Finding the DNA Gap in Arhodomonas Seminole Using DNA Sequencing

Logan Hukill, Jeff Woodruff, Sierra Tessmann, Hunterlyn Alderson, Dr. Patricia Canaan

ABSTRACT

Arhodomonas Seminole is an aerobic, halophilic bacterium enriched from salty, crude-oilimpacted soil in Seminole Co, OK. This is a newly found and researched bacteria. Samples of this bacteria were obtained through the instructor. This samples had been pre-ruptured, pre-purified, and pre-concentrated by the instructor so they were ready to be analyzed. These samples were then sequenced using Blastx at the NCBI website. After sequencing was completed, reeds were produced. By overlapping and correctly ordering these reeds, two contigs were formed and ready to be put together. Predicting adjacent contigs based on what is known in related species and other strategies were used. Ultimately, this was achieved by the PCR process. This process was focused solely on one small portion of the DNA. Designed PCR primers were then created and used to complete the PCR process. After PCR the copied DNA was viewed and examined by gel electrophoresis. After this process the final product was finally gathered, which is what will be presented to you here.

INTRODUCTION

Arhodomonas Seminole is a halophilic bacteria that thrives upon crude oil found in oil filed sites. We are examining this bacteria to fill the gap in its' DNA. We know proteins from similar bacteria that could fill the gap and have tested these to see if they will work for Arhodomonas Seminole. We are using DNA sequencing to determine where the gap is and how big it is. Also we need to find a related protein, from a similar bacteria, to fill the gap. Filling small parts of this gap will lead to filling the whole thing and then we will have the entire genome sequenced for Arhodomonas Seminole.

MAT	ERIALS AN	D METHODS			
dH2O (provided) 70 ü	Contig671_Con	tig465 (Arhodomonas seminole DNA)			
10X Taq buffer 10 Fused Co	ontig sequence 50 ng o	of DNA			
10X dNTP's (all 4) 10	Related Protein Sequence	0.6 μM of each primer			
Arh. sp. Seminole PCR seq	Arh. sp. Seminole PCR sequence 200 µM of PCR nucleotide mix				
gDNA (58 ng/ml) 5	Thermocycler 1.75 m	M MgCl2			
F-primer 2	Gel electrophore	esis Chamber 2.5 units Taq polymerase in Buffer A			
R-primer 2	Running Buffer	Running Gel			
Taq polymerase 1	The 100 µl PCR mix				
 3)Found the length of the gap in our DNA strand 4)A Fused Contig sequence was then used 200 characters to the left of the N's in the sequence and 200 characters to the right of the N's 5)The Fused Contig Sequence was then ran through BlastX and resulted in to many primers 6)The Forward Primer CGGCAGCACCATTCACT was chosen as well as the Reverse Primer TGATTTCCTTGTAGCCGATCC 7)2 microliters of each primer was used in the PCR we prepared using dH2O (provided) 70 ü, 10 microliters of 10X Taq buffer, 10 microliters of 10X dNTP's (all 4), Arh. Seminole, 5 microliters of gDNA(58ng/ml), 1 microliter of Taq polymerase. All substances were placed in a centrifuge tube. 8)We then placed a small mixture (roughly 5 microliters) of our PCR and dye into a DNA Running Gel. 9)We then used BlastX on the Fused Contig sequence with reference to the related protein to find what protein should fill in the gap; which was 2 keto-D-glucanate-dehydrogenase 10)Through the gel we had a working PCR product that gave us the sequence GTTGCCATCTCTTTCCTAGCTGATTGGCTGCCGAGAGAGA					

