Systematic Review and Meta-Analysis

Medicine

OPEN

Using the Sankey diagram to visualize article features on the topics of whole-exome sequencing (WES) and whole-genome sequencing (WGS) since 2012

Bibliometric analysis

Meng-Ju Li, MD[a,b], Tsair-Wei Chien, MBA[c], Kuang-Wen Liao, PhD[b,d], Feng-Jie Lai, MD, PhD[e,f,*]

Abstract

Background: Sequencing technologies, such as whole-exome sequencing (WES) and whole-genome sequencing (WGS), have been increasingly applied to medical research in recent years. Which countries, journals, and institutes (called entities) contributed most to the fields (WES/WGS) remains unknown. Temporal bar graphs (TBGs) are frequently used in trend analysis of publications. However, how to draw the TBG on the Sankey diagram is not well understood in bibliometrics. We thus aimed to investigate the evolution of article entities in the WES/WGS fields using publication-based TBGs and compare the individual research achievements (IRAs) among entities.

Methods: A total of 3599 abstracts downloaded from icite analysis were matched to entities, including article identity numbers, citations, publication years, journals, affiliated countries/regions of origin, and medical subject headings (MeSH terms) in PubMed on March 12, 2022. The relative citation ratio (RCR) was extracted from icite analysis to compute the hT index (denoting the IRA, taking both publications and citations into account) for each entity in the years between 2012 and 2021. Three types of visualizations were applied to display the trends of publications (e.g., choropleth maps and the enhanced TBGs) and IRAs (e.g., the flowchart on the Sankey diagram) for article entities in WES/WGS.

Results: We observed that the 3 countries (the US, China, and the UK) occupied most articles in the WES/WGS fields since 2012, the 3 entities (i.e., top 5 journals, research institutes, and MeSH terms) were demonstrated on the enhanced TBGs, the top 2 MeSH terms were genetics and methods in WES and WGS, and the IRAs of 6 article entities with their hT-indices were succinctly and simultaneously displayed on a single Sankey diagram that was never launched in bibliographical studies.

Conclusion: The number of WES/WGS-related articles has dramatically increased since 2017. TBGs, particularly with hTs on the Sankey, are recommended for research on a topic (or in a discipline) to compare trends of publications and IRAs for entities in future bibliographical studies.

Abbreviations: BS = burst strength, IRA = individual research achievement, MeSH = medical subject heading, NICU = neonatal intensive care unit, RCR = relative citation ratio, TBG = temporal bar graphs, WES = whole exome sequencing, WGS = whole genome sequencing.

Keywords: choropleth map, hT index, research achievement, Sankey diagram, temporal bar graph, whole-exome sequencing, whole-genome sequencing

1. Introduction

The incorporation of sequencing techniques has been rapidly launched into health care,[1] partially attributed to the decreasing costs and clinical utility experience given to clinical settings and some to whole-exome sequencing (WES) and whole-genome sequencing (WGS), which are anticipated to replace conventional genetic tests, such as gene panels and chromosomal microarray analysis.[2]

The authors have no funding and conflicts of interest to disclose.

All data are publicly available in the PubMed library. The datasets generated during and/or analyzed during the current study are publicly available. All data used in this study are available in Supplemental Digital Content.

Supplemental Digital Content is available for this article.

[a] Department of Pediatrics, National Taiwan University Hsin-Chu Hospital, Hsinchu, Taiwan, [b] Department of Biological Science and Technology, College of Biological Science and Technology, National Yang Ming Chiao Tung University, Hsinchu, Taiwan, [c] Department of Medical Research, Chi-Mei Medical Center, Tainan, Taiwan, [d] Institute of Molecular Medicine and Bioengineering, National Yang Ming Chiao Tung University, Hsinchu, Taiwan, [e] Department of Dermatology, Chi Mei Medical Center, Tainan, Taiwan, [f] Center for General Education, Southern Taiwan University of Science and Technology, Tainan, Taiwan.

