ELSEVIER

The design, development, and deployment of a tabletop interface for collaborative exploration of genomic data

O. Shaer[a],*, M. Strait[b], C. Valdes[a], H. Wang[a], T. Feng[a], M. Lintz[c], M. Ferreirae[a], C. Grote[a], K. Tempel[a], S. Liu[a]

[a]*Wellesley College, 106 Central St. Wellesley, MA 02481, USA*
[b]*Tufts University, Medford, MA, USA*
[c]*Olin College of Engineering, Needham, MA, USA*

Abstract

In this paper, we reflect on the design, development, and deployment of G-nome Surfer; a multi-touch tabletop user interface for collaborative exploration of genomic data. G-nome Surfer lowers the threshold for using advanced bioinformatics tools, reduces the mental workload associated with manipulating genomic information, and fosters effective collaboration. We describe our two-year-long effort from design strategy to iterations of design, development, and evaluation. This paper presents four main contributions: (1) a set of design requirements for supporting collaborative exploration in data-intensive domains, (2) the design, implementation, and validation of a multi-touch tabletop interface for collaborative exploration, (3) a methodology for evaluating the strengths and limitations of tabletop interaction for collaborative exploration, and (4) empirical evidence for the feasibility and value of integrating tabletop interaction in college-level education.
© 2012 Elsevier Ltd. All rights reserved.

Keywords: Tabletop interaction; Multi-touch interaction; Bioinformatics; Reality-based interaction; Genome browser; Collaborative learning

1. Introduction

Over the past two decades, Human–Computer Interaction (HCI) research has generated a broad range of interaction styles that move beyond the desktop to new physical and social contexts. Key areas of innovation have been tabletop, tangible, and multi-touch user interfaces. Drawing upon users' pre-existing knowledge and skills of interaction with the real non-digital world such as naïve physics, spatial, social, and motor skills, these interaction styles are often referred to as Reality-Based Interfaces (RBIs) (Jacob et al., 2008). By basing interaction on pre-existing real world knowledge and skills, RBIs offer a more natural, intuitive, and accessible form of interaction that reduces the mental effort required to learn and operate a computational system (Jacob et al., 2008).

The introduction and increasing availability of large, multi-touch, high-resolution displays open the opportunity to design *tabletop* reality-based interfaces that mediate co-located collaboration by allowing multiple users to simultaneously touch and manipulate data representations. Several studies indicate that horizontal interfaces support active reading (Morris et al., 2007) and facilitate external cognition (Patten and Ishii, 2000). However, while these advances in HCI have been applied to a broad range of application domains, little research has been devoted to investigating RBI in the context of scientific exploration. It is particularly important to study RBI in this context, where reducing users' mental workload, supporting collaborative work, and facilitating high-level reasoning could lead to new scientific discoveries. Several RBIs have examined the possibilities of supporting scientific discovery in fields such as molecular biology and chemistry, focusing on the representation of information with inherent physical or spatial structure (e.g., proteins and molecules). We are interested in investigating the application of reality-based interaction to areas where

---

*Corresponding author. Fax: +1 7812833642.
E-mail address: oshaer@wellesley.edu (O. Shaer).

a vast amount of *abstract* information is accessed and manipulated. Examples include phenology, synthetic biology, and genomics.

Advances in genomic technologies have led to an explosive growth in the quantity and quality of biological information. The need to analyze large and complex data sets has in turn driven a change in the tools used in biological research: next to a pipette and a pen, a web browser is currently the most widespread tool available for biologists as it provides access to powerful computational and statistical tools (Skinner et al., 2009). However, existing web-based genomic tools have a high threshold (Shaer et al., 2010) and show severe limitations in terms of persistence (Veretnik et al., 2008), usability (Bolchini et al., 2009), and support of high-level reasoning (Mirel, 2005). Motivated both by biologists' need to access, manipulate, and make sense of the vast amounts of genomic data available and by the potential of tabletop interfaces to enhance collaborative research activities, we developed G-nome Surfer, a tabletop user interface for collaborative exploration of genomic data.

In this paper, we reflect on the design, development, and deployment of G-nome Surfer. We describe our two-year-long effort from design strategy to iterations of design, development, and evaluation. While we review previous versions of G-nome Surfer, limited to the visualization and analysis of *eukaryotic* genomes that were presented at (Shaer et al., 2010, 2011), we also introduce G-nome Surfer Pro, a sister application of G-nome Surfer that provides a suite of *prokaryotic* visualizations and analytical tools. In addition, we describe new findings from evaluating G-nome Surfer in authentic educational laboratory settings. This paper presents four main contributions: (1) a set of design requirements for supporting collaborative exploration in data-intensive domains, (2) the design, implementation, and validation of a multi-touch tabletop interface for collaborative exploration, (3) a methodology for evaluating the strengths and limitations of tabletop interaction for collaborative learning, and (4) empirical evidence for the feasibility and value of integrating tabletop interaction in college-level education.

This paper is organized as follows. We begin with a discussion of related work. We then report our two-year long effort, starting with our design strategy and then describing iterations of design and development. Next we discuss our evaluation framework and results. We end with conclusions and future work.

1.1. Related work

This work draws from prior work in two areas: reality-based interfaces for scientists and interactive surfaces for collaborative learning.

1.1.1. Reality-based interfaces for scientists

A number of systems illustrate the vast possibilities for supporting scientific discovery and higher education through reality-based interaction: Brooks et al. (1990) developed the first haptic display for scientific visualization, used by chemists to investigate docking positions for drugs. Gillet et al. (2004) presented a tangible user interface for molecular biology that used augmented reality technology to view 3D molecular models. Schkolne et al. (2004) 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.

Several projects investigate augmented capture and situated access to biological data. Labscape (Arnstein et al., 2002) is a smart environment for the cell biology laboratory that allows biologists to easily record, relate, and share heterogeneous information. ButterflyNet (Yeh et al., 2006), is a mobile capture and access system for field biologists that integrates paper notes with digital photographs. Similarly, Tabard et al. (2008) explores the integration of biologists' paper notebook with physical and digital sources of information. Mackay et al. (2002) developed a series of augmented laboratory paper notebooks with digital search and links to digital information; Schraefel et al. (2004) developed a tablet-based system for the chemistry lab that replaces paper and supports the execution of experiments. While these systems demonstrate the feasibility of using computation while carrying out experiments and collecting data, we propose a system to support hypothesis formation, reflection, and analysis.

To date, a few systems were developed to facilitate collaboration among scientists across large displays and multi-touch tables. WeSpace (Wigdor et al., 2009) integrates a large data wall with a multi-touch table and personal laptops. It provides a set of services that facilitate spontaneous research meetings. TeamTag (Ryall et al., 2006) allows biodiversity researchers to collaboratively search, label, and browse digital photos. While these three systems informed our design considerations, they target different domains and settings. Finally, Biotisch (Echtler et al., 2010) and CheMo (Song et al., 2011) explore the feasibility of integrating tabletop interfaces as interactive wet lab benches. However, they do not explore the practical implications of deployment in the wet lab environment. The eLabBench (Tabard et al., 2011) is a tabletop system supporting experimental research in the biology laboratory. While the system allows users to interact with heterogeneous information in the lab, it was designed to enhance the work of a single user.

1.1.2. Interactive surfaces for collaborative learning

Several studies have investigated the effects of different interactive surface parameters on collaborative work, including the orientation of the display (Rogers and Lindley, 2004), table size (Ryall et al., 2004) and input techniques (Antle et al., 2009; Hornecker et al., 2008). Marshall et al. (2011) conducted an in-the-wild study of

collaborative interaction with a walk-up-and-use tabletop interface in public settings. However, relatively little work has explicitly examined the effects of interactive surface parameters on collaborative learning. Of the studies directly investigating collaborative learning, most have focused on children (Antle et al., 2009; Fleck et al., 2009; Harris et al., 2009; Rick et al., 2009, 2011) or on mixed groups in informal education settings (e.g., museums) (Antle et al., 2011; Horn et al., 2009). Piper and Hollan (2009) conducted a study with pairs of undergraduate students, comparing the affordance of tabletop displays and traditional paper handouts. However, their study utilized a preliminary tabletop prototype with minimal functionality. We present the evaluation of G-nome Surfer 2.0, a feature-rich interface that supports complex analytical tasks. Schneider et al. (2012) developed Phylo-Genie, a tabletop user interface for collaborative learning of college-level phylogeny. They conducted a comparative study of tabletop interaction and paper-based interaction showing that tabletop implementation fosters collaborative learning by engaging users in the activity. However, to date, this system was not evaluated in authentic classroom settings.

Our evaluation methodology draws upon this body of work but studies tabletop interfaces for complex interaction in the context of college-level collaborative learning.

2. Design

Our goal was to develop a user interface that enhances collaborative exploration of genomic information. Due to the complexity of the genomic research domain, our design strategy combined *rigorous* user-centered and participatory design methods as described below:

1) *Employing users as permanent members of the design team*: Biology students were an integral part of the research team, participating in all stages of the development and evaluation process. In addition, an industry bioinformatics expert served as a consultant for the team.
2) *Establishing partnerships with domain scientists*: We picked two biology research groups in Wellesley College and Boston University as partners in participatory design. We particularly chose research groups that consist of a small but highly collaborative team, are interested in developing computational tools to enhance their research, and are committed to the training of future scientists. We also established a partnership with a team of neuroscience instructors at Wellesley College that teach college-level genomics. We met with our partners on a regular basis throughout development.
3) *Acquiring necessary domain knowledge*: Through extensive training led by our partners, our research team acquired the domain knowledge necessary for understanding users' goals, current practices, and needs.
4) *Conducting interview and focus groups*: To understand current work practices and needs of biologists working with genomic data, we conducted in depth interviews with 38 molecular and computational biologists from Harvard, MIT, Boston University, Wellesley College, and industry. Participants had varied skill levels ranging from world-renowned experts to undergraduate student researchers. In Shaer et al. (2010, 2011) we described findings from a subset of these studies (22 participants) where we investigated the use of bioinformatics tools for research and for college-level teaching of genomics. Recent interviews focused on understanding the workflow of small research groups and identifying collaboration patterns within teams.
5) *Applying ethnographic methods*: To further understand the context within biologists' work, we conducted an 8-week-long ethnographic study where we observed a small team of 9 researchers studying gene interaction in Tuberculosis. We attended formal and informal research meetings where team members reported progress, planned and coordinated activities, received guidance, and discussed research goals and hypotheses. We also conducted observations in research labs at MIT and Wellesley College.
6) *Testing frequently with users*: Because of the scope and complexity of the interface, we performed usability studies on a continual basis throughout the development process. Rather than testing complete versions, we conducted micro-studies, testing the usability of particular features through a series of prototypes of increasing fidelity. We also conducted formal studies that examined the usability of the interface when used to complete a high-level task. Usability studies were followed by experimental and in-situ evaluations.

