Designing for Collaborative Exploration of Pervasive Personal and Environmental Omic Data

Orit Shaer, Jennifer Otiono
Wellesley College
{oshaer, jotiono}@wellesley.edu

Oded Nov
New York University
onov@nyu.edu

Madeleine Ball
Open Humans Foundation
madeleine@personalgenomes.org

Introduction

Recent years have seen a sharp increase in the availability of personal and environmental 'omic' data (e.g. data about genomes or microbiomes) to non-experts. Omic denotes a multicomponent perspective on biology, often based on radical advances in high-throughput DNA sequencing; e.g., personal genomics for the study of many genes, viral metagenomics for pathogen detection, and microbiomics for studying the microbiota associated with digestion, immune response, and other interactions with human health [10].

Popular omic testing services produce data about people's personal genome (e.g. 23andMe), about their microbiome (e.g. skin or gut) and about their living environment (e.g. kitchen counter, bathroom surfaces) [1, 4]. People with no formal training in the life sciences can thereby get access to their omic data by sending a self-collected sample to a direct-to-consumer provider, and results are delivered online. These consumers then need to interpret large amounts of complex data involving sensitive issues such as disease risk, carrier status and potentially meaningful correlations with health and physical traits. The interpretation of the data may impact lifestyle decisions and well-being of users, as well as relevant others (e.g. biological family members or local community members).

In addition, novel methods for rapid DNA testing allow consumers to receive omic data not only about themselves but also about plants and animals they interact with or consume. For example, DNA testing can reveal the presence of dangerous allergens such as almonds, or shellfish [6, 7, 12] in foods, detect undesired ingredients such as horse meat in processed food [11], and identify microbes that cause foodborne illnesses [8]. Rapid DNA tests of surfaces in living spaces (e.g. bed, floor) can also detect bed bugs, and cockroaches [9].

Social relevance

Much of the environmental and nutritional omic data is of high social relevance. Consider, a family or a group of roommates who share a living space. Such groups are likely to share bacteria in their living environment, and their commensal microbiota can be influenced by common lifestyle elements such as nutrition and pets. Within these groups, people might want to share information about the presence of allergens and of particular ingredients, compare personal and environmental data, as well as investigate changes in their own and environmental omic data following lifestyle and/or environmental changes (e.g. diet, new pet, seasons, new furniture). However, tools for aggregating, exploring, sharing and collaboratively making sense of such data are currently not available. The dramatic growth in the availability of omic data to non-experts will only increase the need for such tools.

Opportunities for HCI and CSCW research

The nearly ubiquitous availability of extensive and complex data, in need of understanding by non-experts (over 5 million personal genomic or microbiomics reports provided to individuals as of August 2017 [5]) pose both challenges and opportunities, with substantial societal impact, for HCI and CSCW research, including:

Investigating the motivations, information practices, and needs of non-expert personal users for engaging with omic data. Identifying requirements for tools for self and collaborative exploration within various social contexts including families and co-habiting communities.

Developing tools that integrate omic data sets from different sources and multiple samples - allowing users to aggregate, share, explore, and connect pervasive omic information, and facilitating collaborative sense-making of omic information.

Visualizing the invisible - designing, developing, and evaluating novel interfaces which display timely and actionable omic data in shared living environment and on a user's own body (gut, skin).

Understanding long term user engagement with personal and social omic data as a way to empower people to engage in self driven small-scale citizen science investigations.

Preliminary Platform Design

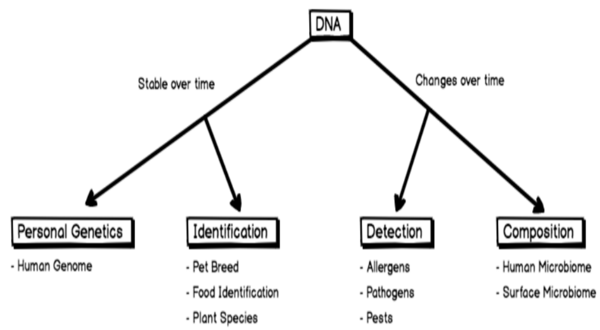

Figure 1. Four data types available from Direct-to-consumer omic testing: personal genetics and identification, both remain stable over time (sufficient to sample once), and detection and composition, which change over time and require multiple samples.

Figure 2. OmiX-AR display for all risk factors. Most recent sample sites are indicated with a circle, and users can tap on the sample site for a popup of the detection and identification results for that sample, and can also customize their list of risk factors for investigation.

