
Understanding Information Practices of Interactive Personal Genomics Users

Orit Shaer

Computer Science Department
Wellesley College
Wellesley, MA, 02481 USA
oshaer@wellesley.edu

Anne West

Wellesley College
Wellesley, MA, 02481 USA
awest2@wellesley.edu

Diana Eastman

Computer Science Department
Wellesley College
Wellesley, MA, 02481 USA

Oded Nov

Polytechnic School of Engineering
New York University
5 Metrotech Center, Brooklyn, NY
11201 USA
onov@nyu.edu

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author. Copyright is held by the owner/author(s).
CHI 2014, April 26 - May 1, 2014, Toronto, ON, Canada.
ACM 978-1-4503-2474-8/14/04.
<http://dx.doi.org/10.1145/2559206.2581145>

Abstract

The dramatic growth in the availability of personal genomic data to non-experts, and the use of digital media for delivering personal genomic information, raise important HCI questions. We seek to advance the theory and practice of HCI by investigating fundamental issues concerning non-expert interaction with complex scientific information. Specifically, we explore how people use interactive tools to engage with their personal genomic information.

We present findings from a study of 63 early adopters of personal genomics, which highlight users' motivations, information practices, and needs.

Author Keywords

Personal genomics; information practices;

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Introduction

Recent years are seeing a dramatic growth in the availability of personal genomic data to non-experts, often online and in interactive forms [8]. As a result, individuals are confronted with an unprecedented amount of sensitive information about themselves,

The Personal Genome Project

The *Personal Genome Project* (PGP) [6] is a global network of research studies spanning 4 countries, which seeks to improve the scientific understanding of the genetic and environmental contributions to human traits through the creation of a genetic public database of 100,000 volunteers.

Participants must be willing to share their genomic sequences as well as health and medical data with the scientific community and the general public.

More than 3000 U.S. citizens are now enrolled there and participants, including whole genome datasets generated as part of the research study, as well as genetic datasets obtained from other companies and uploaded to the site.

which influences their decisions, emotional state, and well-being [10, 13]. Consequently, questions about how people make sense of, and engage with their personal genomic data, are not only of paramount importance for society and policy makers, but also a pressing issue for HCI research.

Advances in genetic research and technologies are likely to further increase the availability of detailed and interactive genomic information to non-experts. This trend raises numerous technological, ethical, and regulatory concerns. A recent example is the FDA 23andMe case [4], in which the FDA ordered that the company stops providing risk assessment reports, claiming that: "serious concerns are raised if test results are not adequately understood by patients or if incorrect test results are reported." [4].

The vast amounts of highly personal and dynamic nature of personal genomic information (which is constantly updated based on new research results) also raise important HCI questions. For example, what are the functional requirements for supporting meaningful engagement of non-experts' with personal genomic information? How can we design effective interaction with personal genomic information? How can potential information overload resulting from the plethora of data be mitigated?

We begin to explore these emerging HCI questions related to users' engagement with their personal genomic information by conducting a qualitative study with 63 Personal Genome Project (PGP) [6] participants, which are early adopters of personal genomic technologies. We seek to understand how users are currently engaging with their personal

genomic information, and what additional information needs are not served by existing interactive tools for exploring personal genomics. Here, we present results from this study, focusing on current information practices and needs of early adopters.

We begin by providing background information on users' engagement with personal genomics.

Background

The Human Genome Project (HGP) was completed a decade ago, publishing the full reference sequence of the human genome in April 2003. This collaborative research program whose goal was the complete mapping and understanding of all the human genes, lasted 13 years and cost \$2.7 billion. Since then, the cost of sequencing a single human genome today has dropped to about \$5,000 in 2013, a cost drop far faster than the rate of Moore's law [8].

Decreasing sequencing costs and technological advances offer the promise of personalized medicine, with genomic information integrated into medical care to provide individualized risk assessment, tailored lifestyle change recommendations, and medications to reduce risk [8].

