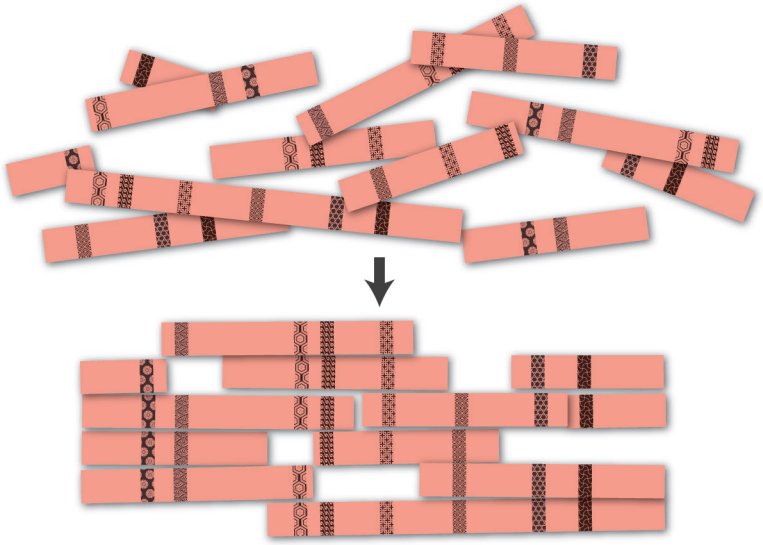
Connect-A-Contig

Paper version

Abstract

Students align pieces of paper "DNA" strips based on the distance between "markers" to generate a DNA consensus sequence. The activity helps students see how fragments of tagged DNA can be used to make a physical map of a genome.



Learning Objectives

To sequence or map long stretches of DNA, researchers first cut the DNA into shorter fragments, then piece them back together.

Overlapping sequences or features on short segments of DNA can be used to assemble much longer contiguous DNA sequences, or "contigs."

Time required

- About 20 minutes

Materials

- Scissors
- Copies (there are two identical sets on each page, and each set includes the contig number)

  Tips: Copy each contig set on a different color of paper to eliminate crossover between the different sets. Sets may be laminated and re-used.

Classroom Implementation

1. Introduce the activity by reviewing relevant content from the Genome Mapping web page: http://learn.genetics.utah.edu/content/cotton/genome/

Be sure to cover the following key ideas:

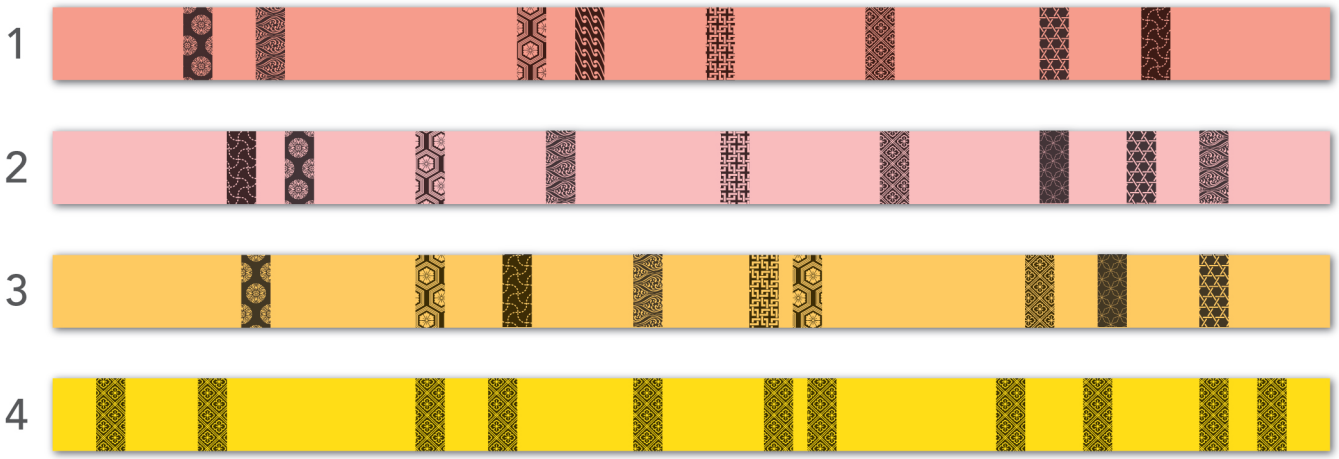
- Whole chromosomes are made of single DNA molecules that are many millions of base pairs long. With current technology, the longest pieces of DNA we can "read" are much shorter than that—up to several thousand base pairs long.
- In order to map a whole chromosome (or other very long stretches of DNA), researchers first collect many copies of it, then they break the copies randomly into smaller fragments. They

