MoClo Planner:
Interactive Visualization for Modular Cloning Bio-Design

Orit Shaer[1], Consuelo Valdes[1], Sirui Liu[1], Kara Lu[1], Traci L. Haddock[2], Swapnil Bhatia[2], Douglas Densmore[2], Robert Kincaid[3]

[1]Wellesley College, [2]Department of Electrical and Computer Engineering Boston University, [3]Agilent Technologies

ABSTRACT

MoClo Planner is an interactive visualization system for collaborative bio-design, utilizing a multi-touch interactive surface. The system integrates the information gathering, design, and specification of complex synthetic biological constructs using the Modular Cloning (MoClo) assembly method. Modular Cloning is a hierarchical DNA construction method that allows for the assembly of multi-part constructs from a library of biological parts in a one-pot reaction. This cutting-edge method facilitates and expedites the assembly of complex biological designs. However, it is an intricate multi-step process, which to date, has not been adequately supported by existing bio-design tools. Novel visual tools are needed in order to make MoClo more tractable and accessible to a broad range of users, to facilitate a less error prone bio-design process, and to improve workflow. MoClo Planner is a result of a participatory and user-centered design process, which included close collaboration with domain experts. Using multi-touch interactions and a rich graphical interface, the system accelerates the MoClo learning process, and reduces design time and errors. In this paper, we present user requirements and describe the design, implementation, and evaluation of MoClo Planner.

Keywords: Synthetic biology, Modular Cloning, Bio-design.

Index Terms: K.6.1 [Management of Computing and Information Systems]: Project and People Management—Life Cycle; K.7.m [The Computing Profession]: Miscellaneous—Ethics

1 INTRODUCTION

Synthetic biology is an emerging research area that couples engineering and biology with the goal of designing organisms with new specified behaviors useful in particular applications. This opens the door to many breakthrough applications, including therapeutics, environmental decontamination, and *in vivo* sensing. For example, Keasling *et al*. engineered *Saccharomyces cerevisiae* to produce a precursor of the anti-malarial drug artimisenin [39]. Anderson *et al*. engineered the bacterium *Escherichia coli* to detect, invade, and destroy tumor cells [2]. These early advances point to the potential of synthetic biology, and the growing size of the annual undergraduate international Genetically Engineered Machine (iGEM) [27] competition indicates that this promise will be explored in the near future.

The field applies engineering principles such as abstraction and modularity into biological research, while solving problems by applying a forward engineering approach: composing a specification of the behavior to be designed into an organism, and then selecting genetic parts and their regulatory architecture to achieve the functional goal. Here, genetic elements are treated as standardized biological parts, and used like "Legos" based on data about their behavior in a simpler context. An implication of this approach is that synthetic biologists must engineer complex systems based on uncertain biological mechanisms.

IEEE Symposium on Biological Data Visualization 2013
October 13 - 14, Atlanta, Georgia, USA
978-1-4799-1659-7/13/$31.00 ©2013 IEEE

The bottlenecks and challenges along the path to realizing the full potential of this field are formidable and numerous. For one, synthetic biology design requires large design spaces to be sampled combinatorially while applying voluminous experimental design for each design candidate. The planning, execution, and the tracking of experimental data and results are currently implemented using ad-hoc processes that limit the scale and complexity of biological design [57]. The cost and skill level required for using biological design technologies pose another challenge, limiting the opportunities for future scientists (a.k.a. students) to engage in experiential learning in synthetic biology.

The goal of this project was to apply key areas of innovation in Human-Computer Interaction such as multi-touch interaction and large interactive surfaces to facilitate problem-driven learning and understanding in synthetic biology. More specifically, this project examines how a visually rich multi-touch interface can be designed to enhance an *Assembly Planning* workflow, which is central to synthetic biology. Generally, the workflow begins with a set of genetic designs to be assembled from a set of available genetic parts, and ends with a set of exact sequences to be stitched together in the laboratory. This is a collaborative workflow that involves investigators, graduate and undergraduate students, and is used in most synthetic biology labs. We focus on a particular case study: Assembly Planning for Modular Cloning.

Modular Cloning (MoClo) [54] is a cutting-edge experimental method that facilitates and expedites the assembly of novel biological designs. It is a hierarchical DNA construction method that allows for the assembly of complex constructs from biological parts in a one-pot reaction. While this method allows for efficient construction, the design of MoClo systems is an intricate multi-step process, which to date, doesn't have adequate computational support. Thus, this process is time-consuming, error-prone, and is difficult for novices to understand. Here, we present the design, implementation, and preliminary evaluation of a novel visual tool—MoClo Planner, which makes MoClo more tractable and accessible to a broad range of users, facilitates a less error-prone bio-design process, and improves workflow.

Our contribution in this paper is threefold. First, we describe the MoClo workflow, followed by an analysis of users, tasks, and potential errors in the MoClo design process. This information was collected through a user-centered design process, which included close collaboration with domain experts. Second, we present the design and implementation of MoClo Planner, an interactive visualization system for collaborative bio-design, utilizing a multi-touch interactive surface. Finally, we present results from two preliminary user studies that assessed the usability and usefulness of MoClo Planner. We discuss our design decisions and lessons learned.

We begin by providing background on the Modular Cloning method and follow with a survey of related work.

2 Background

Modular cloning (MoClo) is an alternative assembly standard based on the Type IIS restriction enzymes BsaI and BpiI, which was introduced in 2011 by Weber et al. [54]. BsaI and BpiI are Type IIS restriction enzymes that have different six base pair non-palindromic recognition sequences and cut at positions +1/+5 and +2/+6, respectively. A key feature for using these enzymes is the ability to remove the restriction sites from the final assembly product, thus eliminating the enzyme site scar in the assembly.

