Proposal for CPET CATALySES Program:

Using a Onlines Databases to Understand Proteins, DNA, and Phylogenetic Relationships Among
Pathogenic Organisms

Matthew Nelson

Biology Teacher

Barron Collier High School

5600 Cougar Drive

Naples, Florida 34109

nelsonma@collierschools.com

Abstract:

Advanced Placement Biology has a curriculum that is built around making connections with content and organizing, analyzing, and synthesizing those connections within the framework of four big ideas. This activity seeks to help students do that by using pathogenic organisms as a vehicle to explore biotechnology, evolutionary relationships, and the central dogma of genetic information transfer. This activity will link information across three of those ideas and tie together many science practices, while simultaneously raising student awareness about infectious diseases. In order to do this, students will apply information about infectious disease pathogens to develop a hypothesis about their phylogenetic relationships and use this to construct a phylogenetic tree. They will utilize their knowledge of shared cellular processes, together with several web based bioinformatics tools to discover more specific information about gene and protein sequencing and use that information to analyze their hypotheses.

Rationale:

Students in Advanced Placement Biology often learn content in isolation, particularly when that content is not expanded upon in the context of laboratory experimentation. A practical, inquiry-based approach to assimilating content has been shown to make learning interesting, challenging, and enjoyable for students (Hofstein, 2004). Furthermore, there is growing research that argues the utility of teaching through inquiry to develop critical student thinking, or minds-on learning, and not just the physical part of learning, or hands-on learning (Abrahams & Millar, 2008) (White, Tasker, & Herrenkohl, 2011).

In order to deepen understanding of proteomics within the context of transcription and translation, as well as to place it into the context and application of phylogenetic relationships, this project will get students to apply information about these concepts to the formulation of a hypothesis about organismal relationships and to synthesize information in the construction of a phylogenetic tree in order to support or reject aspects of their hypothesis. Data integration is becoming more difficult as we catalog an ever expanding amount of biological data. A constantly expanding number of databases are distributed across the internet (Schneider & Jimenez, 2012). Learning how to use these databases and understand the importance of sharing information in science is critical to a good science education. Having a better understanding of data integration and its application to bioinformatics can help students, biologists, and software engineers make the most of their data. Good integration facilitates data sharing between labs, resulting in decreased costs of unnecessary duplication of experiments and integration of multiple data sources helps increase confidence in results if consensus is shown by different experiments (Schneider & Jimenez, 2012).

Intervention:

In order to determine prior knowledge about overarching concepts and the understanding of necessary details, a pre-test will be administered upon commencement of the unit, followed by a post-test at the completion of the unit. Scores will be compared from pre-test to post-test using descriptive and inferential statistics.

Following the pre-test, I will begin the unit with a short video introduction to infectious disease from NIH: Infectious Disease, Then and Now (National Institutes of Health, 1999). Students will complete a 'Causes of Death Quiz' published by NIH to accompany the video (National Institutes of Health, 1999), and then get into groups to classify several diseases as emerging, re-emerging, or endemic. This will give students a good introductory background to infectious diseases as:

- a continuing major cause of human suffering and death around the world.
- the understanding that emerging diseases are diseases that have either not occurred in humans before or have occurred only in small numbers in isolated places.
- re-emerging infectious diseases were once major health problems and then declined dramatically, but are again becoming significant problems for large parts of the population.

Following the completion of this activity, student groups will be given a list of bacterial diseases, along with background information about their cause, vector and, in some cases, the host. They will be asked to hypothesize their evolutionary relationships to one another. Given a general diagram of a phylogenetic tree, they will then have to produce one based on their hypothesis.

Next, as a whole class discussion, I will review the use of the protein phosphofructokinase in the universal process of glycolysis in the cells of organisms and discuss reasons for both its abundance and its prevalence. Together with a discussion about how a common protein can be used for amino acid sequencing and determining common ancestry, this discussion will serve as a starting point to model the use of the sequence query engine at www.uniprot.org (Consortium, 2002-2018), where I will lead them through a search of a disease causing microorganism (Bacillus anthracis) to find a common protein using ATP-dependent+6-phosphofructokinase as an example.

