

LOSCHMIDT
LABORATORIES



Mnohonásobné sekvenční příložení



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

- ❑ Mnohonásobné sekvenční příložení
- ❑ Konsenzuální sekvence
- ❑ Konstrukce příložení
- ❑ Manuální konstrukce příložení
- ❑ Automatická konstrukce příložení
- ❑ Využití mnohonásobného sekvenčního příložení
- ❑ PSI-BLAST
- ❑ Databáze mnohonásobného sekvenčního příložení

Mnohonásobné sekvenční příložen

Mnohonásobné sekvenční příložen je **2D tabulka**, ve které řádky představují jednotlivé **sekvence** a sloupce představují **pozice** aminokyselinových zbytků.

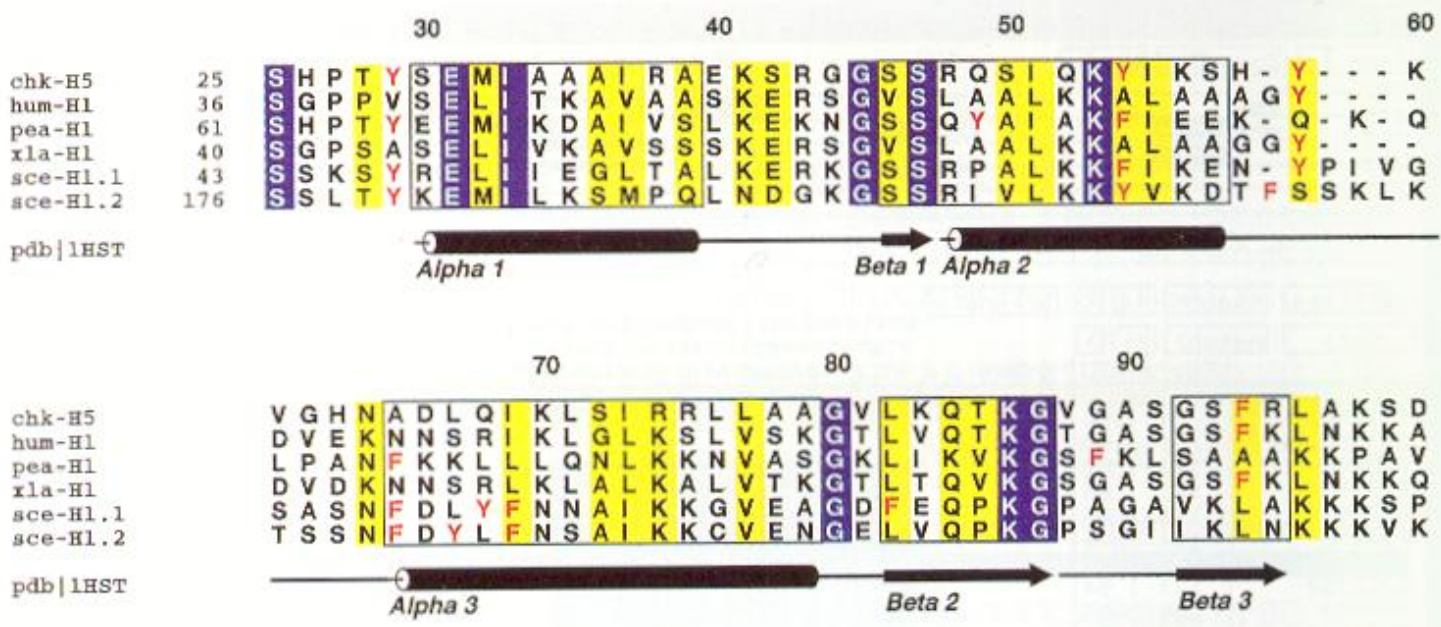
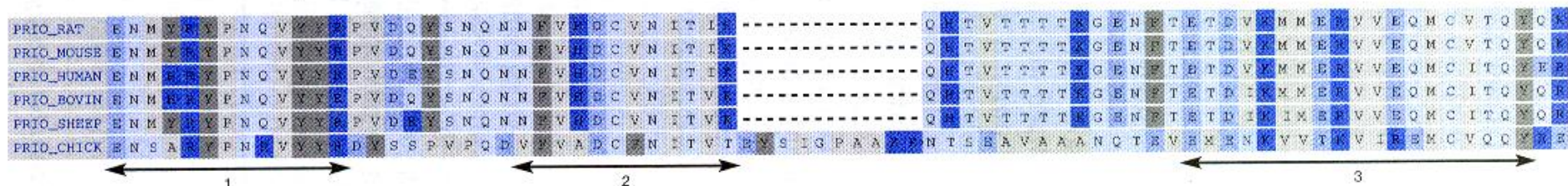


Mnohonásobné sekvenční příložení

Mnohonásobné sekvenční příložení je **2D tabulka**, ve které řádky představují jednotlivé **sekvence** a sloupce představují **pozice** aminokyselinových zbytků.

- ❑ Informativnější než párové příložení
- ❑ Vhodné pro analýzu genových rodin
- ❑ Vhodné pro identifikaci důležitých zbytků
- ❑ Barevně kódované podle vlastností

Mnohonásobné sekvenční přiložení



Konsenzuální sekvence

- **Jediná sekvence** reprezentující mnohonásobné přiložení

	1	2	3	4	5	6	7	8	9	10
I	Y	D	G	G	A	V	–	E	A	L
II	Y	D	G	G	–	–	–	E	A	L
III	F	E	G	G	I	L	V	E	A	L
IV	F	D	–	G	I	L	V	Q	A	V
V	Y	E	G	G	A	V	V	Q	A	L

Konsenzuální sekvence

- **Jediná sekvence** reprezentující mnohonásobné přiložení
 - Ukazuje četnost výskytu aminokyselinových zbytků v každé pozici

	1	2	3	4	5	6	7	8	9	10
I	Y	D	G	G	A	V	–	E	A	L
II	Y	D	G	G	–	–	–	E	A	L
III	F	E	G	G	I	L	V	E	A	L
IV	F	D	–	G	I	L	V	Q	A	V
V	Y	E	G	G	A	V	V	Q	A	L
	y	d	G	G	A/I	V/L	V	e	A	l

