

Blueprint for: GeneBank parsing and visualization

1. GUI mockups

- User experience:
 - Quoting the Mentor (Dr.El Mazouari)

“Practical use case: a team is developing a web app that implements in-house algorithms for annotated in-house proprietary sequences. The web app screens the company sequence database for specific set of features. Sequence hits are then annotated and a Genbank output is generated for annotated sequences.

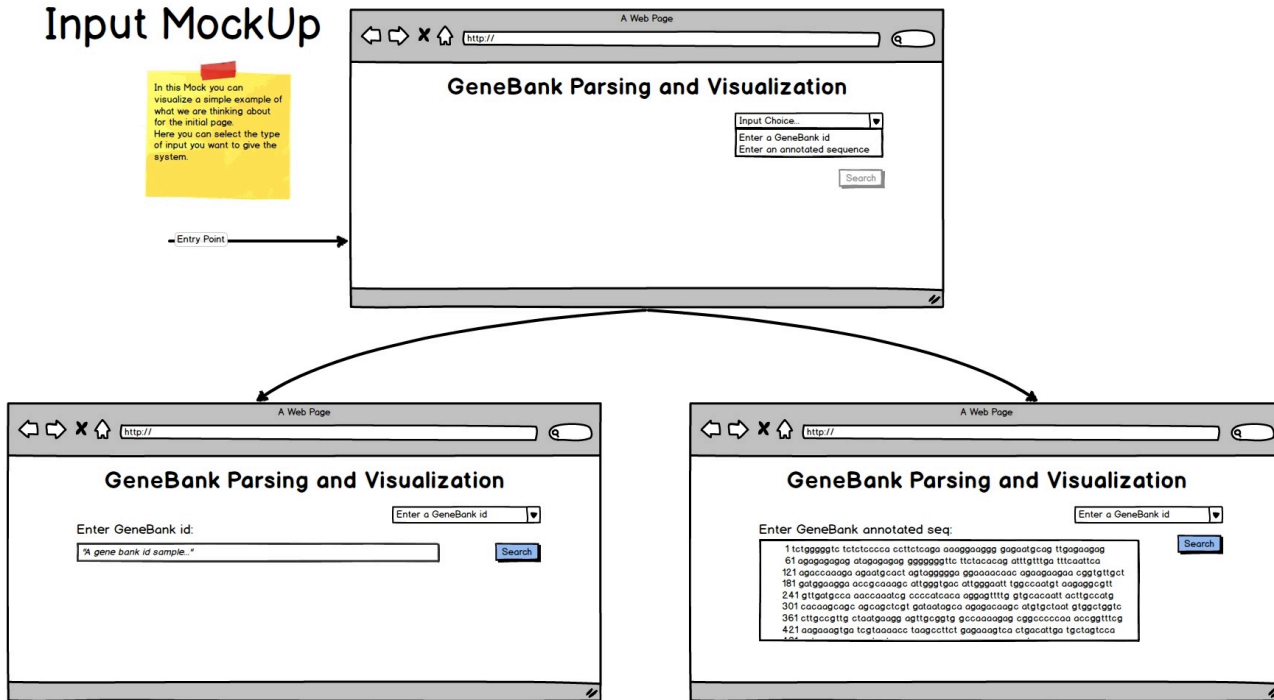
At this time, wet-lab users download the annotated sequence in Genbank format and then open it in VectorNTi or MacVector in order to view the annotation map and features. These extra steps are time consuming... If they can view the sequence directly in their browser from the web, they will be more productive and “happy scientists;”

Something that will help them to view the annotated sequence, select the features they want and export them will be very welcome”

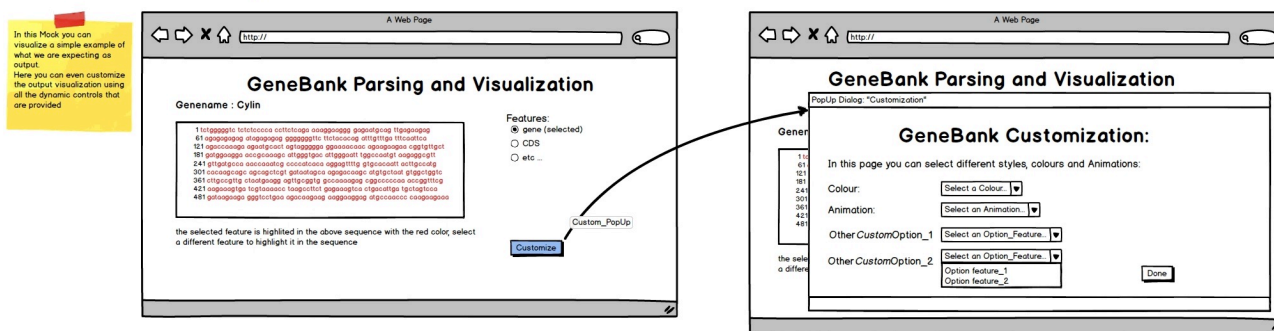
- Functionality
 - Select features from annotated input sequences
 - Parse and Visualize the input sequence in the genbank format
 - Export selected features which should be able to work with later
- Features
 - Easy to use for the end users
 - Highlight and export features in a user friendly
 - Should easily be able to integrated into other web applications.

