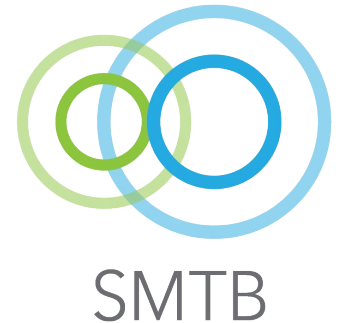


# Эволюция ферментов глиоксилатного цикла

Софья Бичурина, Анастасия Голубева, Пётр  
Екатериночкин, Екатерина Игнатова, Иван Козин,  
Мария Кульбачная, Никита Литвяков, Татьяна  
Мураховская, Екатерина Максимова, Фёдор  
Кондрашов

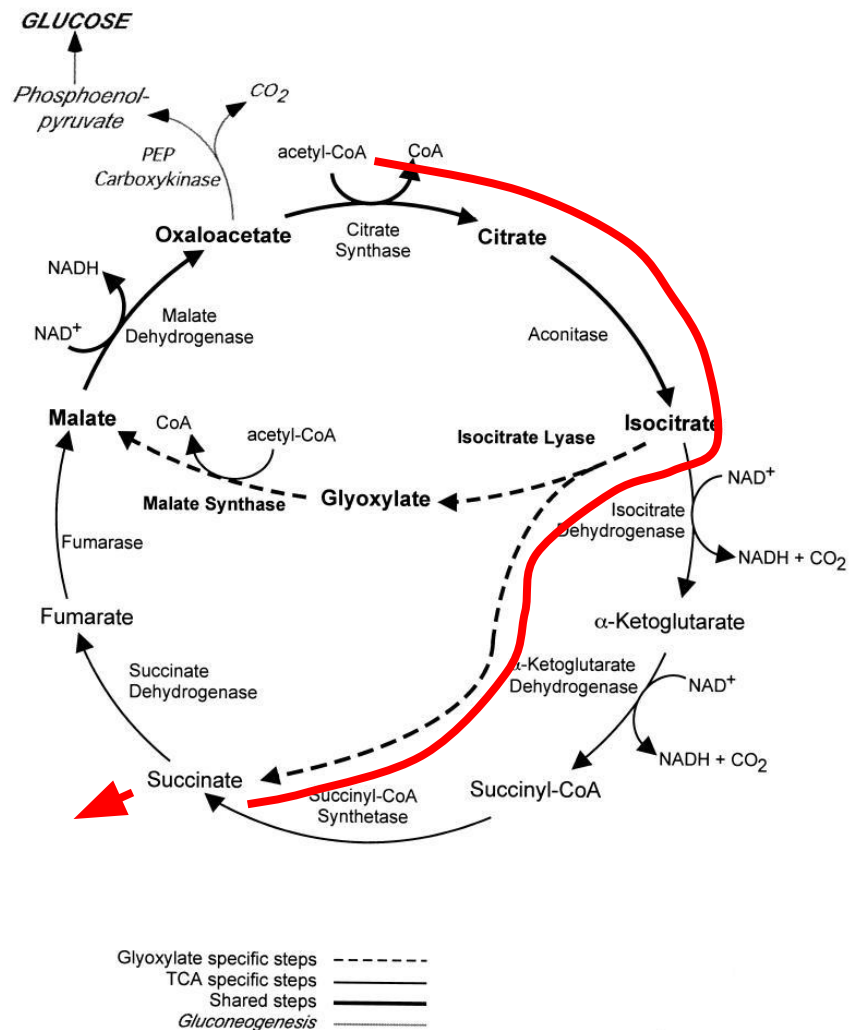


# Глиоксилатный ЦИКЛ

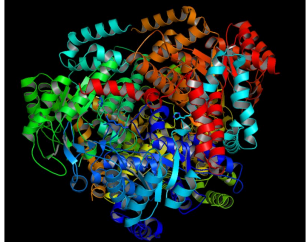
Глиоксилатный цикл – это анаболический путь, служащий для превращения ацетильных групп ацетил-CoA в сукцинат.

Благодаря глиоксилатному циклу, клетка может использовать ацетат в качестве предшественника для синтеза разных молекул.

Используется в богатых липидами семенах в процессе прорастания, когда растение еще не фотосинтезирует, позволяя семенам конвертировать жиры в глюкозу.