Konstrukce příložení

- Hledání korespondence mezi aminokyselinovými zbytky

GVLIQVG

GVLIRQSG

GVPIRQSG

Konstrukce příložení

- Hledání korespondence mezi aminokyselinovými zbytky

GVLI-QVG

| | | | | | |

GVLIRQSG

| | | | | | |

GVPIRQSG

Konstrukce příložení

- Hledání korespondence mezi aminokyselinovými zbytky

GVLI-QVG

||| | |

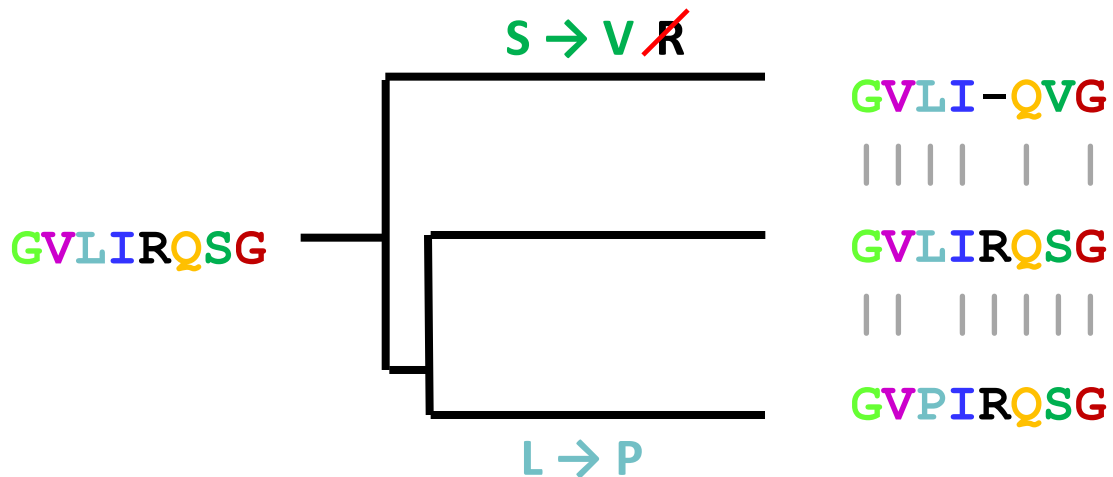
GVLIHQSG

|| ||| |

GVPIHQSG

Konstrukce příložení

- Hledání korespondence mezi aminokyselinovými zbytky



Konstrukce příložení



- ❑ Manuální
- ❑ Automatická
 - ❑ ClustalW
 - Nejznámější a nejpoužívanější
 - ❑ MUSCLE
 - Přesné a velmi rychlé
 - ❑ Toffee
 - Velmi přesné a pomalé

Manuální konstrukce příložen

□ BioEdit

- Editor sekvenčních příložen
- Manuální i automatická konstrukce příložen
- Analýza sekvencí
- Příprava sekvencí



Automatická konstrukce příložení

A _____
B _____
C _____
D _____
E _____

vstupní sekvence



Automatická konstrukce příložení

A _____
B _____
C _____
D _____
E _____

vstupní sekvence

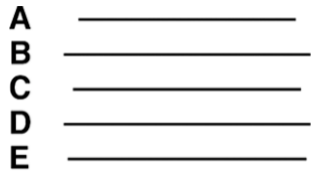


	A	B	C	D	E
A	-				
B	11	-			
C	20	30	-		
D	27	36	9	-	
E	30	33	20	27	-

distanční matice



Automatická konstrukce příložení

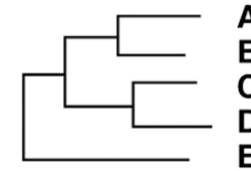


vstupní sekvence



	A	B	C	D	E
A	-				
B	11	-			
C	20	30	-		
D	27	36	9	-	
E	30	33	20	27	-

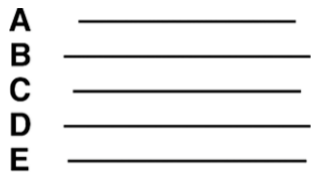
distanční matice



evoluční strom



Automatická konstrukce příložení

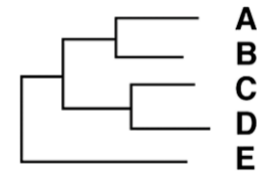


vstupní sekvence

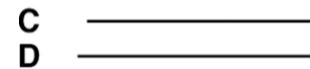


	A	B	C	D	E
A	-				
B	11	-			
C	20	30	-		
D	27	36	9	-	
E	30	33	20	27	-

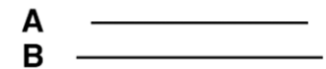
distanční matice



evoluční strom

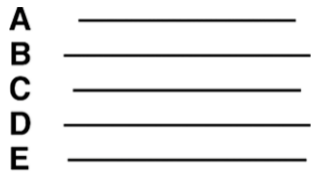


párové příložení
C a D



párové příložení
A a B

Automatická konstrukce příložení

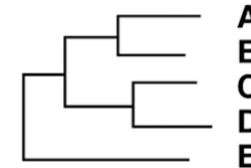


vstupní sekvence

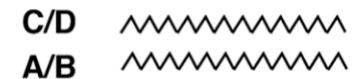
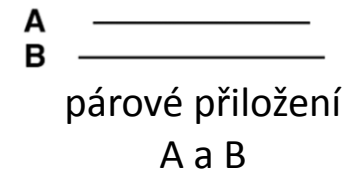
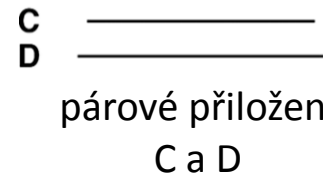


	A	B	C	D	E
A	-				
B	11	-			
C	20	30	-		
D	27	36	9	-	
E	30	33	20	27	-

distanční matice

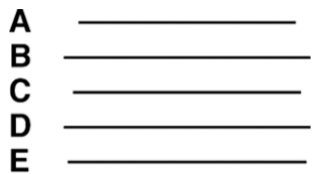


evoluční strom



párové příložení
konsenzu C/D a A/B

Automatická konstrukce příložení

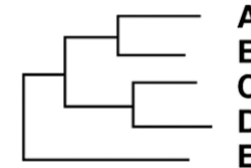


vstupní sekvence

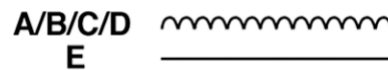
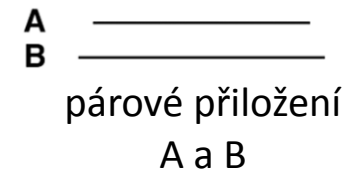
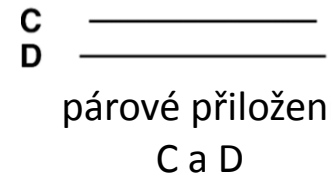


	A	B	C	D	E
A	-				
B	11	-			
C	20	30	-		
D	27	36	9	-	
E	30	33	20	27	-

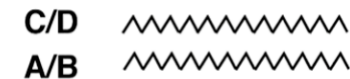
distanční matice



evoluční strom

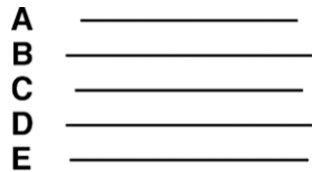


párové příložení E
s konsenzem A/B/C/D



párové příložení
konsenzu C/D a A/B

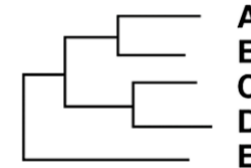
Automatická konstrukce příložení



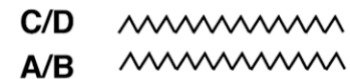
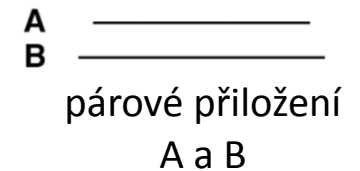
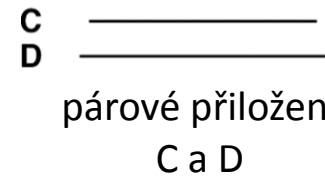
vstupní sekvence

	A	B	C	D	E
A	-				
B	11	-			
C	20	30	-		
D	27	36	9	-	
E	30	33	20	27	-

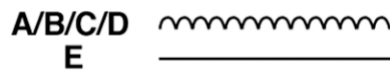
distanční matice



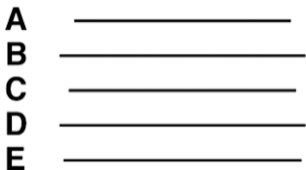
evoluční strom



párové příložení
konsenzu C/D a A/B



párové příložení E
s konsenzem A/B/C/D



finální příložení

Využití mnohonásobného přiložení

- ❑ Konstrukce kvalitního párového přiložení sekvencí
- ❑ Identifikace konzervovaných zbytků
- ❑ Identifikace zajímavých sekvencí
- ❑ Rekonstrukce fylogenetických stromů
- ❑ Citlivé prohledání prohledání sekvenčních databází

