Quality control bioinformatics for recombinant protein production

How might we confirm the purity of biomanufactured proteins to ensure that only the intended recombinant protein is being produced in a given batch?

In this lesson, students will:

1. Analyze the DNA sequencing results to identify the most common recombinant DNA present within each fermentation tank.
2. If your sequence is insulin, check for any mutations that could result in a defective, ineffective, or contaminated batch of protein.

Ohio Standards

Grades 3–12

- Nature of Science

Biology

- B.H.3: Genetic mechanisms and inheritance
- B.H.4: Mutations
- B.H.5: Modern genetics

Student prior knowledge

Students should understand that DNA is a combination of nitrogen bases that code for proteins. The sequences of nucleotides determine the proteins produced. Recombinant DNA techniques are used to create proteins for human uses.

Suggested timeline

One class period

Materials

- A computer with internet access
- Tank sequences:
  - ***ols.plus/fermentation-sequences***
  - ***ols.plus/fermentation-sequences-key***

Teacher preparation

Review the sequences and try out the activity.

Procedure

1. Visit the NCBI website: ***ncbi.nlm.nih.gov***
2. On the homepage, look for the “BLAST” option in the right-hand menu and click it.
3. Select “Nucleotide BLAST” from the list of available tools.
4. Copy and paste sequences one-by-one then follow instructions below to determine what organisms are indicated by each sequence.
5. Under “Database,” ensure the option is set to Standard databases (nr etc.).
6. Scroll down and click the blue “BLAST” button.

7. Wait for the BLAST program to complete the search.
8. On the results page, there will be a section called “Descriptions” with a table listed as “Sequences producing significant alignments”. These are the matches (called “hits”) of your sequence to known nucleotide sequences in the database.
9. Click on the description to see the alignment between your sequence and the database sequence.
10. Record the name of the top hit sequence and other relevant details provided by the alignment shown.
11. Click on the unique “Sequence ID” at the top of the alignment to learn more about the database sequence.

Suggested wrap-up

Discuss the reflection questions with the students, being sure to acknowledge which sequence each group investigated.

Differentiation

Students may need to work in partners with one student directing while the other performs the steps on the computer.

Extensions

Ask students to read an article that is cited on the NCBI database about one of the three recombinant proteins in the activity:

- **Insulin:** A hormone produced by the pancreas that is necessary to regulate sugar in the bloodstream. People with diabetes often need insulin injections to control their disease.
- **Erythropoietin (EPO):** A hormone produced by the kidneys that regulates red blood cell production. Recombinant EPO is used to treat anemia in patients with chronic kidney disease.
- **Tissue Plasminogen Activator (tPA):** A life-saving enzyme used to dissolve blood clots in patients suffering from acute ischemic strokes. Administering tPA shortly after a stroke can restore blood flow and reduce brain damage.

Support information

- People with diabetes used to rely on insulin extracted from cows and pigs to regulate their blood sugar. This is inefficient, costly, and has sometimes triggered immune reactions.
- **Recombinant DNA technology** in the 1980s revolutionized insulin production through a method for genetically modified bacteria produce human insulin.
  - A human gene of interest is isolated—such as insulin, erythropoietin (EPO), or tissue plasminogen activator (tPA)—and inserted into a small circular piece of bacterial DNA called a **plasmid.**
  - The recombinant plasmid is introduced into fast-growing bacteria, typically *Escherichia coli* (E. coli), which serve as protein factories.
  - Once inside the bacteria, the gene is expressed, leading to the mass production of the desired protein.
- **Biomanufacturing,** the process of protein expression by microbes, allowed for safer, purer, and more scalable production of life-saving proteins.
- Purification steps are followed to ensure that only the correct protein is collected for medical use.
- Successful gene insertion does not always guarantee proper protein production.
  - Mutations, contamination with unintended sequences, or loss of the recombinant plasmid can compromise the quality and efficacy of the final product.
- Scientists use bioinformatics tools like NCBI BLAST to verify the accuracy of bacterial DNA. This ensures that the inserted gene is correct and free from harmful alterations.

Career connections

- **Biomanufacturing technician (upstream):** Controls the environment in which cells are grown through monitoring equipment and working with other departments like quality control.
- **Bioinformatics pipeline engineer:** Designs and develops bioinformatics processes for creating new biomanufactured products.
- **Cell culture technician:** Grows living cells in various kinds of containers (i.e. culture flasks, bioreactors, plates etc.).