GeneJax: A Prototype CAD tool in support of Genome Refactoring

Ishan Anand
June 7, 2006

What is GeneJax

- “Google Maps for Genomes”
  - Google Map style navigation
  - JavaScript (aka “AJAX”) Web Browser Application
- Supports the dissection of genomes during genome refactoring process

This Presentation

- Introduction
  - Genome Refactoring
  - Motivatation for GeneJax
- GeneJax Features
  - What worked, what didn't.
- GeneJax Implementation Details
- Future Enhancements

Introduction

Software Refactoring

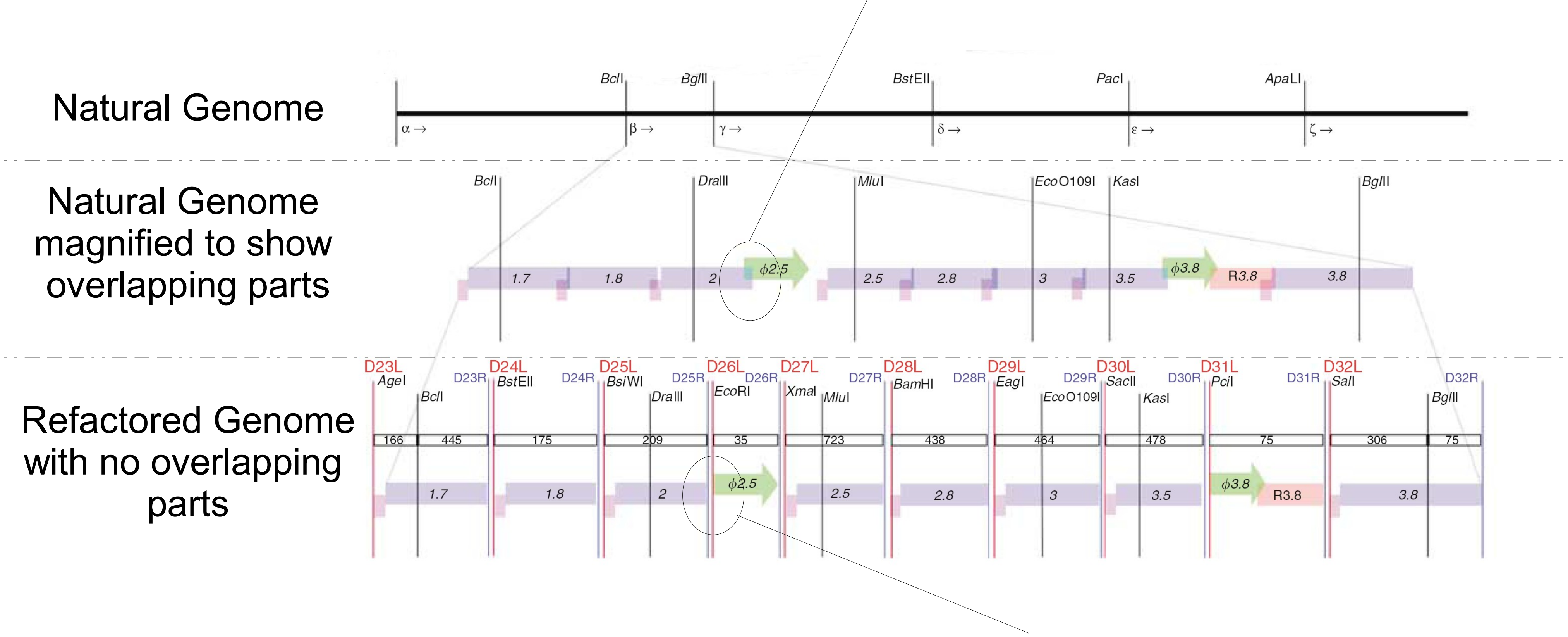
- Refactoring:
  - Redesign of a software program
  - Improves isolation between functional modules
  - Removes unused and redundant modules
  - Adds no new features.
- Benefits:
  - Easier to understand: Each module has less complexity and can be tested inpendently
  - Easier to modify: Modules can be replaced or altered independently

Genome Refactoring

- Refactoring:
  - Redesign of a genome
  - Improves isolation between functional elements (i.e. Genes, promoters, etc.)
  - Removes unused and redundant elements
  - Adds no new features.
- Benefits:
  - Easier to understand: Each genome section has single purpose and can be tested inpendently
  - Easier to modify: Elements can be replaced or altered independently