The four base pair overhang, or *fusion site*, created by either enzyme is user-defined. These fusion sites are used to enable directional assembly of biological parts. By defining non-palindromic fusion sites that flank each part, we can ensure proper order of assembly by assigning the same fusion sites between the 3' and 5' ends of two adjacent parts. MoClo currently exists in three modules: Level 0, Level 1, and Level 2. Level 0 refers to basic biological parts used in synthetic biology, including promoters, ribosomal binding sites, coding sequences, terminators, and so on. Level 1 refers to a concatenation of up to 6 Level 0 Modules, most often arranged in transcriptional units but not necessarily so. Level 2 refers to a concatenation of up to 6 Level 1 Modules. It is theoretically possible to generate an infinite number of MoClo Levels but other assembly constraints restrict the number of levels possible in the laboratory. Another major component to each MoClo Level beyond the biological part is the Destination Vector (DV) that holds the part and contains a handful of key features essential for the MoClo reaction. Figure 1 shows the assembly of a MoClo Level 1 construct.

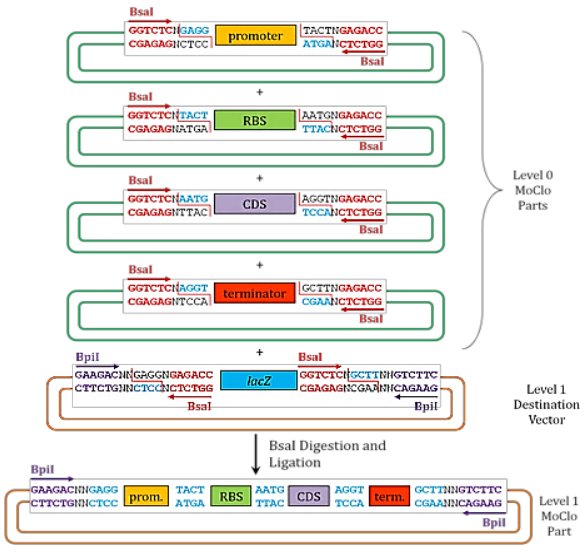

Figure 1: Construction of a Level 1 MoClo Module

3 Related Work

Software Tools for Biological Design

A majority of the current software tools in synthetic biology for supporting design and planning are intended for users with intermediate to advanced skills. Software tools such as Geneious [16] and Ape [6] are popular desktop tools for viewing and editing DNA sequences. Other software tools (e.g., Genome Compiler [17], and Vector Express [53]) work between the sequence and part level of abstraction allowing manual composition of genetic designs. Various primer design tools allow sequence-level engineering of biological parts based on biophysical constraints (e.g. [41, 51]). GenoCAD [13] and Eugene [10] support the constrained combinatorial specification of biological designs. Cello [18] and the Proto BioCompiler [8] are tools for programmable functional specification of biological designs. A number of tools support the curation of biological parts and designs including the MIT Parts Registry [38], and Clotho [57]. Other software tools such as TinkerCell [14] and iBioSim [34] support the modeling and simulation of biological designs. The goal of the Synthetic Biology Open Language (SBOL) [47] consortium is to standardize formats for exchange of data and designs. To date, no software tools support the assembly planning workflow of the MoClo method. Our focus is on providing integrated support for the MoClo design process that makes this process more accessible to novice users.

Reality-Based interfaces for scientific discovery and education

Over the past two decades, Human-Computer Interaction (HCI) research has generated a broad range of interaction styles that leverage users' existing knowledge and skills of interaction with the real, non-digital world such as naive physics, spatial, social and motor skills [28]. Drawing upon users' skills of interaction with the real non-digital world to a greater extent than traditional Graphical User Interfaces (GUI), these interaction styles are often unified under the umbrella of Reality-based Interfaces (RBIs) [28]. By basing interaction on pre-existing real-world skills, RBIs offer a natural and intuitive form of interaction that reduces the mental effort for learning and operating a computational system.

Several RBIs examine the possibilities of supporting scientific discovery and education. Following, we describe those most relevant to our work. Brooks et al. [12] developed the first haptic display for scientific visualization. Gillet et al. [19] presented a tangible user interface for molecular biology that used augmented reality molecular models. Schkolne et al. [42] developed an immersive tangible interface for the design of DNA molecules. While these systems highlight potential benefits of RBIs for scientists, they focus on the representation of objects with inherent physical structure. We are interested in a broader use case, where abstract information is represented and manipulated.

A few systems were developed to facilitate collaboration among scientists across large multi-touch displays (e.g., [40, 48, 49, 55]). However, these systems target expert scientists rather than novices and are not aimed for learning new concepts. The eLabBench [48, 49] is a tabletop system which provides users with situated access to information in the wet lab. Our focus is on enhancing the computational (rather than wet) workbench. Finally, several RBI systems have illustrated the potential to support science education. Most relevant to our work are: PhyloGenie [43]—a tabletop interface for collaborative learning of phylogeny. Involv [24] and DeepTree [11] are tabletop interfaces for exploring the Encyclopedia of Life that are aimed at informal science learning settings. G-nome Surfer is a tabletop interface for collaborative exploration of genomic information [44, 45, 46]. In this work, we go beyond information exploration and delve into the design and specification of new biological systems.

4 User and Domain Analysis

Our first contribution is a domain analysis that includes the identification of users, workflow, tasks, and potential pitfalls in the MoClo design process. This analysis is based on on-site interviews and observations. Also, the fifth co-author is a

synthetic biology researcher, experienced in the MoClo method. Following we describe our findings.

4.1 Users

At least three distinct user groups are involved in the learning and using of the MoClo design process. Each is marked by unique needs and learning goals:

First, *undergraduate biology students* are future scientists early in their career. Core learning goals in the synthetic biology curriculum include gaining experience in defining, specifying, and (whenever possible) implementing biological designs [29]. However, synthetic biology students have limited opportunity to develop design competencies and even fewer chances to implement their designed projects [32]. This can be attributed to three factors: 1) the limited knowledge of how biological parts interact; 2) the lack of mature computational tools for bio-design; and 3) the cost and availability of biological technologies.

