Editorial

Gene Wiki Reviews—Raising the quality and accessibility of information about the human genome

CrossMark

Ginger Tsueng [a], Benjamin M. Good [a], Peipei Ping [b], Erica Golemis [c], Israel Hanukoglu [d], Andre J. van Wijnen [e], Andrew I. Su [a]

[a] The Scripps Research Institute, La Jolla, CA, USA
[b] University of California, Los Angeles (UCLA), Los Angeles, CA, USA
[c] Fox Chase Cancer Center, Philadelphia, PA, USA
[d] Laboratory of Cell Biology, Ariel University, Ariel, Israel
[e] Mayo Clinic, Rochester, MN, USA

ARTICLE INFO

Available online 2 May 2016

ABSTRACT

Wikipedia and other openly available resources are increasingly becoming commonly used sources of information not just among the lay public but even in academic circles including undergraduate students and postgraduate trainees. To enhance the quality of the Wikipedia articles, in 2013, we initiated the Gene Wiki Reviews on genes and proteins as a series of invited reviews that stipulated editing the corresponding Wikipedia article(s) that would be also subject to peer-review. Thus, while the review article serves as an authoritative snapshot of the field, the "living article" can continue to evolve with the crowdsourcing model of Wikipedia. After publication of over 50 articles, we surveyed the authors to assess the impact of the project. The author survey results revealed that the Gene Wiki project is achieving its major objectives to increase the involvement of scientists in authoring Wikipedia articles and to enhance the quantity and quality of the information about genes and their protein products. Thus, the dual publication model introduced in the Gene Wiki Reviews series represents a valuable innovation in scientific publishing and biomedical knowledge management. We invite experts on specific genes to contact the editors to take part in this project to enhance the quality and accessibility of information about the human genome.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The Human Genome Project catalyzed a revolution in terms of access to genomic information, a revolution that now includes public and private efforts to sequence millions of human genomes. While widespread availability of genomic sequence data has enormous potential, this potential is somewhat hindered by our knowledge of each individual gene. Without an accompanying basic "manual" on the structure, function, regulation, and genotype–phenotype relationships of each gene, this information remains mostly inaccessible and unintelligible for the great majority of academic users and society at large.

We are rapidly approaching the age of personalized medicine when individuals will have their genomic information as part of their medical file. This coming of age of precision medicine based on personal genomic information will greatly increase the demand for basic genetic information open to public access. Academic investigators and medical geneticists in the clinic who focus on genes associated with rare diseases are already observing this rush for informed expert knowledge. Specific requests for biomedical knowledge consultation are already fielded by

relatives of patients who are bypassing the traditional route of primary care physician to reach experts from around the world by email.

The enhanced penetration of the open access model in the publishing industry is increasing the accessibility of the scientific articles, yet, these still remain mostly as articles that are difficult to read by (and are often not intended for) lay audiences.

In the age of information revolution, Wikipedia has emerged as the most widely accessed freely available encyclopedia. Recent surveys reveal that Wikipedia articles on biomedical subjects are among the most commonly used educational resources not by just the lay public, but medical students and residents as well (Choi-Lundberg et al., 2016; Herbert et al., 2015; Torous et al., 2015). The WikiProject Medicine group has called Wikipedia "a key tool for global public health promotion" (Heilman et al., 2011).

The Gene Wiki is a growing, informal collection of over 11,000 gene-centric articles in Wikipedia. This initiative was started in 2008, with the aim of having a "continuously updated, community-reviewed, and collaboratively written review article for every human gene" (Huss et al., 2008). Basic articles (called "stubs") were systematically created

based on content extracted from structured databases. These stubs are then edited by the broader Wikipedia community, while "bots" keep the structured content in sync with the source databases.

As of December 2015, the Gene Wiki articles collectively reached 91,609,043 bytes in size and are viewed in aggregate almost five million times per month. The most popular Gene Wiki articles, such as Oxytocin and Insulin, have reached over 55,000 monthly page views (Wikipedia article traffic statistics, 2016). Anecdotally, users (both viewers and editors) come from diverse segments of the society, including students, lay public, and active researchers.

Academic researchers do not typically participate in systematic editing of Wikipedia pages. With few exceptions (Davis, 2016), the editing of Wikipedia pages does not lead to any tangible advance of academic careers. Without offering a means of rewarding researchers for contributing their expertise to Wikipedia, most academics seem unlikely to join the effort.

Seeing this information/knowledge gap, the Gene Wiki project partnered with *GENE* and created the Gene Wiki Review series in (Su et al., 2013). The Gene Wiki Review series utilized a dual publication model in which authors would submit two gene-specific review articles—one peer-reviewed "article of record" that is published in Gene, and one "living article" published in Wikipedia that can continuously evolve.