Genome Refactoring

Region of overlap between parts



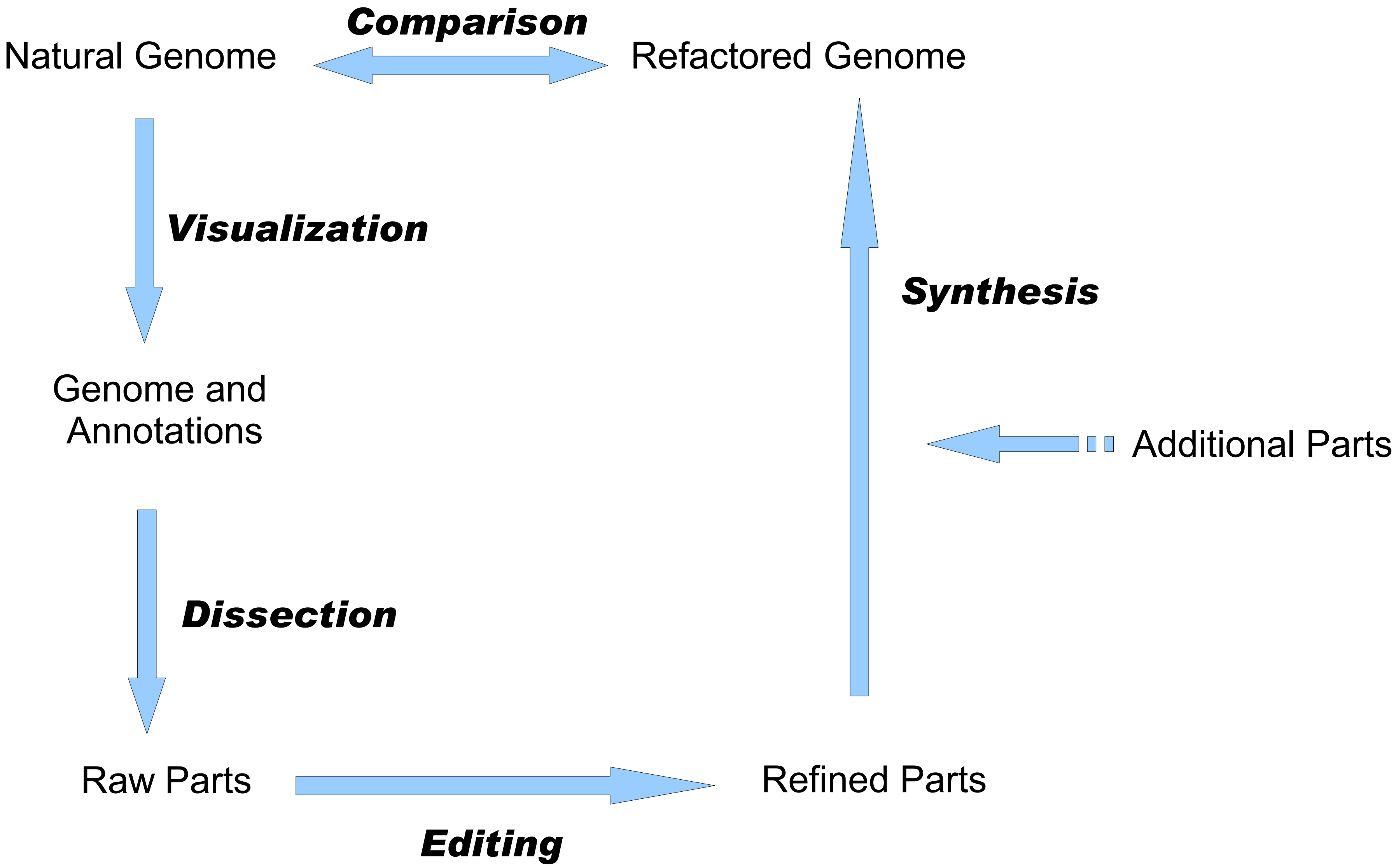
Natural Genome

Natural Genome magnified to show overlapping parts

Refactored Genome with no overlapping parts

Region with overlap between parts removed

Genome Refactoring Process



Natural Genome

Comparison

Refactored Genome

Visualization