Given these challenges and opportunities, our preliminary platform design focuses on supporting social sharing and exploration of omic data. Considering the data types available from DNA testing (Figure 1), and based on our conversations with leading experts from Open Humans [3], American Gut project [1], and NRGene [2], we identify the following design goals: 1) Contextualizing omic data types, which might require direct action, within particular environmental or social settings; 2) Facilitating longitudinal comparison of omic information across data types and sites; 3) Supporting the dynamic construction of various social configurations as a basis for comparison.

Figure 2, shows our preliminary design which overlays identification and detection omic data on shared surfaces, food, living things, and on the user's own body; Figure 3 shows, our design for an online visual interface that supports in-depth collaborative engagement with different types of omic data including both dynamic (e.g. microbiome) and static (e.g. personal genomics), as well as other relevant information such as ongoing research publications that are relevant to users' personal and social data, and afford interaction and co-curation.

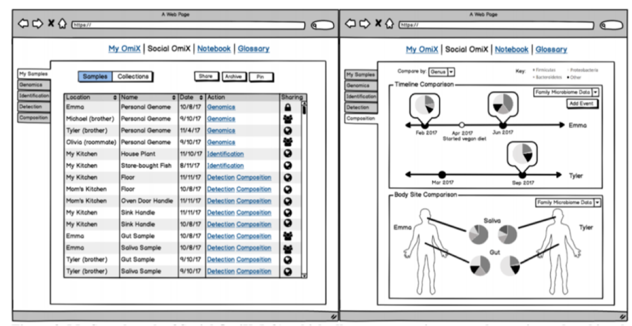

Figure 3. My Samples tab of Social OmiX (left), which allows users to view, sort, share, pin, and archive their samples and the samples of others that have been shared with them, as well as choose an action to investigate these sample in the other four tabs. Composition tab of Social OmiX (right), allows users to explore and compare the bacterial composition of their samples. The timeline comparison allows users to record lifestyle changes to investigate how changes can affect their microbiome. Body site and composition comparison visualizations offer users the opportunity to compare samples from multiple people.

Conclusion

In this paper, we presented research and design goals for supporting collaborative exploration of socially relevant omic data. The abundance and complexity of omic data serves as a challenge and opportunity for HCI research to empower users to engage in small scale citizen science research.

References

1. http://americangut.org/
2. http://www.nrgene.com/
3. http://openhumansfoundation.org/
4. https://ubiome.com/

5. Ancestry (2017). “Ancestry Surpasses 5 Million People in DNA Database, Giving Customers Even More Opportunities to Discover Who They Are and How They Connect to One Another”, Retrieved from https://www.ancestry.com/corporate/newsroom/press-releases/ancestry-surpasses-5-million-people-dna-database-giving-customers-even-more
6. Malcolm Burns, Michael Walker, Timothy Wilkes, Laurie Hall, Kirstin Gray, Gavin Nixon, “Development of a Real-Time PCR Approach for the Specific Detection of Prunus mahaleb,” Food and Nutrition Sciences (2016) 7:703-710, doi:10.4236/fns.2016.78071
7. Anne C Eischeid, “Development and evaluation of a real-time PCR assay for detection of lobster, a crustacean shellfish allergen”, In Food Control, Volume 59, 2016, Pages 393-399, ISSN 0956-7135, https://doi.org/10.1016/j.foodcont.2015.06.013.
8. Andrew D. Huang, Chengwei Luo, Angela Pena-Gonzalez, Michael R. Weigand, Cheryl Tarr, Konstantinos T. Konstantinidis. “Metagenomics of two severe foodborne outbreaks provides diagnostic signatures and signs of co-infection not attainable by traditional methods.” Applied and Environmental Microbiology, 2016; AEM.02577-16 DOI: 10.1128/AEM.02577-16
9. Leigh Kamping-Carder, “Is that a bedbug? A DNA test will tell you in 24 hours”, Retrieved on November 1, 2017 from https://www.brickunderground.com/blog/2014/11/bed_bug_dna_testing
10. Stacey Kuznetsov, Alex S. Taylor, Tim Regan, Nicholas Villar, & Eric Paulos (2012, June). “At the seams: DIYbio and opportunities for HCI.” In Proceedings of the Designing Interactive Systems Conference (pp. 258-267).
11. Liliana Meira, Joana Costa, Caterina Villa, Fernando Ramos, and Isabel Mafra, “EvaGreen real-time PCR to determine horse meat adulteration in processed foods.” Food Science and Technology, Volume 75, January 2017, pages 408-416
12. Mark W Newman, Debra Lauterbach, Sean A Munson, Paul Resnick, and Margaret E Morris, “It's not that i don't have problems, i'm just not putting them on facebook: challenges and opportunities in using online social networks for health.” Proceedings of the ACM 2011 conference on Computer supported cooperative work, 2011. 341-350.