	RESULTS			
S Google ← → C □ www	x & Group2 - BIOC 1990-FRES x Dir IDT - PrimerQuest Details x		Through the N	CPI wabaita
	Order Menu Products & Services + Support & Education + Tools +	Search Go	of our fused fu	ill length cont
	Assay Design / Results	/ Help / About	unknown gap	and 200 amin
	Sequence 1 Assay Set 3 Details		against the rela	ated protein. 7
	Back to Results Parameter Set: General PCR (Primers only) Sequence Name: Sequence 1		the blast were	the forward a
	Amplicon Length: 275 Forward <u>CGGCAGCACCATTCACT (Sense)</u> 50 67	Length Tm GC% 17 62 58.8	our PCR react	ions. Primers
	Reverse <u>TGATTICCTTGTAGCCGATCC (AntiSense)</u> 304 325	21 62 47.6	needed for the	PCR reaction
	1 ACASTGAGAACCATCCCAATCTGCCGGCCTGGATGGTCAAGGGTGTCGG <u>CGGCAGCACCATTCACT</u> GGGCC1 101 GTTCCGGATGCGCACGGTCAACGGCGATATCGACGGGCCCAATCTCCTCGACTGGCCCATCAGCTATGACG 201 NNNNNNNNNNAGAAGAAGATGGGGGTGACCGGCACGAGGAGACCCGGGCTCCCATACCACCAGGGGACAA	JGGGTGGCCTGCGCATGCAGGAGCACGA AGCTCGCCCCTACTACGATCGCGCCGAĢ CTCGTTCAAGGTGCTTGCCGAGGGCGCGA	as the zipper to	o unwind the
	401 GSGCTGCCGC	SCETOCLAGCAGATUGGUTUTECATECA	replicate only	the portion of
		Ţ		Pagant
The p	rimers were only one of the mar	ny ingredients needed to c	complete	dH-O (provided
our P	CR reaction. Among the others a	are water that acted as the	solution	
for the	e reaction to take place, a buffer	solution that provides the	e optimal	
pH in	which the reaction needs to take	e place, several of all four	r amino	Arh. sp. Seminole o
acids	to help zip up the DNA in only the	the specific sections we a	re	(58 ng/□l)
Arhoe	lomonas Seminole as the backh	$n \in a$ and a uncertain and finance of the reaction and f	nally Tag	F-primer
polvn	herase that act like the glue that	helps the primers zin alor	ng the	R-primer
specif	fic segments of DNA.	I Present The Milling	0	<i>Taq</i> polymeras
-	_	~	-	Total
		Crp 2 Crp 2 Crp 5 Crp 8 Crp 14 Crp 14 Pos cont	As this conduction	s model shows eted after perf
			the mo reactio	re vivid the g n matched wi
			primer	s we found an
	-		on the	n the reaction far right to be
	the second s		numbe	r of amino ac
	Gel 1.2		far left	one is not she
One of	f the most important blast exerci	Gel 1.2 Ses our class	and the	e second farth
perform	med was one in which we blaste	the fused contig and	positiv	e control show
receive	ed the related protein 2 keto-glue	conate	★ Lecture Notes - BIOC 1990 × S NCBI BlastCont ← → C	tig671_taïi: ×
dehydi	rogenase. This protein allows us	to observe a	Download GenPept Graphics 2-keto-gluconate dehydrogenase	[Pseudomonas stutzeri]
nartici	llar section of Arhodomonas Ser	ninole's DNA and	Sequence UI: EffWP_014597276.11 L > Seq 1 more title(s) Range 11 2to 520 GenPagt Graphics	ength: 525 Number of Matches: 1
partice	ing to the blast the closest prote		B22 bits(2122) 0.0 Composition Query 89 ATFDVHDDsvivivgsagg ATF=4D0 V4/1+d55A66 Sbjct 2 ATFAQDD06VV105SA66	al matrix adjust. 389/524(74) 455/524(86%) 5/52 IgTLAYELAQKGVKTVVLEAGKRFTMADIENDEWAMFKKIS GTLA LA++++VVLEAGKR+TD IENDEWAMFKKIS GTLANALAKQKIRSVVLEAGKRHTLEDIENDEWAMFKKIS 61
accord	no is the mustain Olista -1	in to this particulate		PAWIVKGVGGSTIHWAGVALRMQEHEFRMRTVNGDIDGAN 448
accord sequer This c	ice is the protein 2 keto-glucona an allow us to make inferences a	te dehydrogenase.	Query 209 MLDRETAGSUMHSENHUM, 100 Sbjtt 62 NLDRETAGSUMHSENHUM, 100 Query 449 LLDRETASDHILLENHUM, 200 LDRETASCH 120 NLDRETASDHILLENHUM, 200	PANLTVKKVGGST+HAGA4LR ++ EFRNR++NGDI GAN PANLTVKKVGGSTVHAGA1LRFR0FFRNRSINGDIAGAN 121 XXXKKMGVTGSKETGLPYHQH+NSFKVLAGAKRIGYKEI K MOVTG TG-YGHYHHNSFKVLAGAKR-GYKEI KHNGVTGP-STGMYHYHPHNSFKVLAGAKR-GYKEI