Využití mnohonásobného přiložení

- Konstrukce kvalitního párového přiložení sekvencí

1	LEAL-LDSL	DVKQ--NITL	VVDWGGMI	GMGFAARHP	ERIKRIVCL	
2	LEAL-LEHVEVRE--NITL	VLHDWGGMI	GMGFAARHP	ERIKRLVLL		
3	LEAL-LDSL	NVKE--NITL	VVDWGGMI	GMGYAARYP	ERIKRLVIL	
4	LEAL-LEHLEVRE--NITL	VVDWGGMI	GMGYAARYP	DRIKRLVIL		
5	LEAL-LEHLDVKE--NITL	VVDWGGMI	GMGYAARHP	ERIKRLVIL		
6	LEAL-LDHL	EVKE--KITL	IVVDWGGMI	GMGYAARHP	ERIKKIVVL	
7	LEAL-LASL	NVTE--KVTL	VVDWGGMI	GMGWATKYP	DAIERLVIL	
8	LERL-LDHL	AIRD--NITL	VVDWGGMI	GMAYAVAH	PAIRRLVVM	
9	LET	L-DHL	QIKE--NITL	VVDWGGMI	GMAYATRYP	ERIKRCVVL
10	LDTL-LRH	LGITG--PVT	LAVVDWGGMI	GFGWALSHHAQ	VKRLVIT	
11	LDAL-LRH	VGIDDVTPL	LAVVDWGGMI	GFGWALAHAVQ	VRRLVMT	
12	LESL-LTW	LDVG---PVD	LVVDWGGAI	GMGWAVRHP	DLVRRIVVL	
13	LFQL-LEHL	QIE---KYS	LVVDWGGAI	GVGCAAFAP	ERVEKLVVL	
14	LIRL-VDEL	DLK---NVIL	IAHDWGGAI	GLSAMHARR	DRLAGIGLL	
15	LERF-VLA	HDLD---RITL	VMHDFGGPV	GMGLAARHP	DRIRRVISA	

Využití mnohonásobného přiložení

- Konstrukce kvalitního párového přiložení sekvencí

1	LEAL	-L	S	L	V	K	Q	--	N	I	T	L	V	V	H	D	W	G	G	M	I	G	M	G	F	A	A	R	H	P	E	R	I	K	R	I	V	C	L		
2	LEAL	-L	E	H	V	E	V	R	E	--	N	I	T	L	V	L	H	D	W	G	G	M	I	G	M	G	F	A	A	R	H	P	E	R	I	K	R	L	V	L	
3	LEAL	-L	S	L	N	V	K	E	--	N	I	T	L	V	V	H	D	W	G	G	M	I	G	M	G	Y	A	A	R	Y	P	E	R	I	K	R	L	V	I	L	
4	LEAL	-L	E	H	L	E	V	R	E	--	N	I	T	L	V	V	H	D	W	G	G	M	I	G	M	G	Y	A	A	R	Y	P	D	R	I	K	R	L	V	I	L
5	LEAL	-L	E	H	L	V	K	E	--	N	I	T	L	V	V	H	D	W	G	G	M	I	G	M	G	Y	A	A	R	H	P	E	R	I	K	R	L	V	I	L	
6	LEAL	-L	H	L	E	V	K	E	--	K	I	T	L	I	V	H	D	W	G	G	M	I	G	M	G	Y	A	A	R	H	P	E	R	I	K	K	I	V	V	L	
7	LEAL	-L	A	S	L	N	V	T	E	--	K	V	T	L	V	V	H	D	W	G	G	M	I	G	M	G	W	A	T	K	Y	P	D	A	I	E	R	L	V	I	L
8	LERL	-L	H	L	A	I	R	D	--	N	I	T	L	V	V	H	D	W	G	G	M	I	G	M	A	Y	A	V	A	H	P	R	A	I	R	R	L	V	V	M	
9	LETL	-L	D	H	L	Q	I	K	E	--	N	I	T	L	V	V	H	D	W	G	G	M	I	G	M	A	Y	A	T	R	Y	P	E	R	I	K	R	C	V	V	L
10	LDTL	-L	R	H	L	G	I	T	G	--	P	V	T	L	A	V	H	D	W	G	G	M	I	G	F	G	W	A	L	S	H	H	A	Q	V	K	R	L	V	I	T
11	LDAL	-L	R	H	V	G	I	D	D	V	T	P	L	L	A	V	H	D	W	G	G	M	I	G	F	G	W	A	L	A	H	A	V	Q	V	R	R	L	V	M	T
12	LESL	-L	T	W	L	V	G	--	--	P	V	D	L	V	V	H	D	W	G	G	A	I	G	M	G	W	A	V	R	H	P	L	V	R	R	I	V	V	L		
13	LFQL	-L	E	H	L	Q	I	E	--	--	K	Y	S	L	V	V	H	D	W	G	G	A	I	G	V	G	C	A	A	F	A	P	E	R	V	E	K	L	V	V	L
14	LIRL	-V	E	L	D	L	K	--	--	N	V	I	L	I	A	H	D	W	G	G	A	I	G	L	S	A	M	H	A	R	R	D	R	L	A	G	I	G	L	L	
15	LERF	-V	L	A	H	D	L	--	--	R	I	T	L	V	M	H	L	F	G	G	P	V	G	M	G	L	A	A	R	H	P	D	R	I	R	I	V	S	A		

Využití mnohonásobného příložení

- Identifikace konzervovaných aminokyselinových zbytků

1	LEAL	-L	S	L	V	K	Q	--	N	I	T	L	V	H	D	W	G	G	M	I	G	M	G	F	A	A	R	H	P	E	R	I	K	R	I	V	C	L				
2	LEAL	-L	E	H	V	E	V	R	E	--	N	I	T	L	V	L	H	D	W	G	G	M	I	G	M	G	F	A	A	R	H	P	E	R	I	K	R	L	V	L		
3	LEAL	-L	S	L	N	V	K	E	--	N	I	T	L	V	H	D	W	G	G	M	I	G	M	G	Y	A	A	R	Y	P	E	R	I	K	R	L	V	I	L			
4	LEAL	-L	E	H	L	E	V	R	E	--	N	I	T	L	V	H	D	W	G	G	M	I	G	M	G	Y	A	A	R	Y	P	E	R	I	K	R	L	V	I	L		
5	LEAL	-L	E	H	L	V	K	E	--	N	I	T	L	V	H	D	W	G	G	M	I	G	M	G	Y	A	A	R	H	P	E	R	I	K	R	L	V	I	L			
6	LEAL	-L	H	L	E	V	K	E	--	K	I	T	L	I	V	H	D	W	G	G	M	I	G	M	G	Y	A	A	R	H	P	E	R	I	K	K	I	V	V	L		
7	LEAL	-L	A	S	L	N	V	T	E	--	K	V	T	L	V	H	D	W	G	G	M	I	G	M	G	W	A	T	K	Y	P	E	A	I	E	R	L	V	I	L		
8	LERL	-L	H	L	A	I	R	D	--	N	I	T	L	V	H	D	W	G	G	M	I	G	M	A	Y	A	V	A	H	P	R	A	I	R	R	L	V	V	M			
9	LETL	-L	H	L	Q	I	K	E	--	N	I	T	L	V	H	D	W	G	G	M	I	G	M	A	Y	A	T	R	Y	P	E	R	I	K	R	C	V	V	L			
10	LDTL	-L	R	H	L	G	I	T	G	--	P	V	T	L	A	V	H	D	W	G	G	M	I	G	F	G	W	A	L	S	H	H	A	Q	V	K	R	L	V	I	T	
11	LDAL	-L	R	H	V	G	I	D	D	V	T	P	L	T	L	A	V	H	D	W	G	G	M	I	G	F	G	W	A	L	A	H	A	V	Q	V	R	R	L	V	M	T
12	LESL	-L	T	W	L	V	G	--	--	P	V	D	L	V	H	D	W	G	G	A	I	G	M	G	W	A	V	R	H	P	L	V	R	R	I	V	V	L	I	L		
13	LFQL	-L	E	H	L	Q	I	E	--	--	K	Y	S	L	V	H	D	W	G	G	A	I	G	V	G	C	A	A	F	A	P	E	R	V	E	K	L	V	V	L		
14	LIRL	-V	E	L	D	L	K	--	--	N	V	I	L	I	A	H	D	W	G	G	A	I	G	L	S	A	M	H	A	R	R	D	R	L	A	G	I	G	L	L		
15	LERF	-V	L	A	H	D	L	--	--	R	I	T	L	V	M	H	D	E	F	G	G	P	V	G	M	G	L	A	A	R	H	P	D	R	I	R	I	V	S	A		