In this project, we are pursuing several objectives:

• to provide an incentive to academic experts to update the Wikipedia pages on the gene(s) that are closest to their expertise;
• to enhance the amount and quality of the articles in Wikipedia;
• to contribute to closing the knowledge gap between the human genome project and publicly available reliable information about human genes.

Here, we present an update on this partnership between *GENE* and Gene Wiki. Since its inception In 2013, the Gene Wiki Review series has steadily grown and now encompasses more than 40 published papers that cover more than 50 genes and Gene Wiki articles. Many additional manuscripts are currently under revision or review, or in preparation. To verify that the partnership between *GENE* and Gene Wiki serves its intended purpose as an incentive for researchers and the community at large to contribute to Wikipedia, we surveyed participating authors.

2. Results

Many corresponding authors (96%) updated the corresponding Wikipedia article or had someone in their research lab perform the update for them. Of this group of authors who participated in updating the Wikipedia page, almost all of them (94%) had never updated an entry in Wikipedia or any other wiki site previously, and few (12%) of those first-time Wikipedia editors had previously considered editing a scientific Wikipedia article prior to the invitation. Most authors cited lack of time/opportunity, knowledge (who can edit, how to edit), or motivation (activation energy, recognition) as reasons regarding why they had not considered editing a scientific Wikipedia article. For most of the authors (82%), the dual publication model was important toward the decision to update a Gene Wiki article (Fig. 1). Furthermore, many authors found the actual process of editing a Wikipedia article to be less difficult than they had expected (Fig. 2).

One of the goals of the Gene Wiki project is to enhance the participation of scientists in routine editing and updating of the Wikipedia entries (or articles), even after publication of their Gene review. To examine the impact of the Gene Wiki series, we surveyed the authors' attitudes towards future involvement in Wikipedia editing. Half (52%) of the responding authors noted that they are open to editing (or having someone in their lab edit) Wikipedia articles in the future. A third (33%)
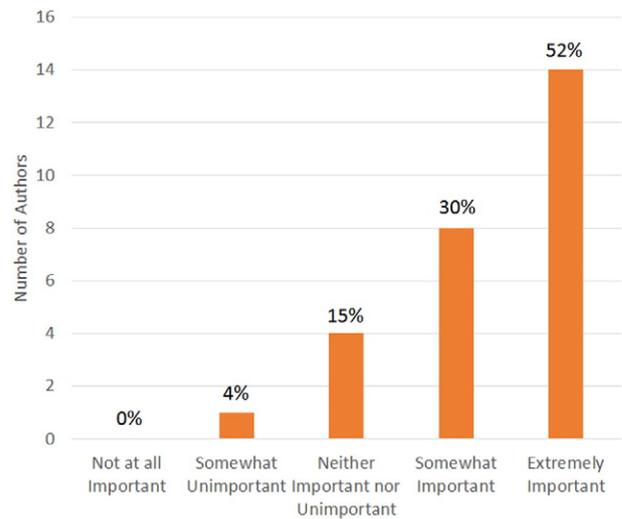

Fig. 1. Importance of the dual publication offer from GENE/Gene Wiki Gene Wiki as a factor in deciding to accept the invitation to improve a gene-specific Wikipedia page, on a scale of from 1 ("not at all important") to 5 ("very important").

of the respondents noted that they would be again involved in editing articles only if the dual publication model was offered. Only 15% of the respondents noted that they would not edit Wikipedia pages in the future. With regards to whether or not contributing an article to the Gene Wiki review series affected the author's perception of Wikipedia as a resource for scientific information: 8% of responding authors indicated they were more skeptical, 46% indicated their perception had not changed, and 47% indicated they had more confidence in the scientific information on Wikipedia.

The genes for the Gene Wiki series were prioritized based on a number of factors. Highest priority was given to genes that were already well-studied in the biomedical literature, and whose Wikipedia articles were both small in size and frequently viewed. For selected genes, we invited authors based on a semi-automated ranking of the number of published articles about the gene of interest.