accord sequer This ca functio	nce is the protein 2 keto-glucona an allow us to make inferences a on of this section of DNA and w	to this particulate te dehydrogenase. bout what is the hat could be the	Query 209 MLDWRTAGSUMHSEINMUN Sbjct 62 NLDWRTAGSUMHSEINMUN Query 449 LLDWRTAGSUMHLEINMUN Query 449 LLDWRTAGSUMHLEINMUN Sbjct 122 LLDWRTAGSUMHLEINMUN Query 449 LLDWRTVDELAPYDRAEQ Query 429 LLDWRTVDELAPYDRAEQ Query 809 IEHDDRGLVTGVVYADKDGN Query 809 IEHDDRGLVTGVVYADKDGN	PAUTUKSVUGST-HAUG-FALR ++ EFRMR-HADDI GAN PAUTUKSVUGST-MAGILAR RAP (EFRMR-HADDI GAN 121 XXXXXIAVAYTGSKETGLPHQAMHGFKVLAGARRADYKEI KMSVTG TG- YT HAMBIGFKVLAGARRADYKEI TG- YMGGC+GANKGT-YTDIPAR-SG(LEVAPS-SHL 10FCVGGCCHAANKST-YTDIPAR-SG(LEVAPS-SHL 10FCVGGCC+GANKST-YTDIPAR-SG(LEVAPS-SHL 10FCVGGCC+GANKST-YTDIPAR-SG(LEVAPS-SHL 10FCVGGCC+GANKST-YTDIPAR-SG(LEVAPS-SHL 10FCVGGCC+GANKST-YTDIPAR-SG(LEVAPS-SHL 2000A0)
accord sequer This ca function	nce is the protein 2 keto-glucona an allow us to make inferences a on of this section of DNA and wi l function of the bacterium if sci	to this particulate te dehydrogenase. about what is the hat could be the entists were able to	Query 269 MLDNRIAGSMHSEINMEN, NUDRIAC Sbitc 62 NLDNRIAGNHLENNNNL Query 449 LLDNPISYELAPYDRAEQ Sbitc 12 LLDNPISYELAPYDRAEQ Query 449 LLDNPISYELAPYDRAEQ Sbitc 12 LLDNPISYELAPYDRAEQ Query 809 TEHDDRGLYTGWYADQGT Query 809 TEHDDRGLYTGWYADQGT Query 809 TEHDDRGLYTGWYADQGT Query 809 VGNIYMTHTTAGYAALAHAPKK Sbitc 237 TEHDDRGKWWYADQGT Query 809 VGNIYMTHTTAGYAALAHAPKK Sbitc 237 VGNIYMTHTTAGYAALAHAPKK Sbitc 29 VGNIYMTHTAGYAALAHAPKK	PAUTUKGVGST-HHAG-ALR ++ EFRMR-HVGDT GAN PAUTUKGVGST-HHAG-ALR ++ EFRMR-HVGDT GAN 121 XXXXXIWAYTGSKETGLPYHQAHIGFKVLAGAKRATGVKEI KNGVTG TG + HV HMBSFKVLATGAKRATGVKEI KNGVTG - TG + HV HMBSFKVLATGAKRAUGYKEI 116FCNGGCC+GANGTHVTDTPRAESGCEURPSSH1 118FCNGGC+GANGTHVTDTPRAESGCEURPSSH1 106FCNGGCK+GANGT+VTDTPRAESGCEURPSSH1 106FCNGGCK+GANGT+VTDTPRAESGCEURPSSH2 106FCNGGCK+GANGT+VTDTPRAESGCEURPSSH2 106FCNGGCK+GANGT+VTDTPRAESGCEURPSSH2 106FCNGGCK+GANGT+VTDTPRAESGCEURPSSH2 106FCNGGCK+GANGT+VTDTPRAESGCEURPSSH2 106FCNGGCK+GANGT+VTDTPRAESGCEURPSSH2 106FCNGGCK+GANGT+VTDTPRAESGCEURPSSH2 106FCNGGCK+GANGT+VTDTPRAESGCEURPSSH2 106FCNGGCK+GANGT+VTDTPRAESGCEURPSSH2 106FCNGGCK+GANGT+VTDTPRAESGCEURPSSH2 106FCNGGCK+GANGT+VTDTPRAESGCEURPSSH2 106FCNGGCK+GANGT+VTDTPRAESGCEURPSSH2 106FCNGGCC+GANGT+VTDTPRAESGCEURPSSH2 106FCNGGC+GANGT+VTDTPRAESGCEU
accord sequer This ca function overall fill in 1	nce is the protein 2 keto-glucona an allow us to make inferences ε on of this section of DNA and w l function of the bacterium if sci nore of the over 700 missing ga	to this particulate te dehydrogenase. about what is the hat could be the entists were able to ps in the genomic	Query 269 MLDRETAGSUHHSEINUT Sbjtt 62 MLDRETAGSUHHSEINUT Query 449 LLDRFA-6 HLDRETAGSUHHSEINUT Query 449 LLDRFA-6 HLDRETSTOELANDHLAEINUT Query 449 LLDRFA-6 HLDRETSTOELANDHLAEINUT Query 629 KAGPMAINTERVISKA- Query Sbjtt 127 LSGPMAINTERVISKA- Query Query 809 TEHODRGLUTGVVVADKOM AGO Query 809 TEHODRGLUTGVVADKOM Sbjtt 237 Query 90 TEHODRGLUTGVVADKOM Sbjtt 237 Query 149 LSAFLOPFORGANSSAN Sbjtt 237 Sbjtt 237 LSAFLOPFORGANSSAN Sbjtt 357 Sbjtt 357 LSAFL PP GORGANSSAN Sbjtt 357 Sbjtt 357 LSAFLPP PORGRAMSSAN Sbjtt Sbjtt Query 148 HRHPHDTSTNERGPRANSCAN Sbjtt Sbjtt Sbjtt Sbjtt	PAUTUKSVUGST-HMAGA-LR ++ EFRWR-HWOT GAN PAUTUKSVUGST-HMAGA-LR +POETRWRSTHKOTGAAN 121 XXXXXIWAYTGSKETGLPYNQAHISFLVLAEGARAUTUKAE K MOYTG TG + TH JMHSTERVLAEGARAUTUKAE K MOYTG TG + TH JMHSTERVLAEGARAUTUKAE ITGFCNGCC-GAAUST-YTDTRAEGSGCCFVPESHALT 10FCNGCC-GAAUST-YTDTRAEGSGCCFVPESHALT 10FCNGCC-GAAUST-YTDTRAEGSGCCFVPESHALT 10FCNGCC-GAAUST-YTDTRAEGSGCCFVPESHALT 10FCNGCC-GAAUST-YTDTRAEGSGCCFVPESHALT 10FCNGCC-GAAUST-YTDTRAEGSGCCFVPESHALT 