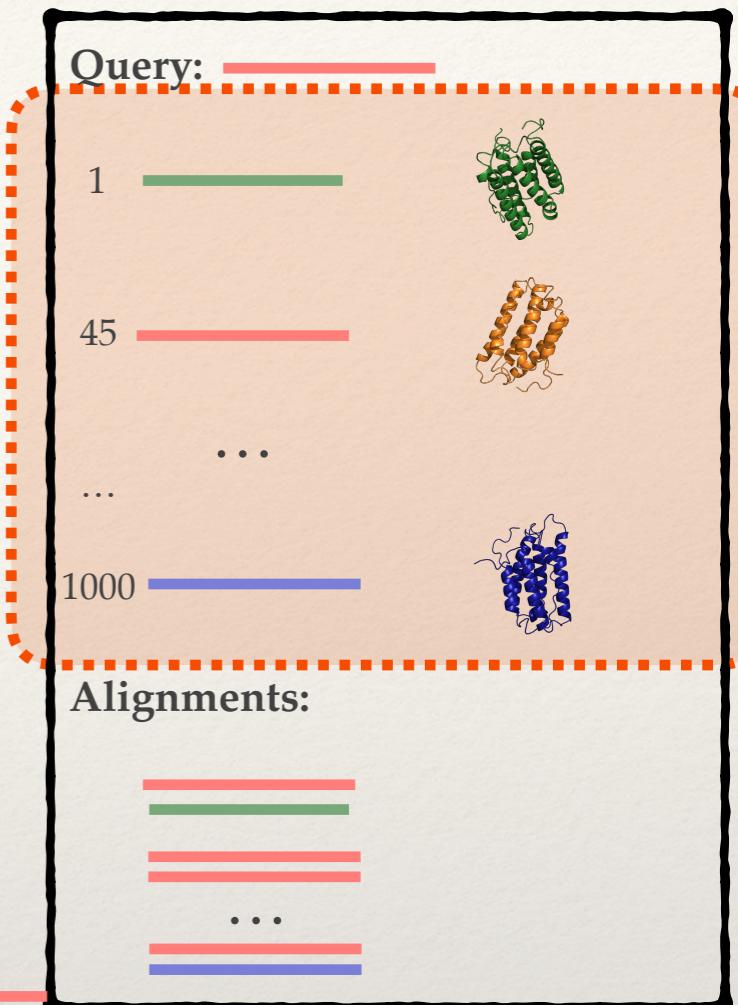


To respect

Can be filled by XXXXs
or other information

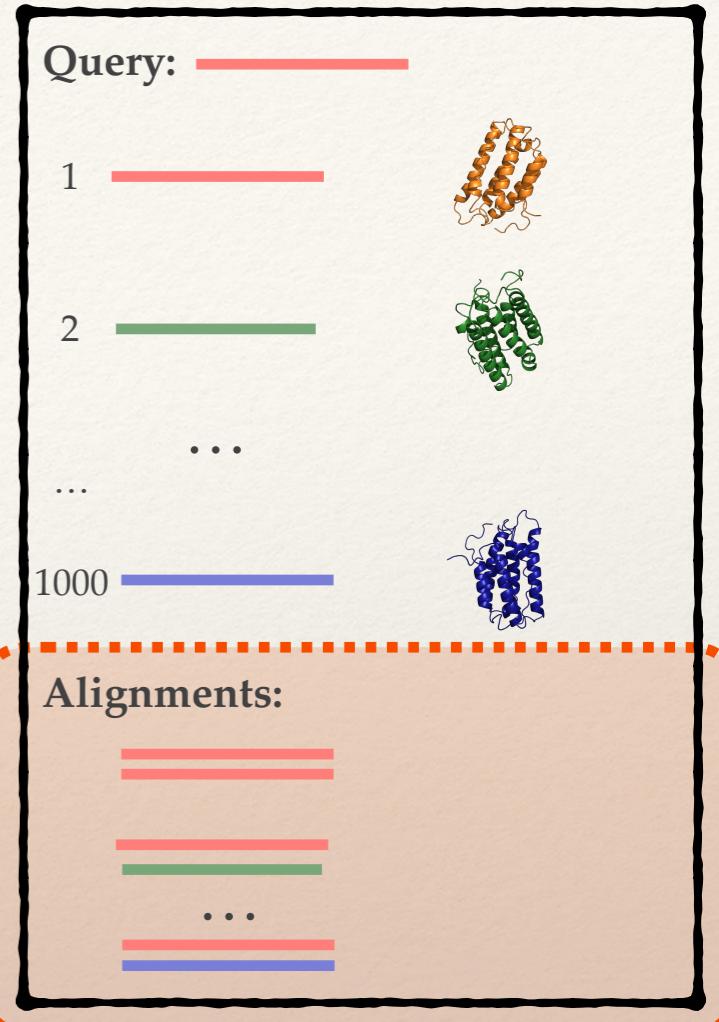
Upstream Output - Downstream Input

*** HITS RANKED ***													
#	Score	Ungaped_score	Pvalue_Q	Pscore	PQTscore	P-Value_T	Q. Length	T. Length	Q. begin-end	T. begin-end	HITS		
1	2678.158	1217.694	3.35E-06	15.148	1.000	1.428 0.00E+00	354	854	1-354	1-854	prc : b.41.1.1		
2	1860.320	914.677	8.49E-04	13.164	0.999	1.207 0.00E+00	354	759	1-354	1-759	Gly_radical : c.7.1.1		
3	1801.664	898.023	1.20E-03	13.026	0.999	1.727 0.00E+00	354	510	1-354	1-510	ATP-sulfurylase : b.122.1.3		
4	1792.538	895.584	1.26E-03	13.005	0.999	1.611 0.00E+00	354	555	1-354	1-555	PAP2 : a.111.1.2		
5	1770.827	895.733	1.43E-03	12.954	0.999	1.466 0.00E+00	354	610	1-354	1-610	CODH : e.26.1.2		
6	1708.751	858.049	2.04E-03	12.808	0.998	1.758 0.00E+00	354	472	1-354	1-472	ghf15 : a.102.1.1		
7	1700.206	876.429	2.14E-03	12.788	0.998	1.885 0.00E+00	354	464	1-354	1-464	Beta_elim_lyase : c.67.1.2		
8	1588.562	810.370	3.94E-03	12.527	0.996	1.773 0.00E+00	354	448	3-354	1-448	lipase : b.12.1.2		
9	1490.716	755.591	6.58E-03	12.300	0.993	0.904 0.00E+00	354	839	1-354	1-839	lipoxygenase : a.119.1.1		