Second, *iGEM teams* consist of students participating in the International Genetically Engineered Machine Competition (iGEM) [27]. The competition provides a project-based research experience driven by real-world problems, in which interdisciplinary teams of undergraduates learn by engaging in collaborative synthetic biology research. Students use a kit of standard biological parts and new parts of their own design, to build biological systems. It is expected that the MoClo method will become a standard assembly technique in iGEM. Our goal is to make MoClo more accessible for these audiences by making the process more tractable and by reducing the workload of managing the large amount of information associated with this process. We also seek to enhance learning by fostering collaboration [1, 3].

Third, *advanced practitioners*, since synthetic biology is an emerging field of research where research methods and techniques are still immature, even advanced practitioners in this field (e.g. graduate students and post-docs) are vacillating between expert and novice roles on various aspects of a research project. Thus, collaboration, particularly in the early stages of a project, is essential for tackling difficult problems and for dividing a project into sub-problems. Our goal for this audience is to make the MoClo design process more tractable and less error prone, as well as to improve workflow while supporting collaboration.

4.2 Workflow and Tasks

A MoClo assembly project is highly collaborative and typically involves investigators, undergraduate, and graduate students working together. The MoClo design process begins when the team creates a collection (i.e., library) of basic biological parts (i.e., Level 0 modules) which includes, but is not restricted to, four basic module types: promoters, 5' untranslated regions (or ribosomal binding sites), coding sequences, and terminators. These parts are typically selected from a local database of parts or from the Registry of Standard Biological Parts [38]. During this phase, researchers use a range of online databases to seek information about the different biological parts. Such information includes related publications, usefulness, assembly compatibility, and chassis. For each selected part, researchers record in a spreadsheet various information including common name, registry ID, category, sub-category, assembly compatibility, and sequence. The researchers then use a separate tool to design primers for each of the selected parts and add an appropriate fusion site based on the type of the part. The primer and fusion site sequences are recorded back in the spreadsheet. Parts from the same type get an identical fusion site if they are going to be used in the same order for the assembly. In addition, researchers need to specify destination vectors for Level 0 that are resistant to a particular antibiotic unique to Level 0.

Compatible sets of Level 0 modules are then combined into transcription units (i.e., Level 1 modules). Since all Level 0 modules from the same type have identical fusion sites, they are freely interchangeable. Thus, researchers often create a large number of design permutations. Since a desired behavior cannot be fully predicted from gene sequences only, the ability to generate multiple combinations of various coding sequences as well as many variants of their regulatory sequences is important [54]. These design permutations need to be validated against a set of rules and constraints. Currently, validation is done manually. Researchers also need to specify destination vectors for Level 1 that are all resistant to an antibiotic unique to Level 1.

Finally, up to 6 transcriptional units can be combined into Level 2 modules. Similar to Level 1, researchers might try a large number of design permutations in order to find a Level 2 construct with optimal performance. Level 2 modules also need to be validated against rules and constraints. Researchers specify destination vectors and antibiotic resistance for Level 2. Following the design of Level 2 modules, researchers need to create a laboratory protocol for manual assembly in the lab or an assembly graph for automated assembly using liquid handling robotics.

The MoClo design workflow relies on several databases and computational tools for retrieving, computing, and organizing the information required for MoClo design. These include public databases such the Registry of Biological Parts, iGEM, Google Scholar, and PubMed as well as web based tools such as PrimerQuest for designing oligonucleotides and ApE for viewing plasmid sequences. Information is retrieved from these databases and collected into a spreadsheet. Designs could be validated using Eugene [10] which is operated from the console, and final designs are specified using SBOL. External design visualization tools (e.g., Pigeon [9]) are used for depicting the final design. Users often switch between different tools, while copying and pasting information from one tool to another. This results in a process that is complicated and error-prone.

Also, while the typical MoClo workflow described above depicts a Bottom-Up Design process, which begins with low-level parts and then creates devices (i.e., complex constructs), an alternative approach to bio-design is Top-Down Design, which begins by specifying devices and then instantiating parts [8]. However, in practice, researchers often switch between Bottom-Up and Top-Down approaches in an iterative, non-linear manner. When considering the design of an interactive visual system for supporting the MoClo design process, it is important to allow users to manage complexity while experiencing the flexible workflow that characterizes authentic scientific problem solving.

Task	Subtask
Search for Level 0 Modules	Filter by type, category, or name.
Retrieve part background information	Filter by various parameters (e.g. name, category, author, sequence, publications).
Create collections (libraries) of Level 0 Modules	Select parts for Level 1
Explore combinatorial design space	Select target MoClo level to permute
Assign non-conflicting fusion sites	Select target MoClo level
Assign antibiotic resistance per level	Test for viability
Assign the restriction sites per level	
Review full sequence for correctness	
Export to create an assembly graph for lab automation/assembly	

Table 1: MoClo Design Task Analysis

We further analyzed this scenario by extracting a list of core MoClo design tasks that need to be supported by our system. This task list is shown in Table 1.

4.3 Errors

In attempting to support an effective MoClo design process, we next discuss potential MoClo design pitfalls. Three types of errors became apparent:

Data:
- Typos or incorrect data in the datasheet.

Design:
- Design permutations with incorrect regulatory architecture.
- Variations on primer design guidelines:
  - Incorrect orientation
  - Primer pairs that:
    - have different annealing temperatures are complementary
    - have a Delta G more positive than -9kca/mole when testing for both self- and hetero-dimers
  - Conflicting and palindromic fusion sites

Biological errors:
- Incorrect design with respect to a desired behavior.
- Use of toxic combinations of small molecules used to induce the devices

4.4 Requirements

Based on the domain analysis above, we identified the following requirements:

R1) Enabling an integrated and fluid workflow - allowing users to progress from the information they currently have (e.g., a database of standardized biological parts) to the information they need (assembly protocols for a set of complex biological constructs) through a series of computational steps while facilitating a flexible non-linear workflow that integrates disparate data sets.

R2) Facilitating a tractable process - providing users with means for searching, comparing, connecting, and organizing the large amount of data collected throughout the MoClo design process, while reducing the mental workload associated with this process. An interface that provides a clear indication of how the design process progresses and highlights the decisions made along the workflow could reduce both the load associated directly with designing complex biological constructs (i.e., task load) and the load keeping track of the process (i.e., syntactic load).

