**How do I identify a programmed translational frameshift?**

Activity Answer Key

This answer key is designed to provide instructors with expected answers and feedback for the activities found within the “Frameshift\_identification\_Activity” document. Instructors can choose to apply points to the activities. In general, I find that low-stakes assessments work small amounts are helpful for this exercise. Instructors can also tailor each of these exercises to the needs of their particular class. For example, they may wish to use examples from the particular cluster for the bacteriophage they are currently annotated, or they may wish to include their phage into the activities to help guide students in their frameshift annotation. Answers for each activity are in bold below:

**Activity 1:**

Analyze each scenario using synteny and determine if a PTFS is likely.

1. A student believes they are dealing with a PTFS in genes 63 and 64. Look at the functions assigned to the surrounding genes to determine if this is likely.

|  |  |
| --- | --- |
| **Gene #** | **Function assigned** |
| 60 | HTH DNA binding protein |
| 61 | Phosphoesterase |
| 62 | No Known Function |
| 63 | No Known Function |
| 64 | No Known Function |
| 65 | DNA primase |

Does synteny support this possibility? Why or why not?

**No – PTFS, when they do occur in bacteriophages, occur in the region near the tape measure protein, and usually have the function of tail assembly chaperone (after minor or major tail proteins). The functions listed do not support a PTFS call.**

1. A student believes they are dealing with a PTFS in genes 10 and 11. Look at the functions assigned to the surrounding genes to determine if this is likely.

|  |  |
| --- | --- |
| **Gene #** | **Function assigned** |
| 9 | Minor tail protein |
| 10 | Tail assembly chaperone |
| 11 | Tail assembly chaperone |
| 12 | Tape measure protein |
| 13 | Minor tail protein |

Does synteny support this possibility? Why or why not?

**Yes – synteny tells us that PTFS occur in the tail assembly chaperone genes (genes 10 and 11) after the minor tail protein and before the tape measure protein.**

1. Given the assigned functions of these genes, which two genes are good candidates for containing a PTFS?

|  |  |
| --- | --- |
| **Gene #** | **Function assigned** |
| 22 | Major tail protein |
| 23 | Tail assembly chaperone |
| 24 | Tail assembly chaperone |
| 25 | Tape measure protein |
| 26 | Minor tail protein |
| 27 | Minor tail protein |

**Synteny tells us that PTFS occur before the tape measure protein in most cases, and after tail proteins. PTFS occur in the tail assembly chaperone genes. Based on this information, we can predict that genes 23 and 24 may contain a PTFS.**

1. The tables below show the gene assignments for 3 different bacteriophages. Answer the following 3 questions about these phages.

|  |  |  |  |
| --- | --- | --- | --- |
| Gene # | Phage #1 | Phage #2 | Phage #3 |
| 22 | Major tail protein | Major tail protein | No Known Function |
| 23 | Tail assembly chaperone | Tail assembly chaperone | Minor tail protein |
| 24 | Tail assembly chaperone | Tail assembly chaperone | Lysin A |
| 25 | Tape measure protein | Tape measure protein | No Known Function |
| 26 | Minor tail protein | Minor tail protein | Holin |
| 27 | Minor tail protein | Minor tail protein | No Known Function |

1. In each bacteriophage where is the frameshift likely to occur (what gene numbers)?

**Phage #1: Genes 23 and 24**

**Phage #2: Genes 23 and 24**

**Phage #3: Synteny does not support a frameshift in the genes listed.**

1. Does bacteriophage #3 have a PTFS? Explain your answer.

**Bacteriophage #3 does not have a PTFS in the genes listed. The tape measure protein is not included in this gene list, so it is possible that it does have a frameshift in a different region of the genome than the one shown.**

1. Based on this set of genes – do you think these 3 phages are from the same cluster? Why or why not?

**No – phages #1 and #2 may be in the same cluster, since the same functions occur in the same order and gene number. Phage #3 is not in the same cluster based on the information provided (the functions associated with genes 22-27 are not structural genes like those in phages #1 and #2).**

**Activity 2:** For each of the following phages there are several gene sequences. Analyze each of these gene sequences using BLAST to assess whether a PTFS may have occurred, and identify which of the two genes are involved.