read the fragments, and then they use computer software to put the pieces together into a contig (short for "contiguous") to get a full-length consensus sequence.

2. Distribute scissors and contig sets: one per individual, pair, or group.
3. If sets are not pre-cut, tell students to cut the strips apart along the doted lines.
4. Explain that the students will act like the software programs that researchers use to assemble genomes. They will need to match up the spacing of the markers (patterned bands) on the overlapping shorter lengths of DNA to build a "contig."
5. If students have trouble, consider offering the following hints:
    - Begin by laying the strips horizontally.
    - Choose a longer strip, then find markers on others that overlap with it.
    - If you don't see where a strip fits, try flipping it around 180 degrees.
6. Students should condense their contig solution into a "consensus sequence." Display the solution to each contig set and ask students to compare their answers to it.

Note: Contig sets 1–3 should each take a few minutes for students to solve. Set 4 is significantly more difficult and it will probably take longer to complete.

Solutions



Discuss

Which pieces of DNA are the most informative? Why? *(Answer: the longer ones, because they contain information about the spacing of more markers.)*

If you have a reference sequence (i.e., a consensus), it's easier to align other tagged sequences to it than to build a reference sequence from scratch. *(Using only the fragments is like sequencing a genome for the first time. Using a consensus sequence is like comparing the genomic sequence of a new individual to one that has been sequenced previously—the sequences will have some differences, but they will be largely identical (about 99.9% for people).*

Explore the concept of "depth of coverage" (the number of fragments that cover a particular span of

the contig). Where is the greatest depth of coverage? Where is the least depth of coverage?

What do the patterned bands represent? *(Answer: a specific DNA sequence. In the case of optical mapping, which this exercise approximates, they represent restriction enzyme sites—specific, short DNA sequences where a restriction enzyme will nick the DNA and add a fluorescent tag.)*

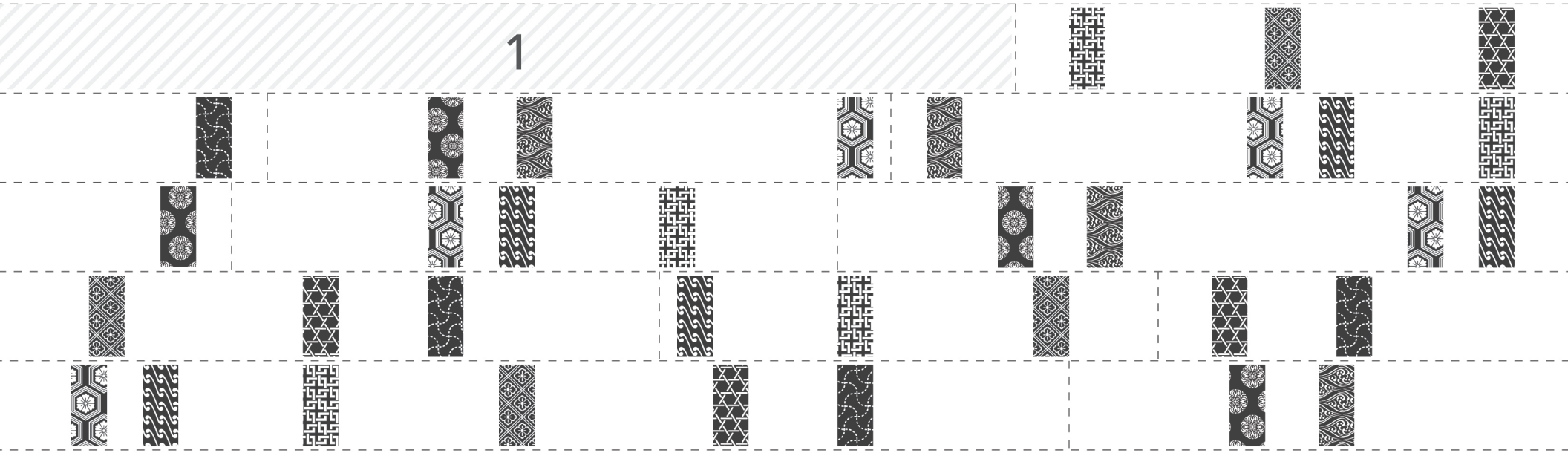
Was it easier to assemble fragments that had multiple types of markers vs. just one type? Why? *(Answer: with just one type of marker, you have only the spacing of the fragments to guide you. With multiple colors, you can use both the color and the spacing.)*