R3) Supporting collaborative and individual work patterns - The collaborative nature and extended duration of a MoClo assembly project (running from several weeks to months) implies that all stakeholders must participate in the design (i.e., planning) process and agree on desired results to ensure that the plan is efficient in time and cost, resilient to failures, and viable given the expertise and resources available. Thus, a visual tool for MoClo design needs to afford effective co-located collaboration. The system also needs to support an efficient workflow for individual users.

R4) Complying with community standards - Standardization is essential for the growth and maturity of the synthetic biology discipline. Sharing data, parts, and techniques across different laboratories is challenging without the use of standards. Likewise, sharing designs is difficult without the use of a standard notation. Thus, our system needs to support emerging synthetic biology standards including SBOL [47] and Eugene [10].

5 The MoClo Planner System

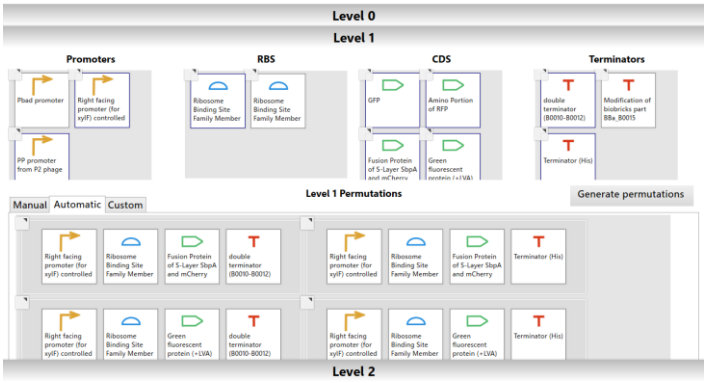

Figure 2: MoClo Planner, Level 1

Informed by these requirements, we designed and implemented MoClo Planner, a collaborative, multi-touch visualization for Modular Cloning Bio-Design. MoClo Planner is intended to be used during the design and specification stages of a synthetic biology experimental cycle as well as in later stages during analysis and reflection. These activities are often collaborative and typically take place in a conference room.

We designed MoClo Planner to support both individual and collaborative work. Co-located collaboration is mediated through multi-touch interaction upon a large interactive surface (SUR40 [31]), while individual work can be carried on any Windows 7 device. Our choice to design and implement MoClo Planner using a large vertical multi-touch surface was informed by current work practices of our users, which typically collaborate through side-by-side work on the whiteboard or a shared screen, as well as by existing research on interactive surfaces. Several studies indicate that large interactive surfaces support collaboration through visibility of actions and egalitarian input [25, 30], facilitate active reading [33], and foster the use of motor-spatial strategies, potentially lowering mental workload [36, 4]. It has also been shown that multi-touch tabletops promote reflection and collaboration in learning environments [21, 37, 45, 52]. Considering these findings as well as the increasing availability and falling prices of commercial hardware platforms, we decided to utilize multi-touch interaction with a large vertical interactive surface to address the requirements discussed above.

Following, we describe the design of the MoClo Planner system. We also discuss the design rationale we followed to support the tasks and meet the requirements that we outlined above.

5.1 Layered Organization

MoClo Planner visualizes the hierarchical MoClo design process using three layered workspaces, each layer corresponds to a level of the MoClo design process and is represented by a “shutter”, an animated vertical sliding panel. We used the metaphor of a shutter because it implies that a particular workspace can be represented as open, partially open, or closed at any given time. One advantage of this three-level design approach is that it enables users to move back and forth between the different shutters in a nonlinear fashion, keeping one or multiple shutters open at any given time visiting and revisiting different stages of the design process (R1). Additional advantage is that the three-layered design maps back to the levels of the MoClo process providing users with means for organizing the data collected throughout the MoClo design process, as well as a clear indication of how the MoClo design process progresses (R3).

The use of multi-touch input enables users to manipulate biological constructs across different stages of the MoClo process using spatial and direct interaction, which could engage the connection between the hand, the eye, and the brain to support users' conceptual understanding and facilitate "thinking through action" [3, 20, 56]. The MoClo Planner interface provides multiple points of entry by allowing graphical elements to be manipulated in parallel across different shutters.

In Level 0, the application allows users to browse a collection of over 2500 biological parts. This collection is drawn in real-time from the Registry of Standard Biological Parts [38]. Users can search the collection directly by part name, or browse by type and categories, such as behavior or function. Users can also add parts to a local database. Parts are visualized using an appropriate SBOL [47] symbol and are color coded based on their type. A context menu associated with each part allows users to view a data sheet and to access a primer designer. The data sheet presents the behavior of a part, its DNA and RNA sequences, and related publications. It is generated in real time, drawing information from the Registry of Standard Biological Parts, the iGEM archive, PubMed, and Google Scholar. Information about biological parts added by users to a local database can be edited manually. We describe the primer designer in the next section.

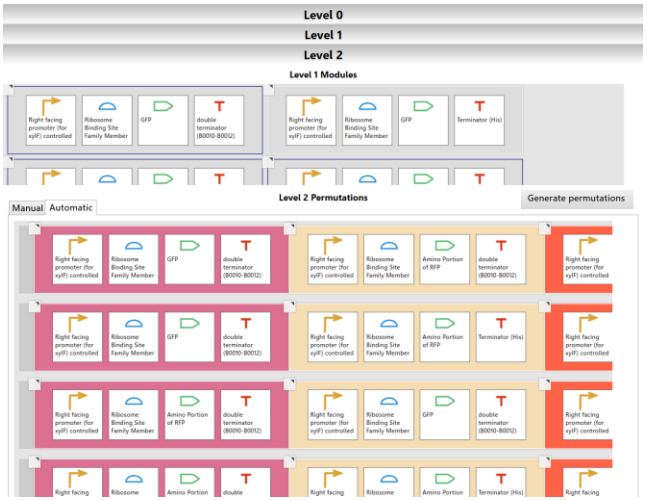

Figure 3: MoClo Planner, Level 2

While working in the Level 0 workspace, users save parts of interest by dragging them into the Level 1 shutter. These parts are automatically organized in the Level 1 workspace into "buckets" according to their function. In Level 1, users design transcription units from the basic parts selected in Level 0. The design of transcription units must satisfy structural constraints that are expressed to users through a template. The application enables users to automatically or manually create and save constraint-based permutations of transcription units. A context menu associated with each transcription unit allows users to view the sequence of this construct or access the primer designer. Transcription units of interest can be dragged to the Level 2 workspace. Figure 2 shows transcription units created in Level 1. In Level 2, users create complex constructs that consist of up to 6 transcription units. Again, users can create and save constraint-based permutations manually or automatically so that all possible constructs can be considered. Figure 3 illustrates permutations created in Level 2.