Groups will then be asked to identify the protein's coding gene and use BLAST to determine its sequence. They will be asked to use the protein sequencing data to analyze their hypothesis about evolutionary relationships. The expected outcome will be to connect amino acid sequencing back to morphological, epidemiological, and circumstantial evidence to either support or reject and modify portions of their phylogenetic tree.

In order for students to make the connection to the central dogma of information transfer in cells (DNA → RNA → protein), student groups will then be asked to use Uniprot to find information about the gene that codes for phosphofructokinase and then use MABL (Marseille-Nice genopole, 2003) to generate a phylogenetic tree using the protein sequence information. They will compare this tree (see Figure 1 below) to their hypothetical tree. A list of the protein sequences for the assigned organisms is shown below in Figure 2, followed by a key for the letters and their corresponding amino acids in Figure 3.

As an interesting extension, student groups will then use genemania.org to investigate gene interactions. Jalview will be used to see common conserved areas to determine the locations of active sites and PyMol viewer will be used to see 3-D rendered proteins. These extension activities will add visual context to student understanding and will help solidify student understanding of the importance of structure in the functioning of proteins.
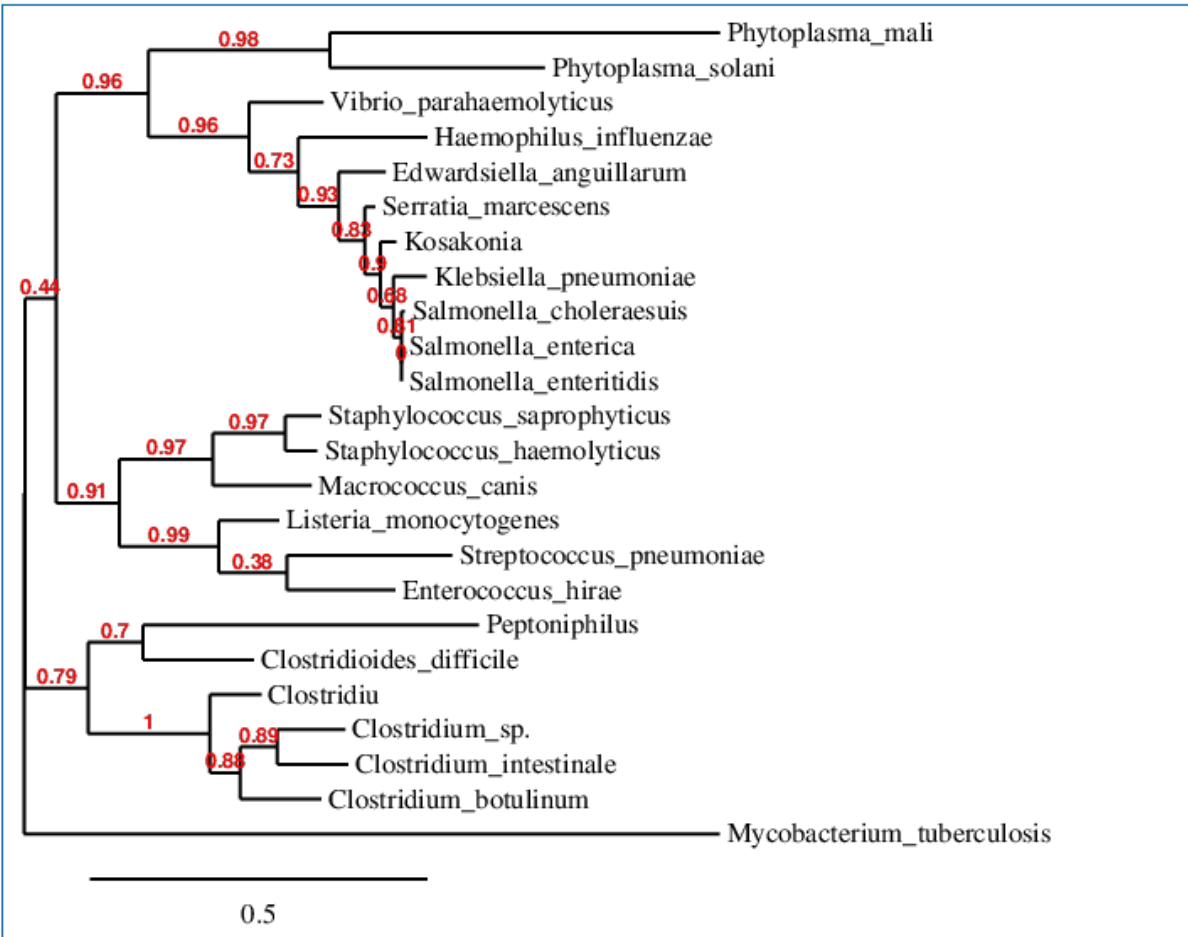
Figure 1:

Tree Rendering results

Figure 1: *Phylogenetic tree.*

Figure 2:

List of gene sequences for the protein phosphofructokinase in various pathogenic organisms:

>Haemophilus_influenzae
MNAAIRGVVRSALAEGLEVFGIYDGYQGLYNNKIKQLNRYSVSDVINRGGTFLGSARFPE
FKDPNIRAKCAEILRSHGIDALVVIGGDGSYMGAKLLTEEHCFPCVGLPGTIDNDVAGTD
YTIGYQTALQTAVDAIDRLRDTSSSHQRISIVEIMGRHCSDLTISAGIAGGCEYIVASEI
EFNREELIQQIERSIIRGKRHAIIAITELLTDVHSLAKEIEARVGHETRATVLGHIQRGG
SPCAFDRILASRMGAYAVDLLLQGKGGYCVGIQNEQLVHHDIIDAINNMQRVFKADWLKV
AKRLE

>Staphylococcus_saprophyticus
MNAAVRAVVRKAIYNNIEVYGVYQGYQGLLNDDIEKLELGSVGDTIQRGGTFLYSARCPE
FKEVEVRKKGIENLRKRGIEGLVVIGGDGSYRGAQRISEECPEIQTIGIPGTIDNDINGS
DFTIGFDTALNTIIECVDKIRDTASSHARTFIIEVMGRDCGDLALWAGLSVGAETIIVPE
VETDIKDIAEKIDNGIKRGKKHSIVMVAEGCMSGEVCAEELTKYINVDARVSVLGHIQRG
GSPSGADRVLASRLGGHAVDLLLDGKSAIGVGIRNNELTDTPFDEIFRSHEHEFDFETYA

LTNELSI

>Staphylococcus_haemolyticus
MNAAVRAVVRKAIYNNIEVYGIYQGYQGLLDDNIHKLELGSVGDTIQRGGTFLYSARCPQ
FKEASVRQKGIENLHKRGIEGLVVVGGDGSYRGAQRISEETKDIKTIGIPGTIDNDINGT
DFTIGFDTALNTIIDSVDKIRDTASSHARTFIVEVMGRDCGDLALWAGLSVGAETILVPE
ANTDIKDIAEKIESGIKRGKKHSIVMVAEGSMSGQECADQLMKYINVDARVSVLGHIQRG
GSPSGADRVLASRLGGYAVDLLMQGRSGQGVGIMNNKLTATPFDEIFNGNDHKFNTDIYE
LAKELSI

>Mycobacterium_tuberculosis
MVGFQNGFRGLLENRRVQLHNDDRNDRLLAKGGTMLGTARVHPDKLRAGLPQIMQTLDDN
GIDVLIPIGGEGTLTAASWLSEENVPVVGVPKTIDNDIDCTDVTFGHDTALTVATEAIDR
LHSTAESHERVMLVEVMGRHAGWIALNAGLASGAHMTLIPEQPFDIEEVCRLVKGRFQRG
DSHFICVVAEGAKPAPGTIMLREGGLDEFGHERFTGVAAQLAVEVEKRINKDVRVTVLGH
IQRGGTPTAYDRVLATRFGVNAADAAHAGEYGQMVTLRGQDIGRVPLADAVRKLKLVPQS
RYDDAAAFFG

