Gelbox — An Interactive Simulation Tool for Gel Electrophoresis

Authors Chaim Gingold chaim.io
Shawn M. Douglas douglaslab.org

Abstract Gel electrophoresis enables separation and visualization of biomolecules such as DNA, RNA, or proteins. Like many powerful tools, effective use of gels can be difficult to learn. Gelbox is a simulation tool to help build intuition for the relationships between experimental input parameters and resulting data output from gel electrophoresis. Our simulation model handles a wide range of settings, including many "suboptimal" values that may be useful for troubleshooting common mistakes by novices.

Introduction

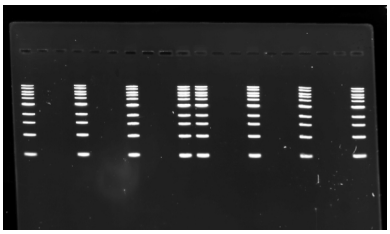
In 1948, when Arne Tiselius was awarded Nobel Prize in Chemistry for his pioneering work on electrophoresis, he summarized his motivations for the research as follows[1]:

> *[In biochemistry] some of the most important problems are to* **isolate** *the substances which are responsible for a specific biological or biochemical effect, and also to* **define** *and* **characterize** *these substances as accurately as possible.*
>
> — Arne Tiselius, 1948 Nobel Lecture

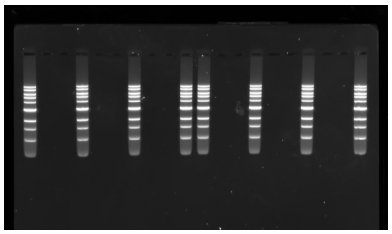
Photo: digitaltmuseum.se

Today, gel electrophoresis is used in biology labs worldwide to isolate, define, and characterize biomolecules. Gel electrophoresis is a powerful and versatile technique that is remarkably precise, yet gentle on the molecules being analyzed. However, the technique can be very difficult to master for a variety of reasons.

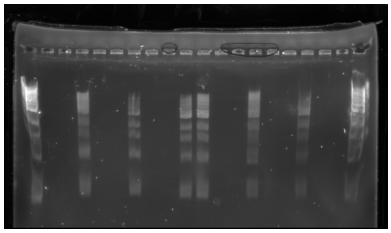
First, gels must be prepared according to detailed protocols. Small variations in the amount of an ingredient or the timing of a step can dramatically affect gel behavior. To illustrate this concept, we ran identical 1kb ladder samples on three agarose gels that were prepared with different buffers. We observed that the degree of band mobility, smearing, and distortion varied greatly for each gel.



400mM Tris, 20mM acetic acid, 1mM EDTA



40mM Tris, 200mM acetic acid, 1mM EDTA

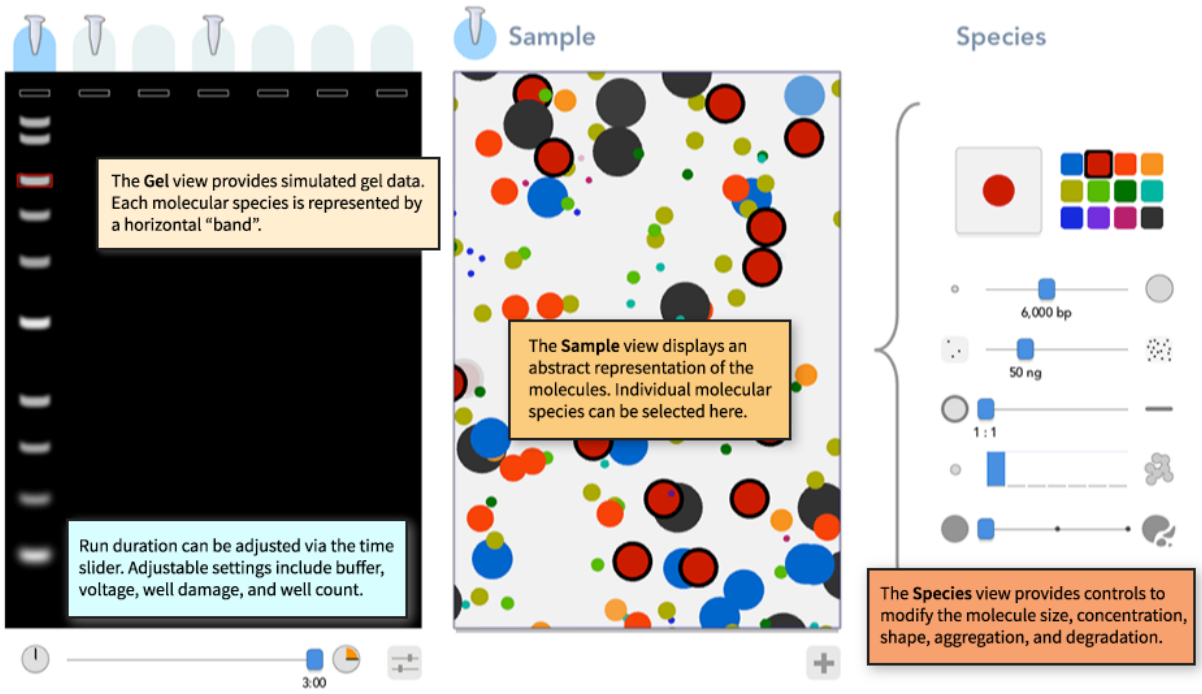


Water only

Even when gels are run successfully, interpreting the data can be tricky because the relationships between "band" patterns and the molecules in a sample are very abstract and often ambiguous. Challenges with preparing and using gels can frustrate budding scientists, interfere with scientific reproducibility, and impede overall research progress.
We believe interactive simulation and visualization tools can help. We created Gelbox, a dynamic “scientific sandbox” for visualizing the relationships between gel bands, sample molecules, and the gel itself. We sought to capture the notion that gels are easy to mess up, but they do so in predictable ways.