Synthesis

Genome and Annotations

Additional Parts

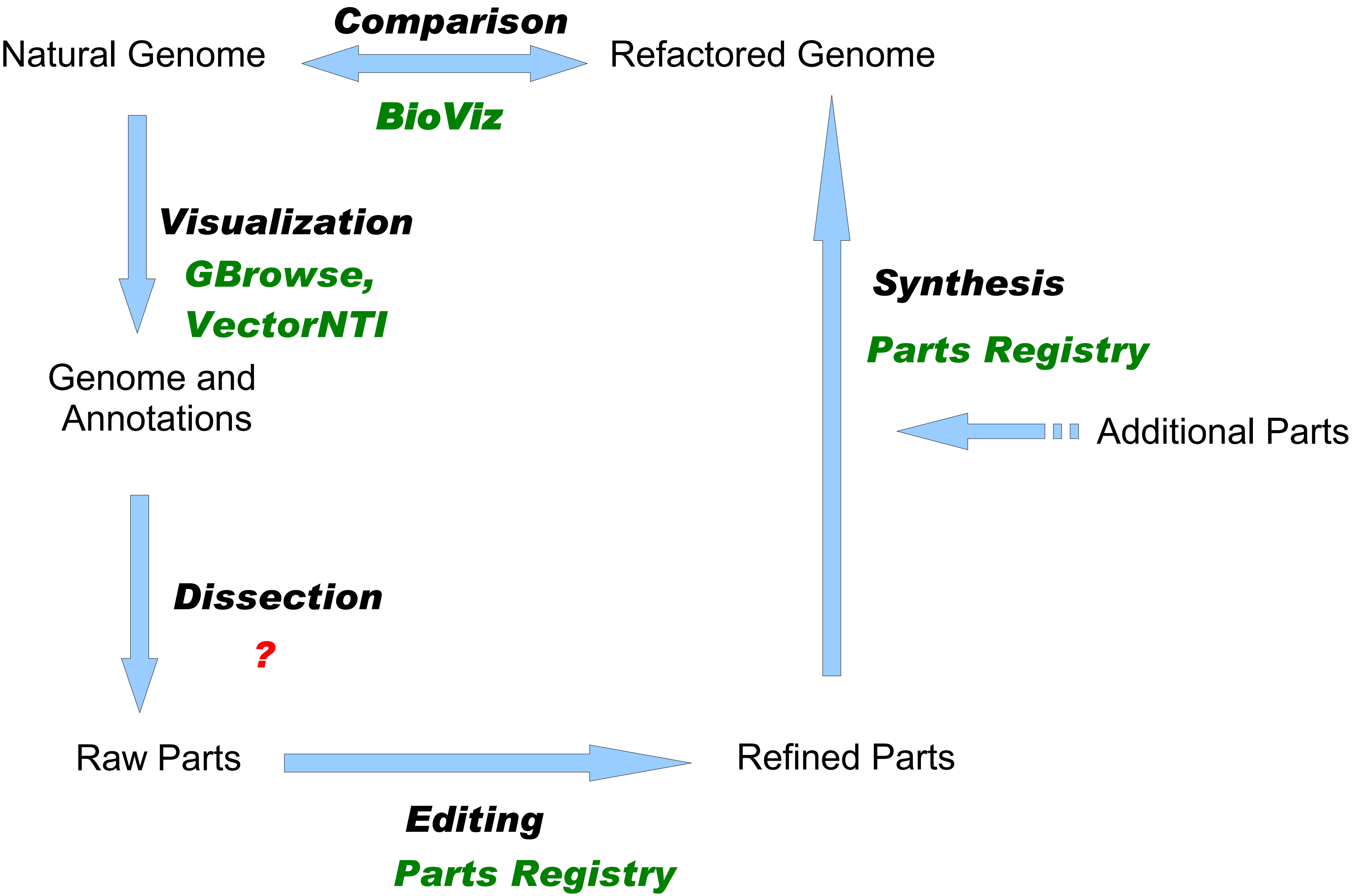
Dissection

Raw Parts

Refined Parts

Editing

Genome Refactoring Tools



Natural Genome

Comparison

Refactored Genome

BioViz

Visualization

GBrowse,

VectorNTI

Genome and
Annotations

Synthesis

Parts Registry

Additional Parts

Dissection

?