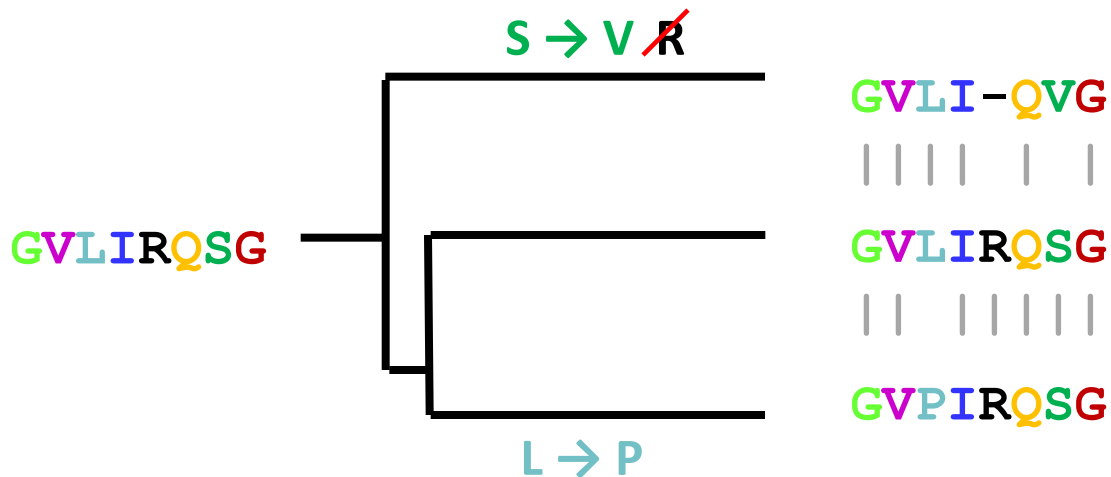
Využití mnohonásobného přiložení

- Identifikace konzervovaných aminokyselinových zbytků
 - Sloupce s **identickými** zbytky ve všech sekvencích
 - Konzervované zbytky jsou **důležité** pro **strukturu** nebo **funkci**

6	L	E	A	L	-	L	D	H	L	E	V	K	E	--	K	I	T	L	I	V	H	D	W	G	G	M	I	G	M	G	Y	A	A	R	H	P	E	R	I	K	K	I	V	V	L	
7	L	E	A	L	-	L	A	S	L	N	V	T	E	--	K	V	T	L	V	V	H	D	W	G	G	M	I	G	M	G	W	A	T	K	Y	P	D	A	I	E	R	L	V	I	L	
8	L	E	R	L	-	L	D	H	L	A	I	R	D	--	N	I	T	L	V	V	H	D	W	G	G	M	I	G	M	A	Y	A	V	A	H	P	R	A	I	R	R	L	V	V	M	
9	L	E	T	L	-	L	D	H	L	Q	I	K	E	--	N	I	T	L	V	V	H	D	W	G	G	M	I	G	M	A	Y	A	T	R	Y	P	E	R	I	K	R	C	V	V	L	
10	L	D	T	L	-	L	R	H	L	G	I	T	G	--	P	V	T	L	A	V	H	D	W	G	G	M	I	G	F	G	W	A	L	S	H	H	A	Q	V	K	R	L	V	I	T	
11	L	D	A	L	-	L	R	H	V	G	I	D	D	V	T	P	L	T	L	A	V	H	D	W	G	G	M	I	G	F	G	W	A	L	A	H	A	V	Q	V	R	R	L	V	M	T
12	L	E	S	L	-	L	T	W	L	D	V	G	--	--	P	V	D	L	V	V	H	D	W	G	G	A	I	G	M	G	W	A	V	R	H	P	D	L	V	R	R	I	V	V	L	
13	L	F	Q	L	-	L	E	H	L	Q	I	E	--	--	K	Y	S	L	V	V	H	D	W	G	G	A	I	G	V	G	C	A	A	F	A	P	E	R	V	E	K	L	V	V	L	
14	L	I	R	L	-	V	E	L	D	L	K	--	--	--	N	V	I	L	I	A	H	D	W	G	G	A	I	G	L	S	A	M	H	A	R	R	D	R	L	A	G	I	G	L	L	
15	L	E	R	F	-	V	L	A	H	D	L	D	--	--	R	I	T	L	V	M	H	D	E	F	G	G	P	V	G	M	G	L	A	A	R	H	P	D	R	I	R	I	V	S	A	

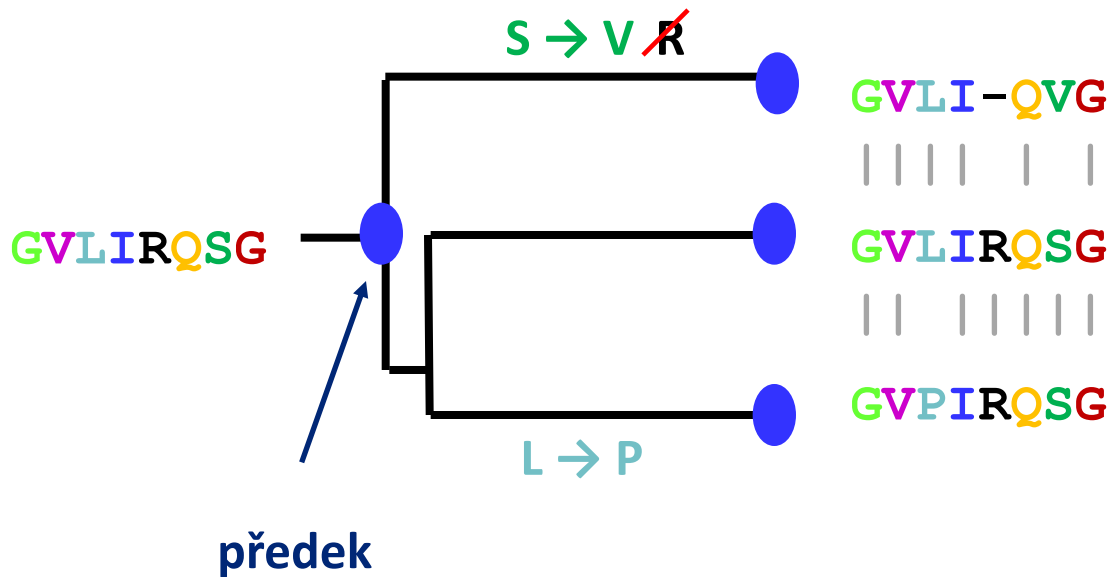
Využití mnohonásobného přiložení

- Rekonstrukce fylogenetických stromů



Využití mnohonásobného přiložení

- Rekonstrukce fylogenetických stromů
 - Předpověď - již neexistujících - sekvencí předků