We also explored the creation of a thematic Gene Wiki Review series focused on cardiovascular diseases. For this 'Cardiac Gene Wiki Review Series', we partnered with domain experts both to identify genes of importance in cardiovascular disease and to identify qualified scientists to invite as authors. As of December 2015, 16 manuscripts covering 19 cardiovascular disease-related genes were accepted for publication, and 6 manuscripts covering 9 genes were in the review process. Based on the early success of the Cardiac Gene Wiki Reviews, we believe that
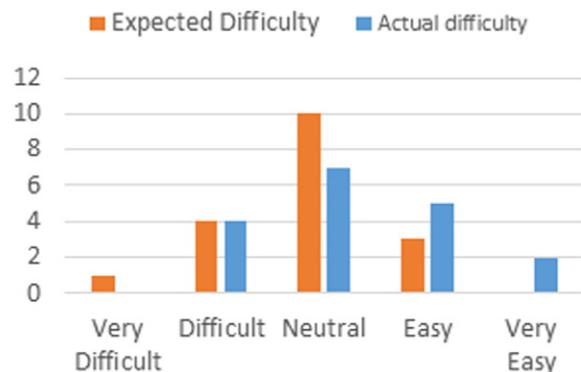

Fig. 2. Editing a Wikipedia article turned out to be less difficult than many authors expected.
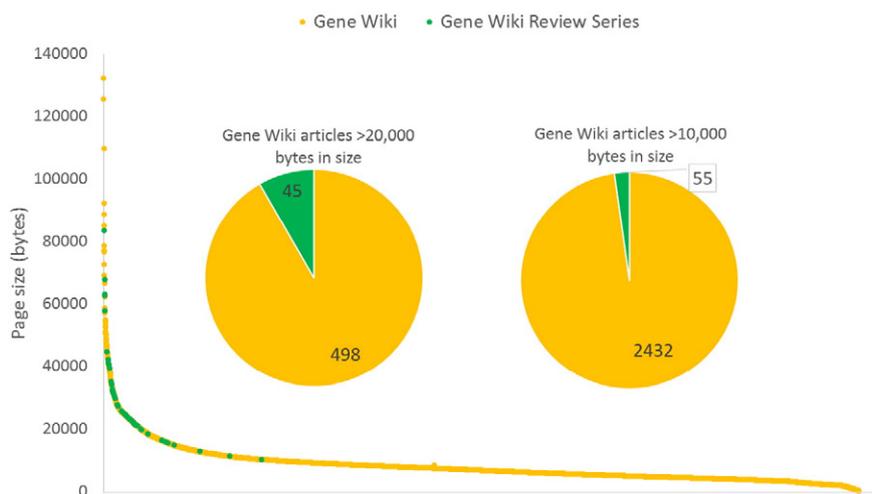
Fig. 3. - Impact of the Gene Wiki Review Series on the Gene Wiki. 11,022 Gene Wiki pages were sorted by page size and plotted with the cumulative number of gene wiki articles on the x-axis.

this is a generalizable pattern that can be applied to other diseases or areas of biology given an appropriate domain expert as a partner.

In addition to introducing experts to improve scientific content in Wikipedia, we hoped that the Gene Wiki review series of articles published in Gene would serve the scientific community. In aggregate, these journal articles have been viewed/downloaded more than 30,000 times between January 2015 and January 2016. Articles that are at least 1 year old have been downloaded on average more than 1000 times over the same period, while articles less than a year old have been downloaded more than 500 times on average. A recently published review on ENaC genes (Hanukoglu and Hanukoglu, 2016) has already been referenced in the HUGO Gene Nomenclature Committee (HGNC) reports for those genes ("Sodium Channels Epithelial (SCNN) Gene Family | HUGO Gene Nomenclature Committee, 2016").

To investigate the impact of the Gene Wiki Review series on the Gene Wiki itself, we compared the sizes and page views of Gene Wiki articles for genes in the Gene Wiki Review series with the rest of Gene Wiki. At 55 genes, the Gene Wiki pages updated via the Gene Wiki review series accounted for less than 0.5% of the estimated 11,022 gene pages in Wikipedia. However, articles edited as a part of this series made up 2% of the nearly 2500 Gene Wiki articles over 10,000 bytes in size and accounted for about 8% of the 543 Gene Wiki articles exceeding 20,000 bytes in size (Fig. 3). Articles updated as part of the Gene Wiki review series were edited about eight times more on average, primarily due to other Wikipedia editors who were helping to format and improve the pages updated by our authors (many of whom had little to no prior experience updating a Wikipedia article).

Improvements to the series since its initial introduction include linking the corresponding Gene Wiki article in an attempt to credit the authors, acknowledging the Gene Wiki project grant in order to ensure the deposition of all articles in this series into Pubmed Central, acknowledging GENE, and the creation of a virtual special issue for this series at http://www.sciencedirect.com/science/journal/03781119/vsi.

3. Conclusions

The author survey results shown above reveal that the Gene Wiki project is achieving its major objectives to increase the involvement of scientists in authoring Wikipedia articles and to enhance the quantity and quality of the information about genes and their protein products. Thus, the dual publication model introduced in the Gene Wiki Reviews series represents a valuable innovation in scientific publishing and biomedical knowledge management.