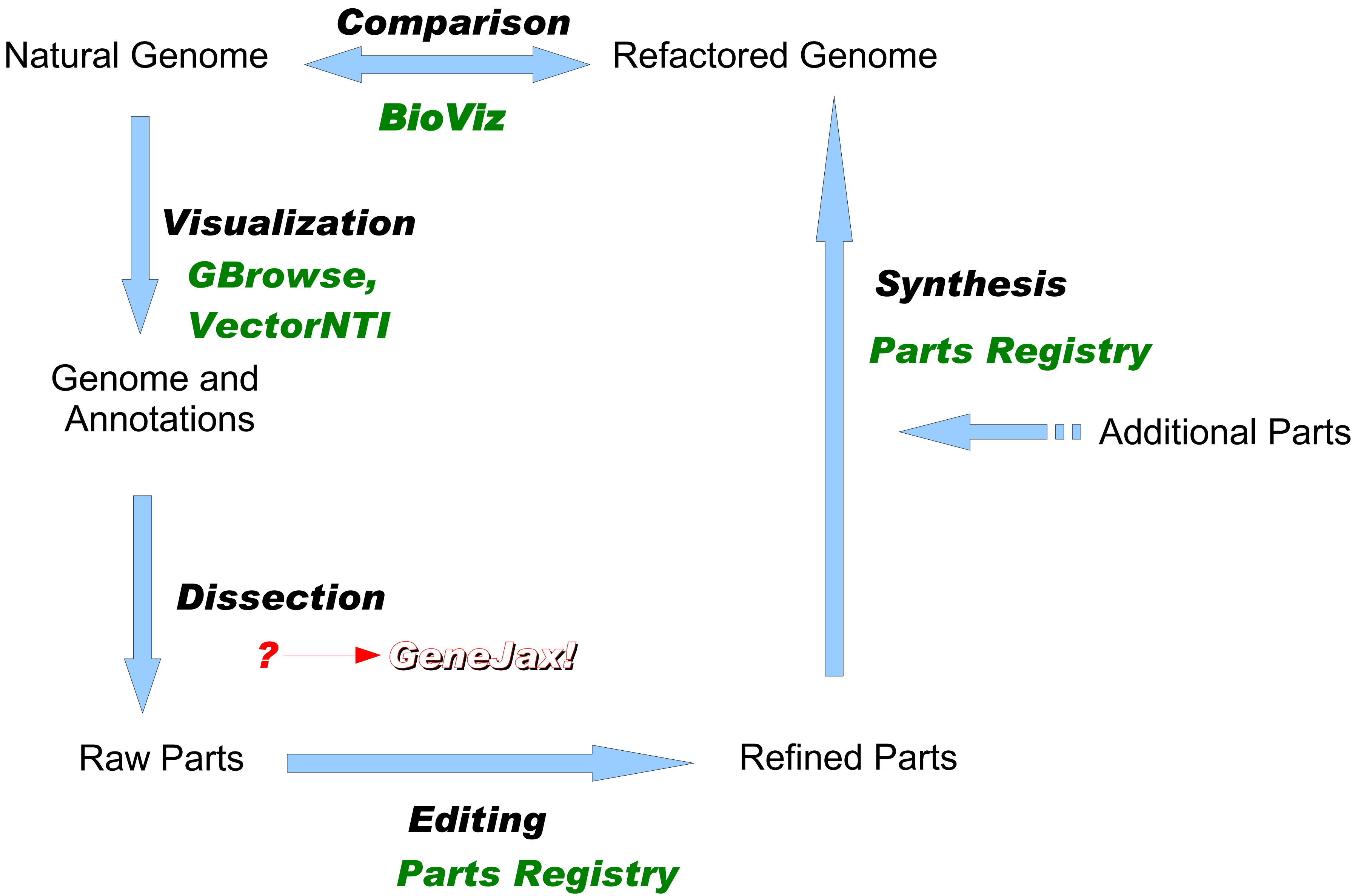
Raw Parts

Refined Parts

Editing

Parts Registry

Dissection Tool Needed



Natural Genome

Comparison

BioViz

Refactored Genome

Visualization

GBrowse,

VectorNTI

Genome and
Annotations

Dissection

? → **GeneJax!**

Raw Parts

Editing

Parts Registry

Refined Parts

Synthesis

Parts Registry

Additional Parts

GeneJax Goals

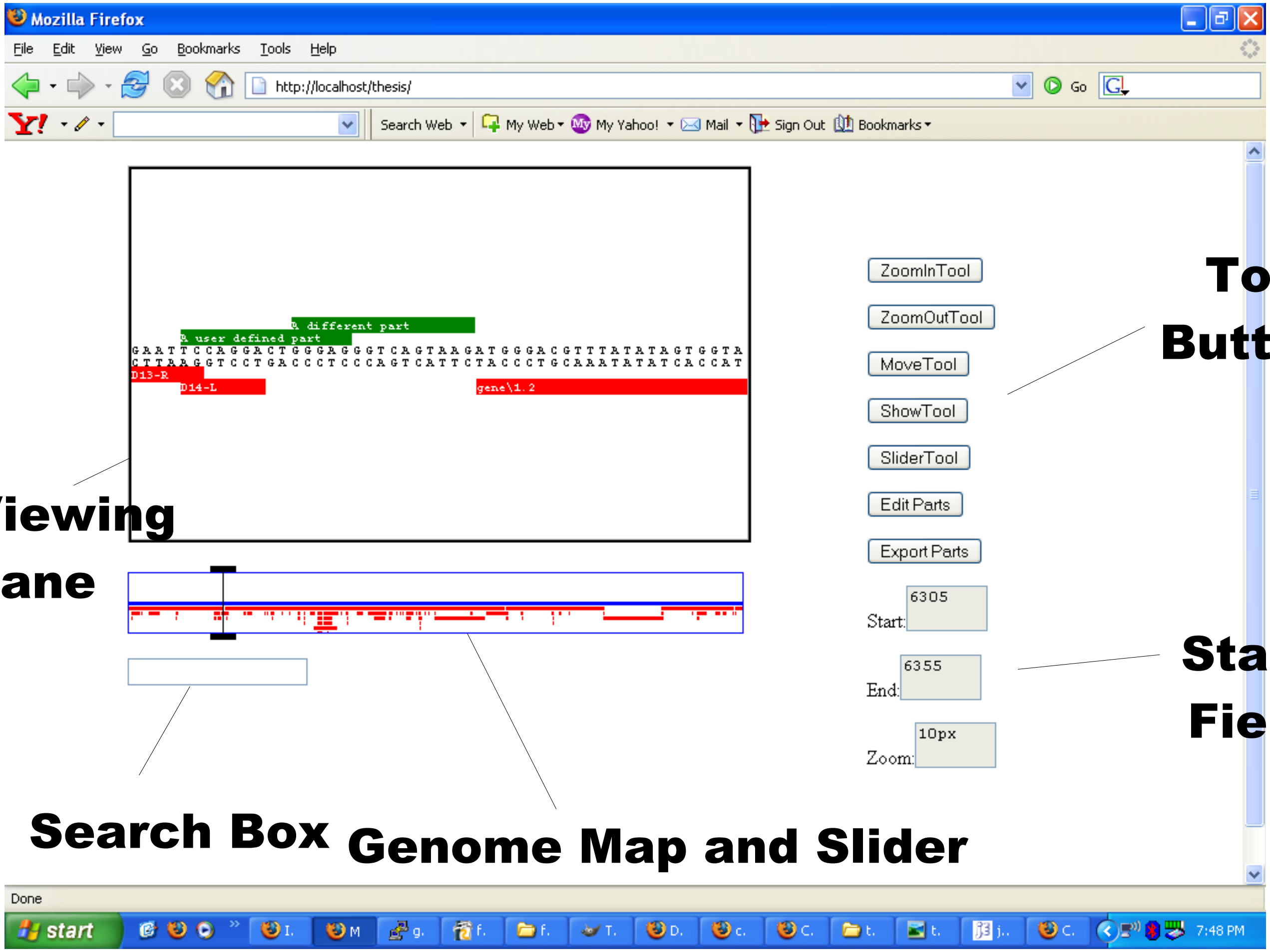
- Genome Dissection tool for Refactoring:
  - Define and extract parts from a natural genome sequence
- Also a Visualization tool by necessity

Inspired by Google Maps

- Similar tasks
  - Map a local region from large data set
  - “Click-drag” navigation
  - Zoom in/out
  - Multiple viewing modes
- Plan
  - Build a Google Maps for T7.1 genome
  - Add part manipulation features

GeneJax Features

GeneJax User Interface



http://localhost/thesis/

Start: 6305

End: 6355

Zoom: 10px

Tool Buttons

Viewing Pane

Status Fields

Search Box

Genome Map and Slider

GeneJax Navigation Features

- Planned Features
  - Click-drag
  - Zoom In/Out
    - Useful range is limited
- Problems
  - Partly needed because zoom didn't work
  - Difficult to move around a genome by click-drag alone.
  - User felt lost. Need to know where in the genome they were.