10	1464.427	776.204	7.52E-03	12.239	0.992	1.537 0.00E+00	354	493	1-354	1-493	RNA_dep_RNA_pol : e.8.1.4		
.													
41	1196.052	623.068	2.71E-02	11.619	0.973	1.334 0.00E+00	354	450	1-354	1-450	voltage_CLC : f.20.1.1		
42	1188.331	635.162	2.81E-02	11.602	0.972	1.745 0.00E+00	354	317	3-354	1-317	CoA_trans : c.124.1.2		
43	1187.870	650.356	2.81E-02	11.601	0.972	1.707 0.00E+00	354	361	1-353	1-361	ghf33 : b.68.1.1		
44	1180.533	622.786	2.91E-02	11.584	0.971	1.497 0.00E+00	354	349	1-350	1-349	NTP_transf_2 : a.160.1.1		
45	1174.603	473.832	7.88E-02	11.032	0.921	2.455 0.00E+00	354	187	160-347	1-187	DNA_polymeraseX : d.218.1.2		
46	1161.402	631.830	3.16E-02	11.540	0.968	1.210 0.00E+00	354	512	1-352	1-512	chorismate_bind : d.161.1.1		
47	1154.196	634.080	3.26E-02	11.523	0.967	1.613 0.00E+00	354	340	1-353	1-340	DAO : c.4.1.2		
48	1136.070	580.075	3.53E-02	11.482	0.965	0.954 0.00E+00	354	555	1-354	1-555	PGI : c.80.1.2		
49	1122.448	608.380	3.74E-02	11.451	0.963	1.704 0.00E+00	354	291	1-353	1-291	tyrosinase : MULTI-DOMAINS , a.86.1.2		
50	1114.715	598.665	3.87E-02	11.433	0.961	1.614 0.00E+00	354	307	1-354	1-307	aakinase : c.73.1.1		



The results are not always ideal!!!