On each of the levels, users can launch a primer designer (see Figure 4), which allows users to convert constructs into concrete genetic elements by generating primers and protocols to facilitate assembly in the lab. Following, we describe the primer designer feature.

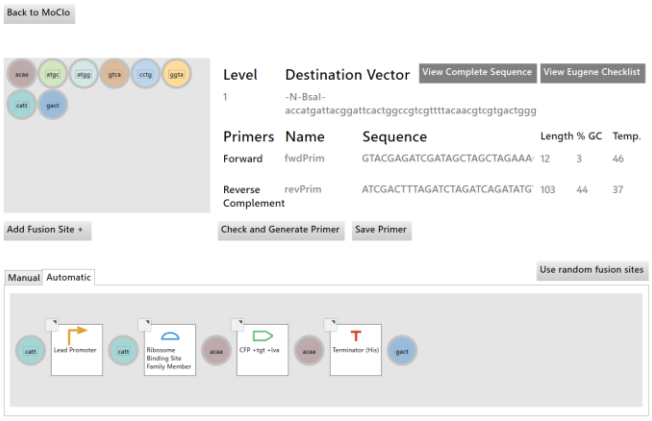

Figure 4: MoClo Planner, Primer Designer

5.2 Primer designer

The primer designer, showed in Figure 4, is accessible from a context menu associated with each basic part, transcriptional unit, and multigene construct. It consists of three views: a fusion site library to which users can add fusion sites; an overview area which displays the modules of the current design; and a primer results view, where users can test, name and save their primers. For each level, the primer designer allows the user to manually choose or automatically assign the fusion sites, the destination vector for that part, as well as view the sequence, the length, GC, and temperature of the primer. Printing the primers is a one step process that exports the entire design sequence to a .csv file for user review and for ordering physical parts.

5.3 Design considerations

The design of MoClo Planner draws on Reality-Based Interaction principles [28] and guidelines provided by the Tangible Learning Design Framework [4], a comprehensive framework that highlights design aspects important for learning and provides guidelines based on cognitive and learning theories. MoClo Planner uses reality-based metaphors such as shutters to visualize a complex hierarchical process and organize large amounts of information. It has been shown that using a conceptual metaphor that is based on image schemas to structure interaction may bootstrap learning of abstract concepts [4]. The organization of the interface using separate "shutters" allows users to use the environment to explore and store information while transitioning back and forth between different stages of the experimental process (R2). MoClo Planner draws upon users' social skills to afford collaborative interaction (R3), which in turn can promote learning and discovery [35]. For example, the system provides multiple points of entry through parallel manipulation of graphical elements (e.g., parts) across different workspaces (i.e., shutters).

While traditional GUIs have well-understood guidelines for interaction design, taxonomies for touch visualizations are still being developed. Our design was informed by the "Data & View" stage of visualization described in Heer & Shneiderman [23] and by the TouchVis gesture taxonomy proposed by Drucker et al. [15]. Filtering data based on categories is achieved through

interaction with touch buttons, selecting is indicated through tapping, navigating is performed through swiping, and zooming (i.e. requesting additional data) is done through holding and using a touch context menu. Organizing the workspace is managed through sliding shutters upwards or downwards and by flicking information artifacts out of the display, while coordinating workspaces is done through dragging across different shutters. The direct manipulation of information artifacts through multi-touch could engage the connection between the hand, the eye, and the brain to support users' conceptual understanding and facilitate "thinking through action" [3, 20, 56]. Also, it has been shown that the use of spatial and physical interaction can reduce cognitive workload [56] and trigger reflection [4, 45].

The design of MoClo Planner also addresses concrete MoClo design pitfalls. It reduces syntax errors by minimizing text entry and eliminating the need to copy and paste information from various databases. Design errors are decreased through the use of templates, redundant coding of biological parts (color, SBOL symbol, text), and computational validation. Biological errors are moderated by providing access to contextual information such as data sheets and related publications.

5.4 Implementation

The MoClo planner is implemented on the Microsoft PixelSense device using MS SDK 2.0 written in C#. Information is drawn from the MIT Registry of Biological Parts, PubMed, Google Scholar, and the iGEM archive. We use the synthetic biology domain specific description languages Eugene [10] and SBOL [47] for validating and specifying new designs (R3). The application can run on Windows machines running Windows 7 with either a keyboard or a multi-touch input.

6 Evaluation

We evaluated MoClo Planner using a tiered method [Shaer et al 2012] with 24 users. Our first study with 12 undergraduate biology students (all female, age 18-23) focused on the *usability* of the system. Our second study with 10 iGEM students (5 female, age 18-26) and 2 instructors (all female, age 27-33) focused on the *usefulness* of the system.

The usability study focused on four dimensions: functionality, learnability, performance, and errors. It was conducted in our lab. We randomly assigned users to work in dyads (overall 6 dyads). In the beginning of the session, we handed users brief reading materials about synthetic biology and the Modular Cloning process (no written or oral tutorial about the system was given). Then, we asked each dyad to use the MoClo Planner to create at least two different multi-gene constructs from a given set of biological parts. This task was selected since it mirrors a real-world research task often conducted by iGEM teams. Users documented their progress, findings, and answered task-related questions on a task form. We collected data through observations and videotaped all sessions. Following the session, we debriefed each dyad.