Based on the information we collected through interviews and ethnographic studies, we identified a set of requirements for supporting collaborative exploration in genomics. Following, we describe these requirements.

2.1. Design requirements

We identified five requirements for supporting collaborative discovery in genomic research:

R1) Alleviating data explosion
R2) Lowering the threshold for using advanced bioinformatics tools
R3) Facilitating an integrated and flexible workflow
R4) Supporting multiple forms of evidence
R5) Fostering collaboration and reflection

Here we discuss each of these requirements:
R1: Alleviating Data Explosion
The recent introduction of next-generation sequencing technology, capable of producing millions of DNA sequences reads in a single run, is rapidly changing the landscape of genomics. In the near future, such sequencing

instruments will become readily available, allowing a single lab to create in one year the same amount of data that was held in all the NIH sequence databases just 5 years ago. In the context of HCI, these advances present the need to provide researchers with means for searching, annotating, comparing, organizing, and sharing this data while reducing the mental workload associated with handling vast amounts of data.

R2: Lowering the threshold for using advanced bioinformatics tools

Genomic research involves a large and diverse community ranging from undergraduate students to world-acclaimed experts. Thus, users of bioinformatics tools differ in their level of expertise both in terms of domain knowledge and of computational experience. We observed that bioinformatics tools in general and genome browsers in particular have a high threshold: they are powerful, but in order for a researcher to take advantage of their power, they must have both broad domain knowledge and extensive training. To support a wider range of users with varying experience levels and skills, there is a need to lower the threshold for using advanced bioinformatics tools.

R3: Facilitating an integrated and flexible workflow

To gain insight into complex biological systems, genomic researchers often link together several data sets, each one handled with a special bioinformatics tool. We observed that biologists often manually fetch data from one bioinformatics tool, reformat the data, apply the next bioinformatics tool, parse the results, reformat the results, and so on. Since the information workflow in genomic research is rarely linear, biologists often repeat these steps. Typically not comfortable with programming, biologists rarely automate a workflow. As genomic sets grow larger, this method of operation becomes more time consuming and error-prone. Thus, there is a clear need for providing means to easily integrate both data and tools in a non-linear and flexible manner.

R4: Supporting Multiple Forms of Evidence

Biologists combine multiple forms of evidence to discover connections and casual relationships, as well as to examine information in different levels of granularity. We found that existing tools often overwhelm users with the amount of data presented on the screen, making it difficult for the user to organize the information in ways that highlight the connections between multiple forms of evidence. This observation is also supported in the literature (Bolchini et al., 2009; Mirel, 2005). We thereby identified a clear need to support organization, manipulation, and comparison of multiple forms of evidence.

R5: Fostering collaboration and reflection

Genomic research is often multidisciplinary and highly collaborative. We observed that collaboration is usually based on emails and shared databases, while research meetings serve as an opportunity to coordinate activities, provide guidance, and discuss hypotheses and research directions. In small labs that consist of faculty and student researchers, several researchers often work together on the

same computer, exploring, analyzing, and discussing biological data. However, current bioinformatics tools utilize traditional GUIs that limit equitable input, constrain gestural abilities, and limit the opportunities for effective face-to-face communication and reflection (Hornecker et al., 2008). This suggests that less constraining human–computer interaction styles, such as tabletop interaction, may be more effective for collaborative exploration. Reflection can be further promoted by providing users with a means for spatially manipulating and easily annotating information artifacts.

2.2. G-nome surfer system: A multi-touch table interface for exploring genomics

Informed by these requirements, we designed and implemented G-nome Surfer, a tabletop interface for collaborative exploration of genomics. G-nome Surfer is intended to be used during early stages of genomic investigations for data exploration and hypothesis formation, 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 G-nome Surfer for small research labs that focus on the training of future scientists, where experienced researchers often work closely with novice researchers to provide guidance.

Our choice to design and implement G-nome Surfer as a multi-touch tabletop interface was informed by existing research, which indicates that tabletop interfaces support collaboration through visibility of actions and egalitarian input (Hornecker et al., 2008; Marshall et al., 2008), facilitate active reading (Morris et al., 2007), and afford distributed cognition, potentially lowering mental workload (Patten and Ishii, 2000). It has also been shown that multi-touch tabletops promote reflection and collaboration in learning environments (Harris et al., 2009; Piper and Hollan, 2009). Considering these qualities of tabletop interaction, as well as the increasing availability and falling prices of commercial hardware platforms, we decided to utilize tabletop interaction to address the requirements discussed above and provide an interface for collaborative exploration of genomic information.

2.3. Summary

When designing interactions for complex domains such as genomics, the lack of formal domain expertise can be a significant hurdle. It has been found that partnership with domain experts is crucial for developing usable and effective interactions in complex domains (Chilana et al., 2010). Including domain experts and users as an integral part of our team was instrumental for formulating requirements, informing our design process, and guiding our evaluation effort.

Through extensive user study, we identified five requirements for supporting collaborative discovery in genomic research: (1) alleviating data explosion, (2) lowering the

threshold for using advanced tools, (3) facilitating an integrated and flexible workflow, (4) supporting multiple forms of evidence, and (5) fostering collaboration and reflection. While there may be multiple ways to address each of these requirements, our challenge was to find a way to address them all together in a coherent design.

Our choice to use tabletop interaction to address these requirements was informed by recent research, which indicates that tabletop interfaces promote collaboration, foster reflection and facilitate active reading. Considering these qualities, we hypothesized that through careful design, a tabletop interface could meet user requirements and facilitate effective collaborative exploration of genomic information.

Following, we describe the development of three G-nome applications: G-nome Surfer 1.0 and G-nome Surfer 2.0 that provide tools for studying *eukaryotic* genomes (Shaer et al., 2010, 2011) and G-nome Surfer Pro that was designed for studying *prokaryotic* genomes and is introduced in this paper.

3. Development

In this section, we describe the development of the three G-nome Surfer applications. For each version, we describe the design rationale that we followed in order to meet the requirements outlined in ''Design requirements'' section

3.1. G-nome Surfer 1.0

The first prototype of G-nome Surfer was designed to support five fundamental information tasks: navigating genomic maps, retrieving genomic sequences, searching for similarity across sequences, searching the literature, and annotating genomic information. To facilitate an integrated experience (R3), G-nome Surfer draws genomic information from several public databases and integrates multiple bioinformatics tools. Following, we describe G-nome Surfer's primary functions and interaction techniques.

G-nome Surfer 1.0 allows users to access a genomic map by specifying a particular chromosome, base-pair range, or gene. A sliding track displays genes on the desired section of the chromosome, represented as arrows. Users are able to pan the chromosome left and right using a flick gesture. Continuous visual feedback helps users to maintain a sense of location (R1). Tapping on a gene displays the structure of a gene in terms of exons and introns in a separate track. A polygon connects the gene and its structure to support the user's sense of location (R1, R2). Selecting a gene or a particular exon displays a menu that allows users to retrieve a genomic sequence (DNA, RNA, or amino acid) or access ontology, selected publications, and gene expression (R4). Each information artifact or genomic sequence is displayed in a new window that references back to the source gene. Users can move resize, orient, and arrange the windows as well as annotate genomic sequences (R1, R2, R5). Fig. 1 shows a screen capture from G-nome Surfer 1.0 that displays a genomic map and related information artifacts.
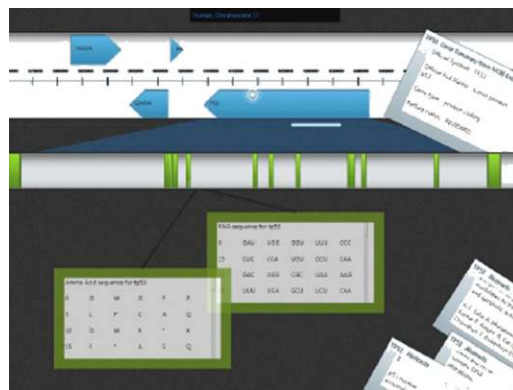

Fig. 1. G-nome Surfer 1.0 displaying the human gene TP53 and related publications.
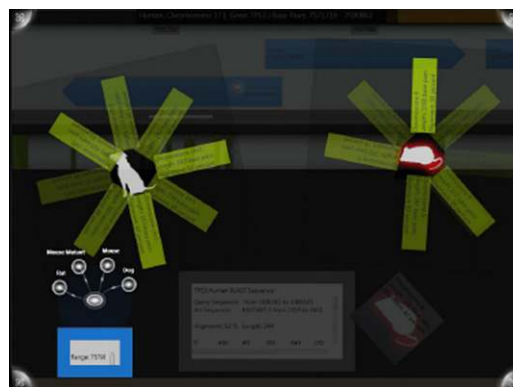

Fig. 2. G-nome Surfer 1.0 BLAST visualization.

G-nome Surfer 1.0 also enables users to search for regions of local similarity between a displayed sequence and other genomes. The user begins the process by placing the tangible BLAST object over a genomic sequence, which displays the BLAST layer—a semi transparent layer that covers the surface and presents the BLAST visualization. We chose to represent the BLAST tool using a tangible object to make this application state change immediate, visible, and easily reversible (R1, R2, R5). We designed a novel visualization for BLAST results that draws upon naïve physics metaphors, using transparency and mass to encode information. Each result is displayed as a rectangle and the degree of similarity is encoded as brightness. The results are organized in a flower-like structure around a target organism. Target organisms with more results are displayed closer to the bottom of the surface. Fig. 2 shows the BLAST visualization.