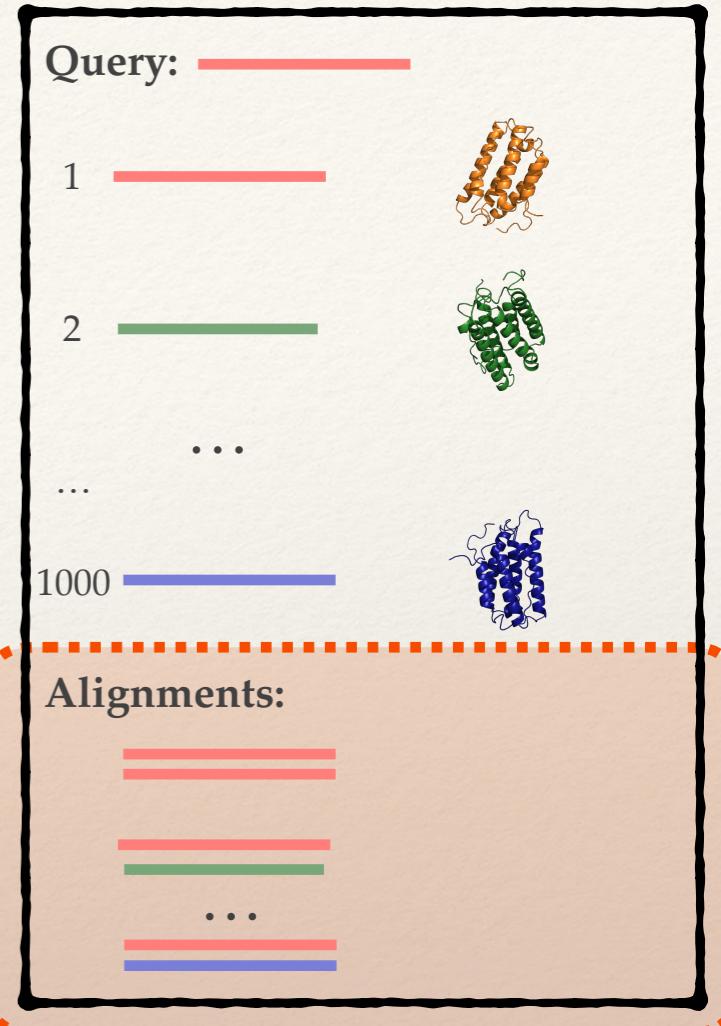
Upstream Output - Downstream Input



Upstream Output - Downstream Input

To respect

Can be filled by XXXXs
or other information



Upstream Output - Downstream Input

→ *** ALIGNMENTS DETAILS ***

No 1
 Alignment : aakinase | P37306.2 | WP_000855379.1 | 98.71, 297 aa. vs aakinase : c.73.1.1
 Score : 567.713 | Normalized score : XXXX | Query coverage : 79.80% | Identity : 30.00% | Gaps : 71.70% | SS Score : XX

Query 1 MKT---LVVALGGNAL---LQRGEALTAENQYRNIASAVPALAR-LARSYRLAIVHGNGPQVGLL-----ALQNL
 Query 1 XXXXXXXXX
 Query 1 XXXXXXXXX

Template 1 GKKMVVAL-----GNAILSNDASAHAQQQALVQTSAVLVHLIKQGHRLIVS-----HGNGPQVGNNLLQQQAADSEKNPAMPDLTCVAMTQGSIGYWLSNALNQELNKAGI
 Template 1 XXXXXXXXX
 Template 1 XXXXXXXXX