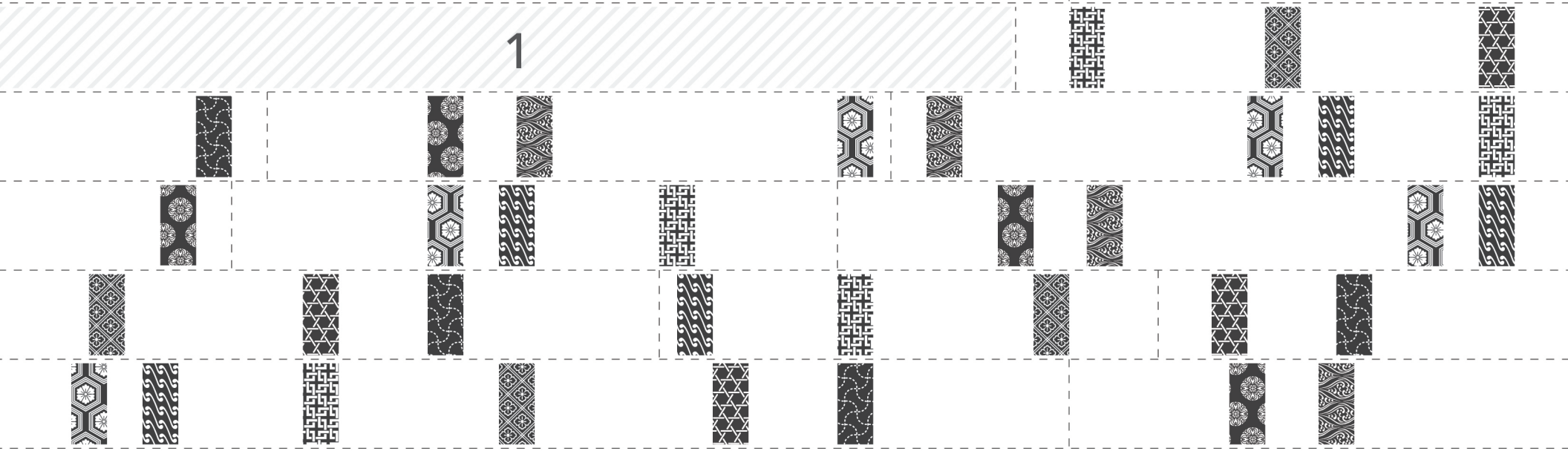
Assembling contigs out of DNA sequences (strings of As, Cs, Gs, and Ts) follows the same principle: instead of using markers, you line up fragments by overlapping DNA sequences.