**Correspondence: Feng-Jie Lai, Chi-Mei Medical Center, 901 Chung Hwa Road, Yung Kung Dist., Tainan 710, Taiwan (e-mail: lai.fengjie@gmail.com).*

How to cite this article: Li M-J, Chien T-W, Liao K-W, Lai F-J. Using the Sankey diagram to visualize article features on the topics of whole-exome sequencing (WES) and whole-genome sequencing (WGS) since 2012: Bibliometric analysis. Medicine 2022;101:38(e30682).

Received: 1 April 2022 / Received in final form: 20 August 2022 / Accepted: 22 August 2022

http://dx.doi.org/10.1097/MD.0000000000030682

Key Points

- We introduced the merit of a Sankey diagram that combines the traditional temporal bar graphs (TBGs) used in bibliometrics.
- TBGs with hot spots and burst strengths were rarely seen in traditional TBGs.
- This paper focuses on the importance of research achievements, such as the impact factor, which cannot be determined solely by the number of publications, and shows how Sankey diagrams can be used and drawn for readers who are interested in bibliometric analysis.

1.1. Advanced techniques in WES and WGS

An increase in the application of sequencing techniques (either exome or genome) has significantly influenced clinical practices and pharmacogenetics. While the genetic disease is suspected in over 50% of the children who are admitted to neonatal intensive care units (NICUs) and die during the first year of life, diagnosis is confirmed in only 20% to 30% of cases, often postmortem.[3,4] As such, multiple factors hinder genetic diagnosis in neonates, including genetic heterogeneity. The disease tends to progress faster in neonates than in other stages of life.[1]

Technological advances in gene sequencing have enabled the rapid reading of any part of the genome at an affordable price.[5,6] NICUs are a key target for the implementation of genomic tools.[7,8] Timely and specific diagnosis of newborns can have critical implications for health and wellbeing for the remainder of an infant's life.[1]

In contrast to traditional molecular diagnostic methods, WES, especially trio-WES (both parents and their affected child sequenced simultaneously), was approved to be associated with a higher molecular diagnostic yield.[2] Broaden diagnostic yield is critical in diagnosing patients with atypical phenotypes, complex phenotypes, or disease-associated hallmark features not yet manifested. Nonetheless, WES does not typically detect intronic variants, trinucleotide repeat expansions, and methylation abnormalities, as well as limited detection of copy-number variants.[9,10] These limitations must be overcome through the WGS technique, which provides coverage of both array and exome targets and further coverage of nonexome regions of the genome.[11]

1.2. Research and publications related to WES and WGS

Unfortunately, as WGS is not as universally pervasive as WES, the reason might be the higher cost and more complex bioinformatics analysis from relevant professionals and qualified lab services required in clinical medicine. However, the issue of financial affordability can be overcome via technique advancement, while numerous WES/WGS-related articles have been published since 2016[12] to address the feasibility and promise of WGS in the foreseeable future.

Which article entities (e.g., publication years, journals, affiliated countries/regions/institutes of origin, document type, and medical subject headings, namely, MeSH terms) contributed most to the fields of WES/WGS is still unknown. We are thus motivated to conduct the bibliometric analysis of articles for a better understanding of the network characteristics of articles on WES/WGS.

1.3. Challenges encountered in bibliometrics

Bibliographic articles place more effort on identifying clusters of topics, literature gaps, and academic silos, as well as the most impactful authors and their research.[13] In contrast to meta-analysis and systematic literature review studies, bibliographic literature reviews use more quantitative and statistical methods to analyze feature characteristics, particularly on the trend of publications and the individual research achievements (IRAs) compared in groups.[14]

Although some studies[15–17] applied line charts or temporal bar graphs (TBGs) to display the trend of publications, the drawback is that there is no such essential information about the hot spots and burst strengths (BS) along the charts/graphs provided to readers.[18–20] Furthermore, how to easily and quickly draw the TBG (or the flowchart) has not been addressed in those studies.[18–20] We are thus going to demonstrate the way to draw the TBG using the Sankey flow diagram (Sankey).[21–23]

Additionally, the IRAs[24–28] should be denoted by bibliometric indices(e.g., the h-index,[29] x-index,[30] g-index,[31] or hT-index[32,33]), taking both publications and citations into account instead of the publications alone to represent the IRA (or contribution to academics), as in these studies.[18,27,34,35] Using the bibliometric indices (e.g., the hT-index) to represent the IRA on a single diagram for comparison among groups is required for development.