10FCNGCC-GAAUST-YTDTRAEGSGCCFVPESHALT 10FCNGCC-GAAUST-YTDTRAEGSGCCFVPESHALT 10FCNGCC-GAAUST-YTDTRAEGSGCCFVPESHALT 10FCNGCC-GAAUST-YTDTRAEGSGCCFVPESHALT 10FCNGCC-GAAUST-YTDTRAEGSGCCFVPESHALT 10FCNGCCH-GAAUST-STFFDGANSSG 988 40RQLARA VCVAORSTESPELLINGSSSFFPDGALNSSGQ 296 WMHRGTT CAGVSDESHMPERGFTGGYLEELLSLQVF 1168 WMHRGTT CAGVSDESHMPERGFTGGYLEELLSLQVF 1168 WMHRGTTCAGVSDESHMPERGFTGGYLEELLAGLPF 556 EVOPMSGWLCCGED.PESHNITLHAEKDQVGLPPV1 1348 EVYMHGSTLCGED.PESHNITLHAEKDQVGLPPV1 1528 KCYEAGGTVCFUPPVSAHMETRMSECFPDGVNKH KCYEAGGTV-LEPPVSAHMETRMSK & KDPONNKH
accord sequer This ca function overall fill in the sequer	nce is the protein 2 keto-glucona an allow us to make inferences a on of this section of DNA and w I function of the bacterium if sci more of the over 700 missing ga ice.	to this particulate te dehydrogenase. about what is the hat could be the entists were able to ps in the genomic	Query 269 MLDRETAGSUMHSENNUM Sbjct 2 MLDRETAGSUMHSENNUM Query 449 LLDRETAGSUMHSENNUM Query 449 LLDRETAGSUMHSENNUM Sbjct 122 LLDRETAGSUMHSENNUM Query 449 LLDRETAGSUMHSENNUM Query 429 LLDRETAGSUMHSENNUM Query 629 KAGMALITIKEVORGRAQQ Sbjct 127 LSGMALITIEKVORGRAQQ Query 809 IEHODRUTGWVARDIG Query 809 IEHODRUTGWVARDIG Query 129 SUMINITITAGAVANINGVVARDIG Sbjct 237 IEHOAKUNGVVARDIG Query 169 LSAFLUPSGNGRAVSAM Sbjct 27 SUMINITITAGAVANINGVARDIG Query 169 LSAFLUPSGNGRAVSAM Sbjct 377 LSAFLUPSGNGRAVSAM Query 1349 HORPD MRGH Sbjct 17 SUBRIVITITAGAVANING Query 1349 HORPD MRGH Sbjct <	PAUTUKGVGST-HHAG-ALR ++ EFRMR-HVBOT GAN PAUTUKGVGST-HHAG-ALR ++ EFRMR-HVBOT GAN 121 XXXXXIFVMGTGS-ETGLPYHQAMIGFKVLAEGARARGYKEI KMEVTG TG-YH HAMBSRVLATGARKNGVKEI TG-YMGGC-HGANKST-WTDIPAG-SGCLEVNPESHLT 136C+0000CH-AGNUST-WTDIPAG-SGCLEVNPESHLT 136C+0000CH-AGNUST-WTDIPAG-SGCLEVNPESHLT 136C+0000CH-AGNUST-WTDIPAG-SGCLEVNPESHLT 136C+0000CH-AGNUST-WTDIPAG-SGCLEVNPESHLT 136C+0000CH-AGNUST-WTDIPAG-SGCLEVNPESHLT 136C+000CH-AGNUST-WTDIPAG-SGCLEVNPESHLT 136C+000CH-AGNUST-WTDIPAG-SGCLEVNPESHLT 136C+000CH-AGNUST-WTDIPAG-SGCLEVNPESHLT 136C+000CH-AGNUST-WTDIPAG-SGCLEVNPESHLT 136C+000CH-AGNUST-WTDIPAG-SGCLEVNPESHLT 136C+000CH-AGNUST-WTDIPAG-SGCLEVNPESHLT 136C+000CH-AGNUST-WTDIPAG-SGCLEVNPESHLT 136C+000CH-
accord sequer This ca function overal fill in the sequer	nce is the protein 2 keto-glucona an allow us to make inferences a on of this section of DNA and w I function of the bacterium if sci nore of the over 700 missing ga ice.	in to this particulate te dehydrogenase. about what is the hat could be the entists were able to ps in the genomic	Query 269 MLDRETAGSUMHSENNUM Sbitc 62 MLDRETAGSUMHSENNUM Query 449 LLDRETAGSUMHSENNUM Query 449 LLDRETAGSUMHSENNUM Sbitc 122 LLDRETAGSUMHSENNUM Query 449 LLDRETAGSUMHSENNUM Query 429 LLDRETAGSUMHSENNUM Query 629 KAGPMAINTENDORACQO Query 809 IEHODRELIVORVADENCO Query 809 IEHODRELIVORVADENCO Query 989 VERNIMITTAGLANDVENDAGQU Query 989 VERNIMITTAGLANDVENDAGQU Query 149 LSAFLUPGSSGREASSAM Query 1429 <td>PAUTUKSUGST-HAGG-ALR ++ EFRIR-HAGT GAN PAUTUKSUGST-HAGG-ALR ++ EFRIR-HAGT GAN LINUXSUGST-MAGGLAR PARCERNRSINGDIAGN LINUXSUGST-MAGGLAR PARCERNRSINGPARCE (NOVTG TO +H HANSERVLA GALRAGYNEL 638 LINUXGG-STGNAUPHANSERVLA GALRAGYNEL 76 LISCYOQGCHAGANST-VTDIPARE 50 CEVPS-SIL LISCYOQGC+GANST-VTDIPARE 50 CEVPS-SIL LISCYOQGC+GANST-VTDIPARE 50 CEVPS-SIL LISCYOQGC+GANST-VTDIPARE 50 CEVPS-SIL LISCYOQGC+GANST-VTDIPARE 50 CEVPS-SIL LISCYOQGCHAGANSTLVTIPAREASOCEVPSOF (NOVAGATAVUANISTESPALLINGSS:FPDGLANSSO UNDYGGAAVC/VANISTESPALLINGSSCPUPUN UNHAGTI AGVCDEDPLESNITTI-HALKDQYGEPPUN UNHAGTI AGVCDEDPLESNITTI-HALKDQYGEPPUN LISKYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNHAGTI AGVCDEDPUN KKYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTING</td>	