GeneJax Navigation Features

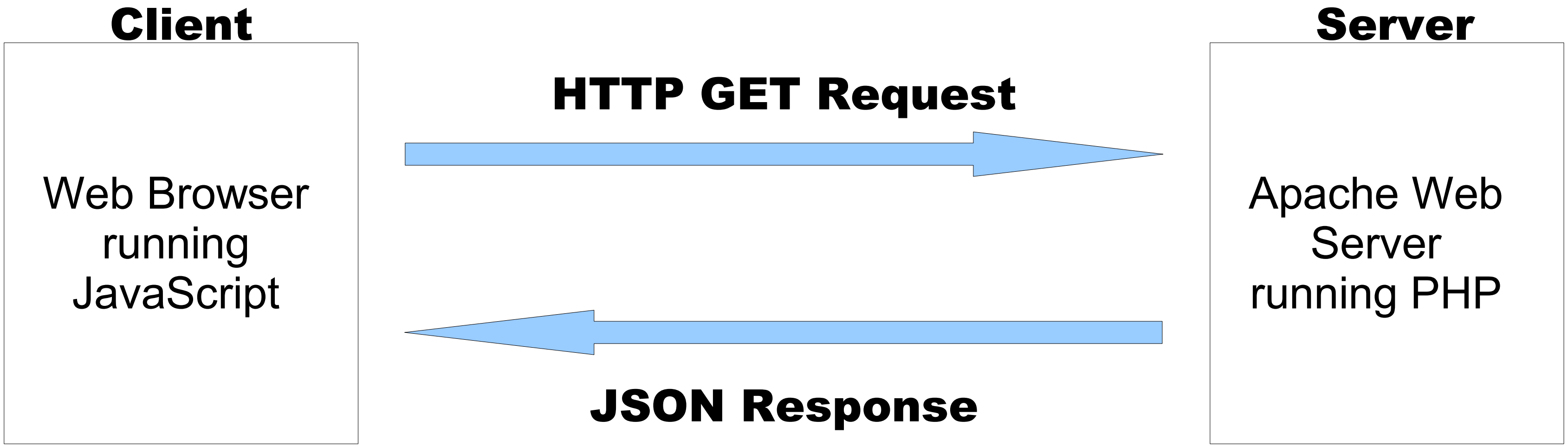
- New Navigation Features
  - Search box
    - Forgot search! Essential part of Google Maps. 'Duh!
    - Jump to any annotation or part
    - Auto suggests parts/annotation
  - Map
    - Map of genome
    - “Slider” indicating current position on map
    - Google Maps later adopted something similar during development of GeneJax

GeneJax Dissection Features

- Part Visualization
  - Parts above sequence
  - Annotations below the sequence
- Part Manipulation
  - Create
  - Rename
  - Delete
  - Export