Finally, like tabletop interfaces in general (Hornecker et al., 2008), G-nome Surfer draws upon users' social skills and existing social protocols to afford collaborative interaction: the system provides multiple points of entry and makes application state visible to all users through the use of visual and physical objects like the BLAST tool (R5).

3.2. G-nome Surfer 2.0

While G-nome Surfer 1.0 was developed to support research activities, G-nome Surfer 2.0 was designed to enhance collaborative inquiry-based learning of college-level genomics (Shaer et al., 2011). Thus, in G-nome Surfer 2.0, we introduced new and explicit support for collaborative learning. Our goal was to help students in linking visual elements to relevant content knowledge, developing both content and process knowledge, bridging across multiple representations and abstraction levels, and finally discussing and reflecting on their findings. Following, we describe G-nome Surfer 2.0's *new* interaction techniques.

G-nome Surfer 2.0 introduced a contextual help tool that provides references to specific visual structures within G-nome Surfer. Placing a tangible flashlight tool on any visual representation of data displays a glossary definition that ties that representation to relevant content knowledge. The definition is supplemented with links to related terms and a scrollable alphabet index (R2). We chose to represent the help tool with a tangible object in order to increase visibility and encourage users to discuss researched terms (R5).

The pair-wise alignment feature allows users to identify similarities and differences between different genomic sequences and between different representations of the same sequence (e.g., DNA, RNA). It also enables users to add free-form annotations (R4). When a user drags a sequence view on top of another sequence view, the two views snap together and create a new visual object displaying the two sequences aligned. This interaction technique allows users to easily and seamlessly align genomic sequences, a task that in current bioinformatics tools involves two different tools and sorting through a long list of results (R2, R3). Fig. 3 shows the pair-wise alignment and contextual help features.

Based on user study results, we redesigned the gene expression tool to include multiple zoom levels and utilize a monochromatic color scheme (R1). The redesigned tools visualize expression levels by presenting a diagram of the organism with different regions highlighted; expression levels are encoded as brightness. Tapping a tissue area

Fig. 4. Exploring mouse gene expression using G-nome Surfer 2.0.

displays a list of sub-regions along with their respective expression levels. Fig. 4 shows users exploring gene expression in mice.

The workflow visualization records and displays all interactions performed upon G-nome Surfer. It was designed to help users to construct process-knowledge in investigations where workflow is rarely linear and involves multiple intermediate data formats (R1, R3). Actions are organized by gene and are displayed as color-coded blocks per gene entry. The color-coding is consistent with the scheme used for differentiating information artifacts so that the computations applied in each stage can be easily distinguished. This visualization facilitates the construction of process knowledge by allowing users to visualize, record, and repeat a workflow.

3.3. G-nome Surfer Pro

While G-nome Surfer 1.0 and 2.0 were developed to support the investigation of *eukaryotic* genomes, G-nome Surfer Pro was designed for *prokaryotic* genomes.

Prokaryotic genomes differ from eukaryotic genomes. They are smaller ($< \sim 5$ Mb) and typically contain a single large circular piece of chromosomal DNA. Many prokaryotic cells contain additional pieces of DNA held on small circular structures called plasmids. Prokaryotic DNA is typically gene rich and, unlike eukaryotic genes, there are no introns in prokaryotic genes. Considering these differences between eukaryota and prokaryota, the development of G-nome Surfer Pro required the design of new visualizations and interaction techniques. G-nome Surfer Pro also introduces support for primer design and primer blast, tasks fundamental to micro- and synthetic biology research. Following, we describe G-nome Surfer Pro.

To search for a prokaryotic genome in G-nome Surfer Pro, users type in a particular organism or navigate a phylogenetic tree, which expands upon request. We chose to utilize an expanding tree visualization to help users navigate the wide range of organisms used in a prokaryotic study (R1). Upon locating a genome to display, users may search for a specific gene.

Fig. 3. G-nome Surfer 2.0 sequence alignment and reference tool.

G-nome Surfer Pro's genomic map consists of a wheel, a circular genome visualization displaying an overview of the genes located on both DNA strands, along with a magnified view of a slice of the chromosome. Users are able to pan the chromosome left and right either by rotating the chromosome wheel (for coarse navigation) or by using a flick gesture on top of the magnified slice (for fine navigation). A visual indicator links the wheel and the slice, helping users to maintain a sense of location (R1). We decided to use a circular genome visualization not only because it mimics the circular structure of prokaryotic chromosomes, but also because of its compact nature that enables users to visualize the entire chromosome within the given space (R1).

When a gene is tapped on top of the chromosomal slice, a menu is displayed, allowing users to retrieve a genomic sequence (DNA or amino acid), access a GenBank note, search for publications, or access the primer designer (R3, R4). Each information artifact or genomic sequence is displayed in a new window that references back to the source gene. Users can move resize, orient, and arrange the windows as well as annotate genomic sequences (R1, R2, R5). Similarly to G-nome Surfer 2.0, users can easily conduct pair-wise alignment of genomic sequences by overlaying two sequence boxes (R2). Fig. 5 shows a screen capture from G-nome Pro that displays a genomic map, an aligned sequence, and related information artifacts.

To allow users to save information artifacts for later review and reflection, G-nome Surfer Pro implements an extended desktop, which behaves like a desk drawer, allowing users to deposit information artifacts. To store information artifacts in the extended desktop, users flick an artifact to the bottom of the surface. Users can then open the "drawer" by "pulling" it, a gesture that covers the surface with a semi-transparent layer. Users can use the space provided by the extended desktop for spatially organizing their information artifacts (R5). Providing a secondary space to hold information also allows users to
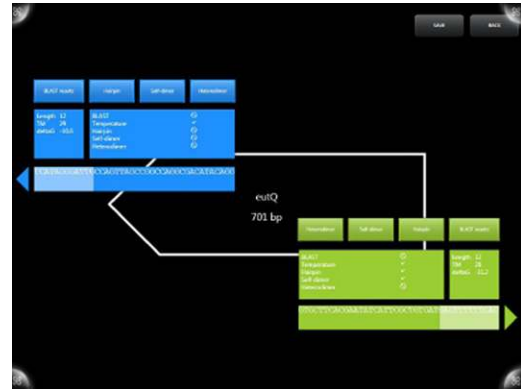

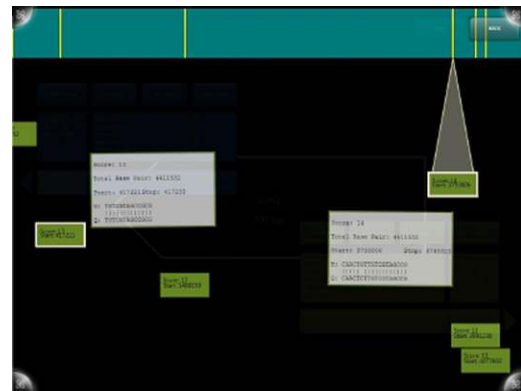Fig. 5. G-nome Surfer Pro displaying the chromosome visualizations, an aligned sequence, and publications.

Fig. 6. G-nome Surfer Pro primer design tool.

Fig. 7. G-nome Surfer Pro BLAST results visualization.

cope with the high information density of genomics by increasing the number of information artifacts a user can access without cluttering the workspace (R1).

After selecting a gene, users can also access the primer design tool in order to design primers, short sequences of DNA marking the beginning and end of a particular region of DNA sequence. Two primers are physically necessary to amplify a particular gene so it can be studied in the lab. To design a primer, users set both the forward and reverse primer sequences (generally the first 18 base pairs and reverse complement of the last 18 base pairs) and then conduct several tests to predict how well a primer will work. Users can modify their primer's sequences and length as needed to satisfy the physical and chemical constraints of gene amplification. We designed G-nome Surfer Pro to allow designers to easily test and manipulate their primer, giving the user more control over their primer design than existing tools (R2, R4). Fig. 6 shows G-nome Surfer Pro's primer design tool.

Once users are satisfied with a primer, they need to use the BLAST tool to search whether the primer's sequence (or parts of it) is present in a different location on the organism's genome. For that purpose, we created a new BLAST visualization, which provides an overview of search results across the entire genome. The visualization enables users to access additional information upon request (R1). Fig. 7 shows G-nome Surfer Pro's primer BLAST

visualization. The visualization allows a user to easily determine if a given search result will affect their primer design (R2, R4).

3.4. Implementation

All three versions of G-nome Surfer are implemented on the commercially available Microsoft Surface using C, WPF, and the Microsoft Surface SDK.

G-nome Surfer 1.0 utilized web services to draw genomic information from various databases, including the UCSC Genome Browser, PubMed and Entrez Gene. The BLAST search was implemented using the Washington University BLAST (WU-BLAST) web service. In G-nome Surfer 2.0, we implemented several architectural changes in order to improve performance, including the integration of a local database of RefSeq genes stored in FASTA files, the implementation of real-time sequence comparison and analysis utilizing the NCBI BLAST+ tool suite, and the integration of data sets from JAX Mice and Allen Brain Atlas. The implementation of G-nome Surfer Pro draws upon that of G-nome Surfer 2.0, but implements additional architectural changes to further improve performance and to support new interaction techniques: we maintain a local database of GenBank files as a source for meta-data of individual genes, use the CGView (Stothard and Wishart, 2005) software for pre-rendering the circular genome visualizations that are stored locally, and use the NCBI BLAST+ and OligoAnalyzer tools (2011) for the implementation of the primer design tool.

3.5. Summary

In this section, we described the development of three distinct versions of G-nome Surfer: G-nome Surface 1.0, which was designed for supporting collaborative research of eukaryotic genomes; G-nome Surfer 2.0, which was designed for fostering collaborative inquiry-based learning of college-level genomics; and G-nome Surfer Pro, which was designed for supporting the research of prokaryotic genomes. Developing these prototypes to meet user requirements, we faced three major challenges: (1) the inherent complexity of the application domain, (2) the lack of standard interaction paradigm for multi-touch tabletops, and (3) the need to trade design guidelines and metaphors against other considerations such as practicality, performance, and physical constraints.