1. **Bacteriophage: Scout**

>Scout\_Draft gp22

MSKIFTLDSFREEVEKEFAPVKIEVDADNSVVLRNLLRIPKGAREEIFGLLERMDKMSEGKSEDEMTVEELEATAGIALRMIELVADTPAGGRILVESLEDDLALTLKVFEAWMEATNPGEAPRSHD

>Scout\_Draft gp23

MDGGHESGGSAALARLIDDYGDAVAADLMETYGVDLRDLFVPESRLTPKWVLVLIKELPVGSRFYSEKRGGPQFRGWDESRYTLAAIVNAVRALQHTYLAAHMKSTPKPPEPYPTPDRNTRKKNNNKPNSFASIAAQMIAAKRAKKARKAAQE

>Scout\_Draft gp24

MAGGGAGGTEVGRISIRVVPNLDNFYRELKTKLEAIEKQLRGNVPIDIDLNARGTRAKMAALMAGLKAQAAQGVDVPVDVNNKGLGAAWREFRAGLADFGRLGKQAAQGVKSYRDEVNRLTLEQQRQRPLLNHTYAWWRSNNIMARRGATILRDFTDALRTQQQWLRQQDRTLTANQARWKSWAMAIRDANVNATNGFRRFRASLQALRGGGGGDDGDGFSRIFGSLGRFGNEAEKAGSQVEHVGKKFLGLTRMGWLVTGVFLAAAPAIALVSGLLAGLPSLIGAFGAGIGAVALGMDGIKAAAEVLMPAFEQMKTAVSSTFQQALVPQFQQLLGLMPMIQTGMQGVAQGMSSMFQGVTDALSKGAGPAQIENLLANTKTFFEQLQPAANQFTQSFLTLASSGSDAFGYLSGSLNTFSTQFNDMVNRVSQNGVMDGAMKGLSQTLDGVTNLFTRLMESGLQAMSQLGGPMNTFLTGIGDLAVALMPALTSLSGLFGNVAGTLGTALAPIVTALTPAFTTLADTLGSLLVPNIQTLGNILTPVATMIGTTLTTALQQIQPMIPGLVESFAQLGSTLVSQLAPHIPALATAMGQMAGAVIKLAPMLISQLVPAFIDLIPSITQLLPHVVSLAESFARMMPTIVPLVSIIFSLIAAFAQAAATIGGVVLGAISSLIGVISEVVAKISEWVSSFAQGVSDIAAKAAELPGMVKSALGDLGSFLVSSGKALVQGFINGIKSMVGAVADAARSVVQAARDFFPFSPAKKGPFSGSGWVDASGQSVGEAFADGLAGTQGKIVETARAIMQAAKDVFGDAANIAFNFNFGQMQSQMASVASSAGDLQRSMSRTVSQSTGSGKIDDETRQMLDQISIRKDELELERQRLQAEKNALDTKDKAGRAALQQRIDELNIQKDQLELQREQLSYQSKYTDSVAQTGAQYDEMFNKLTRMPYDFATANANQFLSDIGISGDGALSQALKEGLKFGEQFIFNVGSMDEAVQGQQTIQNKKSLQFDRR

* + 1. Blast each of these genes and record your results (Does the gene align well with other phages, are the start calls the same?).

**Gp22: Tail assembly chaperone, good alignment with other phages, 1:1 alignment**

**Gp23: Tail assembly chaperone, good alignment with other phages, most aligned 8:120 or 8:142 suggesting the start is incorrect.**

**Gp24: Tape measure protein, good alignment with other phages, 1:1 alignment**

* + 1. Is there a frameshift in this gene sequence? **Yes**
    2. If yes - which two genes are involved? How do you know? If no, what made you reach this conclusion?

**Genes 22 and 23 are involved with the PTFS. Gene 22 has the PTFS that sometimes shifts and creates a longer gene that begins at 22 and ends at the end of 23. BLAST indicates that 22 has the ‘correct’ start, but 23’s needs to be adjusted in the final annotation.**

**Genes 22 and 23 are both assigned the function of tail assembly chaperone. Gene 23 does not start at the same point as other annotated phage genomes, while gene 22 does. This suggests that gene 23 is the gene product that is produced only when the frameshift occurs.**

2. **Bacteriophage Pheobe**

>Phoebe\_Draft gp24

MALNDDAVLTAAVGYVYTAPVGTAAPTPAQLKTLNLTDTGLWTPTGWDSIGHTSRGDMPEFGFDGGDTEVRGSWQKKKLREVTTEDPVDYLTLFLHQFDEQAFELYYGANASTTPGVFGVSAASGDPTEKAFLVVIVDGDERVGFHAHKASVRRDDAIQLPTDDFAALPVRATFLQHNNELLFSWINEDLFNVEEEEE

>Phoebe\_Draft gp25

MSNVFTLDSFREEADREFAPVKLELGGDDAVVLRNVLRIQKTRREEVFQLLEKLDSIAKDDEGKQREEDDLDASEMEAMGDIALRMIELVADNDALGSRLVDELRDDLALTLKVFEAWMNATQPGEAERSPA

>Phoebe\_Draft gp26

MDLRDIYLPESRLSPKLALVLIKELPVGSRFYAEKRGGKQFRGWDESRYALVAIVNAVRALQYTYVAAHSKSKPKPPDPFPTPQRTKARQIRKAGSFAWMAAKQIAAARKRKAQT

1. Blast each of these genes and record your results (Does the gene align well with other phages, are the start calls the same?).