Využití mnohonásobného přiložení

□ Citlivé prohledání sekvenčních databází

			Score	E
			(Bits)	Value
sequences producing significant alignments:				
✓	sp P59336 DHAA_RH0SD	Haloalkane dehalogenase >pdb1BN6 A Chai...	429	1e-118
✓	sp POA3G2 DHAA_RHORH	Haloalkane dehalogenase >sp POA3G3 DHAA...	424	3e-117
✓	pdb1CQW A	Chain A, Nai Cocrystallised With Haloalkane Dehalo...	424	4e-117
✓	sp Q9Z0R0 DHAA_MYCSX	Haloalkane dehalogenase >emb CAA10076.1 ...	422	1e-116
✓	gb AAV70825.1 	HT2 [Expression vector pHT2]	415	1e-114
✓	ref YP_001675030.1 	alpha/beta hydrolase fold [Shewanella hal...	320	8e-86
✓	ref YP_734675.1 	alpha/beta hydrolase fold [Shewanella sp. MR...	318	3e-85
✓	ref YP_001473250.1 	alpha/beta hydrolase fold [Shewanella sed...	317	6e-85
✓	ref ZP_01736514.1 	alpha/beta hydrolase [Marinobacter sp. ELB...	317	6e-85
✓	ref YP_738656.1 	alpha/beta hydrolase fold [Shewanella sp. MR...	316	9e-85
✓	ref YP_001502590.1 	alpha/beta hydrolase fold [Shewanella pea...	316	9e-85
✓	ref NP_717353.1 	hydrolase, alpha/beta hydrolase fold family ...	315	2e-84
✓	ref YP_750057.1 	alpha/beta hydrolase fold [Shewanella frigid...	315	2e-84
✓	ref YP_268879.1 	hydrolase, alpha/beta hydrolase fold family ...	315	2e-84
✓	ref YP_001761524.1 	alpha/beta hydrolase fold [Shewanella woo...	315	3e-84
✓	ref ZP_01841154.1 	alpha/beta hydrolase fold [Shewanella balt...	315	3e-84
✓	ref YP_870347.1 	alpha/beta hydrolase fold [Shewanella sp. AN...	314	4e-84
✓	ref YP_129676.1 	putative haloalkane dehalogenase [Photobacte...	314	7e-84
✓	ref ZP_01221858.1 	putative haloalkane dehalogenase [Photobac...	313	8e-84
✓	ref YP_001365757.1 	alpha/beta hydrolase fold [Shewanella bal...	313	9e-84
✓	ref YP_562379.1 	alpha/beta hydrolase fold [Shewanella denitr...	313	9e-84
✓	ref ZP_01897865.1 	putative haloalkane dehalogenase [Moritell...	313	1e-83
✓	ref YP_001049934.1 	alpha/beta hydrolase fold [Shewanella bal...	313	1e-83
✓	ref YP_943362.1 	alpha/beta hydrolase fold [Psychromonas ingr...	313	1e-83
✓	ref YP_001182970.1 	alpha/beta hydrolase fold [Shewanella put...	312	2e-83
✓	ref YP_001554014.1 	alpha/beta hydrolase fold [Shewanella bal...	312	2e-83
✓	ref ZP_01706252.1 	alpha/beta hydrolase fold [Shewanella putr...	310	7e-83
✓	ref YP_964030.1 	alpha/beta hydrolase fold [Shewanella sp. W3...	310	9e-83
✓	ref YP_510562.1 	haloalkane dehalogenase [Jannaschia sp. CCS1...	308	3e-82
✓	ref ZP_01216824.1 	hydrolase, alpha/beta hydrolase fold famil...	307	8e-82
✓	ref YP_001093840.1 	alpha/beta hydrolase fold [Shewanella loi...	306	1e-81
✓	ref NP_106032.1 	haloalkane dehalogenase [Mesorhizobium loti ...	303	8e-81
✓	gb AAT70109.1 	CurM [Lyngbya majuscula]	303	8e-81
✓	ref ZP_01055470.1 	haloalkane dehalogenase [Roseobacter sp. M...	303	1e-80
✓	ref ZP_01617455.1 	haloalkane dehalogenase [marine gamma prot...	302	2e-80
✓	ref ZP_01592200.1 	alpha/beta hydrolase fold [Geobacter lovle...	300	7e-80
✓	ref ZP_01911259.1 	alpha/beta hydrolase [Plesiocystis pacific...	300	9e-80
✓	ref YP_001230772.1 	alpha/beta hvdrolase fold fGeobacter uran...	300	9e-80

Využití mnohonásobného přiložení

□ Citlivé prohledání sekvenčních databází

- PSI-BLAST – detekce slabých **biologicky významných** podobností

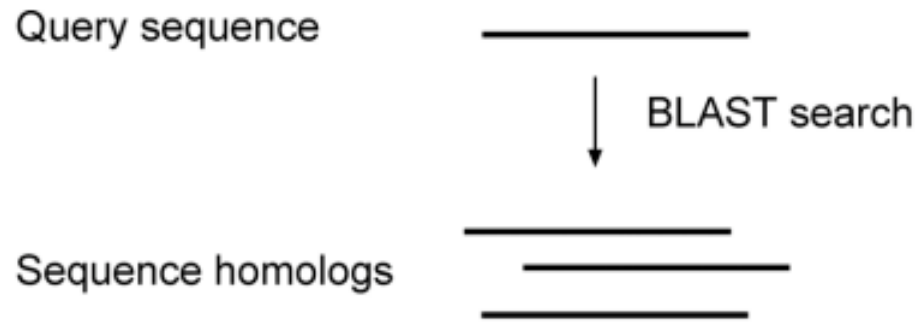
```
quences producing significant alignments:
```