Given the complexity inherent to the domain of genomics and the explosion of data in this field, we applied several design strategies for eliminating complexity including reducing functionality and hiding complexity (Janlert and Stolterman, 2010). In addition, we used reality-based metaphors that draw upon interaction with the non-digital world, leveraging naïve physics concepts such as gravity, transparency, acceleration, and mass in our design as means for reducing complexity (Jacob et al., 2008).

To address the lack of a standard interaction paradigm for multi-touch tabletops, we adopted general design guidelines such as visibility, layout, and feedback but considered them in the context of multi-touch tabletop interaction. We also consulted the guidelines proposed by Wigdor and Wixon (2011) for designing natural user interfaces. For example, in addition to the increased visibility of actions that is inherent to tabletop interaction, we used tangible tokens to make state-changes visible to both those sitting around the tabletop and to those observing the interaction from a distance (e.g., an instructor in a classroom settings). Using tangible objects for invoking state-change also make those changes immediate and easily reversed. In considering layout, we designed to reduce clutter, highlight the connections among information artifacts, and provide space for users to spatially manipulate information artifacts. We also considered the size of user interface elements, optimizing it for touch and ensuring we provide immediate and meaningful feedback for every touch.

Finally, considering the large datasets used by the G-nome Surfer application, we had to make several tradeoffs where reality-based metaphors and design guidelines were traded against other considerations such as practicality, performance and physical constraints. For example, we chose to store several data sets that are relatively stable (e.g., genomic sequences) locally rather than retrieve them in real-time in order to improve the performance of the application. This requires us to update the data manually every few months. Other data sets (e.g., publications) must be retrieved in real-time as their content changes on a regular basis. Retrieving these data sets often takes more than 10 s, during which we provide feedback that indicates progress while taking care not to lock the user interface. When rendering and processing large data sets we had to verify that the application does not slow down the processor to avoid lagging. This required taking a noticeable time frame to start up the application. Also, the limited number of pixels on the Microsoft Surface (1024×768) provided a physical constraint that we needed to address when designing visualizations for large genomic data sets. Jacob et al. provide a useful framework for considering tradeoffs in the design of reality-based interaction (Jacob et al., 2008).

4. Deployment and evaluation

In order to understand the strengths and limitations of G-nome Surfer, we applied a multi-tiered evaluation framework (see Table 1) that consists of three layers. This framework examines the usability, usefulness, and impact of tabletop interaction in collaborative context. Drawing upon the existing body of work in the area, the proposed framework documents a mixed-method approach that aims to provide guidance (rather than an extensive checklist) for the evaluation of collaborative tabletop settings.

Following, we describe our evaluation framework. We then demonstrate its application in the evaluation of G-nome Surfer 2.0.

Table 1
A multi-tier evaluation framework for collaborative tabletop interaction.

	L1: Usability	L2: Usefulness	L3: Impact
Dimensions	Functionality	Efficiency	Performance
	Learnability	Effectiveness	Engagement
	Performance		Collaboration
	Memorability		
	Satisfaction		
	Errors		
Settings	In-Lab Study	Comparative study	In-Situ deployment
			Longitudinal evaluation
Metrics			
Task			
Completion	x	x	X
Workload	x	x	X
Accuracy	x	x	X
Time on task	x	x	
Number and type of errors	x		
Users			
Attitudes	x	x	
Satisfaction ratings	x		
Levels of participation		x	
Equity of participation		x	
Engagement ratings			X
Learning			
Nature of discussion		x	X
Nature of collaboration		x	X
Problem solving strategies			X
Number of hypotheses			X
Methods			
Expert review	x		
Logging information	x		
Observation	x	x	X
Discourse analysis		x	X
Video coding		x	X
Questionnaires	x	x	X
Interviews		x	x
Debrief	x		

4.1. Multi-tiered evaluation framework for tabletop interaction

The proposed framework consists of three layers that examine the usability, usefuless, and impact of tabletop interaction in the context of collaborative learning. The first layer applies micro perspective—focusing on the *usability* of concrete interaction techniques and the effectiveness of individual visualizations. The second layer applies macro perspective—studying the *usefulness* of a system in the context of a full-scale task. Finally, the third layer applies holistic perspective—examining the *impact* of the system on users' performance and practices in-situ. Table 1 provides a summary of our evaluation framework. For each layer, we describe its dimensions, settings, metrics, and methods for data collections.

The first layer, *usability (L1)*, consists of six dimensions that draw upon Schneiderman's definition of usability

(Shneiderman and Plaisant, 2008) and are not specific for collaborative tabletop settings. These include: (1) functionality—the ability of the system to support the user in completing a required set of tasks; (2) learnability—the extent to which it is easy to learn how to use a system; (3) performance—the extent to which the accomplishment of a task satisfies known standards of completeness, accuracy, and speed; (4) memorability—the ability of the user to re-establish proficiency using a system after a period of not using it; (5) errors—the number, kind, and severity of errors as well as how easy it is to recover from errors; and (6) satisfaction—the degree to which a user finds the system pleasant to use. These dimensions are easily quantifiable using mostly task-centered metrics as specified in Table 1.

The second layer examines *usefulness (L2)*, the advantages of a system for accomplishing a particular task in collaborative settings. It consists of two dimensions: (1)

effectiveness—the extent to which users' goals are obtained through an *effective* collaborative process in which group members actively communicate with each other to demonstrate shared effort (Kirschner et al., 2009); and (2) efficiency—the degree to which goals are obtained with the investment of less effort (physical and mental), and time. Effectiveness and efficiency are interdependent and should be considered together.

These dimensions can be quantified by combining various task-, user-, and learning-centered metrics that are calculated using mixed methods. For example, the effectiveness of a collaborative tabletop interface can be quantified by measuring task completion rates and accuracy along with examining the effectiveness of the collaborative process. Collaboration profiles (Shaer et al., 2011) are often useful for describing the nature of a collaboration by highlighting the different roles participants assume throughout the collaborative process. Computing the level of participation per user is helpful for calculating the equity of participation (Harris et al., 2009). Dialog analysis can provide further insight into the nature of discussion carried by users while working on a task. For example, such dialog analysis can reveal the time spent on task-related vs. non task-related talk, or highlight insights gained by the users (Saraiya et al., 2005). Efficiency can be quantified by measuring task completion time as well as mental and physical effort. Subjective mental and physical effort is often measured using the standard NASA TLX questionnaire (Hart and Stavenland, 1988).

Finally, the third layer focuses on studying the *impact (L3)* of a collaborative tabletop system on users' performance and practices in-situ. This layer takes a holistic approach, studying impact on three dimensions: (1) performance—here we consider performance more broadly than in the usability layer, examining not only quantitative task-centered metrics such as time, completion rates, accuracy, and workload, but also learning-centered metrics that focus on *how* users learn together. In particular, we look into how users solve problems and learn in collaborative tabletop practices. We suggest utilizing video and discourse analysis to identify behavioral profiles, problem solving strategies, and the number of hypotheses explored by users; (2) engagement—this dimension goes beyond mere user satisfaction to capture the degree of user's interest, emotional involvement, and dynamic interaction. O'Brien et al. (2008) developed a multi-scale measure for user engagement that considers six attributes of engagement: Perceived Usability, Aesthetics, Focused Attention, Felt Involvement, Novelty, and Endurability; and (3) collaboration—the degree and manner to which users collaborate on a task. Here, we consider various metrics that indicate how users collaborate, what roles they assume during the collaborative process, and whether and when they switch their roles. To rate the effectiveness of the collaborative process, we use the rating scheme created by Meiers' et al. (2007) that considers 5 different dimensions of collaboration: communication, information pulling, coordination, interpersonal relationship, and motivation.

Table 1 provides a summary of our evaluation framework depicting these three layers of evaluation as well as a set of metrics and methods for each layer.

4.2. Evaluating G-nome Surfer

In this section, we describe the application of our evaluation framework to the evaluation of G-nome Surfer.

We applied the first layer of evaluation, *usability (L1)*, on a continual basis throughout the development process. In addition to user testing of each of the complete versions, we often conducted micro-studies studying the usability of particular features through the iterative development and testing of a series of prototypes in increasing fidelity (Shaer et al., 2010).

The second layer, the evaluation of *usefulness (L2)*, was applied through an experimental study with 48 participants that compared undergraduate students' learning of genomics using existing bioinformatics tools and two alternative prototypes of G-nome Surfer 2.0: a collaborative multi-mouse GUI and a tabletop interface (Shaer et al., 2011). We also evaluated the usefulness of G-nome Surfer Pro through a study with 14 student researchers that used the interface for a primer design task.

Finally, we applied the third layer of our evaluation framework to study the *impact (L3)* of G-nome Surfer 2.0 in authentic educational settings, deploying it in an intermediate-level undergraduate Neuroscience laboratory course at our institution. Results from this evaluation provide empirical evidence for the feasibility and value of integrating tabletop interaction in college-level education as well as shed light on *how* users collaborate and solve problems using a tabletop interface in the context of college level inquiry-based learning. In the future, we plan to further evaluate the impact of G-nome Surfer Pro in a longitudinal study deploying it in a synthetic biology research laboratory.

Following, we review our experimental evaluation of G-nome Surfer 2.0 for usefulness (L2) (Shaer et al., 2011). We then describe in detail *new* findings from studying the impact (L3) of deploying G-nome Surfer 2.0 in an intermediate-level undergraduate Neuroscience laboratory course.