**Gp24: mostly major tail protein, aligning 1:1 with other phages**

**Gp25: mostly tail assembly chaperone, mostly aligning 1:1, some aligning 1:22**

**Gp26: Tail assembly chaperone, many hits to other phages, most aligning 1:151**

1. Is there a frameshift in this gene sequence? **Yes**
2. If yes - which two genes are involved? How do you know? If no, what made you reach this conclusion?

**Based on BLAST results and synteny, genes 25 and 26 contain the PTFS. They both have the function of tail assembly chaperone. Gene 25 aligns 1:1 with many other phages, while gene 26 alignment suggests that the start in the draft genome is called incorrectly. Gene 25 shifts into 26 to create the gene 26 product.**

1. **Bacteriophage Quartz**

>Quartz\_Draft gp13

MPLTQWNPATQISRGNVAVGVAPAVVSLDAPALSELTTGIGLDCSITTMNGTSSTDSESIDWLCDPASEQLPGSTTHAMDDLVIKGTGQDDADLIAGLNIGDVVYVWRRDGIPHDTAPAAGQFVWVWKVIITSIDPLEANNTFVGITAHITVLARSKTAVAIAA

>Quartz\_Draft gp14

MSFSSYEELKAAVDERRKDILTIEVDLGARYSQDHEDAKKELAQAEAIQKLAGGGQGFLNDNLEQLKARVAETKPEANSIWLRYGRLQLAEWSMLTKATGLTPIDQYEKVLPQTFKGVYGVDPTAEDEEGNLLHPDAEPLTTDARAVSSRSEETVLPGAMLHVVVNAFMTWQNSSGEISIRPTKSGRV

>Quartz\_Draft gp15

MQAIAIGQDMWKTANKTAIESHQKGNGPDPGMGVYWLSQGEGEVLPTPET

1. Blast each of these genes and record your results (Does the gene align well with other phages, are the start calls the same?).

**Gp13: major tail protein, many alignments 1:1 or 2:3.**

**Gp14: tail assembly chaperone, many alignments 1:1**

**Gp15: tail assembly chaperone, alignments vary – some 1:1, some 1:75, some 1:41.**

1. Is there a frameshift in this gene sequence? **Uncertain. Possibly – but current evidence does not support its annotation.**
2. If yes - which two genes are involved? How do you know? If no, what made you reach this conclusion?

**This one is a challenge. Quartz belongs to a small cluster (EA10 – it only has 3 phages) that does not currently annotate the PTFS. Instead, it annotates two separate tail assembly chaperone genes. The BLAST results indicate that genes 14 and 15 are the tail assembly chaperone genes, and therefore are the most likely to contain a frameshift, with gene 15 being the gene product that only occurs when a frameshift happens. HOWEVER, students may note that there are 1:1 alignment results from gene 15, and that the results that do not align often are not as large of a start change as we normally see with the frameshift (the example above, for instance, had alignments of 1:151 – our results here are much smaller). This indicates that instead of a frameshift we may just be dealing with a gene that has an incorrectly called start (perhaps it just needs to be changed to fill in a gap). This example could generate a good discussion about how we shouldn’t assume that the programmed translational frameshift always exists.**

**4. Bacteriophage Phegasus**

>Phegasus\_Draft gp18

MTQDKQTFPPGPATLLGQELMLEHTDPLIMLTTADRKVTFYLSGGLAAWPRHQDGVNLVEITTPTPEFRNLRAQGARQDGGQTRDTVYDPMQIDAVLLASATTPEGLSRVVSEWIAANDPEQLCRLEWFTFEGGLWWCDVRLEKRFIDRLQQSPRRVKKQVLSTVWMNDLAFWQSVDSTCTWAFSYQTMLDTFKYDTSASKDLGENWPQYRYDGEGGGYWYANGDRAVWRDDPEDPLLTDGVSVLCGPYKDFETATDYQVIDFVIGSFQEITFPDGAENHAWGRLNRDEDGEWAGDGIRASVGPTSAVLHRFNDFEKTRIGLPVPLFPPPFIGEKFRLIIGYQGNPRKYRLLRALTDRSAGVPVLTVTEQGTGSAIGPDHRGIGFGGRAGAALLTQATPASVRKVAAGDNRTETQEGFLTLTNIGERDGWPQIVFEGPGLLEIANGPGSTDMIKFGPLEDGQRVLISTHPRYRAIVDLTQGQVGQQLDGGQKLIDTIVKLLSLGQVPPALQWFESVFGIKPPQGPLYSLLDGRFTRPIPGVRQPRDATTSRIAIRVRDGNANTKVTASVTPMRRWPEAVHD