>Phytoplasma_solani
MKMKNKKEIKKIAVLTSGGDAPGMNAAIRAVVLEGNKQGFEVYGVKDGYLGLYKNEIQLL
EPNSLPRIINISGTFLGTSRFIPFQQDLSIRQQCANNLKKLGIEKLVVIGGDGSYRGAMR
LEEIGIKCVGLPATIDNDINETDFTIGFSTALSNVVDAIEKLRDTSISHRRCSIIEVMGR
HKGDLALYGGVAAGVDLIITPENFLDKQKIFDKIKELNEKKQRHAIVVVTEHLFDVFSLA
KEVELNSGFETRAQILGHIQRGGKPTAEDLVLAARMGSFAVQLLQKNISNCGVCLKGLKL
LDVDFKEIFNSK

>Streptococcus_pneumoniae
MKRIAVLTSGGDAPGMNAAIRAVVRQAISEGMEVFGIYDGYAGMVAGEIHPLDAASVGDI
ISRGGTFLHSARYPEFAQLEGQLKGIEQLKKHGIEGVVVIGGDGSYHGAMRLTEHGFPAI
GLPGTIDNDIVGTDFTIGFDTAVTTAMDAIDKIRDTSSSHRRTFVIEVMGRNAGDIALWA
GIATGADEIIIPEAGFKMEDIVASIKAGYECGKKHNIIVLAEGVMSAAEFGQKLKEAGDT
SDLRVTELGHIQRGGSPTARDRVXILEQFKGPIPVIIRYEEEQKTIVSPHHFVAKSNELE
EKLNEIVMKTIYR

>Peptoniphilus
MRLAILTSGGDAPGMNATIRAVTRTCIFEGIEVYGINGGYEGLIDAKFEKLVQSSVADII
QRGGTILSTSRSGRFMTEEGLEKAINSLKIFEIDALIVLGGDGSLKGAKKLRDRGITVIG
IPCSIDNDLAYTDFTIGFMTAVETVTEAISHIRDTTDAHCRANVVEVMGRECGDIALYSG
VSSGAESIIVPEIETNINEIAEKALMGKNRGKRHHIIIMAEGCGSAFDFAKDFQDLTGIE
TRVTVLGYIQRGGAPSIFDRILASQFGYLAVKEAKKGNSRVLGFKDGRAMTMTFEEMFEA
KKVFREDLLDILKTVSI

>Clostridium sp.
MKKIAVLTSGGDAPGMNAAIRAVVRTALYNGIEVMGVQRGYSGLINGELFKMDRNSVSDI
IQRGGTILRTARCPEFKNEEVRKKAAKILQAYGIEALVVIGGDGSFTGAKLLSKLGIKTI
GLPGTIDNDLAYTDFTLGFDTALNTVVDAINKIRDTSTSHERVSIVEVMGRDCGDLSLYA
GICGGAEAIIIPEMEFDRDELCKTILEGKNKGKMHNLIILAEGIGGAFELAKYVEEITGI
ETRATILGHIQRGGSPSATDRLLASRMGAKAIEVLLEGKTSRVIGIRNNQIIDQDIDEAL
DMESKFDKELYDVAQILS

