

# Active learning strategies for bioinformatics teaching

Lex Nederbragt

Centre for Ecological and Evolutionary Synthesis (CEES), Dept. of Biosciences, University of Oslo, Oslo, Norway

Contact: [lex.nederbragt@ibv.uio.no](mailto:lex.nederbragt@ibv.uio.no) [@lexnederbragt](https://twitter.com/lexnederbragt) [flixblog.com](https://www.flixblog.com) [wordpress.com](https://www.wordpress.com)

CEES  
Center for Ecological and Evolutionary Synthesis

ifb  
Institute for Informatics



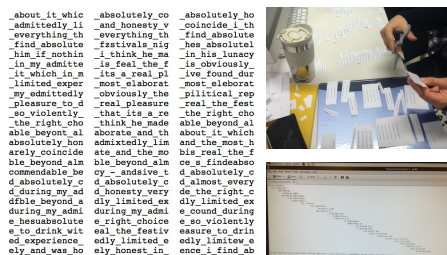
## Background

- After implementing active learning techniques, I find my teaching has become more interesting, more active, more interactive, and thus hopefully better
- This poster summarizes and demonstrates the techniques I have implemented my bioinformatics curriculum for analysis of high-throughput sequencing data (part of the INF-BIO 5121 and 9121 courses)

## The assembly exercise

### Pen and paper

- Students receive printed sheets with short 14 character snippets ('reads') randomly drawn from a real English text
- In groups of 4 to 5 students, students try to reconstruct the source text.
- Students are very active and really make a collaborative effort, and seem to be enjoying it
- Enables a discussion of many important aspects of the challenges of genome assembly

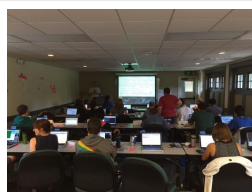


Some of the 'reads' the students were given

Students putting the pieces together

Credit: [ivory.idyll.org/blog/the-assembly-exercise.html](https://www.ivory.idyll.org/blog/the-assembly-exercise.html)

More information: [flixblog.wordpress.com/2014/11/04/on-the-benefits-of-open-for-teaching](https://www.flixblog.wordpress.com/2014/11/04/on-the-benefits-of-open-for-teaching)



A class of active learners

Source: <https://twitter.com/stealjapanda/status/636233683444350976>

## Building small graphs

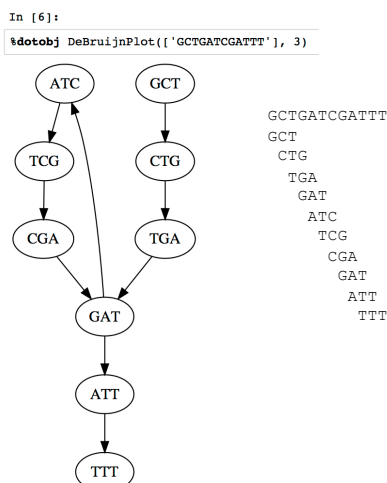
### Interactive notebook

- Students explore small De Bruijn graphs in an interactive electronic notebook (Jupyter) with python code
- The code visualizes the nodes and the effect of choosing different kmer sizes on the graph's structure

Exercise 1: simple De Bruijn graph  
Do the exercise by execute the code cell below.

Let's first generate a really simple De Bruijn graph of a short read 'GCTGATCGATTT' with kmer size 4

Study the graph and see if it indeed encodes the read now change the kmer size to 3 and again trace the read: what happened?

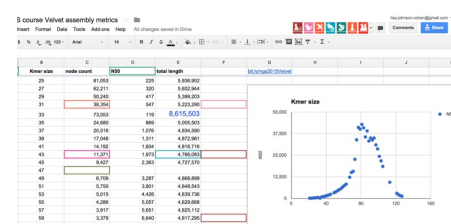


More information: [flixblog.wordpress.com/2014/11/04/on-the-benefits-of-open-for-teaching](https://www.flixblog.wordpress.com/2014/11/04/on-the-benefits-of-open-for-teaching)

## Finding the optimal settings

### Google spreadsheet

- Students collaboratively explore parameter space by each picking a number (kmer size), and running a velvet assembly using that number (kmer)
- Students log the resulting assembly metrics on an editable google spreadsheet, which has a graph showing the assembly results for the different settings
- The graph is shown on the presenter's screen in the classroom and updates in real time
- The goal is to teach the effect of choosing different kmer sizes for De Bruijn Graph assembly



Final spreadsheet at [bit.ly/INFBIO1](https://bit.ly/INFBIO1)  
Screenshot credit: Lisa Cohen

A really cool learning experience [...] a class of people all assembling data then populating table with results to make an x-y graph of N50 vs. k!

## Peer instruction

### Google form

- I pose a multiple choice question
- Students discuss the question in pairs
- Each pair gets to vote through an online system, I use a simple google form for this
- I show the tally of the votes, using the 'summary' function of the google form
- Student pairs discuss the tally and consider whether they want to change their opinion
- Students vote again and the tally is shown again
- The answers are discussed by me

#### Question

Why did the contig N50 distribution show a peak?

- There is an optimal kmer size for resolving all repeats
- Higher and lower kmers result in fewer overlaps between the reads
- At higher kmers, more repeats are resolved, but overlaps between the reads are lost
- At higher kmers, more repeats are resolved, but errors in the reads fragment the assembly more

#### Answer form

Multiple choice questions

Current question

- ☐ Option 1  
☐ Option 2  
☐ Option 3  
☐ Option 4

Submit

#### First responses

20 responses

Summary

Current question

Option 1: 10%  
Option 2: 20%  
Option 3: 10%  
Option 4: 20%

#### Second responses

12 responses

Summary

Current question

Option 1: 10%  
Option 2: 20%  
Option 3: 10%  
Option 4: 20%

Credited to Eric Mazur: 'ConcepTests'

#### Further reading

<https://flixblog.wordpress.com/2015/08/31/active-learning-strategies-for-bioinformatics-teaching-2/>

'Active learning' techniques benefit learning. Freeman et al (2014) *PNAS* 10.1073/pnas.1319030111

Get a copy of this poster through figshare: <https://dx.doi.org/10.6084/m9.figshare.1541056> or scan the QR code



## Conclusions

Does it work?

I have no proof, but I think so!

Does it take time to prepare?

Yes, more than lecturing

Will you keep doing it?

Absolutely

## Quiz!

Are you using active learning in your teaching or are you planning to?

Yes, and I will do more of it

Yes

No, but your poster inspired me

No