GeneJax Implementation

GeneJax Architecture



Client

Web Browser running JavaScript

HTTP GET Request

JSON Response

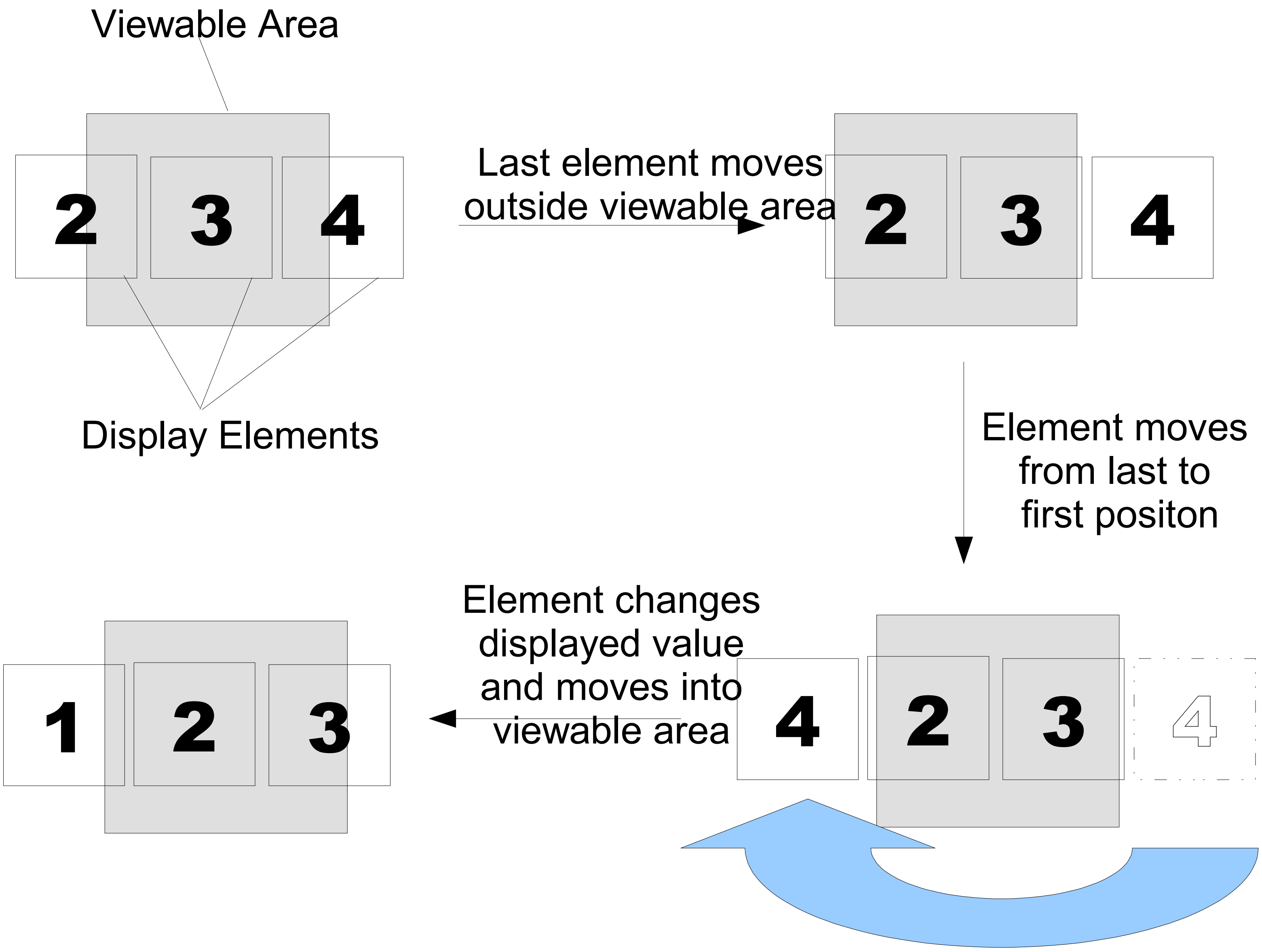
Server

Apache Web Server running PHP

Click-drag animation

- Basic display elements are text elements
  - Google Maps, etc. display elements are images
  - Leverage font resizing for zoom features
- JavaScript animation
  - Sluggish at first.
  - Biggest improvement: Caching DOM Access
- Two algorithms
  - Base pairs (fixed number of display elements)
  - Parts (variable number of display elements