# Изоцитратлиаза

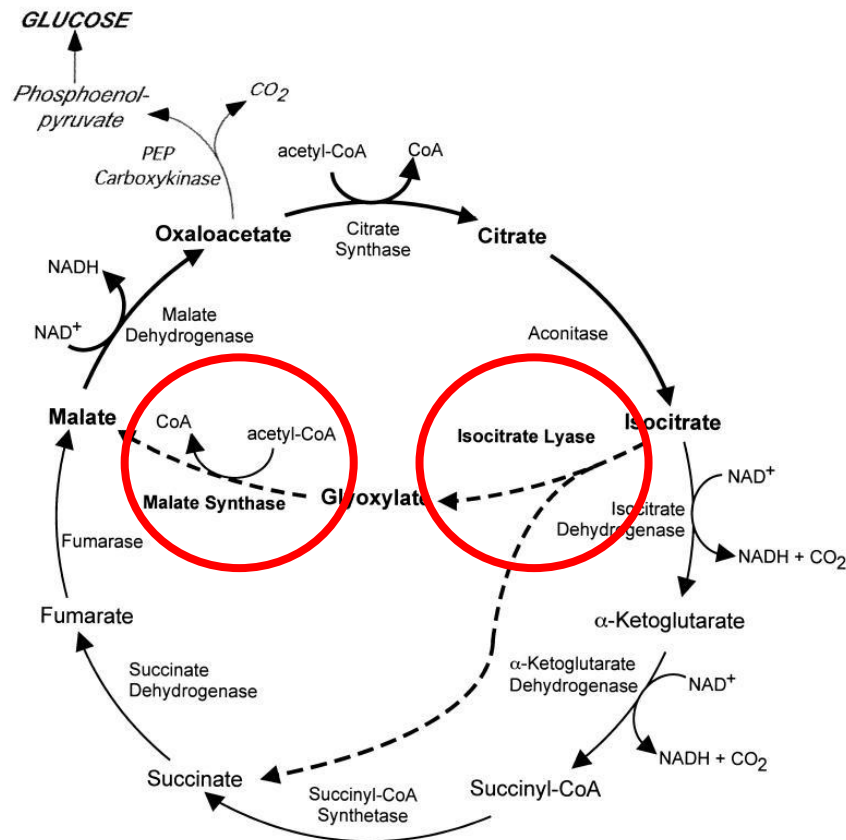


Глиоксилатный цикл определяют два ключевых фермента: изоцитратлиаза и малатсинтаза.

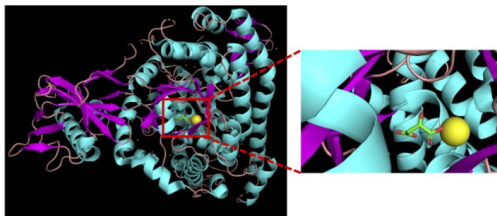
Изоцитратлиаза превращает изоцитрат в сукцинат и глиоксилат.