Online Interaction with Personal Genomics

The rapid decline in the costs of DNA sequencing over the past decade has led to widespread access to personal genomics data. Several companies now offer personal genomic services directly to consumers. For example, Illumina's Understand Your Genome conference provides consumers with an opportunity to sequence their genome and learn about their personal genomic information as part of a two-day educational

Personal Genomic data	Users
23andMe	38
Complete genomics	11
Family Tree DNA	11
Ancestry DNA	5
NG Genographic	2
Microbiome	2
Sorenson	1
Medical tests	1
Exome	1

Table 1. Number of users per personal genomic data source..

event. Test results are returned digitally on an Apple iPad and are reviewed using Illumina MyGenome application [5].

Direct-to-consumer genetic testing (DTCGT) is a relatively new and developing online service, which enables individuals to acquire genetic information without the mandatory involvement of a healthcare provider by sending a saliva sample to a DTCGT company—at the cost of a few hundred dollars. Several popular DTCGT services offer interactive online reports of non-health related information including traits and ancestry information (e.g. AncestryDNA [2] and FamilyTreeDNA [3]). 23andMe [1] also provided risk assessment results for about 250 conditions. However, the reporting of health-related information directly to consumers has been stopped while it is undergoing FDA review, which examines whether tests results are accurate and are adequately communicated to and understood by consumers [4].

The highly personal and complex nature of personal genomic information raises important HCI questions. Our goal is to contribute towards developing interactive systems, which help people engage with their personal genomic information and act upon it, while at the same time we advance the theory and practice of HCI by investigating fundamental issues such as information presentation and non-expert user interaction with complex scientific information.

Related Work

Little empirical data exists about the attitudes and motivations of personal genomic users [10]. Few studies have recruited participants who had actually

used DTCGT. In these studies, curiosity was mentioned as participants' primary motivation for undergoing genomic testing [10, 13]. Participants also stated that they would use information gained from the test to take personal responsibility for their future health [9]. Other themes include fascination with genealogy, contribution to research, and recreation [10, 13]. Studies also identified several concerns among DTCGT users, including privacy [8], the nature of the results and their future impact [9, 10, 13]. Because only a small number of users around the world have had their entire genome sequenced and returned to them [8], to our knowledge no studies have investigated the perspectives of such users. While these studies shed some light on the motivations and concerns of early DTCGT users, they do not provide insight regarding the information practices and needs of personal genomic users.

To date, little HCI research has focused on users' engagement with personal genomic information. Lachance et al. [12] examined the features of websites in which consumers could directly purchase and receive genetic testing. They found that the quality of informational content, literacy demands, and usability across health-related DTC websites varied widely. Their findings indicate that most users would struggle to find and understand the important information on the majority of sites. Other researchers have considered user engagement with genomic and biological information more broadly, focusing mainly on novel interaction techniques for large biological data sets for learning [7, 14, 15, 16] rather than understanding *personal* genomic users' specific needs.

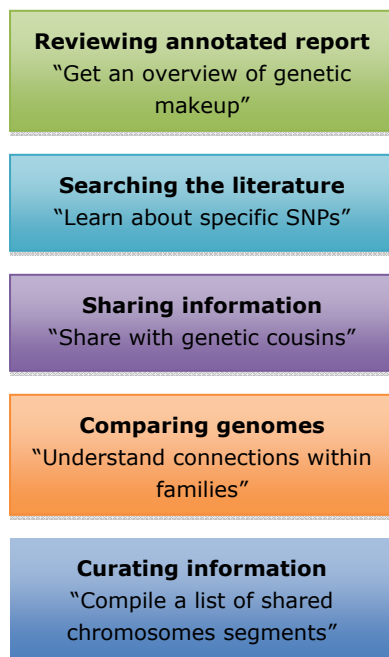


Figure 1. Common information tasks in the exploration of personal genomic information.