Base Pair Animation



Viewable Area

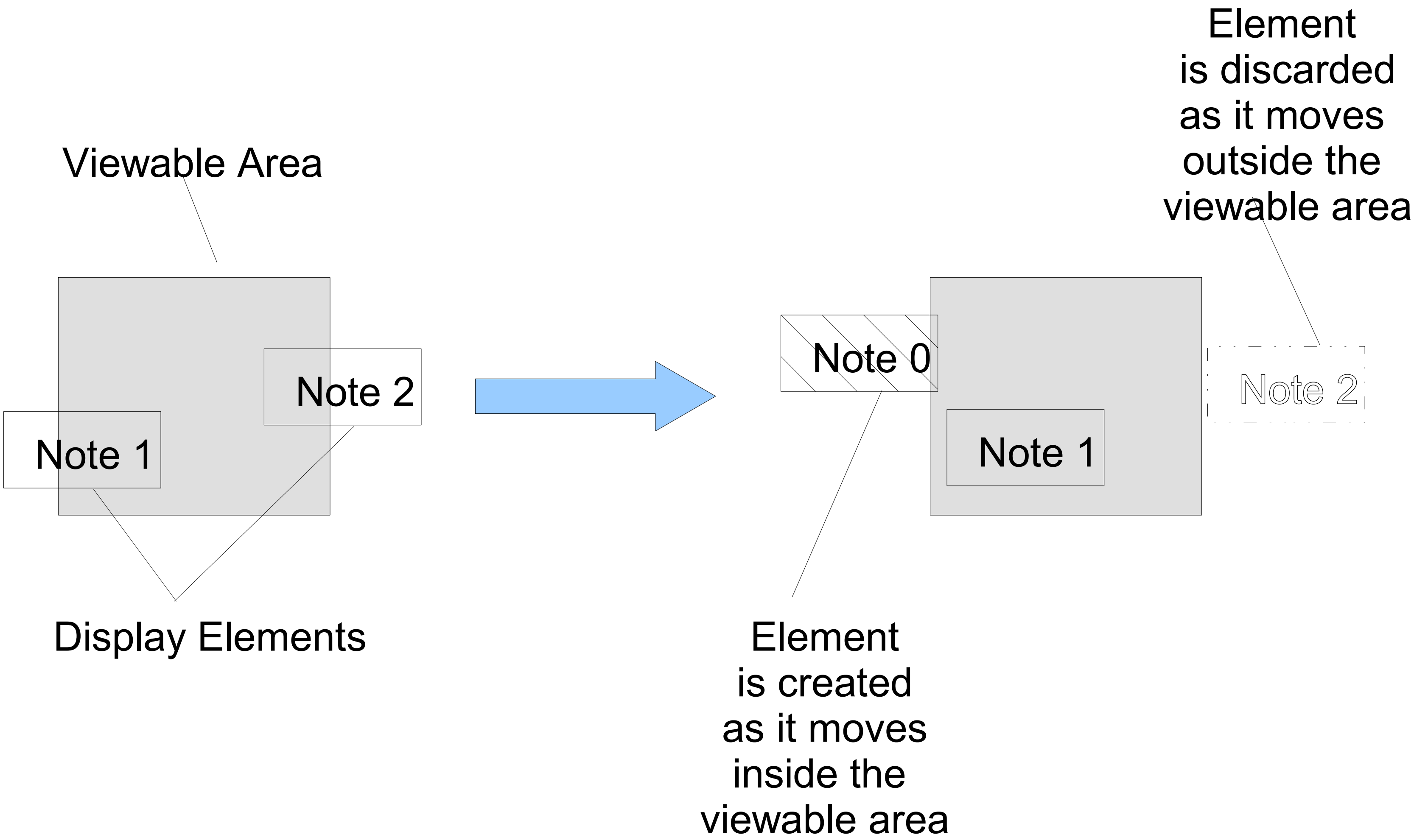
Display Elements

Last element moves outside viewable area

Element moves from last to first positon

Element changes displayed value and moves into viewable area

Part/Annotation Animation



Viewable Area

Note 2

Note 1

Display Elements

Note 0

Note 1

Note 2

Element
is discarded
as it moves
outside the
viewable area

Element
is created
as it moves
inside the
viewable area

GeneJax Server

- Server
  - Loose collection of PHP scripts
  - Not major focus of development effort
  - Just wanted something that works
- Base Pair Lookup
  - Flat text file
  - n^{th} base pair is at n^{th} byte of file
- Annotation Lookup
  - PHP file evaluated by server to get annotation information

Comparison with GBrowse

- Another “Google Maps for Genomes” project
- http://genome.biowiki.org/
- Independently proposed
- Very cool. Very impressive.
- Different approach
  - No on-the-fly rendering by client or server
  - Display elements are images
  - Only visualization, no genome dissection

Comparison with GBrowse

- Advantages
  - More efficient than GeneJax for server and client
  - Can display any type of annotation
- Disadvantages
  - No part manipulation features yet, but they have thought about it.
  - Difficult to independently manipulate parts
    - Base pairs and annotations are inseparably commingled together into static image.

Future Enhancements

Future Navigation Enhancements

- Zoom needs to be fixed
  - Dynamically switch to “functional” view
- Possible map/slider improvements
  - Zoom within map
  - Multiple maps/sliders at different scales
  - Map is always “zoomed” out by fixed ratio
- Web style navigation
  - Forward, back button
  - Bookmarks

Future Enhancements

- Integration with MIT Biological Parts Registry
- Scaling to support larger Genomes
  - Client limited by
    - “Local” number of annotations. Will scales
    - “Global” number of parts. May scale. Can get help from server if it does not.
  - Server
    - Base pair retrieval should scale
    - Annotation retrieval may not scale.
    - Hard to maintain. Migrate to a real database?