4.3. Evaluating usefulness: experimental study

To investigate G-nome Surfer's *usefulness* (L2) in the context of collaborative inquiry-based learning of genomic concepts, we conducted a between-subjects experiment with 48 undergraduate students. We compared the system to both current state-of-the-art tools with traditional GUI and to a collaborative multi-mouse GUI. We considered both the effectiveness and the efficiency of the collaborative interaction facilitate by G-nome Surfer, and utilized mixed methods such as observation, discourse analysis, video coding, and post-task questionnaires to compute various metrics such as verbal and physical participation, equity of participation, performance, task workload, and the nature of collaboration and discussion. Table 1

summarizes the metrics and methods we used in this evaluation of usefulness (L2). The results of this study and the statistical analysis methods used are described in detail in (Shaer et al. (2011)). Here we review our main findings.

Findings from this study indicate that G-nome Surfer reduces users' stress levels and workload compared to current state-of-the-art tools as well as improves students' attitude (i.e., confidence and enjoyment) and performance. Fig. 8 shows subjective workload and attitude data. We found the two G-nome Surfer conditions (i.e., tabetop and multi-mouse GUI) resulted in significantly less workload and stress, and significantly more enjoyment and confidence than the traditional state-of-the-art tools. Table 2 shows the score participants received on their lab-report based on correctness. Correctness scores were statistically significantly higher in the two G-nome Surfer conditions than in the traditional GUI bioinformatics tools. No significant differences were found between the two G-nome Surfer conditions in respect to stress, workload, attitude, or performance (i.e., score).
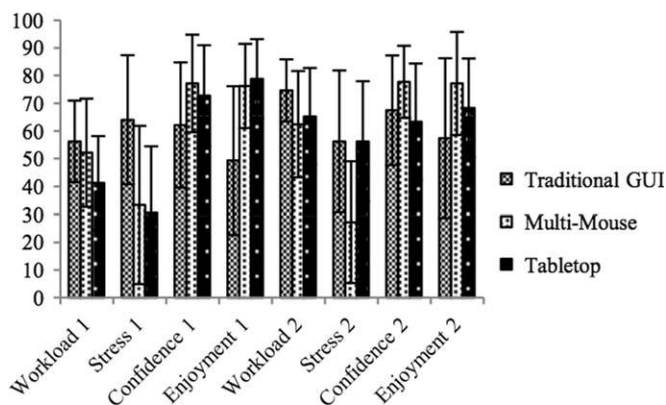

Fig. 8. G-nome Surfer 2.0 subjective workload and attitude data.

We did find some significant differences between the tabletop and multi-mouse conditions that highlight four ways in which tabletop interaction benefits learning:

1) *Physical participation:* Participants in the tabletop condition exhibited significantly higher levels of physical participation, expressed by increased spatial manipulation of information. Our metric of physical participation combined both touch events (and mouse events) and "off-line" gestures—movements of the hands and other body parts. Table 2 shows levels of verbal and physical participation per each experimental condition. We found that the tabletop condition exhibited statistically significant higher levels of physical participation compared to both the traditional GUI and the multi-mouse GUI. These high levels of physical participation can be attributed to high touch rates. While the significantly higher touch rates in the muti-touch tabletop were expected compared to the traditional single-touch GUI, the comparison with the multi-mouse GUI suggests that direct touch combined with a horizontal display promotes touch. We observed that in the tabletop condition participants manipulated information artifacts–moving, resizing, and rotating-to a greater extent than in the two other conditions. Often, users aligned information artifacts side by side for comparison and then moved them around the table to share with their partner or to place them in an area of the tabletop for later use. In the context of learning, increased physical participation is a positive indicator, as several theories of embodied cognition suggest that spatial manipulations can help reasoning about abstract concepts (Klemmer et al., 2006).

2) *Encouraging reflection*: In the tabletop condition, participants spent significantly more time on reflection activities and articulated a larger number of insights than in the

Table 2
Descriptive statistics of experiment results. Score is calculated on a scale of 0–100, time is measured in minutes and seconds, verbal participation is calculated as the mean number of utterances per minute per user, physical participation is calculated as the mean number of "offline gestures" and touch events per user per minute, equity is calculated using Gini co-efficint which produces a value between 0 and 1, the closer the value to 0 the higher the equity.

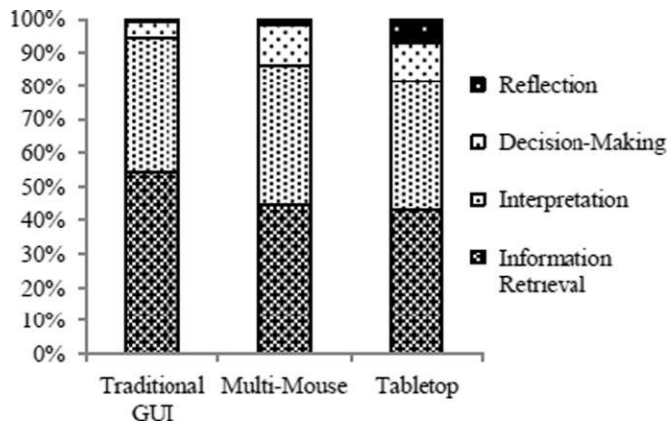|	Traditional GUI M (SD)	Multi-mouse GUI M (SD)	Tabletop M (SD)
Score	66.37 (10.8)	83.87 (9.13)	85.75 (11.8)
Time			
Task 1	27:24 (9:35)	27:58 (10:55)	21:25 (2:41)
Task 2	36:13 (10:25)	28:15 (11:01)	27:38 (5:10)
Participation			
Verbal	6.63 (2.2)	6.25 (2.0)	7.25 (1.8)
Physical	2.38 (0.9)	4.74 (2.2)	14.9 (7.5)
Touch rate	1.2 (0.56)	3.4 (1.96)	13.3 (7.3)
Equity			
Verbal	0.05 (0.04)	0.1(0.1)	0.07 (0.1)
Physical	0.14 (0.05)	0.12 (0.1)	0.15 (0.06)
Touch	0.27 (0.12)	0.15 (0.1)	0.17 (0.06)

Fig. 9. Time per activity relative to total session time.

other conditions. We found that while there is no statistically significant difference in the total amount of time taken to complete the experimental tasks between the different conditions, there were differences in how much time participants spent on each activity type depending on condition. Table 2 specifies the time-on-task for each condition. Fig. 9 shows breakdown for each type of activity relative to overall session time. We found that more time was spent on reflection in the tabletop condition than in the traditional GUI and in the multi-mouse GUI. Through a discourse analysis we found that in the tabletop condition the number of articulated insights (Saraiya et al., 2005) was significantly higher than in both the traditional tools and the multi-mouse GUI. In the tabletop condition users tend to articulate and reflect on their insights more than in the traditional GUI and multi-mouse conditions, where users often quietly write an answer and move on, keeping discussion brief. Since research indicates that student's understanding of the nature of science is enhanced through reflection (Singer et al., 2005), increasing reflection is an important strength.

3) *Fostering effective collaboration:* In the tabletop condition, participants were engaged in a more effective collaborative learning process than in the other conditions. This was evident from the turn-taking collaboration style exhibited by most tabletop pairs, compared to the driver-navigator, driver-passenger, and independent work styles that were more prevalent in the other conditions. Discourse analysis data revealed that in the tabletop condition there were a significantly higher number of coordination utterances, and a significantly lower number of disengagement utterances compared to both the traditional GUI and the multi-mouse GUI conditions. We observed that in the tabletop condition, more so than in other conditions, users established task division through continuous discussion—defining and taking on individual subtasks as they went through the task. Taken together, turn-taking collaboration style, higher number of coordination utterances, and lower

number of disengagement utterances provide an indication of effective collaboration.
4) *Facilitating intuitive interaction:* The tabletop condition facilitated more intuitive interaction. This is evident from a statistically significantly lower number of utterances related to interaction syntax found in the tabletop condition compared not only to the traditional bioinformatics tools but also to the multi-mouse version of G-nome Surfer that exhibits similar features and visual design. Furthermore, we found that in this condition users spent less time finding information and more time discussing it (see Fig. 9).

These findings provide empirical evidence for the *usefulness* of a multi-touch tabletop interaction in the context of college-level collaborative learning of abstract concepts. Results indicate that tabletop interaction benefits collaborative learning by facilitating a more *effective* collaborative learning process in terms of physical participation, reflection, and collaboration style. However, while some findings indicate that multi-touch tabletop interfaces are more intuitive to use, the study does not provide conclusive evidence that the tabletop condition facilitates a more *efficient* learning process than a multi-mouse GUI (in terms of invested mental effort).

4.4. Evaluating impact: in-class deployment

Following the experimental study evaluating the usefulness of G-nome Surfer, we evaluated the *impact* (L3) of G-nome Surfer 2.0 in authentic classroom settings, deploying the system in an intermediate-level undergraduate Neuroscience laboratory course in our institution. We were particularly interested in studying G-nome Surfer's impact on four dimensions: performance, collaboration, user engagement, and user satisfaction. Next, we describe the evaluation setup and findings in respect to these dimensions.

The study was conducted as part of a five-module laboratory series titled "From Behavior to Molecule". In the beginning of this laboratory series, students are presented with mice exhibiting an unclassified motor mutation and a list of 5 genes for which there are known motor mutations. Throughout the five modules, students investigate the behavior and anatomy of the mice as well as the motor mutations associated with each of the candidate genes. By the end of the laboratory series, students are required to conclude which of the candidate genes is responsible for the motor mutation exhibited by the mice and the physical nature of the mutation in the DNA. G-nome Surfer 2.0 was used in the fourth session (Bioinformatics) to investigate each of the 5 candidate genes.

4.4.1. Procedure and participants
The study took place in two laboratory sections, held on different days and lasted 3 h each. Two Microsoft Surface devices were set up in the Neuroscience laboratory room.

Three stools were arranged next to each device. Three researchers and the two NEUR200 course instructors were present in the room during the study but did not interact with students while they were working with G-nome Surfer 2.0.