>Clostridium_botulinum
MKTIAVLTSGGDAPGMNAAIRAVVRTGLDKGLKVMGIQRGYSGLINGEIFEMDRQSVADI
IHRGGTILRTARCEEFKTEAGRKKGVGILKAFGIDGVVVIGGDGSFQGAQLLSKLGVKTI
GIPGTIDNDLAYTDYTIGFDTATNTVLDAINKLRDTSSSHERVSIVEVMGRCGDLALFA
GIGGGAESVIVPEKEFNEDELCKKILEGKLRGKLHNLIILAEGVGGGEALTKKVEEVTGI

```
QTRLTTLGHIQRGGSPSAFDRVLACRLGAKAVELLIEGKSSRVVGLRNNKVVDDDIDEAL
SMESKFDDDLYDIAEILSY

>Clostridioides_difficile
MKTIGLLTSGGDAPGMNAAIRAVVRSAIYYGCKVYGINRGYKGLLEEDLTEMNLSSVGDI
IHRGGTILKSSRCEEFKTEEGRLKAVKILKKYKIDCLVVIGGDGSFAGAQKLSDLGFPAI
GIPGTIDNDLAYTDYTIGFDTAMNTIIDAIGKIRDTSSSHERVNIVEVMGRHCGDLALYA
GLAGGAETIIVPEVEITVDEVALRLKTTQKRGKRHSIIVLAEGVGSASDLEKELKKESGA
DLRVTVLGHVQRGGSPTVSDRILASRLGVRAVELLLDGKSARVVGIKENKIIDLEISEAL
AQKKVFDKEAYEMAKILSI


>Clostridium_botulinum
MRTIAVLTSGGDAPGMNAAIRAVVRTGLEKGLKVMGIQRGYNGLINGEIFEMDTHSVSDI
IQRGGTILRTARCEEFRTEQGREKAAKILKAFGIDGLVVIGGDGSFHGAQLLSKLGINTV
GLPGTIDNDLAYTDYTIGFDTSINTVLDAINKLRDTSTSHERVSVVEVMGRNCGDIALYT
GVAGGAESIIIPEKEYNADKLCKQILQGKLKGKMHNLVLLAEGVGGANELAKYIEEVTGI
ETRSTILGHIQRGGSPTCMDRILASRMAYKAVELLISGKSSRVVGIKNGEIIDMDIDEAL
AVERSFDQELYDIATILSK
>Listeria_monocytogenes
MKRIAILTSGGDAPGMNAATRAVVRKAIYEGLEVYGINYGFLGLVNGDIRKLELGSVGDL
LHRGGTFLYSARYPEFATEEGQLKGIEQLKKHQIDGLVVIGGDGSYHGAEALTKRGFPTI
GIPGTIDNDISGTDFTIGFDTALNTVLDALDKIRDTATSHERTFIIEVMGRDAGDIALWS
GLAGGAEAIIVPEESFNMDDVVDRLNKGRERGKKHSIIVVAEGVMSGNEFAKQLAEYGDY
HARVTVLGHVQRGGSPTAFDRVLASRLGARSVELLLENRGGLAVGIRENRIVENNIGEIL
KEKHTLDQKLFDLASILSI


>_Clostridium_intestinale
MKKIAILTSGGDAPGMNAAIRAVVRTALHNGLEVMGVERGYAGLISGELHPMDRRSVSDI
IQRGGTVLRTARCNEFKTEEGREKGAKILKAYGVDALVVIGGDGSFTGAKLLSKLGVKTV
GLPGTIDNDLAYTDYTIGFDTALNTVLDAINKLRDTSTSHERVSVVEIMGRNCGDLALYA
GLAGGAESIVVPEKGFDKDELCKTILEGKVKGKMHNLILVAEGVGGAEELASYIEEVTGI
ETRATILGHIQRGGSPSASDRVLASKLGHRAVEVLLEGKTARVIGIRNNQVIDVDIDEAL
AIERKFDEKLYDIATEISS

>Phytoplasma_mali
MNIAVLTSGGDSPGMNAAIRSIVKVGIQNNISIYGIKYGYLGLYQNQMDLLKLEDVENII
RMGGTFLGTSRFIEFIENTNIIQKSVQNLKNKDISVLIIIGGNGSYKGALKLIDFGVKCI