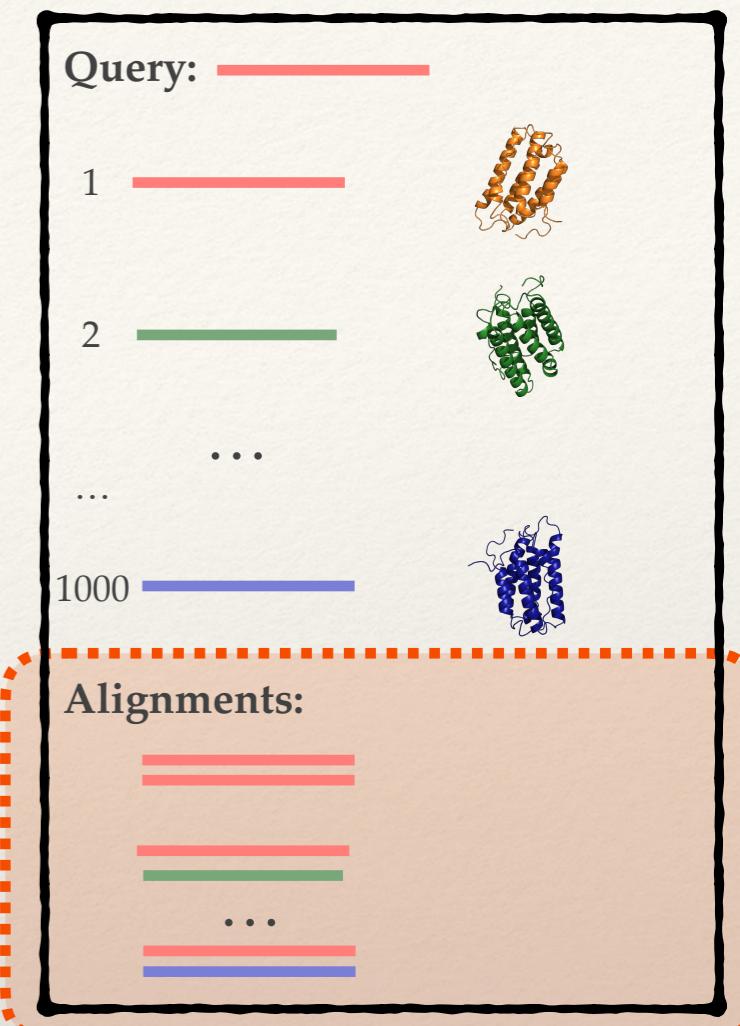
XXXXX XXXXX

No 2
 Alignment : aakinase | P37306.2 | WP_000855379.1 | 98.71, 297 aa. vs ADP_PFK_GK : c.72.1.3
 Score : 512.939 | Normalized score : XXXX | Query coverage : 100.00% | Identity : 10.62% | Gaps : 73.51% | SS Score : XX

Query 1 MKTLVVALGGNALLQRGEALTAENQYRNIASAVPALARLARSYRLAIVHGNGPQVGLLALQNLAWKEVEPYPLDVLVAESQGMIGYMLAQSLAQPMPPVTTLTRIEVSPDDPAFLQPE
 Query 1 XXXXXXXXX
 Query 1 XXXXXXXXX

Template 1 MKE-----SLKDRIRLWKRLYVNAFENALNAIPNVKGULL-----AYNTNID-----
 Template 1 XXXXXXXXX
 Template 1 XXXXXXXXX