1.4. Aims of this study

We applied the Sankey to investigate the evolution of article entities in WES/WGS fields using publication-based TBGs and compare the IRAs among entities using the hT-indices shown on the Sankey.

2. Methods

2.1. Data sources

We programmed Microsoft Excel's Visual Basic for Applications modules to extract abstracts from icite analysis[36] as of March 12, 2022 and matched them to topical entities, such as article identity numbers, citations, publication years, affiliated countries/regions of origin, and medical subject headings (MeSH terms), in PubMed. The search string was defined as the following: (((((“whole”[All Fields] OR “wholeness”[All Fields] OR “wholes”[All Fields]) AND “genome”[MeSH Major Topic]) NOT ((“whole”[All Fields] OR “wholeness”[All Fields] OR “wholes”[All Fields]) AND “exome”[MeSH Major Topic])) AND “sequence analysis, dna”[MeSH Major Topic] AND “humans”[MeSH Terms]) OR ((“whole genome sequencing”[MeSH Major Topic] OT “whole exome sequencing”[MeSH Major Topic]) AND “humans”[MeSH Terms])) AND ((humans[Filter]) AND (2012:2021[pdat])). A total of 3599 abstracts were collected and deposited in Supplemental Digital Content 1, http://links.lww.com/MD/H378.

The relative citation ratios (RCRs) for each article were extracted from the icite analysis, implying that the RCR values measure the scientific influence of each paper by field- and time-adjusting the citations it has received and benchmarking to the median for National Institutes of Health publications.[37]

Ethical approval is not necessary for this study because all the data were obtained via the Internet.

2.2. Three parts divided in this study

Based on the study goals, 3 phrases (i.e., data arrangement in Section 2.1, data analysis, and representation in this section) were divided to achieve the study goals.

2.2.1 Productive countries of articles on WES/WGS. Traditionally, the geographical distribution of countries/regions in publications is applied to bibliographical studies.[24,30,38] In this study, we separately display the distribution of articles on WES and WGS for countries/regions using choropleth maps.

2.2.2. TBGs using the Sankey to draw. Traditionally, TBGs are displayed in rows with identical width bars, each one by one from top to bottom, even if hot spots and BS are marked on the TBGs.[18–20] The drawback of the traditional TBG is that no such information about data distribution has been displayed over the years. The enhanced TBG is thus proposed to combine the stacked bar chart with the TBG into the Sankey. Four visualizations of the enhanced TBG were made for displaying the comparisons of countries, journals, research institutes, and MeSH terms in publications on WES/WGS.

2.2.3. hTs for entity comparisons in IRAs on the Sankey. The hT index[32,33] was applied to evaluate the contributions to the fields of WES/WGS for each entity. The most prominent countries/regions with higher hT indices on the topics of WES and WGS were highlighted on the enhanced TBGs.

The hT index based on the RCR[37] is computed by equation 1, where j is the j-th article in descending order and n is the RCRs of an article for an entity. The residual weights in excess and tail citations are presented in equation 2.

$$\mathrm{h-index}_{(top\quad j)} = \frac{j}{2j-1}, \quad n_j \le j \tag{1}$$

$$\mathrm{hT}_{(j)} = h + residual \quad weights = \frac{j}{2j-1} + \sum_{i=1}^{n} \frac{1}{2i-1}, \quad n_j > j \tag{2}$$

$$hT = \sum_{j=1}^{N} hT_{(j)}, \tag{3}$$

Now, if an author, for example, has N cited articles with associated citations n1, n2, n3, …, nN (ranked in descending order), the hT score for any single paper ranked j in the list (with nj RCRs), denoted hT(j) in equation 2. The hT-index for the whole list of cited articles is then calculated by summing all weights (allocated in equations 1 and 2) through equation 3. As such, the resulting sum is the hT index. Readers are invited to practice it at the link[39] (e.g., input 10 to obtain 2.13 for a single article with 10 citations, 3.28 for 100 citations, 4.44 for 1000 citations, 5.59 for 10,000 citations, or a vector (e.g.,) RCRs composed of {6,4,2,2,1} in 5 articles to obtain hT = 4.03).