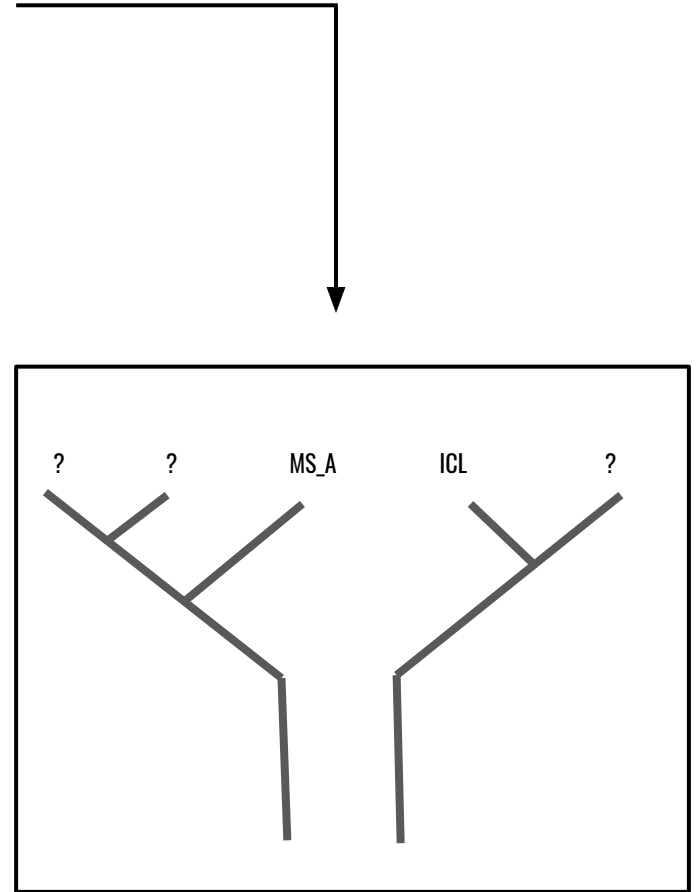
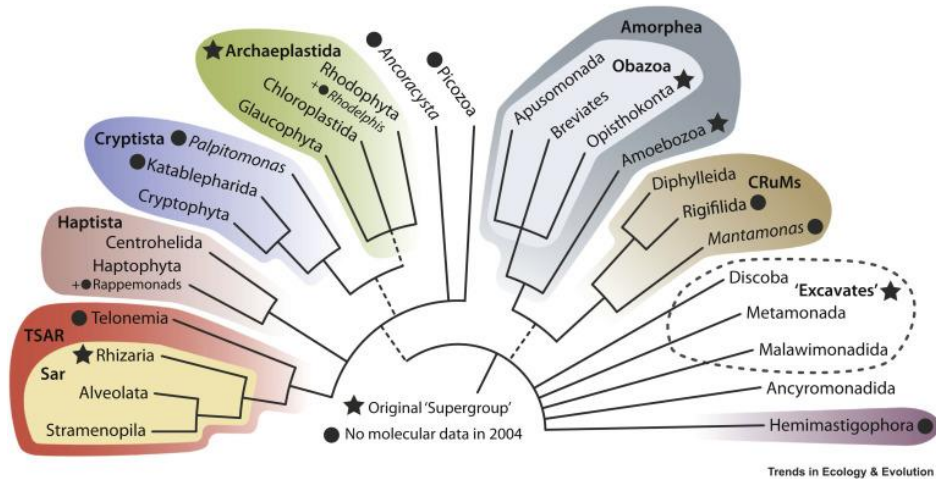
Малатсинтаза превращает глиоксилат и ацетил-CoA в малат.

Изучая эволюцию глиоксилатного цикла, стоит обратить внимание на эти ключевые ферменты, изоцитратлиазу и малатсинтазу.



# Малатсинтаза





В научной литературе есть работы, подтверждающие гипотезу горизонтального переноса в эволюции глиоксилатного цикла.

Целью данного проекта является исследование полной эволюции глиоксилатного цикла (насколько это возможно) с применением всех последних данных полногеномного секвенирования.

# Ход работы

При просмотре деревьев мы могли понять какие белки близкородственные, а какие нет

Заключение

## Выводы

Мы хотели изучить как белки глиоксилатного цикла пропадали или приобретались у разных типов организмов.

В заключительном этапе мы анализировали строение филогенетических деревьев

Анализ

## Деревья

Выравнивания используются для построения филогении.

Мы искали белки сходные по последовательности с малатсинтазой и изоцитратлиазой в различных организмах



данные

Найденные последовательности выравниваются.

BLAST – это сайт позволяющий находить сходные последовательности нуклеотидов.

# Методы

**BLAST для поиска ILC и MS у: Бактерий, Архей и Эукариот.**  
Искали рекурсивно, начиная с последовательностей из E. coli (IC1L: MHO05562 MSa: MHO05561).

**Muscle имплементированный в MEGA:**  
Выравнивание и построение первого филогенетического дерева методом Maximum Likelihood.

**MrBayes:**  
Второе построение дерева через командную строку с параметрами по умолчанию.

**Mesquite:**  
Визуализация данных из MrBayes в более удобном виде.

## malate synthase [Phycisphaeraeae bacterium]

Sequence ID: MAE64273.1 Length: 627 Number of Matches: 1

Range 1: 6 to 627 GenPept Graphics

Score	Expect	Method	Identities	Positives	Gaps
899 bits(2324)	0.0	Compositional matrix adjust.	427/622(69%)	514/622(82%)	0/622(0%)

Query	5	SIIRDNLKEQDYITEEVLALANTMAHFNKSVKHLMNERIKRRNDRHNSKRIFYDPP	64
Sbjct	6	ALQTDRLDQTYPDVLTEDALRALDALAPLNDVREEMAAIQRRRRRAAEGERTFLDP	65

Query	65	ESLTPGSSIKVKDAREGRFESVIPIHDLQRQWIGTGPAAKPNASLESIRNAAYALLSG	124
Sbjct	66 <th>++I+D L++ Y D+ TE+ L AL+ +A N + +M RI+RR R +RI F DP</th> <th>125</th>	++I+D L++ Y D+ TE+ L AL+ +A N + +M RI+RR R +RI F DP	125