			Score	E
			(Bits)	Value
✓	sp P59336 DHAA_RH0SD	Haloalkane dehalogenase >pdb1BN6 A Chai...	429	1e-118
✓	sp POA3G2 DHAA_RHORH	Haloalkane dehalogenase >sp POA3G3 DHAA...	424	3e-117
✓	pd 1CQW A	Chain A, Nai Cocrystallised With Haloalkane Dehalo...	424	4e-117
✓	sp Q9ZERO DHAA_MYCSX	Haloalkane dehalogenase >emb CAA10076.1 ...	422	1e-116
✓	gb AAV70825.1	HT2 [Expression vector pHT2]	415	1e-114
✓	ref YP_001675030.1	alpha/beta hydrolase fold [Shewanella hal...	320	8e-86
✓	ref YP_734675.1	alpha/beta hydrolase fold [Shewanella sp. MR...	318	3e-85
✓	ref YP_001473250.1	alpha/beta hydrolase fold [Shewanella sed...	317	6e-85
✓	ref ZP_01736514.1	alpha/beta hydrolase [Marinobacter sp. ELB...	317	6e-85
✓	ref YP_738656.1	alpha/beta hydrolase fold [Shewanella sp. MR...	316	9e-85
✓	ref YP_001502590.1	alpha/beta hydrolase fold [Shewanella pea...	316	9e-85
✓	ref NP_717353.1	hydrolase, alpha/beta hydrolase fold family ...	315	2e-84
✓	ref YP_750057.1	alpha/beta hydrolase fold [Shewanella frigid...	315	2e-84
✓	ref YP_268879.1	hydrolase, alpha/beta hydrolase fold family ...	315	2e-84
✓	ref YP_001761524.1	alpha/beta hydrolase fold [Shewanella woo...	315	3e-84
✓	ref ZP_01841154.1	alpha/beta hydrolase fold [Shewanella balt...	315	3e-84
✓	ref YP_870347.1	alpha/beta hydrolase fold [Shewanella sp. AN...	314	4e-84
✓	ref YP_129676.1	putative haloalkane dehalogenase [Photobacte...	314	7e-84
✓	ref ZP_01221858.1	putative haloalkane dehalogenase [Photobac...	313	8e-84
✓	ref YP_001365757.1	alpha/beta hydrolase fold [Shewanella bal...	313	9e-84
✓	ref YP_562379.1	alpha/beta hydrolase fold [Shewanella denitr...	313	9e-84
✓	ref ZP_01897865.1	putative haloalkane dehalogenase [Moritell...	313	1e-83
✓	ref YP_001049934.1	alpha/beta hydrolase fold [Shewanella bal...	313	1e-83
✓	ref YP_943362.1	alpha/beta hydrolase fold [Psychromonas ingr...	313	1e-83
✓	ref YP_001182970.1	alpha/beta hydrolase fold [Shewanella put...	312	2e-83
✓	ref YP_001554014.1	alpha/beta hydrolase fold [Shewanella bal...	312	2e-83
✓	ref ZP_01706252.1	alpha/beta hydrolase fold [Shewanella putr...	310	7e-83
✓	ref YP_964030.1	alpha/beta hydrolase fold [Shewanella sp. W3...	310	9e-83
✓	ref YP_510562.1	haloalkane dehalogenase [Jannaschia sp. CCS1...	308	3e-82
✓	ref ZP_01216824.1	hydrolase, alpha/beta hydrolase fold famil...	307	8e-82
✓	ref YP_001093840.1	alpha/beta hydrolase fold [Shewanella loi...	306	1e-81
✓	ref NP_106032.1	haloalkane dehalogenase [Mesorhizobium loti ...	303	8e-81
✓	gb AAT70109.1	CurM [Lyngbya majuscula]	303	8e-81
✓	ref ZP_01055470.1	haloalkane dehalogenase [Roseobacter sp. M...	303	1e-80
✓	ref ZP_01617455.1	haloalkane dehalogenase [marine gamma prot...	302	2e-80
✓	ref ZP_01592200.1	alpha/beta hydrolase fold [Geobacter lovle...	300	7e-80
✓	ref ZP_01911259.1	alpha/beta hydrolase [Plesiocystis pacific...	300	9e-80
✓	ref YP_001230772.1	alpha/beta hvdrolase fold fGeobacter uran...	300	9e-80