We found that all users were able to complete the task successfully, producing biologically correct and valid designs (typically with some difficulty that was resolved through collaborative work). On average, dyads spent 42:50 minutes working on the task (SD=13:13). Dyads' answers on the task form indicate that they understood the MoClo process as well as related concepts such as permutation and fusion sites. We observed that users understood the "shutter" metaphor and were able to easily transition back and forth between stages of the MoClo process. These findings indicate good learnability and performance as well as show that the interface facilitates a flexible and integrated workflow (R1). In terms of functionality and errors, the study highlighted several problems that led to a design iteration in which we modified the keyword search, added color coding per construct and per part category, added information to the Level 0 data sheet, and redesigned the primer design feature to provide users with more control over their primer design.

Following this design iteration, we conducted an additional study with 12 participants (10 iGEM students, 2 iGEM instructors; 7 female), which focused on *usefulness*. In particular, we examined *performance*, *engagement,* and *collaboration*. We conducted the study in users' work environment. We asked users to work in dyads to complete the experimental task: design and specify at least two new multi-gene constructs from the biological parts they had been using in their lab. This task was chosen since it mirrors the real-world research task of iGEM teams, which is typically carried out with the help of an instructor using various ad-hoc computational tools. Users documented their progress and answered task-related questions on the task form. We collected data through observations and videotaped all sessions. Following the session, users filled out questionnaires.

All five student dyads were able to complete the task successfully, designing biologically correct constructs, within an average time of 45:32 minutes (SD=10:47). In their post task questionnaire, users indicated that they gained a good understanding of the MoClo process (5.20 on a scale of 7, SD=1.03) and were confident in their designs (5 on a scale of 7, SD=1.5). Through discourse analysis, we found that during the session, all 5 student dyads used MoClo process terminology (e.g. CDS, RBS, permutation, fusion site, promoter, primer) with a mean of 7.2 MoClo terminology utterances per dyad. The instructors' dyad used MoClo terminology in 32 utterances. We also found evidence for peer teaching within dyads, which indicates autonomous and effective collaborative learning. Taken together, these findings provide evidence that *novice* users were able to use the system to successfully specify new biological designs.

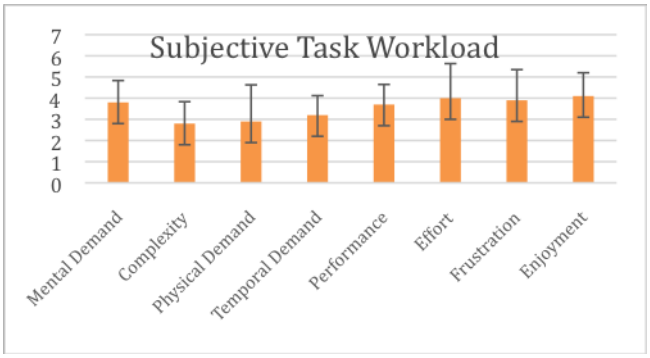

Figure 3: NASA TLX results from the MoClo Planner usefulness study along three additional dimensions of complexity, frustration, and enjoyment.

Our design aimed to reduce the mental workload associated with operating the interface (i.e., syntactic load) while also reducing the load associated directly with designing complex biological constructs (i.e., task load) (R2). We used the NASA TLX [22] post-task questionnaire to measure subjective task workload. We found that, on average, students rated their task success (i.e., performance) as moderate (3.90 SD=1.45) as well as their mental demand and effort (3.8 SD=1.03, 3.9 SD=1.45, 4 SD=1.63). These results are positive considering the high intrinsic task load. Users rated physical and temporal demand as moderately low. We also

asked users about frustration levels, which they rated as moderate (3.9, SD=1.45). We observed that one cause for frustration was the oversensitivity of the Microsoft PixelSense device. Users also reported that the complexity of the interface was moderately low (2.83, SD=1.03) and that overall they enjoyed using the system (4.1 on a scale of 7, SD=1.10). Figure 5 shows the results.

Considering the importance of collaboration in the MoClo design process, we also studied the nature of collaboration afforded by the MoClo Planner interface (R3). We observed that in the beginning of a session one user was often unsure about trying out the new technology and thus waited until the partner took physical control of the system. When the timid user became comfortable with the technology, they began interacting with the system. In two dyads this resulted in turn-taking collaboration for the rest of the session, where both users participated physically and verbally. In the other three dyads, one user assumed the role of a driver (who participates both physically and verbally) for the rest of the session while the other participated as navigator (who guides the driver through verbal and gestural cues). Both collaboration styles resulted in effective, task-focused collaboration.

Finally, in-depth interviews with the instructors post-task completion revealed high satisfaction with the collaborative learning process mediated by the system and of the integration with standards such as the Registry of Biological Parts, Eugene, and SBOL (R4). The instructors expressed their will and commitment to support future design iterations of this system with the goal of deploying it for longitudinal use by iGEM teams. The instructors highlighted two additional functional requirements to be addressed in future design iterations: integration with the Clotho data model, which handles all data generated by the wet lab, and support for the specification of target behavior using Eugene.

7 Discussion

Our evaluation with novice users showed that overall, MoClo Planner met our four design requirements: 1) Enabling an integrated and fluid workflow; 2) Facilitating a tractable process; 3) Supporting collaborative and individual work patterns; and 4) Complying with community standards. MoClo Planner supported experiential and collaborative learning of the MoClo method, which resulted in the design of biologically correct multi-gene constructs. The system facilitated a flexible workflow that complies with emerging standards in Synthetic Biology. Our findings also show clear gains resulting from the implementation of MoClo Planner as a multi-touch interface rather than a traditional GUI. This is indicated by the direct and fluid interaction afforded by the interface as well as by the effective collaboration mediated by the large multi-touch display.