18 students total (17 female, 1 male; ages 18–21) from two laboratory sections participated in the study. Each section was split randomly into three groups of two, and one group of three, for a total of four groups per lab section (eight groups total). One week prior to the Bioinformatics laboratory session, students were asked by their instructors to use online bioinformatics tools to investigate each of the candidate genes. At the beginning of the lab session students were asked to cast a vote indicating which of the candidate genes they identified as the source of the mutation based on their investigation so far. Then the class was divided so that two groups used G-nome Surfer 2.0 while the other two received a tutorial on a new topic (unrelated to the investigated genes). A researcher introduced each group to the Microsoft Surface and G-nome Surfer 2.0 application, and then students proceeded to investigate the five mouse gene mutations. When the instructors finished the tutorial, the four groups switched. Once all four groups had completed both parts, the students voted again on the mouse gene mutation. The instructors then led a discussion of the votes, revealing the correct answer.

4.4.2. Results

Our evaluation focused on four dimensions: performance, engagement, collaboration, and satisfaction. These dimensions and their metrics are discussed in our evaluation framework and are summarized in Table 1. Here, we present the results for each of these dimensions.

4.4.3. Performance.

To gain insight into how users work together using a multi-touch tabletop to complete a task, we considered performance more broadly, examining both quantitative measures such as time-on-task, accuracy and workload, and qualitative indicators. In particular, we looked into *how* users solve problems using a collaborative multi-touch tabletop interface, and how such interface facilitates work in university lab settings.

Time-on-task:

The 8 groups spent on average 32 min using the system (SD=11:51 min) investigating an average of 6.5 hypotheses (SD=1.20) per group during that time. We considered each gene search a hypothesis, disregarding mis-searches due to typos. We attribute the relatively large standard deviation in time-on-task to natural variations among students and groups.

Accuracy:

One week prior to the Bioinformatics laboratory session, students were asked by their instructors to use online bioinformatics tools to study the candidate genes and determine the correct mouse gene mutation. Students then voted at the beginning of the session. Of the 18 students, only 1 (5.6%) correctly identified the mouse gene mutation upon arrival.

After working with the G-nome Surfer system, 5 students (total from both sections) correctly identified the mutation (27.8%). Comparatively, in a previous semester where G-nome Surfer 2.0 was not used and instead students were asked to work in teams to consider their notes while consulting online bioinformatics tools, only 1 out of a total of 8 students (12.5%) correctly identified the mutation. While these numbers are not sufficient for quantifying how much was learned by the students in each of the semesters and may be influenced by natural variations between students and groups, they provide an indication that students were able to complete their task while exploring multiple hypotheses, and that the interface led to some improvement in performance.

Task Workload:

We used the NASA TLX (Hart and Stavenland, 1988) post-task questionnaire to measure subjective task workload. We found that, on average, participants rated their task success (i.e., performance) moderately high while rating the task workload across its different dimensions relatively low. Fig. 10 shows results per dimension. Considering the moderate mental workload and effort reported by students and the relatively high intrinsic cognitive load of the task (as evident by the relatively low success rates), these results indicate that the mental workload associated with operating the horizontal multi-touch interface was low.

Problem Solving Strategies:

To study how a horizontal multi-touch surface facilitates collaborative problem solving, we conducted in-class observations as well as video analysis. In particular, we were interested in learning how users utilize the horizontal multi-touch surface to organize, relate, and share information. We identified that each of the groups applied one of three distinct problem-solving strategies: (1) comparison, (2) sequential-comparison, and (3) sequential-redundancy. Table 3 describes each of these strategies. It also specifies how many of the groups applied each of these strategies. Participant responses from post-task interviews further highlight how users utilized the horizontal surface for problem solving: "G-nome Surfer really helped in examining
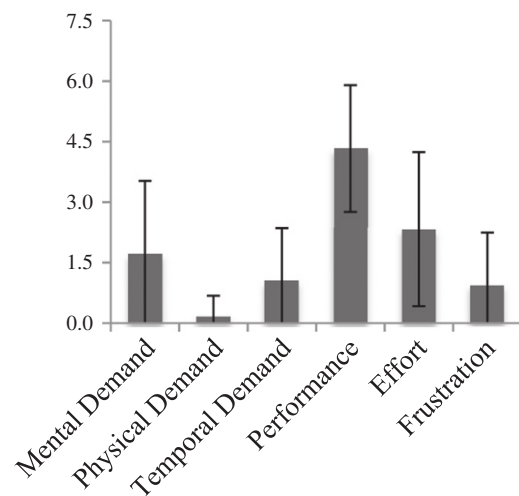

Fig. 10. Results from the NASA TLX questionnaire (N=18).

Table 3
Problem-solving strategies.

Strategy	Description	Examples
Comparison (1)	Information objects are retained once retrieved and continuously compared to new information objects. The group uses the surface to accumulate and spatially arrange information.	A: Whoa, this is really different from the others… B: I just think this looks different- A: They do not have any purkinje cells. B: This looks bigger… A: Wait…Ok, yeah.
Sequential-comparison (5)	Task is segmented: group investigates one gene at a time until all genes have been examined. Information objects are discarded after each segment, but a subset of objects is retrieved for comparison after all genes have been investigated once. Users use the surface to spatially arrange the information and for side by side comparisons.	A: I think it's one of the first three that we had…either *wobbler* or *weaver*. B: Ok, do we want to look at those again? A: Yeah, we have time.
Sequential-redundancy (2)	Task is segmented: pair investigates one gene at a time and eliminates each as they go. Objects are casually discarded during or after each segment. Information objects are re-retrieved and re-examined a total of two or more times, but only one gene is investigated at a time.	A: Ok, do you want to throw this one away then? B: Yeah. If we throw this away, are we going to look up another? B: Are we done with this mouse? A: Yeah A: And this is *weaver*. B: I think this is the same thing. A: Yeah, so we're totally tossing this…

Table 4
Interaction with paper-based materials.

Style	Description	Examples
Central	Binders sit open in laps or are perched directly in front of users on the Surface edge. There are 1–2 *active* papers that are open and spread out in the user's immediate view and are accessed continuously. Additional papers (typically 2–3) are also frequently accessed. Information from the papers and notes is used to steer direction and activity on the Surface.	A: (reads from paper) A: We did not see a staggering gait. B: Okay, let's keep this in mind. A: Yeah, the only thing is that it has to be homozygous. B: And now we want to search GRID2? A: Yeah.
Referential	Binders sit either open or closed in laps or perched on the nearby edges of the Surface, with 1–2 frequently accessed papers resting on top of the binder. Binders are accessed regularly throughout the session, but information is derived from only a few (1–2) papers within. Paper access is driven by activity on the Surface, but a piece of information in the papers and notes sometimes spurs the Surface retrieval of more or previously seen information for further examination.	A: Virtually no purkinje cells…Did we see any purkinje cells? B: (reads from paper) B: Ours were deficient. A: Hmm.
Supplemental	Binders containing papers and notes rest typically closed in laps, on the floor, or perched on the far edges of the Surface. They are used only a handful of times to look up spellings of gene names, for example, or in lieu of technical and physical limitations of the Surface.	A: 'I need a pen…' B: [while A grabs pen and notebook] 'The two options were homozygous recessive and heterozygous.' A: 'So this would be like:' [shows B drawing on paper] B: 'Just one.' A: 'Like that?' B: [takes notebook and pen] 'So, like RR, so it'd be like:' [shows A drawing]

the strengths and weaknesses of your own hypothesis with those of others in the group. Being able to see and compare researched data side-by-side was particularly useful." Another student described: "It was very helpful in the sense that information was very easy to find and could be organized into places that are easily seen and used."

Use of paper-based material:

We were particularly interested to learn how G-nome Surfer 2.0 fits into the teaching laboratory ecology and impacts current practices. In this course, students document their work in the lab using a paper-based lab notebook and notes. During the lab session we neither encouraged nor discouraged participants from referring to their paper-based materials. We observed three different ways in which participants interacted with physical information artifacts such as lab notes and papers while working with G-nome Surfer 2.0. Table 4 describes these different ways and shows an example for each style.

In general, students were able to effectively integrate digital and physical information artifacts. We noticed that

students often used the surface to spatially organize both digital and physical information. One participant described the use of digital and physical (paper-based) information: "We used our lab notes and the information that was provided on the G-nome Surfer together to come up with better hypotheses."

4.4.2.2. Engagement. Individual engagement was assessed using O'Brien's engagement questionnaire on scale of 1 to 5 (1 is strongly disagree, 5 is strongly agree) (2008). This questionnaire considers six dimensions of user engagement. Five of the six dimensions – perceived usability, involvement, endurability, novelty, and aesthetics – ranked highly, while focused attention fell about neutral (likely due to being situated in a classroom environment). Fig. 11 shows the results of the engagement questionnaire.

4.4.2.3. Collaboration. To understand *how* users collaborate when using G-nome Surfer 2.0 we conducted both in-class observations and video analysis. We identified that
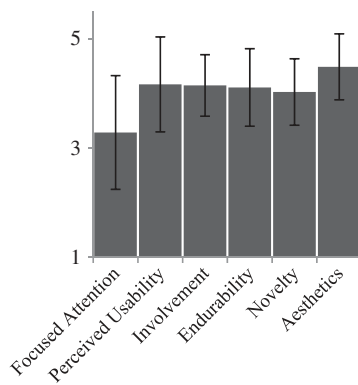

Fig. 11. Results of engagement questionnaire.

while using G-nome Surfer 2.0 individual participants typically assume one of five roles (i.e., collaboration profiles):

- *Independent*: user is absorbed in her own activity; there is a minimal verbal communication.
- *Driver-Passenger*: the driver is fully engaged in the task and operates the interface; the passenger is not focused on the task.
- *Driver-Navigator*: both users are engaged. The navigator contributes with suggestions and observations, the driver operates the interface and sometime follow suggestions.
- *Turn-Taker*: both users operate the interface as well as make and accept suggestions and observations.

Table 5 describes each of these roles and specifies the number of students that fit into each of these profiles. These collaboration profiles are based on those we defined in Shaer et al. (2011). We observed that users did not switch roles during the session. The navigator and the passenger did "warm up" to the interface after a short while, but their physical participation was still minimal. For example, they would move an information artifact or remove it from the table. Overall, 5 groups could be described as turn-takers, 1 group as driver-navigator, 1 group as driver-passenger, and 1 group as independent.