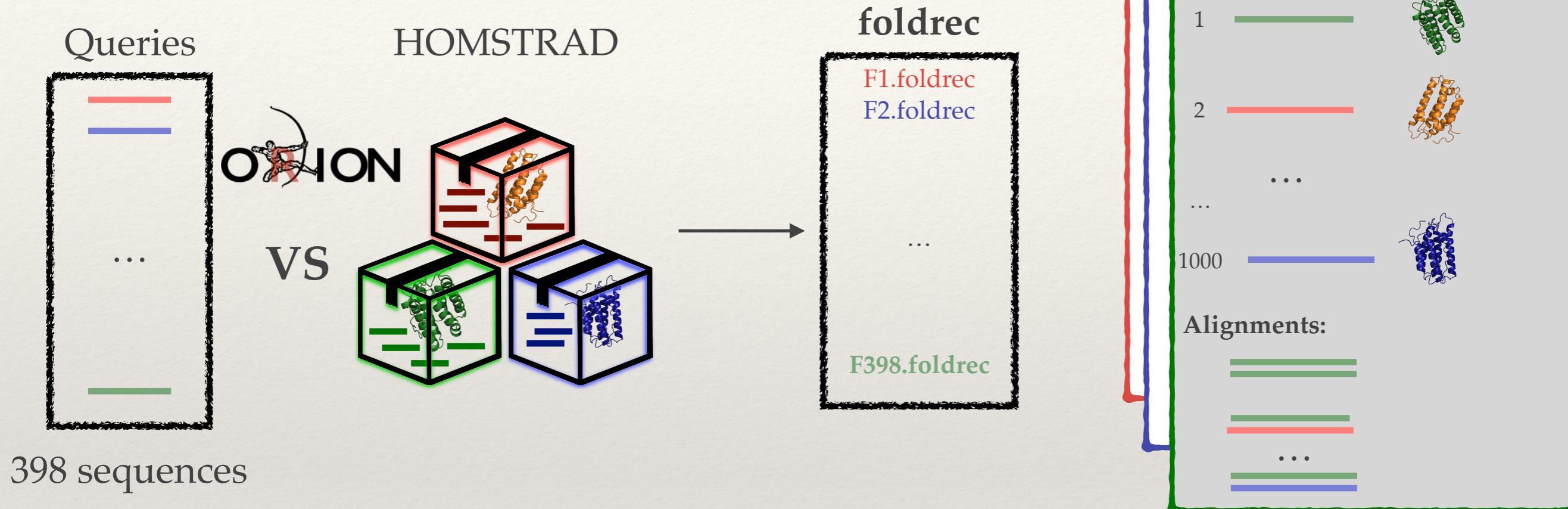
XXXXX XXXXX



To respect

Can be filled by XXXXs or other information

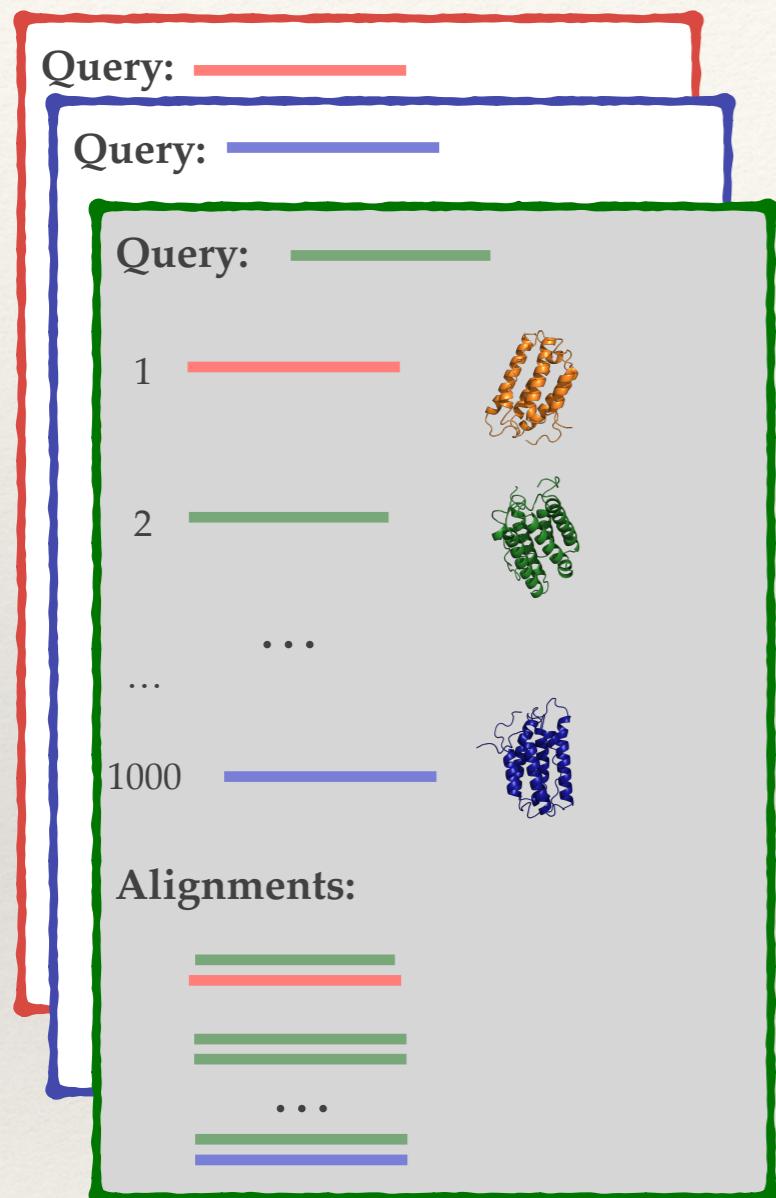
Downstream Input



Ghouzam Y, Postic G, de Brevern AG, Gelly JC. Improving protein fold recognition with hybrid profiles combining sequence and structure evolution. Bioinformatics. 2015 Aug 7