- MockUps:
 - These are the original MockUps done before contacting the client:

Input MockUp



Output MockUp

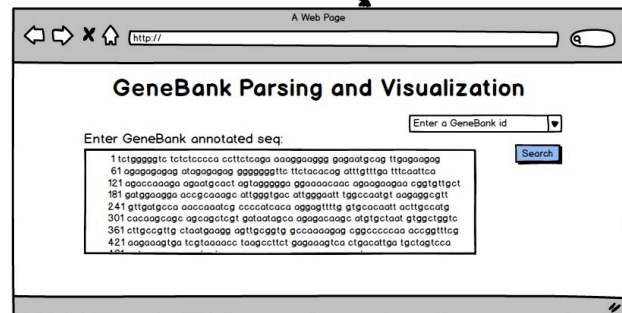
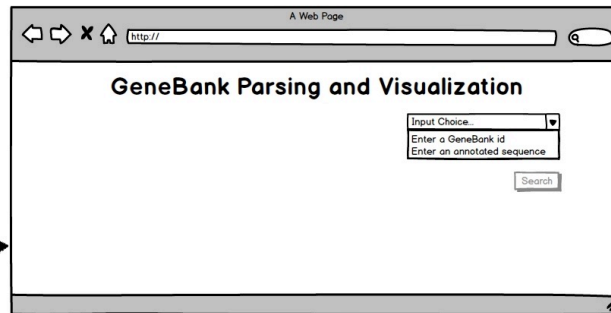


- These are the refactored MockUps:

Input MockUp

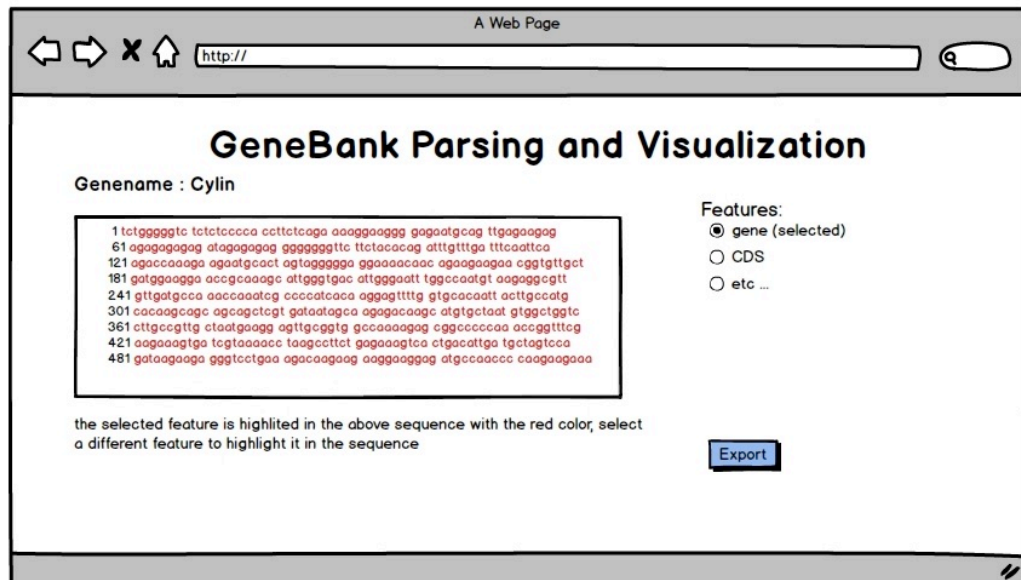
In this Mock you can visualize a simple example of what we are thinking about for the initial page. Here you can select the type of input you want to give the system.

Entry Point



Output MockUp

In this Mock you can visualize a simple example of what we are expecting as output. Here you can even customize the output visualization using all the dynamic controls that are provided



2. Application design

- Expected technical difficulties
 - Implementing the parser
 - Selecting and exporting features dynamically
 - Highlighting multiple features
- Fancy libraries you plan to use
 - JQuery
 - D3 (if necessary)
 - BioJava
 - BioJS(?)

3. Your data

- Remarks about your input format
 - The input is going to be annotated sequence and it should be in the Genbank Format

4. Roadmap

1. Understand the application domain and the logic that we are supposed to implement
2. Understand the input we need to work (how to convert sequence in genbank format to genbank file)
3. Parse the genbank file and extract the features dynamically into Javascript objects.
4. Visualize the results in a user friendly in the browser
5. Make the features exportable!