Each student group was also rated immediately post-task on Meier's nine dimensions of computer-supported collaboration (2007) using a scale of 1 (low) to 5 (high). The nine dimensions consider five aspects of collaboration: communication, information processing, coordination, interpersonal relationship, and motivation. Seven of the eight groups rated highly for all of the five dimensions, reflecting the distribution of collaboration profiles presented above. The wide variance is due to the group classified as

Table 5
Collaboration profiles.

| Profile	Description	Example
Driver (2)	Participates by physically acting out the actions required to acquire information: touching the surface, reading aloud, annotating, and (sometimes) waiting for navigator's instructions.	A: So the cerebellum is highly affected… A: (reading) Mice hemizygous for this… A: What does hemizygous mean?
Navigator (1)	Participates physically by pointing, taking notes, or reading out important information. The driver and navigator delegate sub-tasks to each other. The navigator is involved in forming the group hypotheses.	A: Can we do a search? A: Oh…go down maybe? A: I know we definitely noticed a lack of granule cells.
Passenger (3)	Only moderately physically or verbally involved in the activity. The passenger might comment and touch the surface to do a menial task (e.g., throwing an information artifact away).	A: So if we did not find those differences in the brain… A: but it depends on to what extent… A: because if we are just comparing the two… A: so this one could still be a possibility. B: Ok.
Independent (2)	An independent user takes notes individually; keeps information artifacts isolated to her "space" and does not share insights with partner.	A: What are you looking at? B: Uhhh, expression.
Turn-taker (10)	A turn taker is engaged in the activity. While turn-takers, may not be touching the surface simultaneously they actively participate in each sub-task and take turns in physical interactions or notes.	A: I think the main thing about this one is- B: One or more limbs. A: I do not think they were paralyzed. B: Yeah. A: Should we try another one?
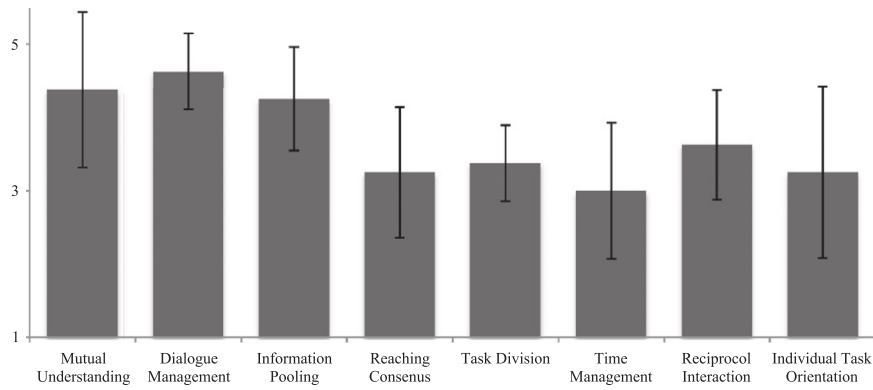
Fig. 12. Ratings on Meier's nine dimensions of computer-supported collaboration.

independent, which correspondingly received an average rating of 1.89 for each of the dimensions. The measures for Meier's nine dimensions are illustrated in Fig. 12.

Finally, in post-task interviews students commented on the value of G-nome Surfer in fostering collaboration and discussion, writing, "G-nome Surfer contributed most to my satisfaction because while I usually prefer to work alone, the teamwork felt effortless, natural, and fun," and "discussions were able to get further with the help of the G-nome Surfer."

4.5. Discussion

These findings provide empirical evidence for the feasibility and value of using multi-touch tabletop interaction in college-level educational settings. We assessed the impact of using G-nome Surfer 2.0 in collaborative inquiry-based learning using mixed methods, focusing on three dimensions: performance, collaboration, and engagement.

We found that G-nome Surfer 2.0 was successfully integrated into the teaching laboratory *ecology*. Students were able to effectively integrate digital and physical information artifacts. We noticed that students used the surface to spatially organize both digital and physical information. From observations and post-task interviews, we found that the ability to accumulate information on the surface and compare information side by side had an important role in students' problem-solving process. Considering the moderate mental workload and effort reported by students and the relatively high intrinsic cognitive load of the task, this study indicates that the mental workload associated with operating the horizontal multi-touch interface was low. The collaborative activity around the tabletop did not disrupt the parallel tutorial activity led by an instructor. This setting, in which students rotate between different activities during a single lab session, is typical to a college level science laboratory.

Our findings also indicate that G-nome Surfer 2.0 leads to high levels of engagement, fosters effective collaboration, and facilitates student learning. While our results are not sufficient for quantifying how much participants actually learned using the system, they indicate reasonable and even improved performance compared to a previous

semester in which G-nome Surfer 2.0 was not used. These results highlight several features such as the exploration of a relatively large number of hypotheses, the utilization of reflective problem solving strategies, and effective collaboration that are considered beneficial for learning in the sciences. Several studies indicate that discussion, reflection, and exploration are crucial for learning in the sciences (Okada and Simon, 1997; Roth, 2001; Singer et al., 2005; Tanner et al., 2003).

This study has several limitations that point towards future work. First, we studied one-time use in educational settings. Additional studies of longitudinal use are necessary in order to determine whether findings are affected by novelty. We are planning to deploy G-nome Surfer Pro in an advanced biology course for a period of 4 weeks. Second, our measurement of student performance is based on a single data point that does not necessarily assess individual learning or participants' ability to apply their learning. In the future, we plan to use additional assessment instruments to further measure student performance.

4.6. Summary

In this section we have presented a multi-tiered evaluation framework for tabletop interaction and described its application in the evaluation of G-nome Surfer 2.0. The proposed framework consists of three layers that examine the usability, usefulness, and impact of tabletop interaction in the context of collaborative learning and discovery. This framework takes a holistic approach for gaining an understanding of the strengths and limitations of tabletop interaction in collaborative settings by utilizing a variety of quantitative measures and qualitative indicators. Combined together, the dimensions and metrics proposed by this framework highlight multiple facets of the collaborative process mediated by a particular interface.

We applied this framework in the evaluation of G-nome Surfer 2.0: we continuously tested new prototypes for usability; we studied the usefulness of G-nome Surfer 2.0 in a between-subjects experiment that compared G-nome Surfer 2.0 to traditional GUI bioinformatics tools and to a multi-mouse interface; and evaluated G-nome Surfer's

impact in authentic educational settings by deploying it a college-level neuroscience laboratory course. Our findings highlight several advantages of tabletop interaction for collaborative learning including increased reflection and physical participation, intuitive interaction, and effective collaboration. We also highlighted *how* users collaborate and solve problems using a tabletop interface in the context of college level inquiry-based learning. These findings provide empirical evidence for the feasibility and value of integrating tabletop interaction in college-level education.

5. Conclusions and future work

This paper presents four main contributions: (1) a set of design requirements for supporting collaborative exploration in domains where vast amount of information is accessed and manipulated, (2) the design, implementation, and validation of a multi-touch tabletop interface that enhances collaborative learning and discovery, (3) a methodology for evaluating the usability, usefulness, and impact of tabletop interaction, and (4) empirical evidence for the feasibility and value of integrating tabletop interaction in college-level education. While the domain of genomics provides the frame for this work, this paper provides several lessons relevant to tabletop interaction and computer-supported collaborative learning in general:

5.1. The value of a multi-touch table

The multi-touch tabletop in G-nome Surfer has proven effective for complex interaction with large amounts of data. In particular, we found that the multi-touch interaction seemed natural to users while the horizontal surface facilitated problem solving by allowing users to accumulate, spatially arrange, and compare digital and physical information artifacts. The multi-touch tabletop also fostered effective collaboration where group members were actively engaged in interaction and discussion, sharing and exchanging information while demonstrating mutual understanding.

5.2. The value of strategies for reducing complexity

Given the explosion of data and computational tools in genomics, we sought to apply design strategies for eliminating complexity. In addition to traditional strategies for eliminating complexity (Janlert and Stolterman, 2010) such as reducing functionality and hiding complexity, we chose to utilize reality-based interaction and metaphors in our design as a means for reducing complexity (Jacob et al., 2008). We found that utilizing multi-touch interaction on a horizontal surface reduced the complexity associated with indirect interaction and the use of multiple mice (Shaer et al., 2011). We also used reality-based metaphors in the design of visualizations and interaction, leveraging naïve physics concepts such as gravity, transparency, and mass.

As we expected, we found that compared to interaction with state-of-the-art bioinformatics tools, interaction with G-nome Surfer caused users to exhibit significantly lower levels of mental workload (Shaer et al., 2011).

5.3. The value of rigorous user-centered and participatory design methods

Applying rigorous methods such as employing users as a permanent part of the team, undergoing extensive domain training, establishing partnership with domain scientists, and continuously testing the system, has proven essential in the development of an interface that facilitates complex interaction with large amounts of data. Through frequent interactions and active involvement in all stages of development, users developed a sense of ownership. This helped not only in designing effective interaction but also in gaining access to data sets, recruiting study participants, and coordinating in-situ evaluations.

5.4. The value of multi-layered evaluation

Finally, to gain an understanding of the strengths and limitations of tabletop interaction in the context of collaborative discovery, we utilized a multi-tiered evaluation framework that investigates the usability, usefulness, and impact of an interface. This framework utilizes a variety of quantitative measures, qualitative indicators, and evaluation settings that, when combined together, highlight multiple facets of the collaborative learning and discovery process mediated by a particular interface.

In the future, we intend to continue to integrate G-nome Surfer into college-level biology courses and further evaluate its use in a longitudinal study. We also plan to study its impact on scientific discovery by deploying it in research labs. Our long-term goal is to understand how tabletops and interactive surfaces can enhance scientific discovery and learning in data-intense areas.

References

Antle, A.N., Droumeva, M., and Ha, D., 2009. Hands on what? Comparing children's mouse-based and tangible based interaction. Proc. of IDC 2009. ACM Press, 80–88.