GIPGTIDNDISGTEFTIGFHTALNTIVESIDKIKDTSKSHKRCTLVEVMGRQKGDLALYS
SICNENNLVIIPEYNNSKLNVLKKMQSFRNQEQSHAIVILTERQFDINVLAKEIELVSGF
ETKITILGHTQKGGIPLPQDRILGTLMGFYSIEMIRKNFFNIALGIRNNKIISYPFTEVL
ESKKDDIFQSLWEINNQLI


>Macrococcus_canis
MKKIAVLTSGGDAPGMNAAVRAVVRKAMFHGIEVYGVYQGYQGLINNNIKKMERGTVGDK
IQRGGTFLQSARCPEFKDPEVRKKAIANLNNLGIEGLVVIGGDGSYRGAQRLNEEGIKTI
GIPGTIDNDINGTDFTIGFDTALNTIVEAIDKIRDTASSHERTFIVEVMGRDAGDLALWS
GLAGGAETVLCPEHRKDVKVIADKIQQGIERGKKHSIIVVAEGVMTGEECGQELKKYINV
DTRISVLGHMQRGGSPSGMDRVLASRLGGYAVELLMNDETGLAVGIQNNSLSKTKFEDIF
TSVHHLDEKMYDLSNELSI

>Edwardsiella_anguillarum
MIKKIGVLTSGGDAPGMNAAIRGVVRAALSKGLDVYGIHDGYLGLYEGRIEQLDRYSVSD
VINRGGTFLGSARFPEFRDEKVREKAVENMREFGLDALVVIGGDGSYLGAKRLTEMGFPC
IGLPGTIDNDVAGTDYTIGYFTALETVVEAIDRLRDTSSSHQRISIVEVMGRYCGDLTLA
```

```
AAIAGGCEYIVLPEVEFKREELVEEIQAGITKGKKHAIVAITEHICDVDELAKFIEQETG
RETRSTVLGHIQRGGAPVAYDRILASRMGAYSIELLLQGYGGRCVGIQNEKLVHHDIIDA
VENMKRPFKGDWLDTAKILF

>Salmonella_enterica
MIKKIGVLTSGGDAPGMNAAIRGVVRAALTEGLEVMGIYDGYLGLYEDRMVQLDRYSVSD
MINRGGTFLGSARFPEFRDENIRAVAIENLKKRGIDALVVIGGDGSYMGAKRLTEMGFPC
IGLPGTIDNDIKGTDYTIGYFTALGTVVEAIDRLRDTSSSHQRISIVEVMGRYCGDLTLA
AAIAGGCEFIVVPEVEFNREDLVAEIKAGIAKGKKHAIVAITEHMCDVDELAHFIEKETG
RETRATVLGHIQRGGSPVPYDRILASRMGAYAIDLLLEGHGGRCVGIQNEQLVHHDIIDA
IENMKRPFKSDWMECAKKLY

>Serratia_marcescens
MIKKIGVLTSGGDSPGMNAAIRGVVRSALSEGLEVFGIYDGYLGLYEDRMEQLDRYSVSD
MINRGGTFLGSARFPEFRDENVRAKAIENLKNRGIDALVVIGGDGSYMGAKRLTEEGFPC
IGLPGTIDNDVAGTDYTIGYFTALETVVEAIDRLRDTSSSHQRISIVEVMGRYCGDLTLA
AAIAGGCEFIVLPEIEFNREDLVCEIKAGIAKGKKHAIVAITEHICDIDELARHIEQETK
RETRATVLGHIQRGGSPVAYDRILASRMGAYAIELLLQGYGGRCVGIQNEKMVHHDIIDA
IENMKRPFKGDWLETAKKLY
>Salmonella_enteritidis
MIKKIGVLTSGGDAPGMNAAIRGVVRAALTEGLEVMGIYDGYLGLYEDRMVQLDRYSVSD
MINRGGTFLGSARFPEFRDENIRAVAIENLKKRGIDALVVIGGDGSYMGAKRLTEMGFPC
IGLPGTIDNDIKGTDYTIGYFTALGTVVEAIDRLRDTSSSHQRISIVEVMGRYCGDLTLA