This project builds on concepts from Earth Primer, an interactive geology science book by Gingold[2]. We have been greatly influenced by Bret Victor's work, especially Up and Down the Ladder of Abstraction[3]. We also drew inspiration from Parable of the Polygons by Vi Hart and Nicky Case[4], and many other works[5–7].

The Gelbox Interface



Commands & Shortcuts

Bands

Duplicate	Hold Option, click & drag, then release.
Move band	Click & drag. Hold Shift to snap size.
Select next	Tab key; Shift+tab to reverse.
Delete	Select and press delete or backspace.

Samples

Duplicate	Hold Option, click & drag sample.

Move sample	Click tube, drag to an empty well.
Delete sample	Click tube, drag up, and release.
Add species	Click on + button with a sample selected.

Loupes

Show loupe	Hold ⌘ key when cursor is over gel.
Fixed loupe	Hold ⌘ key and click on gel. Draggable.
Delete loupe	Click on loupe and press delete.

Files

Save gel	⌘-S will save a file with extension ".json"
Load gel	⌘-O, or drag file into window. Or use app menu.
Save Sample	⌘-Option-S will save a sample file with extension ".json"
Load Sample	Drag a gel ".json" file onto the window or drag a sample ".json" file into the desired lane.
Load image	Drag image (file extensions: png, jpg, gif) file onto the window.

Discussion

We built Gelbox as a learning tool to complement the ad hoc learning methods routinely used in laboratory settings. While working with real gels provides valuable hands-on experience, the learning process is constrained by a slow feedback loop. Several hours can pass between making a mistake in preparing a gel and observing its effects. Fortunately, computer simulation can reduce the time between modifying a gel parameter and observing the effects to fractions of a second, making cause-and-effect relationships visible and explicit.

While other gel simulators do exist, they only provide one static graphical representation of predicted DNA fragments under perfect experimental conditions[8–11]. These tools are typically included with software for gene editing or DNA cloning. In contrast, Gelbox provides multiple dynamic representations of a gel under a wide range of conditions. Molecules can vary in size, concentation, shape, or degree of aggregation and degradation. To our knowledge, Gelbox is the only gel simulation tool that attempts to capture ways in which gels can be imperfect or fail.

To build the simulation model, we started with the goal of writing code that would convert DNA length (bp) to y-axis band mobility in the gel. When starting out we found it useful to reference expressions from Van Winkle, Beheshti, and Rill[12], and the dissertation of A. Beheshti[13], but we ultimately developed our own model directly from gel image data for a 1kb ladder. If you wish to understand the simulation model in the Gelbox source code, a good starting point is GelSim.cpp, which converts the parameters of a sample species and its environment to band geometry. Many of the simulation parameters are hot-loaded from json files in the tuning folder.

Gelbox has several limitations. First, the simulation model is relatively narrow in scope compared to a real gel. We spent most of our effort on creating the various representations

and did not have time to account for many parameters and molecule types. We focused on agarose gels and DNA fragments ranging in size from 100 to 14000 base pairs. Gelbox does not currently support RNA or protein molecules. Second, the simulation could benefit from further calibration. We did not systematically validate all parameters with experimental data, thus the predicted band mobility and behavior may not match real-world results. Third, our models do not have any physical basis, that is, we do not account for kinetic or thermodynamic properties of the DNA molecules. However, Gelbox might be used as a visualization tool for other physically based simulations if they were adapted to output our text-based file format.

In summary, we have created a dynamic, interactive gel electrophoresis simulator with multiple linked representations of the system. Gelbox may help learners establish some basic intuition for interpreting gel data and for troubleshooting problematic experimental conditions. It has several limitations, many of which could be addressed with further development, including more experimental validation and integration with other simulation tools. We hope our work inspires others to build additional interactive simulations for other aspects of scientific research.

Acknowledgements

We thank J Brown for assistance with the app icon and introductory video, and P Nafisi for running several gels used to calibrate the simulation model. We thank P Rothemund and N Case for helpful feedback and suggestions. This work was supported by Pew-Stewart Scholars Program for Cancer Research, the National Science Foundation Center for Cellular Construction (DBI 1548297), and the UCSF Program for Breakthrough Biomedical Research, which is partially funded by the Sandler Foundation.

References

1. A. Tiselius, *Electrophoresis and Adsorption Analysis as Aids in Investigations of Large Molecular Weight Substances and Their Breakdown Products*, 1948 Nobel Lecture.
2. C. Gingold, *Earth Primer*, 2015.
3. B. Victor, *Up and Down the Ladder of Abstraction*, 2011.
4. V. Hart and N. Case, *Parable of the Polygons*, 2014.
5. N. Case et al., *Explorable Explanations*, 2015.
6. The Concord Consortium, *Molecular Workbench*, 2013.
7. Desmos *Calculator*, 2011.
8. DNAstar, *Agarose Gel Simulations*, 2017.
9. Snapgene, *Simulate an Agarose Gel*, Retrieved 2018.
10. Genome Compiler, *Virtual Gel Simulation*, 2016.
11. MacVector, *Agarose Gel Simulation*, Retrieved 2018.
12. D.H. Van Winkle, A. Beheshti, R.L. Rill, *DNA electrophoresis in agarose gels: A simple relation describing the length dependence of mobility*, 2002 Electrophoresis 23, 15–19.
13. A. Beheshti, *DNA electrophoresis in agarose gels: A new mobility vs. DNA length dependence*, 2002 Dissertation, The Florida State University.