Several lessons learned from the development of MoClo Planner could be applied in the development of other tools. First, the fluid user interface design, which enabled users to move back and forth between different MoClo levels, provided necessary support for the iterative and non-linear MoClo design process. When considering the design of an interactive visual system for Bio-design it is important to allow users to manage complexity through flexible and fluid interaction. Second, while automatic generation and validation of combinatorial designs can help users cope with complexity, it is important to provide users a sense of control. It became clear that users' desired not only to be the initiators of actions but also to be able to refine and control different aspects of their designs. Finally, drawing on reality-based and embodied cognition design principles led to an intuitive design that fosters collaborative learning and facilitates spatial problem solving.

Our work has several limitations that point towards future work. First, we studied one-time use in experimental settings. Additional studies of longitudinal use are necessary in order to determine to what extent the MoClo Planner system supports authentic MoClo assembly projects. Second, our evaluation focused on novice users (i.e., biology students and iGEM researchers) but more studies are needed to assess the extent to which the MoClo Planner system supports practitioners in the field. In terms of functionality, two additional requirements were highlighted by users including integration with the Clotho data model and support for the specification of target behaviour using Eugene. In the future, we will address these requirements as well as deploy MoClo Planner for longitudinal use by iGEM teams and by practitioners.

8 Conclusion and Future Work

In summary, this paper makes three contributions: 1) We provide an analysis of users, workflow, tasks, and potential errors of a particular Assembly Planning MoClo design process; 2) We present the design and implementation of MoClo Planner, an interactive visualization system for collaborative bio-design utilizing a multi-touch interactive surface; and 3) We present results from two preliminary user studies that showed that overall, MoClo Planner facilitates an integrated, collaborative, tractable, and fluid workflow while complying with community standards. We also discuss our design decisions and lessons learned including the importance of providing a fluid user interface, balancing between automation and user control, and on reality-based and embodied cognition design principles.

9 Acknowledgements

We thank Nahum Seifeselassie, Casey Grote, and Taili Feng, Natalie Kuldell, and the BU and MIT iGEM teams. This work was partially funded by NSF grant no. IIS-1149530 and by Agilent Technologies.

10 References

[1] Anderson, J. (2013). Learning Trails.

[2] Anderson, J. C., et al. (2006). Environmentally Controlled Invasion of Cancer. *Journal of Molecular Biology*, 619-327.

[3] Anderson, M. L. (2003). Embodied cognition: A field guide. *Artificial Intelligence*, 91-130.

[4] Antle, A. N., & Wise, A. (2013). Getting down to details: Using learning theory to inform tangibles research and design for children. *Interacting with Computers*.

[5] Antle, A. N., Droumeva, M., & Ha, D. (2009). Hands on what?: comparing children's mouse-based and tangible-based interaction. *International Conference on Interaction Design and Children* (pp. 80-88). New York: ACM.

[6] *ApE: A plasmid editor*. (n.d.). Retrieved from ApE: http://biologylabs.utah.edu/jorgensen/wayned/ape/

[7] Beal, J., Lu, T., & Weiss, R. (2011). Automatic Compilation from High-Level Biologically-Oriented Programming Language to Genetic Regulatory Networks. *Plos One 6(8)*.

[8] Beal, J., et al. (2012). An End-to-End Workflow for Engineering of Biological Networks from High-Level Specifications. *ACS Synthetic Biology 1 (8)*, 317-331.

[9] Bhatia, S., & Densmore, D. (2013). Pigeon: a design visualizer for synthetic biology. *ACS Synthetic Biology*.

[10] Bilitchenko, L., et al. (2011). Eugene--a domain specific language for specifying and constraining synthetic biological parts, devices, and systems. *PLoS One*.

[11] Block, F., et al. (2012). The DeepTree Exhibit: Visualizing the Tree of Life to Facilitate Informal Learning. *IEEE Transactions on Visualization and Computer Graphics.* Atlanta: IEEE.

[12] Brooks, F. P., Ouh-Young, M., & Kilpatrick, P. J. (1990). Project GROPEHaptic displays for scientific visualization. *SIGGRAPH* .

[13] Cai, Y., Wilson, M., & Peccoud, J. (2012). GenoCAD for iGEM: a grammatical approach to the design of standard-compliant constructs. *Nucelic Acids Research*, 2637-3644.

[14] Chandran, D., Bergmann, F., & Sauro, H. (2009). TinkerCell: modular CAD tool for synthetic biology. *Journal of Biological Engineering* .

[15] Drucker, S., et al. (2013). TouchVix: A Case Study Comparing Two Interfaces for Data Analytics on Tablets. *CHI.* Paris: ACM.

[16] *Geneious*. (n.d.). Retrieved from Geneious: http://www.geneious.com

[17] *Genome Compiler*. (n.d.). Retrieved from Genome Compiler: http://www.genomecompiler.com

[18] Ghamari, R., Stanton, B., Haddock, T., Bhatia, S., Clancy, K., Peterson, T., . . . and Densmore, D. (2011). Applying Hardware Description Languages to Genetic Circuit Design. *International Workshop on Bio-Design Automation.*

[19] Gillet, A., Sanner, M., Stoffler, D., Goodsell, D., & Olson, A. (2004). Augmented Reality with Tangible Auto-Fabricated Models for Molecular Biology Applications. *Visualization.* Austin: IEEE.

[20] Glenberg, A. M. (2008). Embodiment for education. *Handbook of Cognitive Science: An Embodied Approach*, 355-372.

[21] Harris, A.,et al. (2009). Around the table: Are multiple-touch surfaces better than single-touch for children's collaborative interactions? *ACM CSCL* (pp. 335-344). ACM Press.

[22] Hart, S., & Staveland, L. (1988). Development of NASA-TLX (Task Load Index) : Results of Empirical and Theoretical Research. *Human Mental Workload* .

[23] Heer, J., & Schneiderman, B. (2012). Interactive Dynamics for Visual Analysis. *Queue*, 30-56.

[24] Horn, M., Tobiasz, M., & Shen, C. (2009). Visualizing Biodiversity with Voronoi Treemaps. *International Symposium on Voronoi Diagrams* (pp. 265-270). Copenhagen: IEEE.

[25] Hornecker, E., Marshall, P., Dalton, N., & Rogers, Y. (2008). Collaboration and Interference: Awareness with Mice or Touch Input. *CSCW* (pp. 167-176). New York: ACM.