PAUTUKSUGST-HAGG-ALR ++ EFRIR-HAGT GAN PAUTUKSUGST-HAGG-ALR ++ EFRIR-HAGT GAN LINUXSUGST-MAGGLAR PARCERNRSINGDIAGN LINUXSUGST-MAGGLAR PARCERNRSINGPARCE (NOVTG TO +H HANSERVLA GALRAGYNEL 638 LINUXGG-STGNAUPHANSERVLA GALRAGYNEL 76 LISCYOQGCHAGANST-VTDIPARE 50 CEVPS-SIL LISCYOQGC+GANST-VTDIPARE 50 CEVPS-SIL LISCYOQGC+GANST-VTDIPARE 50 CEVPS-SIL LISCYOQGC+GANST-VTDIPARE 50 CEVPS-SIL LISCYOQGC+GANST-VTDIPARE 50 CEVPS-SIL LISCYOQGCHAGANSTLVTIPAREASOCEVPSOF (NOVAGATAVUANISTESPALLINGSS:FPDGLANSSO UNDYGGAAVC/VANISTESPALLINGSSCPUPUN UNHAGTI AGVCDEDPLESNITTI-HALKDQYGEPPUN UNHAGTI AGVCDEDPLESNITTI-HALKDQYGEPPUN LISKYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNHAGTI AGVCDEDPUN KKYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTINGSKCPUPUN UNCYEAGGAATUVILPPVPASHWCTING
accord sequer This ca function overall fill in the sequer	Algoments < Clustative - x Algoments < Clustative - x contraction of the over 700 missing ga nore of the over 700 missing ga nore.	in to this particulate te dehydrogenase. about what is the hat could be the entists were able to ps in the genomic	Query 269 MLDRETASSIHHSEINMU Sbjtt 62 MLDRETASSIHHSEINMU Query 449 LLDPF+veLPHVRL Sbjtt 122 LLDP+veLPHVRL Sbjtt 122 LLDP+veLPHVRL Query 629 KAOPMINTERVOR Sbjtt 127 LSGPNAINTERVOR Sbjtt 227 ISPNAINTERVOR Query 809 IEHORGLUTVVARDEN Sbjtt 227 ISPNAINTERVOR Sbjtt	PARTICKUGGST-HHARGALR, PAR EPRIME-HKOT GAN PARTICKUGGST-HHARGALR, PAR EPRIME-HKOT GAN PARTICKUGST-HHARGALR, PAR EPRIME-HKOT GAN LINCKUGST-HARGALR, PAR EPRIME-HKOT GAN KNOYT GT CT VH MENT GT STEMAHHARGALR, PAR EPRIME-HKOT GAN LINCKUGST-HARGALR, PAR EPRIME-HKOT GAN LINCKUGST-HARGALR, PAR EPRIME-HKOT GAN LINCKUGST-HARGALR, PAR EPRIME LINCKUGST, PAR EPRIME PAR EPRIME LINCKUGST, PAR EPRIME
particle accord sequer This ca function overal fill in the sequer $e \Rightarrow c www.$	Ace is the protein 2 keto-glucona an allow us to make inferences a on of this section of DNA and w I function of the bacterium if sci nore of the over 700 missing ga nce. * Algments < ClustalW2 * * * * * * * * * * * * * * * * * * *	in to this particulate te dehydrogenase. about what is the hat could be the entists were able to ps in the genomic	Query 269 MLDRETASSIHESEMENT bjtct 62 MLDRETASSIHESEMENT Query 449 LLDRETASCHILENPRALE Sbjtct 122 LLDRPH-teAPWTALE Sbjtct 127 LSGNPMALTERVORACQ Guery 629 KAGPMALTIERVORACQ Query 629 IHODSGLIVTOVVAKDE Sbjtct 277 LSGNPMALTERVORACQ Query 809 IHODSGLIVTOVVAKDE Sbjtct 277 LSGNPMALTERVORACQ Query 189 LSGNPMALTERVORACQ Query 189 ISHONGKIVADAGE Sbjtct 277 LSGNPMALTERVORACQ Query 189 ISHONGKIVADAGE Sbjtct 277 LSGNPMALTERVORACQ Query 189 ISHONGKIVADAGE Sbjtct 277 SGNPMALTERVORACQ Query 139 HIRDPHOTERREAVED Sbjtct 477 GSNPMALTERVORACQ Query 139 HIRDPHOTERREAVED Sbjtct 477 GSNPMALTERVORACQ TSNPMALTERVORACSGRETTS Sbjtct 477 GSNPDMALTSSSFT Sbjtct 477 GSNPDMALTSSSSFT Sbjtct 477 GSNPDMALTSSSSSFT Sbjtct 477 GSNPDMALTSSSSFT Sbjtct 477 GSNPDMALTSSSSSFT Sbjtct 477 GSNPDMALTSSSSST Sbjtct 477 GSNPDMALTSSSSSFT Sbjtct 477 GSNPDMALTSSSSSFT Sbjtct 477 GSNPDMALTSSSSS Sbjtct 477 GSNPDMALTSSSSST Sbjtct 477 GSNPDMALTSSSSST Sbjtct 477 GSNPDMALTSSSSST Sbjtct 477 GSNPDMALTSSSSS Sbjtct 477 GSNPDMALTSSSSST Sbjtct 477 GSNPDMALTSSSSST Sbjtct 477 GSNPDMALTSSSSST Sbjtct 477 GSNPDMALTSSSSST Sbjtct 477 GSNPDMALTSSSSSST Sbjtct 477 GSNPDMALTSSSSSS Sbjtct	PAULYKVIGST-HARAGALR +- EFRIK-HKOT GAN PAULYKVIGST-HARAGALR +- EFRIK-HKOT GAN III XMEVTG - TO +H MARGALR +- KERNERKOT GAN III KMEVTG - TO +H MARGALR +- EFRIK-HKOT GAN III KMEVTG - TO +H MARGALR +- EFRIK-HKOT GAN III IIICCONCENDANT III IIICCONCENDANT IIICCONCENDANT IIIICCONCENDANT IIIICCONCENDANT IIIICCONCENDANT IIIICCONCENDANT IIIICCONCENDANT IIIICCONCENDANT IIIICCONCENDANT IIIIICCONCENDANT IIIIICCONCENDANT IIIIICCONCENDANT IIIIICCONCENDANT IIIIICCONCENDANT IIIIIICCONCENDANT IIIIICCONCENDANT IIIIIICCONCENDANT IIIIIICCONCENDANT IIIIIICCONCENDANT IIIIIICCONCENDANT IIIIIICCONCENDANT IIIIIICCONCENDANT IIIIIICCONCENDANT IIIIIICCONCENDANT IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