Qualitative Study: Understanding Users

Methodology

To better understand information practices and needs of personal genomic information users, we conducted an exploratory qualitative study. 63 study participants were recruited from the Personal Genome Project volunteer community. Participants filled in an online questionnaire consisting of 10 open-ended questions about their experiences engaging with personal genomics. We asked participants about their motivations for exploring personal genomic information, what tools they use to explore and make sense of their information, and what information needs could not be fulfilled by existing tools. We also asked about the impact their discoveries had on their life in terms of behavior change, and about concerns regarding personal genomics.

Data was analyzed using content analysis methods. First level codes were developed from preliminary review of the data by two independent coders and were then collapsed into advanced categories based on frequency. Categories were analyzed for the identification of themes. In what follows we report results regarding the information practices and needs of users.

Results

63 participants filled out the online questionnaire (29 women), aged between 21 and 71, with an average age of 47. All participants had prior access to their personal genomic data. Some participants had personal genomic information from more than one service provider. Table 1 shows the number of users for each type of personal genomic data.

82% of the participants hold undergraduate or graduate degrees, 32% of the participants hold doctoral degrees, and 30% of the participants work in life sciences related fields.

Tools and information tasks for personal genomics

We asked participants which websites and computational tools they used for engaging with their personal genomic information, and how they used these tools to learn from their data.

Only 11% of participants have used tools beyond those offered by their personal genomics provider (e.g. 23andMe, PGP, AncestryDNA, fTree). Additional 10% of the participants have mentioned that they tried to explore their genomes using the tools provided by their service provider or using other tools but found the tools to be too complicated, overwhelming, or “not user friendly”. Table 2 presents a summary of tools and websites mentioned by more than 2 participants. For each of these tools we summarize its main features.

Overall, we found that although our study participants were motivated by a diverse set of questions, ranging from understanding traits, to identifying health risks, and to learning about their ancestry, they used existing tools to perform five common information tasks: **Reviewing an annotated report; Searching the literature; Sharing information; Comparing genomes; and Curating information.** Figure 1 depicts these information tasks, and provides an example quote motivating each task.

Information needs

We also asked users what could help them learn more from their personal genomic data information. The following needs have emerged from user responses:

Tool	Users	Main features
23andMe (Testing service, interpretative tool)	22	-Health report on over 240 conditions with external links - Visualized ancestry info -Download of raw data file
Promethease (Interpretative tool)	8	- Annotated report with latest information from SNPedia -Comparison of two genomes -Family report
SNPedia (Database)	5	- Medical, genealogical, and phenotypic variant associations - SNP articles with links to publications and summary -Facility for sharing data -Access to shared data
PGP- (Testing service, database)	7	- Annotated report of variants sorted by clinical importance and impact. - Article for each variants includes summary and links to relevant publications -Facility for sharing data -Access to shared data
Gedmatch (Database, Interpretive tool)	7	-Comparison of user's data with public results -Visualized information about selected matches - Genetic distance calculator -Relationship calculator
AncestryDNA (Testing service, interpretive tool)	7	-Information on ancestry -Updates of new matches -Access to family tree -Download of raw data file
PubMed (Database)	6	- Access to books and journals about all aspects medicine and life sciences.

Table 2. Interactive tools for exploring personal genomics.

- 1) **Integrated resources** – multiple users highlighted a need for seamless integration of data resources including annotated genomes, publications, various public databases, and health related data. In the words on one participant: “Integrated databases of published research that allow the end user, through a seamless interface, to connect personal data with any possibly relevant literature and public data.”
- 2) **Visualization and information presentation** – users also commented on the need to visualize their genomes. In the words of one user: “I’d be interested in seeing a graphic illustration of my chromosome sets.” Another user wrote: “It would be great to show the SNPs by chromosomal location and in relation to other genes.” Other users commented on the difficulty to interpret tabular and dense reports.
- 3) **Data triangulation** – several users asked for the ability to triangulate data from several individuals in order to understand connections within families.
- 4) **Content** – multiple users asked for adapting the content of personal genomic reports towards lay people. Users also asked to integrate educational materials within the reports: “every time I try to understand something, I have to educate myself via Google, instead of the interface that gives me my genetic data educating me. The research it takes holds me back from using my info more.” Finally, users asked for highlighting actionable information: “Features that show more clearly what reasonable actionable options there might be for dealing with or preventing various illnesses.”
- 5) **Sharing** – users also highlighted the need for tools that facilitate and encourage information sharing. In the words of one user: “The thing that would help the most would be for people to be willing to share more information.” Another user asked for: “easy to use, at home programs, will be needed to compare one’s data with those of friends.”