[26] IDTDNA. (n.d.). *PrimerQuest*. Retrieved from Integrated DNA Technologies.

[27] *International Genetically Engineered Machine Competition*. (2013). Retrieved from International Genetically Engineered Machine : http://igem.org/Main_Page

[28] Jacob, R. J.,et al. (2008). Reality-Based Interaction: A Framework for Post-WIMP Interfaces. *ACM CHI.* Florence: ACM.

[29] Kuldell, N. (2007). Authentic Teaching through Synthetic Biology. *Journal of Biological Engineering*.

[30] Marshall, P. (2007). Do tangible interfaces enhance learning? *TEI* (pp. 163-170). LA: ACM Press.

[31] Microsoft. (n.d.). *Microsoft PixelSense*. Retrieved from Microsoft PixelSense: http://www.microsoft.com/en-us/pixelsense/default.aspx

[32] Mitchell, R., & Yehudit, J. D. (2011). Experiential Engineering Through iGEM -- An Undergraduate Summer Competition in Synthetic Biology. *Journal of Science Education and Technology*, 156.

[33] Morris, M., Brush, A., & Meyers, B. (2007). Reading revisited: evaluating the usability of digital display Surfaces for Active Reading Tasks. *IEEE Tabletop* (pp. 79-86). IEEE.

[34] *Myers Research Group: iBioSim*. (n.d.). Retrieved from iBioSim: http://www.async.ece.utah.edu/iBioSim/

[35] Okada, T., & Simon, H. (1997). Collaborative discovery in a scientific domain. *Cognitive SCience*, 109-146.

[36] Patten, J., & Ishii, H. (2000). A comparison of spatial organization strategies in graphical and tangible user interfaces. *Designing Augmented Reality Environments.*

[37] Piper, A., & Hollan, J. (2009). Tabletop displays for small group study: affordances of paper and digital materials. *Human Factors in Computing.* ACM.

[38] *Registry of Biological Parts*. (n.d.). Retrieved from Registry of Biological Parts: http://partsregistry.org/Main_Page

[39] Ro, D.-K., Paradise, E. M., Ouellet, M., Fisher, K., j Newman, K. L., Ndungu, J. M., . . . Keasling, J. D. (2006). Production of the antimalarial drug precursor artemisinic . *Nature*, 440.

[40] Ryall, K., Morris, M. R., Everitt, K., Forlines, C., & Shen, C. (2006). Experiences with and observations of direct-touch tabletops. *Tabletop.* Cambridge, MA: IEEE.

[41] Salis, H. (2011). The Ribosome Binding Site Calculator. (V. C.A., Ed.) *Methods in Enzymology (498)*, 19-42.

[42] Schkolne, S., Ishii, H., & Schroder, P. (2004). Immersive Design of DNA Molecules with a Tangible Interface. *Visualization* (pp. 227-234). Washington: IEEE.

[43] Schneider, B., Strait, M., Muller, L., Elfenbein, S., Shaer, O., & Shen, C. (2012). Phylo-Genie: Engaging Students in Collaborative 'TreeThinking' through Tabletop Techniques. *Human Factors in Computing Systems* . Austin: ACM.

[44] Shaer, O., Kol, G., Strait, M., Fan, C., Catherine, G., & Elfenbein, S. (2010). G-nome Surfer: a Tabletop Interface for Collaborative. *Conference on Human Factors in Computing Systems.* 2010: ACM.

[45] Shaer, O., Strait, M., Valdes, C., Feng, T., Lintz, M., & Wang, H. (2011). Enhancing Genomic Learning through Tabletop Interaction. *Human Factors in Computing Systems* . Vancouver: ACM.

[46] Shaer, O., Strait, M., Valdes, C., Wang, H., Feng, T., Lintz, M., . . . Liu, S. (2012). The design, development, and deployment of a tabletop interface for collaborative exploration of genomic data. *International Journal of Human-Computer Studies*, 746-764.

[47] *Synthetic Biology Open Language*. (2011). Retrieved from http://www.sbolstandard.org/

[48] Tabard, A., Hincapie-Ramos, J. D., & Bardram, J. E. (2012). The eLabBench in the Wild – Supporting Exploration in a Molecular Biology Lab. *Human Factors in Computing Systems.* Austin: ACM.

[49] Tabard, A., Hincapie-Ramos, J.-D., Ebensen, M., & Bardam, J. E. (2011). The eLabBench: an interactive tabletop system for the biology laboratory. *Interactive Tabletops and Surfaces* (pp. 202-211). New York: ACM.

[50] Tabard, A., Mackay, W., & Eastmond, E. (2008). From individual to collaborative: the evolution of prism, a hybrid laboratory notebook. *Computer supported cooperative work* (pp. 569-578). New York: ACM.

[51] *The UNAFold Web Server*. (n.d.). Retrieved from The UNAFold Web Server: http://mfold.rna.albany.edu/

[52] Valdes, C., Ferreirae, M., Feng, T., Wang, H., Tempel, K., Liu, S., & Shaer, O. (2012). A Collaborative Environment for Engaging Novices in Scientific Inquiry. *Interactice Tabletops and Surfaces.* Cambridge: ACM.

[53] *Vector Express*. (n.d.). Retrieved from Vector Express: http://www.vectorexpress.com

[54] Weber, E., Engler, C., Gruetzner, R., Werner, S., & Marillonnet, S. (2011). A Modular Cloning System for Standardized Assembly of Multigene Constructs. *PLoS ONE*.

[55] Wigdor, D., Jiang, H., Forlines, C., Borking, M., & Shen, C. (2009). The WeSpace: The Design, Development, and Deployment of a Walk-Up and Share Multi-Surface Visual Collaboration System. *Human Factors in Computing Systems.* Boston: ACM.

[56] Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review*, 625-636.

[57] Xia B, B. S. (2011). Clotho: A Software Platform for the Creation of Synthetic Biological Systems, A Developer's and User's Guide for Clotho v2.0. *Methods in Enzymology*.