Query	125	ADGWMFDGEDALGQ++MSLDN RLKLA I +D +FL+VAEQVAEMNRW E F G+ IV	184
Sbjct	126 <th>ADGWMFDGEDALGQVSTMSLDNRLKLAIDRDTMFLEVAEQVAEMNRWGEPPFRP+IV</th> <th>185</th>	ADGWMFDGEDALGQVSTMSLDNRLKLAIDRDTMFLEVAEQVAEMNRWGEPPFRP+IV	185

Query	185	KDWKSQLDFTTKIFRCRGLHLDRRHIDBADVALAASIDLCLVYVNNYRLKSGSSIV	244
Sbjct	186 <th>+HWSQLDFTTKIR RGLHLDRH+R DG +ASI D LYVNNH+RL+ G+SIV</th> <th>245</th>	+HWSQLDFTTKIR RGLHLDRH+R DG +ASI D LYVNNH+RL+ G+SIV	245

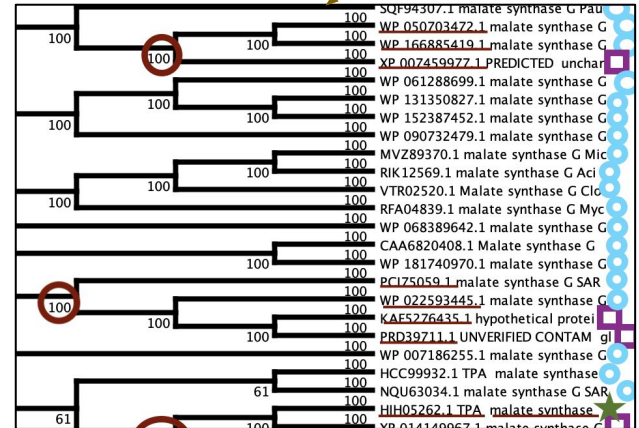
  

Query	245	LYLPKIQTAGEAALNMSLGALEDHLGENTGIKVVYLVLEQ EATYQLMEIRAVLKGHFV	304
Sbjct	246 <th>LYLPKIQTAEALFNDILAERHLGLAVGTIKVVYLVLEQVEASQLMEIRAALAPHFV</th> <th>305</th>	LYLPKIQTAEALFNDILAERHLGLAVGTIKVVYLVLEQVEASQLMEIRAALAPHFV	305

Query	305	GYNTGRWDYINSVDAMAWDKGFINPNIESVTMTYGYMRNVEDRVRRAVNTPDINGNCAI	364
Sbjct	306 <th>G+NTGRWDYINSVDAMAWNRDFVNPNDIAIVMTYGYMRHYEDRVRRAVNTPDNRGRCAL</th> <th>365</th>	G+NTGRWDYINSVDAMAWNRDFVNPNDIAIVMTYGYMRHYEDRVRRAVNTPDNRGRCAL	365

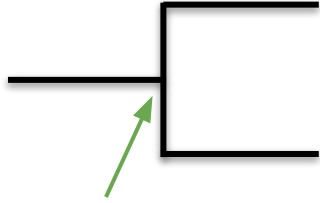
Species/Abbrv	L	S	E	D	D	F	W	M	S	L	E	N	I	V	D	E	F
1. MBT17576.1 malate synthase G Dehalococcoidia bacterium\	L	S	E	D	D	F	W	M	S	L	E	N	I	V	D	E	F
2. MBU27748.1 malate synthase G Flavobacteriales bacterium\	I	S	E	D	H	F	W	T	S	L	E	N	I	I	K	E	F
3. MSU49663.1 malate synthase G Opitutus sp.)	A	P	E	K	F	W	A	T	L	A	V	L	V	R	E	L	
4. MTI33903.1 malate synthase G Luteivirga sdotyamensis\	V	E	Q	D	A	F	W	A	G	L	A	A	V	E	D	L	
5. CAA6820408.1 Malate synthase G (EC uncultured Thiotricha	V	E	A	D	K	F	W	A	G	L	E	G	I	V	N	D	L
6. HCC99932.1 TPA: malate synthase G Planctomycetaceae ba	L	D	P	A	V	W	T	A	F	G	A	I	V	E	T	L	
7. HEV57096.1 TPA: malate synthase G Phycisphaerales bacter	I	D	P	A	F	W	H	A	L	A	D	L	H	R	G	F	
8. MAL98640.1 malate synthase G Alteromonadaceae bacterium\	V	D	K	E	Q	F	W	T	G	L	D	T	L	V	H	D	L
9. MBK68909.1 malate synthase G Legionellales bacterium\	I	S	T	K	K	F	W	E	G	F	D	R	V	T	H	E	L
10. MBP08944.1 malate synthase G Acidiferrobacteraceae bact	V	P	D	T	F	W	H	S	F	A	A	V	I	D	E	L	
11. MSO48278.1 malate synthase G Thermoleophillia bacterium	V	D	T	R	F	F	I	G	L	S	E	M	I	H	R	F	
12. MSO48278.1 malate synthase G Thermoleophillia bacterium	V	D	T	R	F	F	I	G	L	S	E	M	I	H	R	F	
13. MVZ89370.1 malate synthase G Microbacter sp. ANSKLABC	V	D	P	A	F	W	A	G	A	S	E	I	F	T	E	L	
14. NCA68620.1 malate synthase G Sphingobacteriia bacterium\	I	D	P	D	Q	L	W	Q	S	F	G	E	M	V	R	E	L