Antle, A.N., Tanenbaum, J., Seaborn K., Bevans, A., Wang, S., 2011. Futura: Design for collaborative learning and game play on a multi-touch digital tabletop. In Proceedings of Conference on Tangible,

Embedded and Embodied Interaction (TEI'11), ACM Press (Portugal, 2011). 93–100.

Arnstein, L., Hung, C.-Y., Franza, R., Zhou, Q.H., Borriello, G., Consolvo, S., Su, J., 2002. Labscape: a smart environment for the cell biology laboratory. In IEEE Pervasive Computing Magazine 1 (3), 2–13.

Bolchini, D., Finkelstein, A., Perrone, V., Nagl, S., 2009. Better bioinformatics through usability analysis. Bioinformatics 25 (3), 406–412.

Brooks, F.P., Ouh-Young, M., Batter, J.J., and Jerome Kilpatrick, P., 1990. Project GROPEHaptic displays for scientific visualization. In Proceedings of the 17th Annual Conference on Computer Graphics and interactive Techniques, Dallas, TX, USA.

Chilana, P.K., Wobbrock, J.O. and Ko, A.J., 2010. Understanding usability practices in complex domains. Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '10). Atlanta, Georgia (April 10–15, 2010). New York: ACM Press, pp. 2337–2346.

Echtler F., Haussler M., and Klinker G., 2010. BioTISCH: the interactive molecular biology lab bench. In Proceedings of the international conference extended abstracts on Human factors in computing systems (CHI EA '10).

Fleck, R., Rogers, Y., Yuill, N., Marshall, P., Carr, A., Rick, J., Bonnett, V., 2009. Actions speak loudly with words: unpacking collaboration around the table. Proc. of Interactive Tabletops and Surfaces.

Gillet, A., Sanner, M., Stoffler, D., Goodsell, D., Olson, A., 2004. Augmented reality with tangible auto-fabricated models for molecular biology applications. In Proceedings of the Conference on Visualization '04. IEEE Visualization.

Harris, A., Rick, J., Bonnett, V., Yuill, N., Fleck, R., Marshall, P., and Rogers, Y., 2009. Around the table: are multiple-touch surfaces better than single-touch for children's collaborative interactions? Proc. of CSCL ACM Press, 335–344.

Hart, S.G., Stavenland, L.E., 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research, in Human Mental Workload. Elsevier 139-183.

Horn, M.S., Tobiasz, M., Shen, C., 2009. Visualizing biodiversity with voronoi treemaps. In Proc. Sixth International Symposium on Voronoi Diagrams in Science and Engineering.

Hornecker, E., Marshall, P., Dalton, N.S., Rogers, Y., 2008. Collaboration and Interference: Awareness with Mice or Touch Input. Proc. of CSCW, ACM Press.

Jacob, R.J., Girouard, A., Hirshfield, L.M., Horn, M.S., Shaer, O., Solovey, E.T., and Zigelbaum, J., 2008. Reality based interaction: a framework for post-WIMP interfaces. Proc. of Human F actors in Computing Systems, ACM Press.

Kirschner, F., Paas, F., Kirschner, P.A., 2009. A cognitive load approach to collaborative learning: united brains from complex learning. Educational Psychology Review 21 31 + 42.

Klemmer, S.R., Hartmann, B., and Takayama, L., 2006. How bodies matter: five themes for interaction design. Proc.of DIS 2006, ACM Press.

Janlert, L.E., Stolterman, E., 2010. Complex interaction. ACM Transactions on Computer–Human Interaction 17, 2 Article 8.

Mackay W.E., Pothier G., Letondal C., Boegh K., Sorensen H.E., 2002. The missing link: augmenting biology laboratory notebooks. In Proceedings of the ACM symposium on User interface software and technology (UIST '02).

Marshall, P., Hornecker, E., Morris, R., Dalton, N., Rogers, Y., 2008. When the fingers do the talking: a study of group participation with varying constraints to a tabletop interface. Proc. of Tabletop 08, 37–44.

Marshall, P., Morris, R., Rogers, Y., Kreitmayer, S., Davies, M., 2011. Re-thinking 'multi-user': an in-the-wild study of how groups approach a walk-up-and-use tabletop interface. Proceedings of the 29th International Conference on Human Factors in Computing Systems. ACM CHI 2011, ACM Press, to appear.

Meier, A., Spada, H., Rummel, N., 2007. A rating scheme for assessing the quality of computer-supported collaboration processes. International Journal of Computer- Supported Collaborative Learning 2, 63–86.

Mirel, B., 2005. Supporting cognition in systems biology analysis: findings on users' processes and design implications. Journal of Biomedical Discovery and Collaboration 27, 2.

Morris, M.R., Brush, A.J.B., Meyers, B., 2007. Reading revisited: evaluating the usability of digital display surfaces for active reading tasks. Proc. of IEEE Tabletop, 79–86.

O'Brien, H.L., Toms, E.G., Kelloway, E.K., and Kelley, E., 2008. Developing and evaluating a reliable measure of user engagement. Proceedings of the American Society for Information Science and Technology 45, 1, 1–10.

Okada, T., Simon, H.A., 1997. Collaborative discovery in a scientific domain. Cognitive Science 21 (2), 109–146.

OligoAnalyzer tools, 2011, 〈http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/〉.

Patten, J., Ishii, H., 2000. A comparison of spatial organization strategies in graphical and tangible user interfaces. Process of Designing Augmented Reality Environments, 41–50.

Piper, A.M., Hollan, J., 2009. Tabletop displays for small group study: affordances of paper and digital materials. Proc. of CHI 2009, ACM Press.

Rick, J., Rogers, Y., Haig, C., Yuill, N., 2009. Learning by doing with shareable interfaces. Children, Youth and Environments 19 (1), 321–342.

Rick, J., Marshall, P., and Yuill, N., 2011. Beyond one-size-fits-all: how interactive tabletops support collaborative learning. In Proceedings of IDC 2011, pp. 109–117. ACM Press.

Rogers, Y., Lindley, S., 2004. Collaborating around large interactive displays: which way is best to meet? Interacting with Computers.

Roth, W.M., 2001. Gestures: their role in teaching and learning. Review of Educational Research 71 365 + 392.

Ryall, K., Forlines, C., Shen, C., Morris, M.R., 2004. Exploring the effects of group size and table size on interactions with tabletop shared-display groupware. Proc. of CSCW 2004. ACM Press, 284–293.

Ryall, K., Forlines, C., Shen, C., Morris, M.R., Everitt, K., 2006. Experiences with and observations of direct-touch tabletops. In Proc. IEEE Tabletop'06.

Saraiya, P., North, C., Duca, K., 2005. An insight-based methodology for evaluating bioinformatics visualizations. Visualization and Computer Graphics 11 (4), 443–456.

Shaer, O., Kol, G., Strait, M., Fan, C., Grevet, C., Elfenbein, S., 2010. G-nome surfer: a tabletop interface for collaborative exploration of genomic data. Proc. Of Human Factors in Computing Systems. ACM Press.

Shaer, O., Strait, M., Valdes, C., Feng, T., Lintz, M., Wang, H., 2011. Enhancing genomic learning through tabletop interaction. Proc. Of Human Factors in Computing Systems. ACM Press.

Schkolne, S., Ishii, H., Schroder, P., 2004. Immersive design of DMA molecules with a tangible interface, visualization. IEEE, 227–234.

Schneider B., Strait M., Muller L., Elfenbein S., Shaer O., Shen C., PhyloGenie: Engaging Students in Collaborative Tree-Thinking through Tabletop Techniques. ACM CHI 2012 Human Factors in Computing Systems Conference (in print).

Shneiderman, B., Plaisant, C., 2008. Designing the User Interface, 5th Ed. Addison-Wesley.

Schraefel, M.C., Hughes, G.V., Mills, H.R., Smith, G., Payne, T.R., Frey, J., 2004. Breaking the book: translating the chemistry lab book into a pervasive computing lab environment. In Proceedings of Human factors in computing systems (CHI '04).

Skinner, M.E., Uzilov, A.V., Stein, L.D., Mungall, C.J., Holmes, I.H., 2009. JBrowse: a next-generation genome browser, {\em Genome Research} 19 (9), 1630–1638.

Singer, S.R., Hilton, M.L., Schweingruber, H.A., 2005. America's lab report: investigations in high school science. National Research Council.

Song, K., Kim, G., Han, I., Lee, J., Park, J., Ha, S., 2011. CheMO: mixed object instruments and interactions for tangible chemistry experiments. In Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems (CHI EA '11).

Stothard, P., Wishart, D., 2005. Circular genome visualization and exploration using CGView. Bioinformatics 21 (4), 537–539.

Tabard A., Mackay W.E., Eastmond E., 2008. From individual to collaborative: the evolution of prism, a hybrid laboratory notebook. In Proceedings of the ACM conference on Computer supported cooperative work (CSCW '08).

Tabard A., Hincapié-Ramos J.D., Esbensen M., Bardram J.E., 2011. The eLabBench: An Interactive Tabletop System for the Biology Laboratory. ACM International Conference on Interactive Tabletops and Surfaces 2011. Nov 13–16, 2011, Kobe, Japan.

Tanner, K., Chatman, L.S., Allen, D., 2003. Cooperative learning in the science classroom—beyond students working in groups. Cell Biology Education 2, 1.

Veretnik, S., Fink, J., Bourne, P., 2008. Computational biology resources lack persistence and usability. PloS Computational Biology 7 (4) July.

Wigdor, D., Jiang, H., Forlines, C., Borkin, M., Shen, C., 2009. The wespace: the design, development and deployment of a walk-up and share multi-surface visual collaboration system. Proceedings of CHI 2009, Boston, MA.

Wigdor, D., Wixon, D., 2011. Brave NUI World : Designing Natural User Interfaces for Touch and Gesture. Morgan Kaufmann.

Yeh, R., Liao, C., Klemmer, S., Guimbretiere, F., Lee, ' B., Kakaradov, B., Stamberger, J., Paepcke, A., 2006. ButterflyNet: a mobile capture and access system for field biology research. In Proceedings of the SIGCHI conference on Human Factors in computing systems,571–580.