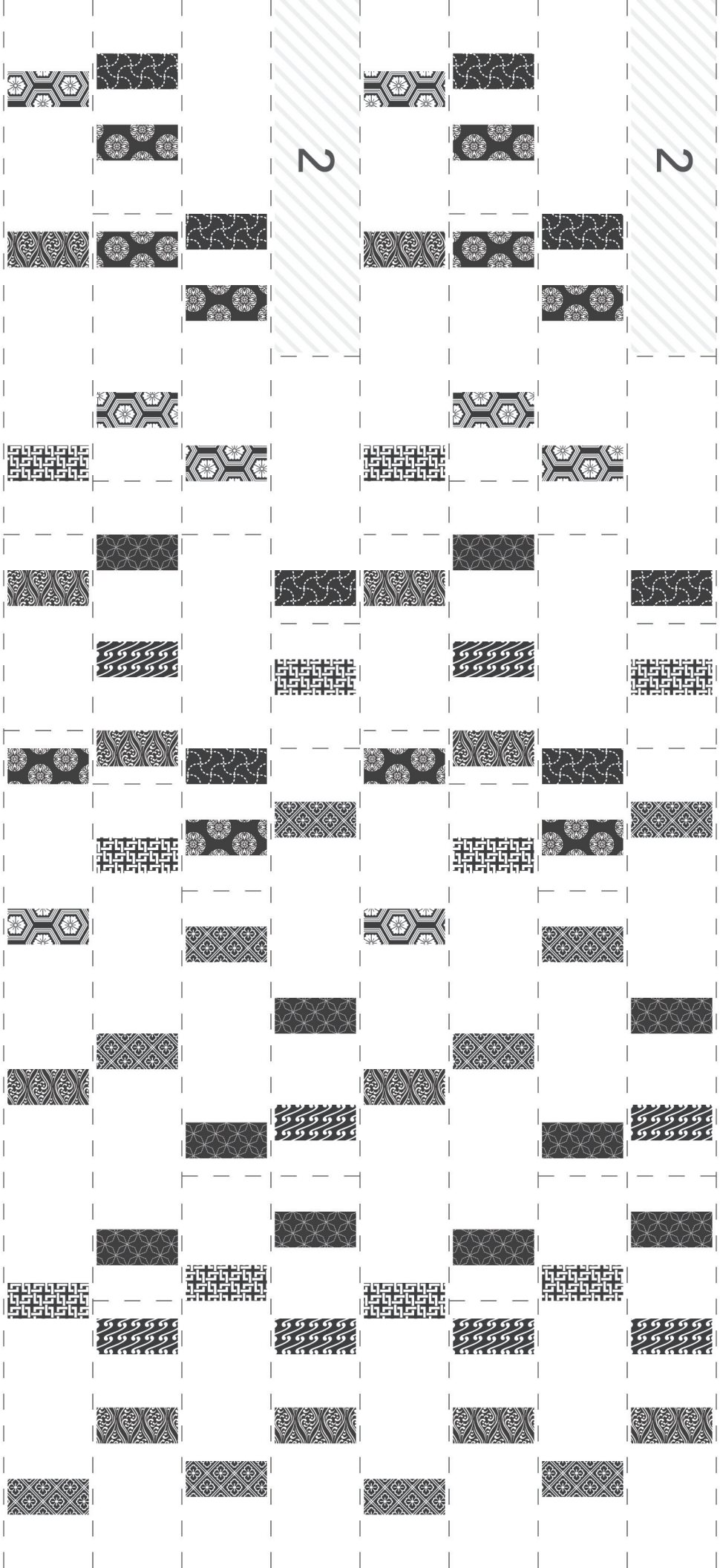
Credits

Joshua Udall and Lindsay Chaney, Brigham Young University
Josh Stowers, Lakeridge Junior High School
Matt Woolley, Mountain Ridge Junior High
Genetic Science Learning Center

This material is based upon work supported by the National Science Foundation under Program Award #1339412. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.


1


1


2
2

3
3

4
4