Although the use of Wikipedia pages as an educational resource is rapidly growing, the reliability of these pages is being questioned in some academic circles (Egle et al., 2015). Even an active biomedical Wikipedian has emphatically noted that Wikipedia entries should not be cited as a source of information as these represent a summary of other sources (Rasberry, 2014). The dual publication model we instituted provides an in-depth and well-referenced scientific article that can be cited as a reliable peer-reviewed reference that is linked to each respective Wikipedia page.

Having covered less than 0.3% of human genes so far, we are looking forward to a (very) long and productive future for this initiative. We invite scientists to participate as authors and as Wikipedia editors in this important effort to render genomic information more accessible to both researchers and society at large. If you would like to submit a review to this series, please contact one of the editors. If you would like to contribute immediately to this effort, just find the Wikipedia article for your gene of interest and start editing.

References

Choi-Lundberg, D.L., Low, T.F., Patman, P., Turner, P., Sinha, S.N., 2016. Medical student preferences for self-directed study resources in gross anatomy. Anat. Sci. Educ. 9, 150–160. http://dx.doi.org/10.1002/ase.1549.

Davis, L., 2016. Tenure awarded based in part on Wikipedia contributions – Wikimedia blog. [online]. (Blog.wikimedia.org. Available at: http://blog.wikimedia.org/2011/04/06/tenure-awarded-based-in-part-on-wikipedia-contributions/ Accessed 23 Mar. 2016).

Egle, J.P., Smeenge, D.M., Kassem, K.M., Mittal, V.K., 2015. The Internet School of Medicine: use of electronic resources by medical trainees and the reliability of those resources. J. Surg. Educ. 72, 316–320. http://dx.doi.org/10.1016/j.jsurg.2014.08.005.

Hanukoglu, I., Hanukoglu, A., 2016 Apr 1. Epithelial sodium channel (ENaC) family: phylogeny, structure–function, tissue distribution, and associated inherited diseases. Gene 579 (2), 95–132. http://dx.doi.org/10.1016/j.gene.2015.12.061 (Epub 2016 Jan 7.Review. PubMed PMID: 26772908; PubMed Central PMCID: PMC4756657).

Heilman, J.M., Kemmann, E., Bonert, M., Chatterjee, A., Ragar, B., Beards, G.M., Iberri, D.J., Harvey, M., Thomas, B., Stomp, W., Martone, M.F., Lodge, D.J., Vondracek, A., de Wolff, J.F., Liber, C., Grover, S.C., Vickers, T.J., Meskó, B., Laurent, M.R., 2011. Wikipedia: a key tool for global public health promotion. J. Med. Internet Res. 13, e14. http://dx.doi.org/10.2196/jmir.1589.

Herbert, V.G., Frings, A., Rehatschek, H., Richard, G., Leithner, A., 2015. Wikipedia—challenges and new horizons in enhancing medical education. BMC Med. Educ. 15, 32. http://dx.doi.org/10.1186/s12909-015-0309-2.

Huss III, J.W., Orozco, C., Goodale, J., Wu, C., Batalov, S., Vickers, T.J., Valafar, F., Su, A.I., 2008 Jul 8. A gene wiki for community annotation of gene function. PLoS Biol. 6 (7), e175. http://dx.doi.org/10.1371/journal.pbio.0060175 (PubMed PMID: 18613750; PubMed Central PMCID: PMC2443188).

Rasberry, L, 2014. Citing wikipedia. Don't Do It—Wikipedians Wouldn't BMJ. 348, p. g1819. http://dx.doi.org/10.1136/bmj.g1819.

Sodium channels epithelial (SCNN) Gene Family | HUGO Gene Nomenclature Committee, 2016. Genenames.org. Retrieved 22 March 2016, from http://www.genenames.org/cgi-bin/genefamilies/set/185.

Su, A.I., Good, B.M., van Wijnen, A.J., 2013 Dec 1. Gene Wiki Reviews: Marrying Crowdsourcing with Traditional Peer Review. Gene 531 (2), 125. http://dx.doi.org/10.1016/j.gene.2013.08.093 (Epub 2013 Sep 5. PubMed PMID: 24012870).

Torous, J., Franzan, J., O'Connor, R., Mathew, I., Keshavan, M., Kitts, R., Boland, R., 2015. Psychiatry Residents' Use of Educational Websites: A Pilot Survey Study. Acad. Psychiatry 39, 630–633. http://dx.doi.org/10.1007/s40596-015-0335-8.

Wikipedia article traffic statistics, 2016. *Stats.grok.se*. Retrieved 27 January 2016, from http://stats.grok.se/.