Conclusions and Future Work

We contribute findings from a study of 63 users, which are early adopters of personal genomics. Our findings identify common information tasks, and highlight information needs to be addressed by interactive tools for personal genomics. To our knowledge this is the first HCI study that focuses on information practices and needs for non-expert engagement with personal genomics. We draw on these findings in our current effort towards the design of new interaction techniques and tools for personal genomics. Thus, our future work will involve more in-depth interviews with users and designers of personal genomic tools, as well as quantitative studies of the relationships between users' motivations and their information behavior.

Acknowledgments

We thank Jason Bobe and the PGP volunteers. This work is partially funded by NSF grant no. IIS-1017693

References

- [1] (n.d.). Retrieved from 23andMe: <https://www.23andme.com/>.
- [2] (n.d.). Retrieved from AncestryDNA: <http://dna.ancestry.com/>.
- [3] (n.d.). Retrieved from Family Tree DNA: <http://www.familytreedna.com/>.
- [4] (n.d.). Retrieved from FDA: <http://www.fda.gov/oc/2013/enforcementactions/warningletters/2013/ucm376296.htm>
- [5] (n.d.). Retrieved from Illumina: <http://www.illumina.com/company/events/>
- [6] (n.d.). Retrieved from Personal Genome Project: <http://www.personalgenomes.org/>.
- [7] Block, F., Horn, M. S., Phillips, B. C., Diamond, J., Evans, E. M. and Shen, C. (2012). The DeepTree Exhibit: Visualizing the Tree of Life to Facilitate Informal Learning. *IEEE Visualization and Computer Graphics*, 18(12).
- [8] Davies, K. (2010). *The \$1,000 genome: the revolution in DNA sequencing and the new era of personalized medicine*. New York, NY: Free Press.
- [9] Goldsmith, L., Jackson, L., O'Connor, A. and Skirton, H. (2012). Direct-to-consumer Genomic Testing. *European Journal of Human Genetics*, 20.
- [10] Gollust, S.E., Gordon, E.S., Zayac, C., Griffin, G., Christman, M.F., Pyeritz, R.E. et al. (2012). Motivations and Perceptions of Early Adopters of Personalized Genomics: Perspectives from Research Participants. *Public Health Genomics*, 15(1), 22-30.
- [11] Kuznetsov, S., Taylor, A. S., Regan, T., Villar, N. and Paulos, E. (2012). At the Seams: DIYbio and Opportunities for HCI. In *Proc of ACM DIS*.
- [12] Lachance, C.R., Erby, L.A., Ford, B.M., Allen, V.C. and Kaphingst, K.A. (2010). Informational Content, Literacy Demands, and Usability of Websites Offering Health-related Genetic Tests Directly to Consumers. *Genetics in Medicine*, 12(5), 304-312.
- [13] McGowan, M.L., Fishman, J.R. and Lambrix, M.A. (2010). Personal Genomics and Individual Identities: Motivations and Moral Imperatives of Early Users. *New Genetics and Society*, 29(3).
- [14] Schneider, B., Strait, M., Muller, L., Elfenbein, S., Shaer, O., Shen, C. (2012). Phylo-Genie: engaging students in collaborative 'tree-thinking' through tabletop techniques. In *Proc. ACM CHI*.
- [15] Shaer, O., Mazalek, A., Ullmer, B. and Konkell, M. (2013) From Big Data to Insights: Opportunities and Challenges for TEI in Genomics. In *Proc. ACM TEI*.
- [16] Shaer, O., Strait, M., Valdes, C., Feng, T., Lintz, M. and Wang, H. (2011). Enhancing Genomic Learning through Tabletop Interaction. In *Proc. ACM CHI*.