□ První iterace





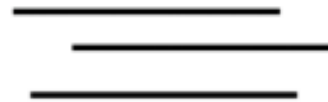
□ První iterace

Query sequence



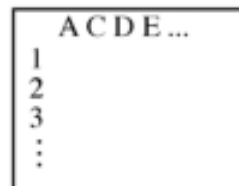
BLAST search

Sequence homologs



Multiple sequence alignment

Profile



- Druhá iterace



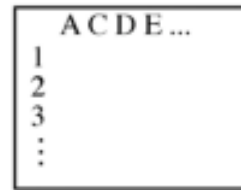
Profile

	A	C	D	E	...
1					
2					
3					
⋮					

□ Druhá iterace



Profile



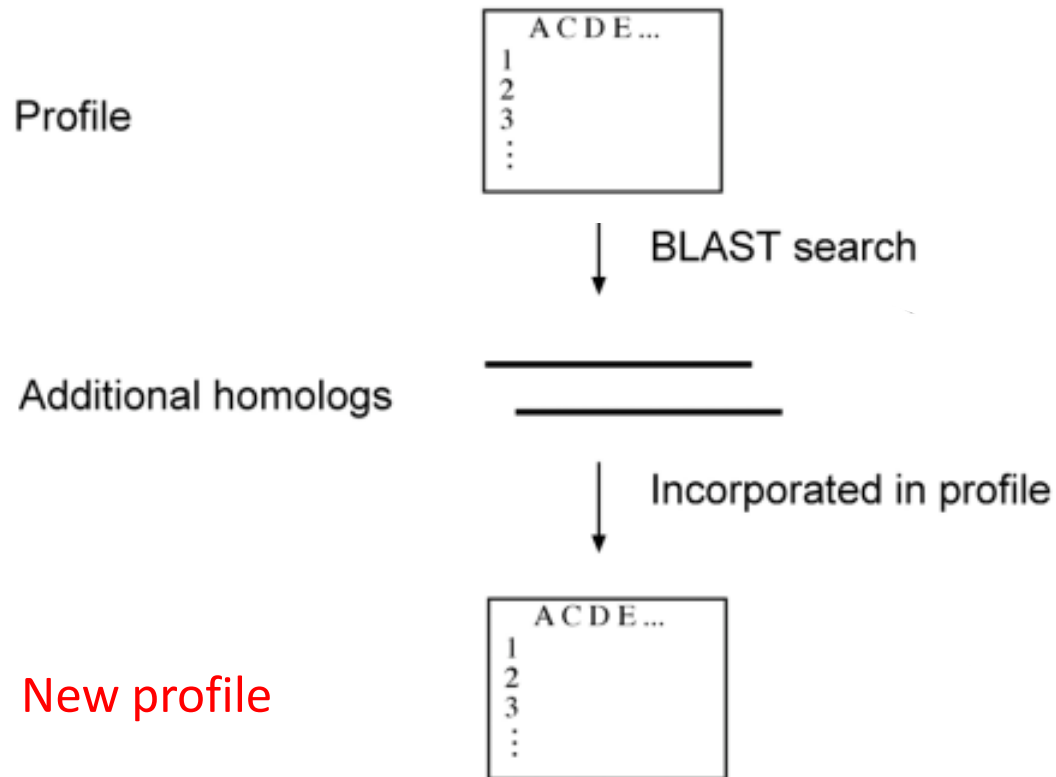
↓ BLAST search

Additional homologs



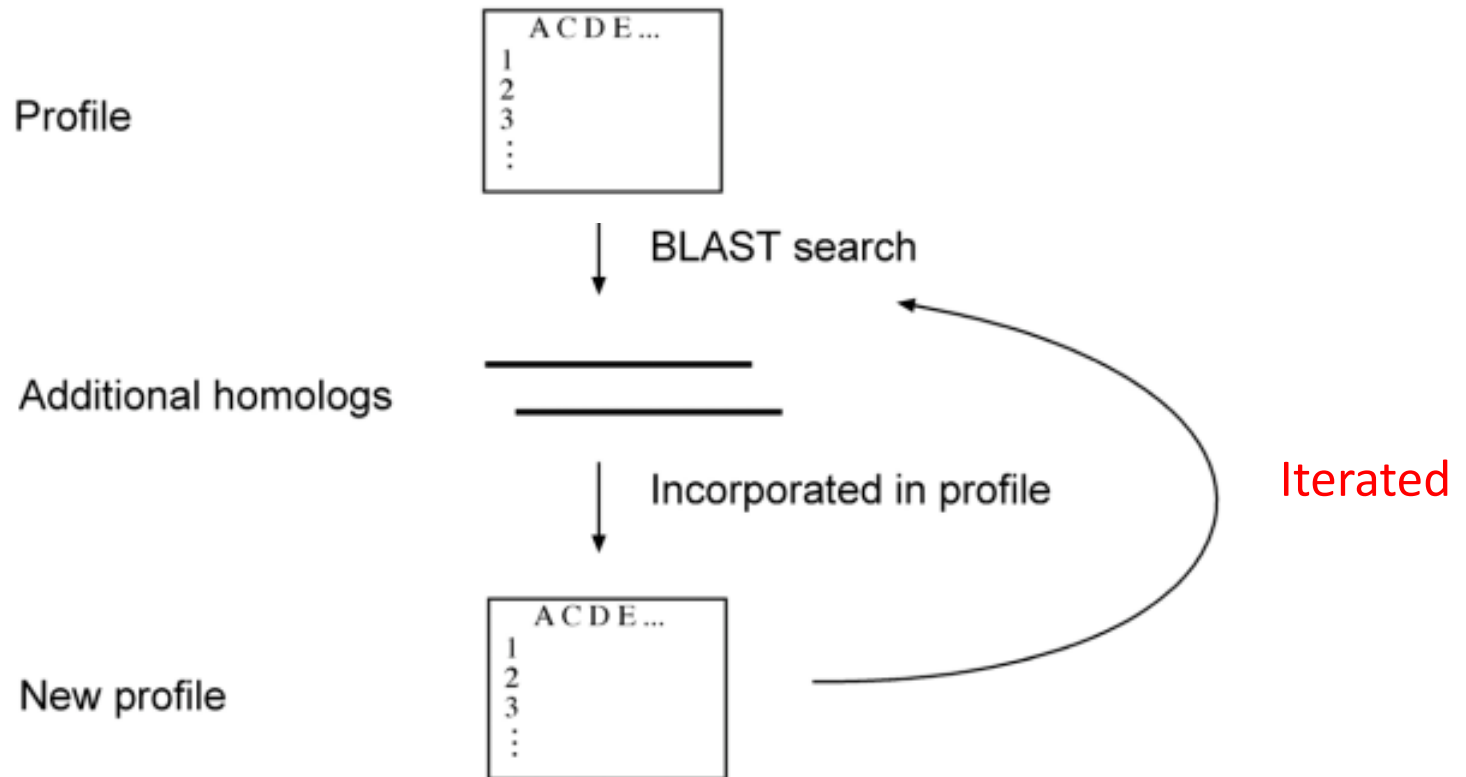


□ Druhá iterace



PSI-BLAST

□ Druhá iterace



PSI-BLAST

□ Vstup

BLAST BETA

Home Recent Results Saved St

MSLGAKPFGEKKFIEIKGRRMAYIDEG
TGDPILFQHGNTSSYLWRNI

▶ NCBI/BLAST/blastp suite: BLASTP programs search protein databases using a protein query. [more...](#) [Reset page](#) [Bookmark](#)

Enter Query Sequence

Enter accession number, gi, or FASTA sequence [Clear](#)

Query subrange [?](#)

From

To

LGAKPFGEKKFIEIKGRRMAYIDEGTGDPILFQHGNTSSYLWRNIMPHCAGLGRLIACDLIGM

Or, upload file [?](#)

Job Title

Enter a descriptive title for your BLAST search [?](#)

Choose Search Set

Database [?](#)

Organism Optional Any Human *A.thaliana* Mouse Custom... [?](#)
Search only sequences from selected organism [?](#)

□ Výsledek

Skóre E-hodnota



Sequences producing significant alignments:

			Score (Bits)	E Value	
<input checked="" type="checkbox"/>	sp P59336 DHAA_RHOSD	Haloalkane dehalogenase >pdb 1BN6 A Chai...	429	1e-118	
<input checked="" type="checkbox"/>	sp POA3G2 DHAA_RHORH	Haloalkane dehalogenase >sp POA3G3 DHAA_...	424	3e-117	
<input checked="" type="checkbox"/>	pdb 1CQW A	Chain A, Nai Cocrystallised With Haloalkane Dehalo...	424	4e-117	S
<input checked="" type="checkbox"/>	sp Q9ZERO DHAA_MYCSX	Haloalkane dehalogenase >emb CAA10076.1 ...	422	1e-116	
<input checked="" type="checkbox"/>	gb AAV70825.1 	HT2 [Expression vector pHT2]	415	1e-114	
<input checked="" type="checkbox"/>	ref YP_001675030.1 	alpha/beta hydrolase fold [Shewanella hal...	320	8e-86	G
<input checked="" type="checkbox"/>	ref YP_734675.1 	alpha/beta hydrolase fold [Shewanella sp. MR...	318	3e-85	G
<input checked="" type="checkbox"/>	ref YP_001473250.1 	alpha/beta hydrolase fold [Shewanella sed...	317	6e-85	G
<input checked="" type="checkbox"/>	ref ZP_01736514.1 	alpha/beta hydrolase [Marinobacter sp. ELB...	317	6e-85	
<input checked="" type="checkbox"/>	ref YP_738656.1 	alpha/beta hydrolase fold [Shewanella sp. MR...	316	9e-85	G
<input checked="" type="checkbox"/>	ref YP_001502590.1 	alpha/beta hydrolase fold [Shewanella pea...	316	9e-85	G
<input checked="" type="checkbox"/>	ref NP_717353.1 	hydrolase, alpha/beta hydrolase fold family ...	315	2e-84	G
<input checked="" type="checkbox"/>	ref YP_750057.1 	alpha/beta hydrolase fold [Shewanella frigid...	315	2e-84	G
<input checked="" type="checkbox"/>	ref YP_268879.1 	hydrolase, alpha/beta hydrolase fold family ...	315	2e-84	G
<input checked="" type="checkbox"/>	ref YP_001761524.1 	alpha/beta hydrolase fold [Shewanella woo...	315	3e-84	G
<input checked="" type="checkbox"/>	ref ZP_01841154.1 	alpha/beta hydrolase fold [Shewanella balt...	315	3e-84	
<input checked="" type="checkbox"/>	ref YP_870347.1 	alpha/beta hydrolase fold [Shewanella sp. AN...	314	4e-84	G
<input checked="" type="checkbox"/>	ref YP_129676.1 	putative haloalkane dehalogenase [Photobacte...	314	7e-84	G
<input checked="" type="checkbox"/>	ref ZP_01221858.1 	putative haloalkane dehalogenase [Photobac...	313	8e-84	
<input checked="" type="checkbox"/>	ref YP_001365757.1 	alpha/beta hvdrolase fold [Shewanella bal...	313	9e-84	G

❑ Skóre

- ❑ Normalizované skóre = součet substitucí a penalizací za mezery
- ❑ Nedostatečně kvantifikuje významnost (**vyšší** je lepší)