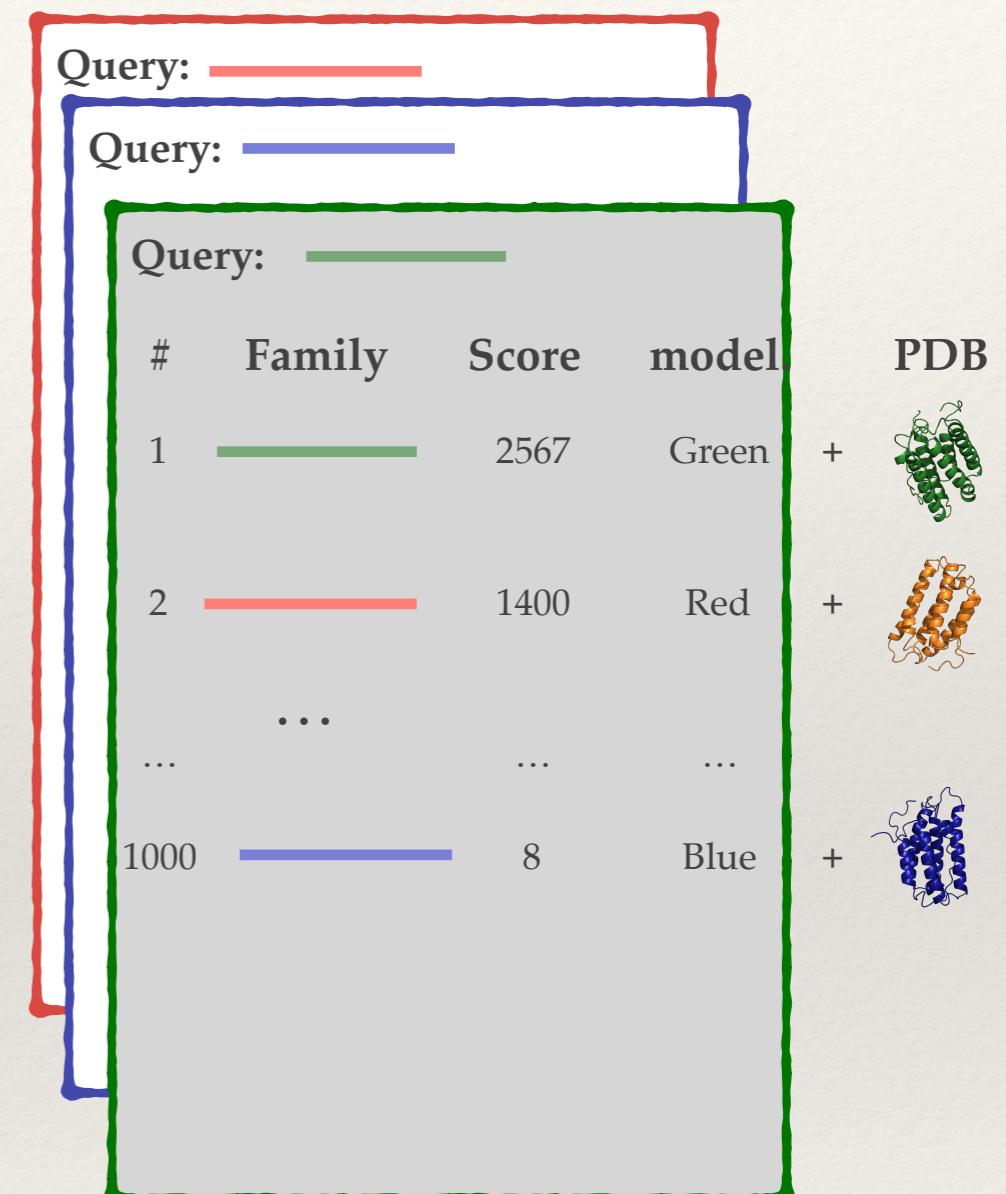
Downstream Input - Output

INPUT



Foldrec file

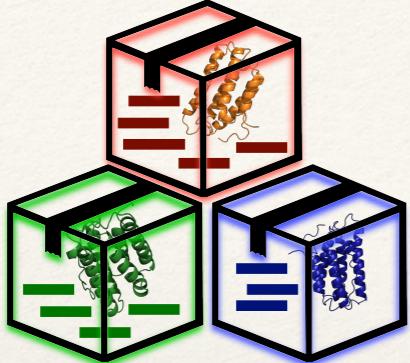
OUTPUT



List rescored

Data Resume

All



HOMSTRAD database (1010 protein families):

- 1 PDB structure (Family Master)
- 1 Fasta sequence of the Master
- Scop ID of the family
- 1 Multiple Sequences Alignment file