particle accord sequer This ca function overal fill in the sequer $e \rightarrow e e e e e e e e e e e e e e e e e e$	Ace is the protein 2 keto-glucona an allow us to make inferences a on of this section of DNA and w I function of the bacterium if sci nore of the over 700 missing ga tee.	in to this particulate te dehydrogenase. about what is the hat could be the entists were able to ps in the genomic	Query 209 MLDRETAGSUMESEMENT Spict 62 HLDRETAGSUMESEMENT Query 409 LLDPTSDELAPUPMENT Spict 122 LLDPTSDELAPUPMENT Query 629 K-ADPUATITURY OR PACO Spict 127 LSPMAINTERVISE Query 1289 LSPL PP 6068 Spict 127 LSPL PP 6068 Query 1129 LSPL PP 6068 Spict 127 GSPENTIFICESORGAUSANT Query 1239 GRIPPhOTTERUPHONENT Spict 127 CSPENTIFICESORGAUSANT Spict 127 GSPENTIFICESORGAUSANT Spict 127 GSPENTIFICESORG	PAILTVKVGST-HAAGA-LR -+ EPRIN-HADE GAN III XXXXXIVVGST-HAAGA-LR -+ EPRIN-HADE GAN IIII XXXXIVVGST-FELPYNDANGE/VLAGARG-WY III XMVTG TO +Y HANG-KARGYKEI IIIC XMVTGTO-AUXIVIGST-FELLINGSGC/EVRPS-FAL IIICC/XXXXIXISTYTDIPAE 5G CEVRPS-IL IIICC/XXXXIXISTYTDIPAE 5G CEVRPS-IL IIICC/XXXXIXISTYTDIPAE 5G CEVRPS-IL IIICC/XXXXIXISTYTDIPAE 5G CEVRPS-IL IIICC/XXXXIXISTSFRLLINGSSCFFDGIANS5G IIIICC/XXXXIXISTSFRLLINGSSFFDGIANS5G IIIICC/XXXXIXISTSFRLLINGSSFFDGIANS5G IIIIICC/XXXXIXISTSFRLLINGSSFFDGIANS5G IIIIIICC/XXXXIXISTSFRLLINGSSFFDGIANS5G IIIIIICC/XXXXIXISTSFRLLINGSSFFDGIANS5G IIIIIIICC/XXXXIXISTSFRLLINGSSFFDGIANS5G IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
particle accord sequer This ca function overal fill in the sequer $(contiger)_{ta}$ $(contiger)_{ta}$ $(contiger)_{ta}$ $(contiger)_{ta}$ $(contiger)_{ta}$ $(contiger)_{ta}$ $(contiger)_{ta}$ $(contiger)_{ta}$ $(contiger)_{ta}$ $(contiger)_{ta}$	Ace is the protein 2 keto-glucona an allow us to make inferences a on of this section of DNA and w I function of the bacterium if sci more of the over 700 missing ga nce. ***********************************	in to this particulate te dehydrogenase. about what is the hat could be the entists were able to ps in the genomic	Query 269 MLDRETAGSUHSEENHU Spite 62 MLDRETAG HI + 81H-NL Query 409 LLDRETAGSUHSEENHU Query 409 LLDRETAGSUHSEENHU Query 629 KAGPMAINTEVYGBACQ Query 629 KAGPMAINTEVYGBACQ Query 629 FEROGGUT VOXOCQ Query 629 FEROGUT VOXOCQ Query 629 FEROGGUT VOXOCQ Query 629 FEROGGUT VOXOCQ Query 629 FEROGUT VOXOCQ Query 629 FEROGGUT VOXOCQ Query 629 FEROGUT VOXOCQ Query 629 FEROGUT VOXOCQ Query 629 FEROGGUT VOXOCQ Query 629 FEROGUT VOXOCQUE VOXOCQ	PARTYLVSVGST-HABGE-ALB +- EFRING-MODI GAN PARTYLVSVGST-HABGE-ALB +- EFRING-MODI GAN 121 2000XIV/VGVGST-HABGE-ALB +- EFRING-MODI GAN 2000 2000 2000XIV/VGVGST-EFRICAL HABGE-ALB +- EFRING-MODI GAN 2000XIV/VGVGST-EFRICAL HABGE-ALB +- EFRING-MODI GAN 2000 2000XIV/VGVGST-EFRICAL HABGE-ALB +- EFRING-MODI GAN 2000XIV/VGVGST-EFRICAL HABGE-ALB +- EFRING-ALB +- EFRING-MODI GAN 2000XIV/VGVGST-EFRICAL HABGE-ALB +- EFRING-MODI GAN 2000XIV/VGVGST-EFRICAL HABGE-ALB +- EFRING-ALB +- EFRING-MODI GAN 2000XIV/VGVGST-EFRICAL HABGE-ALB +- EFRING-ALB +- EFRING +- EFRING-ALB +- EFRING-ALB +- EFRING-ALB
Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta	Ace is the protein 2 keto-glucona an allow us to make inferences a on of this section of DNA and w I function of the bacterium if sci more of the over 700 missing ga ace.	in to this particulate te dehydrogenase. about what is the hat could be the entists were able to ps in the genomic	Query 209 LDRETAGSUMESEMENT Spite 62 HUDRETAGSUMESEMENT Query 409 LUDPTSOFELAPORTAR Query 409 HUDPTSOFELAPORTAR Spite 127 LSPHATHER Query 409 HUDPTSOFELAPORTAR Spite 127 LSPHATHER Query 109 KAGONANTITAG-A-AUPR Spite 127 LSPHATHER Query 1199 LSAFLOPPORTROPASEDT Spite 127 LSPHATHER Query 1199 LSAFLOPPORTROPASEDT Spite 127 SSPHATHER Query 1129 CARPONTROPASEDT Spite 127 SSPHATHER Query 1129 CARPAR Spite 127 SSPHATHER Spite 127	PARTYCKKIGST-HARAGARA +- EFRIK-HADI GAN PARTYCKKIGST-HARAGARA +- EFRIK-HADI GAN III XIXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta	Acce is the protein 2 keto-glucona an allow us to make inferences a on of this section of DNA and w I function of the bacterium if sci more of the over 700 missing ga nce.	in to this particulate te dehydrogenase. about what is the hat could be the entists were able to ps in the genomic	Query 289 IDDRESSIGNMESTING IN + ENHANCE Spite 62 IDDRESSIGNMESTING IN + ENHANCE Query 490 IDDRESSIGNMESTING IN + ENHANCE Query 400 IEED 65 (* 100000 VT00000 VT000000 VT000000 VT000000 VT00000 VT00000 VT000000 VT00000 VT00000 VT00000 VT00000 VT00000 VT000000 VT00000 VT00000 VT00000 VT000000 VT00000 VT00000 VT00000 VT00000 VT00000 VT00000 VT000000 VT00000000	PARTY WORKST HARAGEAR AT A FURTHER HARD I GAN PARTY WORKST HELPYNAMEALAR FAR PERMISSION I 121 CONSERVING STATEMENT AND A CONSERVING I 121 CONSERVING STATEMENT I 176 I MEY GIT TO HILL HARD STATEMENT I 176 I TO HILL AND A CONSERVING I 126 I TO HILL AND A CONSERVING I 127 I TO HILL AND A CONSERVING I 127 I TO HILL AND A CONSERVING I 128 I TO HILL AND A CONSERVING I 128 I TO HILL AND A CONSERVING I 128 I TO HILL AND A CONSERVING I I I TO HILL AND A CONSERVING I I I TO HILL AND A CONSERVING I I I I I I I I I I I I I I I I I I
Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta	ace is the protein 2 keto-glucona an allow us to make inferences a on of this section of DNA and w I function of the bacterium if sci more of the over 700 missing ga ace.	In to this particulate te dehydrogenase. about what is the hat could be the entists were able to ps in the genomic	Query 209 MDREFAGEMENT AND HE SEMAN DURING AND HE SEMAN Query 400 ILDEPTSOEL APPRAINE Query 620 KAOPMENT AND HE SEMAN Solid 27 ILDEPTSOEL APPRAINE Query 620 KAOPMENT AND HE SEMAN Solid 27 ILDEPTSOEL APPRAINE Query 620 KAOPMENT AND HE SEMAN Solid 27 ILDEPTSOEL APPRAINE Query 100 ILDEPTSOEL APPRAINE Query 1109 LASTLEPTSOELAND HE SEMAN Query 1109 LASTLEPTSOELAND HE SEMAN Query 1109 LASTLEPTSOELAND HE SEMAN Query 1109 ILSEMAN HE SEMAN Query 1109 ILSEMAN Query 1109 ILSEMAN Query 1109 ILSEMAN Query 1109 ILSEMAN Query 1109 ILSEMAN Query 1109 ILSEMAN QUERY 1100 ILSEMAN QUERY 11	PARTY WORKST CHARACTER AT A FURTHER WORK IN I III
Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta Contig671_ta	ace is the protein 2 keto-glucona an allow us to make inferences a on of this section of DNA and w I function of the bacterium if sci more of the over 700 missing ga ace.	in to this particulate te dehydrogenase. about what is the hat could be the entists were able to ps in the genomic	Query 200 IDDRIAGSUMESEMENT Spite 21 IDDRIAGSUMESEMENT Query 400 IDDRIAGSUMESEMENT Spite 127 ISSMENT Query 400 IEBORGY TOWNROUGH Spite 127 ISSMENT Query 100 IEBORGY TOWNROUGH Spite 127 ISSMENT Query 1100 IEBORGY TOWNROUGH Spite 127 ISSMENT Query 1100 ISSMENT Spite 127 ISSMENT Query 1200 ISSMENT Spite 477 ISSMENT Spite 477 ISSMENT Spite 477 ISSMENT ISSMENT	PARTICINUSS THAGA AND AT LEPRING AND 111 XXXXXIVVXGST THAGA AND AND AND AND AND AND AND AND AND AN