AAIAGGCEFIVVPEVEFNREDLVAEIKAGIAKGKKHAIVAITEHMCDVDELAHFIEKETG
RETRATVLGHIQRGGSPVPYDRILASRMGAYAIDLLLEGHGGRCVGIQNEQLVHHDIIDA
IENMKRPFKSDWMECAKKLY

>Enterococcus_hirae
MKRIGILTSGGDAPGMNAAIRAVVRKGIYEGLEVYGINYGFAGLVAGDIRRLDVADVGDK
IQRGGTFLYSARYPEFATEEGQLKGIEQLKKFGIEGLVVIGGDGSYHGAMALTRRGFPAV
GVPGTIDNDIPGTDFTIGFDTAINTVLESIDRIRDTATSHVRTFVIEVMGRNAGDIALWS
GVAGGADEIIIPEHDFDMASVAKKIQEGRDRGKKHCLIILAEGVMGGNEFAEKLSEYGDY
HTRVSILGHVVRGGSPSARDRVLASKFGAYAVDLLRAGKGGLCVGIRDNEMVAADIIETL
ESNKHKPDLSLYDLNNSLSF

>Klebsiella_pneumoniae
MIKKIGVLTSGGDAPGMNAAIRGVVRAALTEGLEVFGIYDGYLGLYEDRMVQLDRYSVSD
MINRGGTFLGSARFPEFREEHIRAVAIENMKKRGLDALVVIGGDGSYMGAMRLTEMGFPC
IGLPGTIDNDIKGTDYTIGFFTALSTVVEAIDRLRDTSSSHQRISVVEVMGRYCGDLTLA
AAIAGGCEFIMVPEVEYTRDDLVAEIKAGIAKGKKHAIVAITEHMCDVDELASYIEKETG
RETRATVLGHIQRGGSPVPYDRILASRMGAYAIELLLQGHGGRCVGIQNEKLVHHDIIDA
IENMKRPFKNDWLDCAKKLY

>Vibrio_parahaemolyticus
MIKKIGVLTSGGDAPGMNAAIRGVVRTALSEGLEVFGVYDGYLGLYEGRIEKLDRSSVSD
VINKGGTFLGSARFPEFKQVEVREKAIENLKKHGIDALVVIGGDGSYMGAKKLTEMGYPC
IGLPGTIDNDIAGTDYTVGYLSALNTVIDAIDRLRDTSSSHQRISIVEIMGRHCGDLTLM
SAIAGGCEYIITPETGLDKDKLISNIQDGIAKGKKHAIIALTELMMDANELARDIEAATG
RETRATVLGHIQRGGRPTAFDRVLASRMGNYAVHLLLEGHGGRCVGIVKEQLVHHDIIDA
IENMKRPVRNDLYKVAEELF

>Salmonella_choleraesuis
MIKKIGVLTSGGDAPGMNAAIRGVVRAALTEGLEVMGIYDGYLGLYEDRMVQLDRYSVSD
MINRGGTFLGSARFPEFRDENIRAVAIENLKKRGIDALVVIGGDGSYLGAKRLTEMGFPC
IGLPGTIDNDIKGTDYTIGYFTALGTVVEAIDRLRDTSSSHQRISIVEVMGRYCGDLTLA
AAIAGGCEFIVVPEVEFNREDLVAEIKAGIAKGKKHAIVAITEHMCDVDELAHFIEKETG
RETRATVLGHIQRGGSPVPYDRILASRMGAYAIDLLLEGHGGRCVGIQNEQLVHHDIIDA
```

```
IENMKRPFKSDWMECAKKLY

>Kosakonia
MIKKIGVLTSGGDAPGMNAAIRGVVRAALTEGLEVAGIYDGYLGLYEDRMIQLDRYSVSD
MINRGGTFLGSARFPEFREEHVRAVAIENMKKRGIDALVVIGGDGSYMGAKRLTEMGFPC
IGLPGTIDNDVAGTDYTIGYFTALHTVVEAIDRLRDTSSSHQRISIVEVMGRYCGDLTLA
AAIAGGCEFIVLPEVEFSREDLVAEIKAGIAKGKKHAIVAITEHICDIDELAKYIEAETQ
RETRATVLGHIQRGGSPVPYDRILASRMGAYAIELLLQGHGGRCVGIQNEKMVHHDIIDA
IENMKRPFKGDWLECAKKLY
```