Upstream

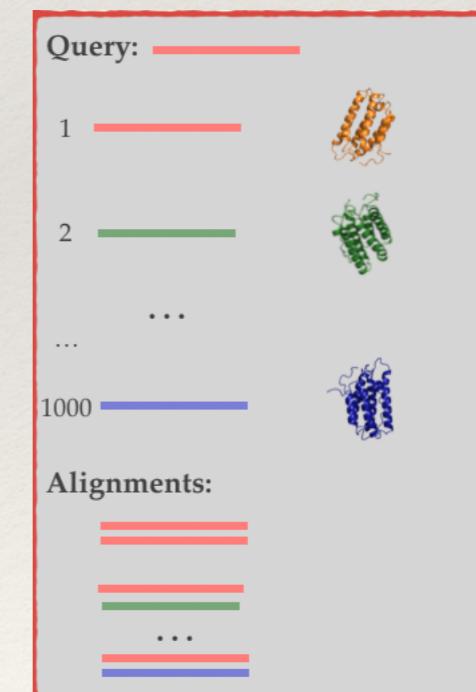
398 queries in Fasta format

queries398.multifasta

```
>mdm2 | P31015.4 | WP_011194573.1 | 24.24%
MPKGEFKIKMVEPIRLIPREDREAAIKAHHYNPFLLRSDDVYIDLLTDSTGAMSQFQWSAMML
GDESYAGASSYYRLEKEAVTDITGYEYVLPTHQGRGAESAFQLITRPGMYVLSNMFFTRGHV
QLAGGRPIDLLLDPTEEYHPFKGNMDTARLEAFIQEHGAEINACIVMTVTNSAGGQPVSMANI
RETSRIARKYGILLLFDVARYAENCHFIRMEEGYADKAPIDIAREMFSYGDGLMMSAKKDALVN
IGLLAFLKDEEELYTRVGGTVVPFEGFLTYGGLAGRDLEAMAVGLREALDPDYLAYRVQVEYLN
LLRSAGIPIQWPVPGHHAVIDAKFLPHIPWDQFPFHALTVALYQEGGVRTVEVGSLVMGRDPET
GENVRSPFETRLAIPRRVYTNLHLEDVAETVINAFQREQIRGVKFTREPKVLRHFTAHFDLV
>TNFR_c6 | P68636.1 | NP_570600.1 | 98.56%
MKSVLYSYIILFLSCIIINGRDTVPYAPSNGKKCDNEYKRHNLCCLSCPPGTYASRLCDSKTNQC
TPCGSGTFTSRNNHLPACLSCNRCDSNQVETRSCNTTHNRICECSPGYYCILKGSSGCACVSQ
TKCGIGYGVSGHTSAGDVICSPCGLGTYSRTVSSADCKEPCPNTNFYIDVEINLYPVNDTCTR
TTTGISESISTSELTITMNHKDCDPVFREEYFSVLNKVATSGFFTGENRYQNISKVCTLNFEIK
CNNKGSSSKQLTKAKNDDGIMPHSETVTLVGDCLSSLVDIYILSNTNTQDYETDTISYHAGNVL
VDSHMPGSCDIIHKLITNSKPTHFL
>FN1 | Q5R8J0.1 | NP_001126112.1 | 98.0%
MNAMKRGGLCVLLCGAVFALPSQEIHARVRRGARSQYQVICERDKTOMIYQOHQSWRPVLRSNR
VEYCWCNSGRAQCHSVPRSCSEPRCFNGGTQCQASYFSDFVQCPEGFAGKCEIDTRATCYED