>Phegasus\_Draft gp19

MPRVENIGDYLDLAEIQRKLLSHNPHEVMDAARQVAEVDATPSGEVTCTVRTNTYKLAGEASNRKSLQVSWPRLAVPTGKLVLDGDDTLADVVLNCHETVVPVVVDCGPLRWSGRVDVAHDKFGDPNEPDTIECELIHDKVWLTRVVAFPWWFMPLQWQGPPTRGVAFGNAISVIKYLFTSQFMRIQLGLWELVNNLLSLNLDWRSYFSTLLMQNPGEELELRDIVQMATTPVYVVPSAGWNDTSPFISLNWRMDELLQLVTKTCEDNGLTIEVYLWEPGMPQPDPFAEATNLLRVPTVVVDVKDRMQVTGITGTAFDGLQRTFVDLLGSMFGEALKPFLDPNNEAAYAPDGVNIAPALGVHAIKPWCVFNADHPRSGVRGVVSHHHPIAWRTITGGKSPAWLNSLVDATLAWLIDMITIVLGVTGVPGTILDGAFHDIAFAFQQTDNFDRRLKLGPYGLPEVFIPTGSGSYTLEAFFQQKSAQYDTRGYVSGQLVVDNCFPYELGRDTFPGALATFIRRGRVVTDFIENATLIETRGEPTQVQFQIGDGKAEEAPAAKLQRRFGDLQAGVNIALMAS

>Phegasus\_Draft gp20

MAIIVDEDKGTISFTECTVTFPYGFSVSSGVGTIVITPAGGVASFPLAIQGASGLPPNITMAFHVIGPDDPLPDPNPEMTVIDEGGPGEAAHYHYDCYVQKGDKGDAASFNFLDADDLEDGEDLAEGDVGTNGYVLSYAYEGTGTPGIRFIPQKTGDIRGPSAIAATAWSNTAIRLLCAVTLEAKPFPRKVLPTGSVVVTGSADTRVDLVAYLGDPDDGGVEIGRAFGQAGAAPPPLVMAGGPPATTAGGNANYAIVPAGQSATVYFRAEQKASSSNNWATAGAPDGARAGVVVVAV

1. Blast each of these genes and record your results (Does the gene align well with other phages, are the start calls the same?).

**Gp18:Minor tail protein, many 1:1 alignments**

**Gp19: Minor tail protein, many 1:1 alignments**

**Gp20: Minor tail protein, many 1:1 alignments**

1. Is there a frameshift in this gene sequence? **No.**
2. If yes - which two genes are involved? How do you know? If no, what made you reach this conclusion?

**All of the genes have good alignments (1:1) with many phages. This suggests that they all have ‘correct’ start calls and function as their own gene. None of the genes indicate a frameshift with an overlapping start with a different gene. All genes have the function of minor tail protein in other annotated genomes.**

**Activity 3: \*\* For this activity answers will vary. Instructors should choose phage clusters that they feel will most benefit their class.**  For each phage cluster, assess whether a PTFS is likely, and where in the genome it may occur, based on comparisons across phage genomes in that cluster.

1. Choose 5 genomes from the A1 cluster, including 1 draft genome.
   1. Which genomes did you choose?

**This answer will vary by student. You can check whether the phages chosen belong to the cluster that you specify (in this case A1)**

* 1. Examine the genomes – does it look like a PTFS occurs in these phage genomes? How do you know?

**Yes – A1 phages do have a PTFS. You can see the frameshift in the annotated genomes as two overlapping genes that begin at the same point in the genome. These genes are located directly before the tape measure protein.**

* 1. If a PTFS occurs – where does it occur?

**Genes 23 and 24 are associated with the frameshift.**

* 1. If a PTFS occurs - On your draft genome, which two genes are involved in the frameshift but have not yet been annotated correctly?

**This will depend on the draft genome chosen. It should be the genes directly in front of the tape measure protein, that align well with the chaperone proteins in the annotated genomes.**