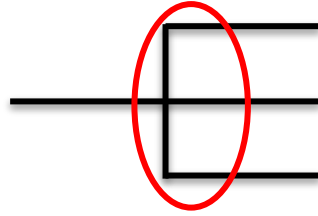
# ICL и MS, которые мы нашли

	ICL1	ICLz	ICLh	MSa	MSg	MSi	MSh	MSd
<b>Эукариоты</b>	+	+	-	+	+	+	-	-
<b>Бактерии</b>	+	-	-	+	-	-	-	+
<b>Археи</b>	-	-	+	-	-	-	+	-

1. У эукариот редко встречаются ICLz, MSg, MSi, сравнительно часто встречается ICL1 и MSa.
2. У многих эукариот не встречаются оба фермента одновременно.
3. Ферменты ICL1, MSa встречаются у бактерий повсеместно, а MSd в редких случаях.
4. У архей редко находятся эти ферменты, встречается изоформа ICLh, которой нет у бактерий.

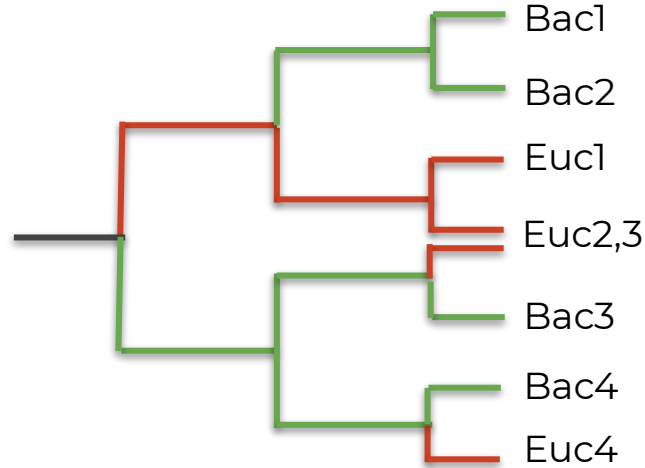
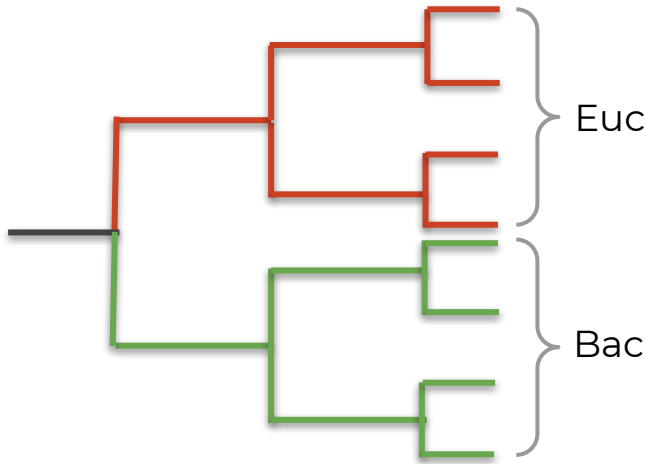


Эволюционное событие



Неразрешенность дерева

- Каждое разветвление на филогенетическом дереве означает одно эволюционное событие.
- Невозможно одновременное существование нескольких эволюционных событий.
- Обычно, белки организмов разных групп находятся отдельно друг от друга



Контаминация

VS

Горизонтальный перенос



# Контаминация

Контаминации - это загрязнения, которые обычно происходят при секвенировании. Например, иногда в ДНК организма при секвенировании может попасть фрагмент ДНК бактерии.