Figure 3:

The Single-Letter Amino Acid Code

G	Glycine	Gly	**P**	Proline	Pro
A	Alanine	Ala	**V**	Valine	Val
L	Leucine	Leu	**I**	Isoleucine	Ile
M	Methionine	Met	**C**	Cysteine	Cys
F	Phenylalanine	Phe	**Y**	Tyrosine	Tyr
W	Tryptophan	Trp	**H**	Histidine	His
K	Lysine	Lys	**R**	Arginine	Arg
Q	Glutamine	Gln	**N**	Asparagine	Asn
E	Glutamic Acid	Glu	**D**	Aspartic Acid	Asp
S	Serine	Ser	**T**	Threonine	Thr

Data collection and analysis:

I will collect data in two ways. First, in order to determine improvement in student retention, conceptual understanding, depth of knowledge, and application, I will administer a pre-test and post-test. The same test will be used, but the pre-test will determine their mastery from prior classroom instruction, while the post-test will determine improvement in these areas following completion of this unit. The questions on the test will be based on AP Biology's essential knowledge and science practice components related to the topics in this unit.

The second set of data I will use is more qualitative. I will administer a Stem Semantics Survey to assess student attitudes toward STEM subjects and STEM related careers to see if these attitudes improve as a result of this project. This survey has been analyzed and found to have respectable to excellent internal consistency reliability, as well as good content, construct, and criterion-related validity for the areas assessed (Tyler-Wood, Knezek, & Christensen, 2010).

Connections:

This lesson and its associated research will connect to several aspects of the CATALySES Summer Institute. Students will be getting some general knowledge of emerging, re-emerging, and endemic. They will use pathogens for protein, genetic, and phylogenetic analysis using several web-based

databases that were demonstrated at the Institute. Finally, they will be making connections within the context of diseases that should improve their understanding of complex biological concepts, such as the central dogma of genetic information transfer, the importance of nucleic acid and protein sequencing, and the significance of DNA mutations and protein structure.

Works Cited

Abrahams, I., & Millar, R. (2008). Does practical work really work? A study of the effectiveness of practical work as a teaching and learning method in school science. *International Journal of Science Education, Volume 30*, 14.

Consortium, U. (2002-2018). *Uniprot*. Retrieved from Uniprot: https://www.uniprot.org

Hofstein, A. S. (2004). Providing high school chemistry students with opportunities to develop learning skills in an inquiry-type laboratory: a case study. *International Journal of Science Education and Technology*, 47–62.

Marseille-Nice genopole. (2003). *MABL*. Retrieved from Methodes Algorithmes Bio-informatique LIRMM: http://www.phylogeny.fr/simple_phylogeny.cgi

National Institutes of Health, N. I. (1999). *Emerging and Re-Emerging Infectious Diseases*. Retrieved from Emerging and Re-Emerging Infectious Diseases: NIH Curriculum Supplement Series, Grades 9-12: https://science.education.nih.gov/supplements/webversions/InfectiousDiseases/activities/activity1.html

Schneider, M. V., & Jimenez, R. C. (2012). Teaching the Fundamentals of Biological Data Integration Using Classroom Games. *Plos Computational Biology*.

Tyler-Wood, T., Knezek, G., & Christensen, R. (2010). Instruments for Assessing Interest in STEM Content and Careers. *Journal of Technology and Teacher Education, 18*(2), 345-368.

user76844. (2015, 01 23). *How do I compare student pre-test scores with post-test scores to evaluate whether or not they "learned"?* Retrieved from Mathematics Stack Exchange: https://math.stackexchange.com/q/1115634

White, B., Tasker, T., & Herrenkohl, L. (2011). Pedagogical practices to support classroom cultures of scientific inquiry. *Cognition and Instruction, 29*, 1-44.