❑ E-hodnota

- ❑ Rovná počtu BLAST příložení s příslušným skóre, která mohou vzniknout **náhodou**
- ❑ Kvantifikuje významnost příložení (**nižší** je lepší)
- ❑ Významné jsou hity s E-hodnotou **<0.01**



□ Výsledek

Příložen

Sequences producing significant alignments:		Score (Bits)	E Value
>gb AAT70109.1 CurN [Lyngbya majuscula] Length=341			
Score = 303 bits (777), Expect = 8e-81, Method: Composition-based stats. Identities = 148/297 (49%), Positives = 188/297 (63%), Gaps = 8/297 (2%)			
Query	2	SEIGTGFPPDPHYVEVLGERMHYVDVGPDRDGPVFLFHGNPTSSYLWRNIIPHV-APSHR	60
		I + FPF VEV G + YVD G G PVLFLHGNPTSSYLWRNIIPV A +R	
Sbjct	41	LPISSEFPFAKRTVEVEGATIAYVDEG--SGQPVLFLHGNPTSSYLWRNIIPYVVAAGYR	98
Query	61	CIAPDLIGMGKSDKPDLDYFFDDHVRYLDAFIEALGLEEVVLVIHDWGSALGFHWAKRNP	120
		+APDLIGMG S KPD++Y DHV Y+D FI+ALGL+++VLVIHDWGS +G A+ NP	
Sbjct	99	AVAPDLIGMGDSAKPDIEYRLQDHVAYMDGFI DALGLDDMLVVIHDWGSVIGMRHARLNP	158
Query	121	ERVKGIACMEFIRPI----PTWDEWPEFARETFQAFRTADVGRELIIDQNAFIEGVLPK-	175
		+RV +A ME + P P+++ F+ RTADVG ++++D N F+E +LP+	
Sbjct	159	DRVAAVAFMEALVPPALPMPSEYAMGPQLGPLFRDLRTADVGEKMLVDGNFFVETILPEM	218
Query	176	CVVRPLTEVEMDHYREPFLKPVDPREPLWRF PNEIPIAGEPANIVALVEAYMNLHQSPVP	235
		VVR L+E EM YR PF R P ++P E+PI GEPA A V WL SP+P	
Sbjct	219	GVVRSLSAEAMAAYRAPFPTRQSRLLPTLQWPREVPIGGEPAFABAEVLKNGEWMASPIP	278
Query	236	KLLFWGTPGVLIPPAEARLAESLPNCKTVDIGPGLHYLQEDNPDIGSEIARWLPG	292
		KLLF PG L P L+E++PN + +G G H+LQED+P LIG IA WL	
Sbjct	279	KLLFHAEPGALAPKPVVDYLSENVPNLEVRVFGAGTHFLQEDHPHLIGQGIADWLRR	335
ref NP_106032.1	haloalkane dehalogenase [Mesorhizobium loti ...	303	8e-81 G
gb AAT70109.1	CurN [Lyngbya majuscula]	303	8e-81
ref ZP_01055470.1	haloalkane dehalogenase [Roseobacter sp. M...	303	1e-80
ref ZP_01617455.1	haloalkane dehalogenase [marine gamma prot...	302	2e-80
ref ZP_01592200.1	alpha/beta hydrolase fold [Geobacter lovle...	300	7e-80
ref ZP_01911259.1	alpha/beta hydrolase [Plesiocystis pacific...	300	9e-80
ref YP_001230772.1	alpha/beta hvdrolase fold [Geobacter uran...	300	9e-80 G

Databáze mnohonásobného příložen

□ Pfam

wellcome trust
sanger
institute

HOME | SEARCH | BROWSE | FTP | HELP

Pfam
keyword search Go

96 architecture 9801 sequence 4 interactions 1013 species 126 structures

Family: *Abhydrolase_1* (PF00561)

Summary
Domain organisation
Alignments
Trees
Curation & models
Species
Interactions
Structures

Alignments

There are various ways to view or download the sequence alignments that we store. You can use a sequence viewer to look at either the seed or full alignment for the family, or you can look at a plain text version of the sequence in a variety of different formats. [More...](#)

View options

Alignment: Seed (48) Full (9801)

Viewer: jalview

View

Formatting options

Databáze mnohonásobného přiložení

□ Pfam

wellcome trust sanger institute

HOME | SEARCH | BROWSE | FTP | HELP

Pfam keyword search Go

Family: *Abhydrolase superfamily*

Summary
Domain organisation
Alignments
Trees
Curation & models
Species
Interactions
Structures

Alignments

Currently showing rows 1 to 30 of 48 rows in this alignment. Show rows of alignment Go

P07000/82-321	FDVLIIDHRGQGRSG.RLLAD.....	PHLGHVNRFNQYVDDLAAFWQQ
P53264/170-440	WCIHAIIDLPGYGFSSRPKFP.....	FEYPRDNIHSVQDFHFERIHY
P65824/134-506	FDLVGFDPKGVASRR.PAIWCNSDADMDRLRAEPQVDYSREGVAHIENETKQFVGRVCVDMKGNFLAHVGTVNVAKDLDA	
P53750/56-285	FHIIAPDLPFGFTE.T.....	PENYKFSFDLCEISIGY
P53208/67-318	ADIFSVVVRNHGISP.....	KAIPYDYTTLTNDLIY
P27747/162-370	YTVVALDLPCHGQSS.....	PRLAGTTLAQMAGFVAF
P42786/58-307	FRIVIIDRQGRSH.P.....	YACAEENTTQDLVADIEP
P07383/63-269	KRYLALDLRCHGGTS.....	IPKCCYYVSDFAEDVSI
P65822/138-415	FDLVGFDPKGVGHST.PALRCRTDAEFDAYRRDPMADYSPAGVTHVEQVYRQLAQDCVDRMGFSFLANIGTASVARDMDY	
P46547/82-313	FRVLLLDQRGTGHST.PIHAELL.....	AHLNPRQQADYLSHFRAQSIVRDAEI
P24640/93-308	YHLIIPDLLGFGNS.K.....	PMTADYRADAQAATRLHF
P26174/63-280	YRVIVPDLPGHGSR.S.....	TARNRFGLKPMaedLMP
P66777/56-296	FRIVRYDNRGVGRSSVP.....	KPIsAYTMAHFADDfDA
P52705/31-252	HKVTALDMAASGIDP.R.....	QIEQINSFDEYSEPLL

Reference

- ❑ Claverie, J-M., & Notredame, C. (2006) **Bioinformatics for Dummies** (2nd ed.) Wiley Publishing, Hoboken, p. 436.
- ❑ Xiong, J. (2006) **Essential Bioinformatics**, Cambridge University Press, New York, p. 352.
- ❑ **MUSCLE**: <http://www.drive5.com/muscle/>
- ❑ **TCoffee**: <http://tcoffee.vital-it.ch/cgi-bin/Tcoffee/tcoffee.cgi/index.cgi>
- ❑ **ClustalW**: <http://www.ch.embnet.org/software/ClustalW.html>
- ❑ **BioEdit**: <http://www.mbio.ncsu.edu/BioEdit/bioedit.html>
- ❑ **PSI-BLAST**: <http://blast.ncbi.nlm.nih.gov/Blast.cgi>
- ❑ **PhyML**: <http://www.phylogeny.fr/>
- ❑ **PHYLIP**: <http://evolution.genetics.washington.edu/phylip.html>
- ❑ **MPI Toolkit**: <http://toolkit.tuebingen.mpg.de/>
- ❑ **Pfam**: <http://pfam.sanger.ac.uk/>