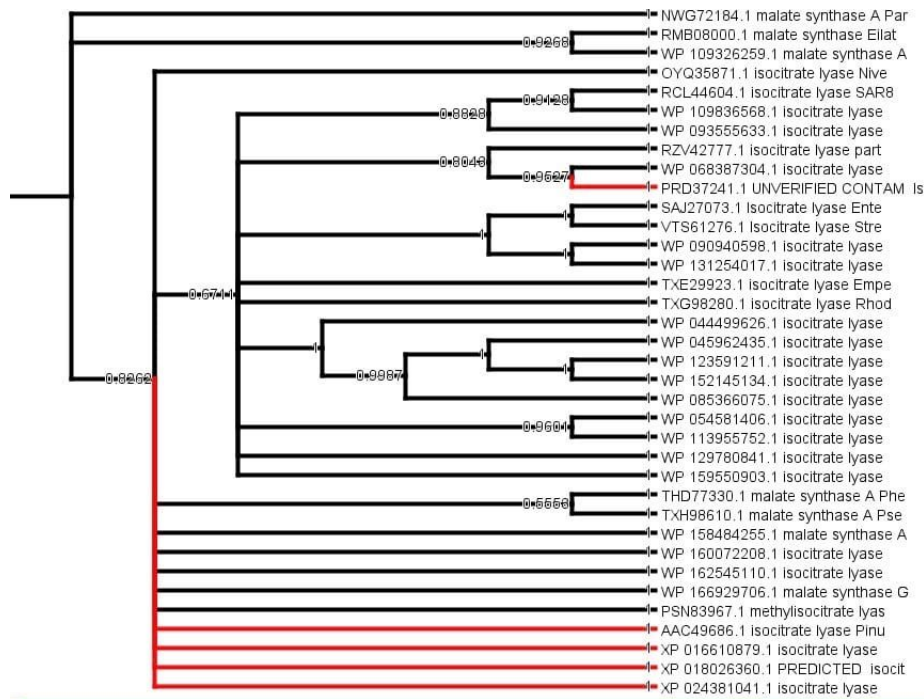
Мы боролись с ними на этапе работы с бластом, исключая подозрительные сиквенсы в виде:

- неполных геномов
- некрасивых и коротких выравниваний
- подозрительно хороших выравниваний

Для этого использовали методы обратного бласта и попарного сравнения, но некоторые контаминации всё же проникли в итоговую таблицу.

- бактерии  
 - эукариоты

# Пример контаминации



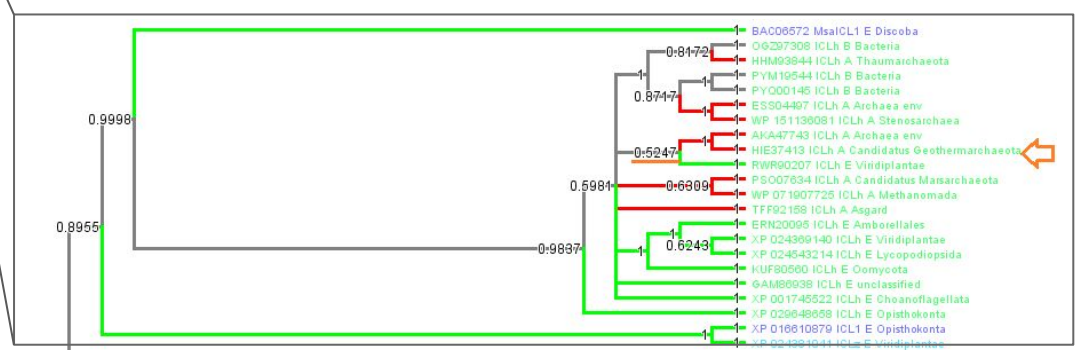
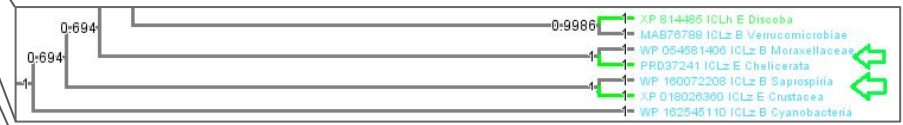
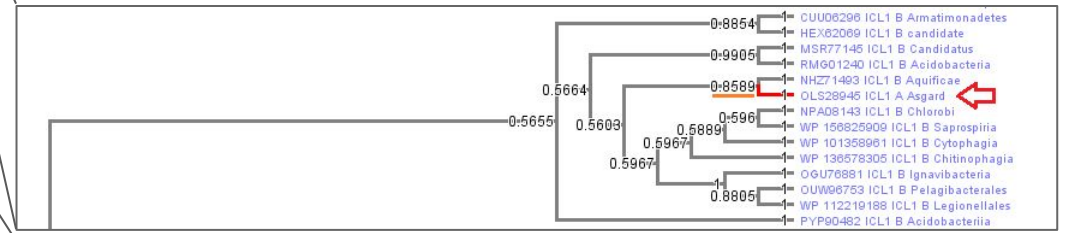
Но проверка попарным выравниванием подтверждает, что это контаминация и никакого горизонтального переноса мы тут не нашли