ere able to blast a portion O amino acids before the Is after the unknown gap sults we received from verse primers we used in esential to the ingredients e successful. Primers act e helix of the DNA to we are trying to sely.

olume (□l)	Tube Color	Added
70		
10		
10		
5		
2		
2		
1		
100	-	-

electrolysis which we ag our PCR experiments, apple the better the PCR forward and reverse mixtures we made to re is a molecular standard pared in terms of the he two columns on the his is the negative control the left column is the ry vividly.



lded a more detailed sequence when we nst the forward reaction. By come before and after determine the ugh related proteins

DISCUSSION

Arhodomonas Seminole is a bacterium found initially in Seminole County whose function is unclear due to the incomplete genomic sequence. It was the task of Biochemistry 1990 to continue the work of our professor Dr. Canaan in her quest to fill in these missing DNA fragments. Through processes such as PCR and gel electrolysis, four out of the fourteen groups were able to receive forward and reverse primers from their work, or simply one of the two primers. From our mentors going over the remaining groups work three more groups' revealed primers from the PCR and electrolysis experiments. Although Dr. Canaan believes there may be over 700 gaps in Arhodomonas Seminole's DNA sequence, the fourteen groups that participated in this study did not close the case but helped to further her work. Simply because primers were found does not mean they fill the gap but through this we were able to conduct more blast exercises from the NCBI website to find the proteins that most closely relate to this gap. By finding proteins that relate to DNA sequences in Arhodomonas Seminole we can better identify its end function which is the overall goal of this project. Our group identified the protein 2 Keto D Gluconate dehydrogenase another common name gluconate 2-dehydrogengenase and it acts on the CH-OH group of donors. Facts like these can help us identify, can this bacterium be helpful to humans in some way? Is it potentially harmful? Arhodomonas Seminole was found near oil drilling sites and the initial function was thought to be this bacterium could eat oil remnants out of the soil, a toxin that is not properly removed any other way. If when the genomic sequencing is done and it proves this could be true of Arhodomonas Seminole, scientists are then presented with what kinds of soil can it thrive in and what levels will prove too toxic? The work is just never finished.

REFERENCES

•1."Basic Local Alignment Search Tool." Blastx: Search Protein Databases Using a Translated Nucleotide Query. N.p., n.d. Web. 11 Nov. 2014.

•2."American Society for MicrobiologyApplied and Environmental Microbiology." Isolation of a Novel Arhodomonas Sp. Strain Seminole and Its Genetic Potential to Degrade Aromatic Compounds at High Salinity. N.p., n.d. Web. 11 Nov. 2014.

GRP#2 #701, 01