QGISYRTWSTAESGAECTNWNSALAAQPKPSGRRPDAIRLGLGNHNYCRNPDRDSKPWCYVFKA
GKYSSEFCSTPACSEGNSDCYFGNGLAYRGTSHLTESTGASCLLWNMSMILIGKVTAQNPNAQALG
LGKHNCRNPDGDAKPCWHLKNRRLTWYCDVPSCTCGLRQYSQPQFRIGGLFADIASHPWQ
AAIFARHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHTLVILGRTRYRVPGEEEQKFEVE
KYTVHKEFDDDTYDNIDALLQLKSFSRCAQESSVVRTVCLPPADLQLPDWTECELSGYGKHEAL
SPFYSERLKEAHVRLYPSSRCTSQHLLNRTVADNMLCAGDTRSGGPQANLHDACQGDGGPLVCL
NDGRMTLVGIIISWGLGCGEKDVPGVYTKVNTYLDWIHDNMRP
>Me-amine-deh_L | P22641.2 | WP_036750427.1 | 98.4%
MLGNFRDDMVEKLSRRVAGRTSRRGAIGRLGTVLAGAALVPLLVDRRGRVSRNAAGPAEGVD
PRAKWQPQNDIQACDYWRHCSIDGNICDCSGGSLTNCPPGTLATASWVASCYNPTDGQSYLIA
YRDCCGYVNSGRCPCLNTEGELPVYRPEFANDIWCFAEADDAMTYHCTISPIVGKAS
>DISIN | Q10743.2 | NP_062127.1 | 43.42%
MVLPTVLILLLSWAAGLGGQYGNPLNKYIRHYEGLSYNVDSLHQKHORAKRAVSHEDFLLDFH
AHGRQFNLRMKRDTSLFSDEFKVETSNKVLDDYDTSHIYTGHYGEEGSFSHGSVVDGRFEGFIQT
RGGTIFYIEPAERYIKDRILPFHNSVIYHEDDINYPHKYGPQGGCADHSVFERMRKYQMTGVEEGTR
```

Downstream

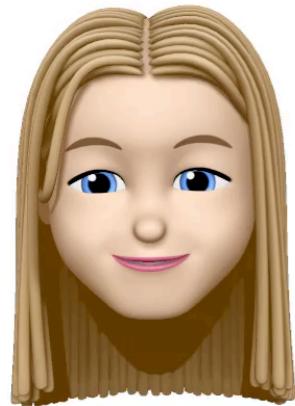
398 rankings in Foldrec format



Do not hesitate to ask us!!!

Meet-U

edition 20²⁰



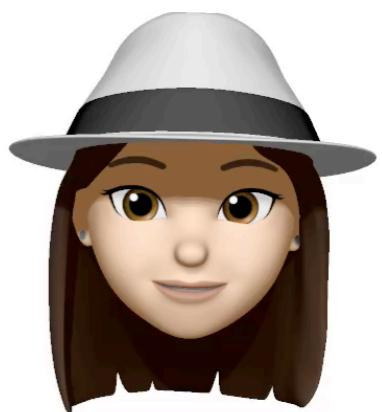
Anne Lopes



Elodie Laine



Chris Papadopoulos



Maureen Muscat



Juliana Bernardes



Yasser Mohseni