	Score	Expect	Method	Identities	Positives	Gaps
	768 bits(1982)	0.0	Compositional matrix adjust.	368/524(70%)	429/524(81%)	3/524(0%)
Query	9		HVRAVKEKLGGT--WGAIKRPEDAARMVQNRFHGTGLDIAKYTAAIMRKDMAEYDADSSKY H++A+ E G W AI PE AARM +QNR +GLDIA+YTAAMR+DMAEYD D+S+Y			66
Sbjct	7		HIKALTELAGKNPNWAAINPEYAARMRLQNRKFSGLDIARYTAAIMRRDMAEYDQDASRY			66
Query	67		TQSLGCWHGFIQAQQMIANKKYFGTTSKRYIYLSGWMVAALRSEFGPLPQDSMHEKTSVP TQSLGCWHGFIQAQQ+IA KK+ +T KRY+YLSGWM+AALRSEFGPLPQDSMHEK+VP			126
Sbjct	67		TQSLGCWHGFIQAQQLIATKKHQKSTEKRYLYSGWMTAALRSEFGPLPQDSMHEKTAVP			126
Query	127		KLIAETIYTLRQADAKELNDLFRALQKAEQAGDSAKAAEIIQSQINDFESHVVPITADIDA LI EY FLRQAD +EL+ LF AL KA AG++ + +++QI+ F++HVPITADIDA			186
Sbjct	127		ALIGETIYDFLRQADTRELDDLFSALDKARAAGETREETALLAQIEGFQTHVVPITADIDA			186
Query	187		FGFNEEATYLLTKQMIIEAGACAIQIENQVSDAKQCQGHQGVKVTYPHEDFLAKINAVRYAF FGFN EATYLL K+MIEAGAC IQIENQVSD KQCQGH QKVTYPH DFLAKINAVRYAF			246
Sbjct	187		FGFNAEATYLLAKKMIIEAGACCQIENQVSDKQCQGHQGVKVTYPHADFLAKINAVRYAF			246
Query	247		LELGDVDEGIVARTDSEADLTQKIPVSKKEGDLASQYISYLDTKEIDISEASDEILIK LELGDV+GVIVARTDS GA LT++I VS++ GDL QY +LD EI +E +E+I			306
Sbjct	247		LELGDVGGIVARTDSDLGAGLTKQYAVSEKPGDLGDQYNRFLDCDEIHPAEIQNGEVLTIS			306
Query	307		RDGKLRHPRTLASGLYQFREGTQHDRVLDVCTS LQNGADMIWIETPTPDVAGIAGFVND RDGKL RP RLASGL+QF++G+ DR VLDC+TSLQNGADMWIEET P + I G++			366
Sbjct	307		RDGKLRPKRLASGLFQFQKGSGEDRCVLDCTISLQNGADMWIEETEKPHIQIKGLMLDR			366
Query	367		IKKQVPNAKLVYNNPSPFNWTLNFRQQYDRWVAEGKDVSGYDRAKLMSAEYDNLSEAD I++ VPNAKLVYNNPSPFNWTLNFRQQ +D W AEG+D++ Y R LMSAEYD +EL			426
Sbjct	367		IRVFNPAKLVYNNPSPFNWTLNFRQQVDAWQAEGRDLAAYQRDALMSAEYDTELKGV			426
Query	427		ADEKIRTFQADAAREAGVFHHLITLPTTYHTAALSTHELAKYGFGEGLMAYVAGVQRKEI AD KI++FQ DA+R+AGVFHHLITLPTTYHTAALST L++YGF+ MLAYV GVQR+EI			486
Sbjct	427		ADAKIQSFQRDASRDAGVFHHLITLPTTYHTAALSTDLMSRGYFGSDAMLAYVKGVRQREI			486
Query	487		RGGIACVKHQAMAGSDIGDDHKEIFAGENALKAGDASKNTMNOF 530 R GIACVKHQ MAGSDIGD+HKE FAG ALKA NTH+QF			529
Sbjct	487		RRGIACVKHQMAGSDIGDHNHKEIFAGVALAKAS -GKDNTMHOQF 529			

На этом дереве видно, что эукариоты выделены отдельно от бактерий, но один организм находится около бактерий. На первый взгляд кажется, что это горизонтальный перенос генов

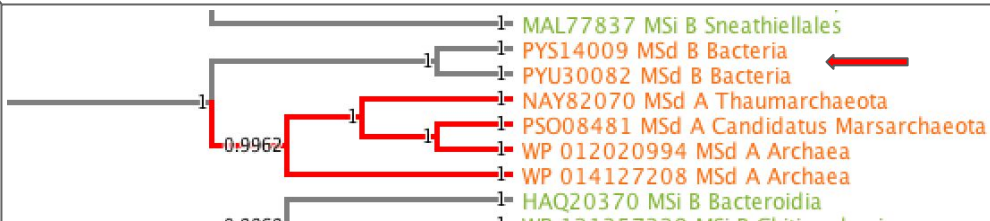
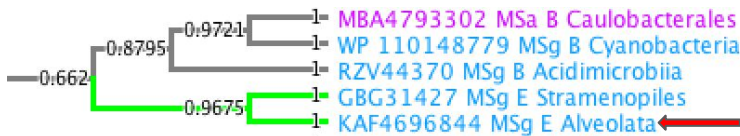
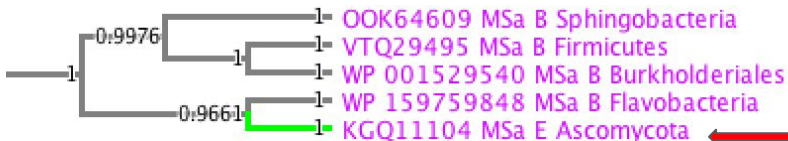
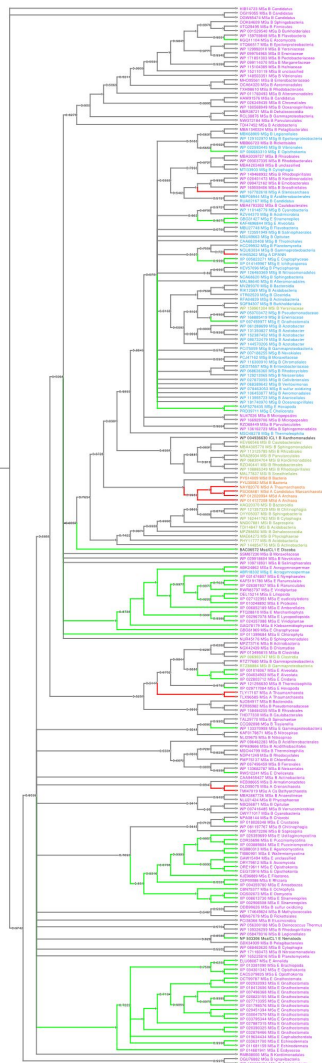
# ICL

Некоторые предварительные выводы по построенным деревьям



# MS

Некоторые  
предварительные  
выводы по построенным  
деревьям



# Заключение

1. Мы нашли две новые изоформы изоцитратлиазы и три новые изоформы малатсинтазы
2. Мы выяснили, что у большинства эукариот ферменты глиоксилатного цикла кластаризуются вместе с архейными. Это указывает на их происхождение от общего предка и последующую потерю во многих кладах эукариот.
3. Некоторые случаи подозрительно похожи как на случаи горизонтального переноса так и на контаминацию.
4. В некоторых случаях существование глиоксилатного цикла не доказывается наличием ферментов

# Над чем ещё можно работать?

1. Увеличить точность выравнивания и деревьев.
2. Доказать наличие глиоксилатного цикла экспериментально и ознакомиться со статьями об изученных случаях.
3. Изучить закономерности сочетания изоформ в разных организмах.
4. Корректировать результаты после появления новых данных о живых организмах.