Social network analysis[40,41] was performed on cowords of entities (e.g., years, countries, institutes, journals, MeSH terms, and document types) in articles. The relation weights between entities were computed based on the 3599 articles. The relationship (denoted by arc) was applied to draw the Sankey for each neighbor entity (called a node in the diagram).

2.3. Tools used in this study

We applied the author-made modules in MS-Excel to arrange and analyze the data. The ways to draw the enhanced TBGs and the hTs comparison among entities using the SankeyNATUC[42] are linked in reference.[43]

The pages of hypertext markup language (HTML) used for Google Maps were created for choropleth maps. All relevant information was linked to dashboards on Google Maps. This study flowchart is shown in Figure 1.

3. Results

3.1. Geographical distribution of publications on WES/WGS

The 3 countries (the US, China, and the UK) have occupied most articles in the WES/WGS fields since 2012; see Figure 2.

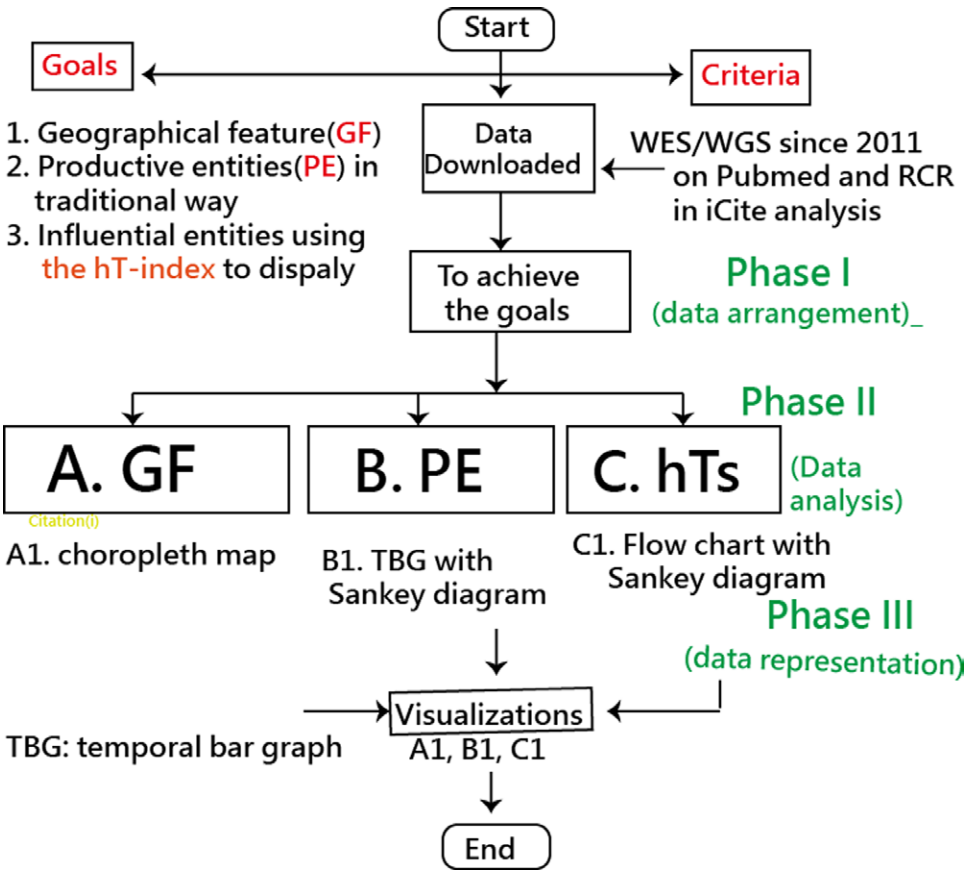

Figure 1. Study flowchart with three phases and three goals to achieve.

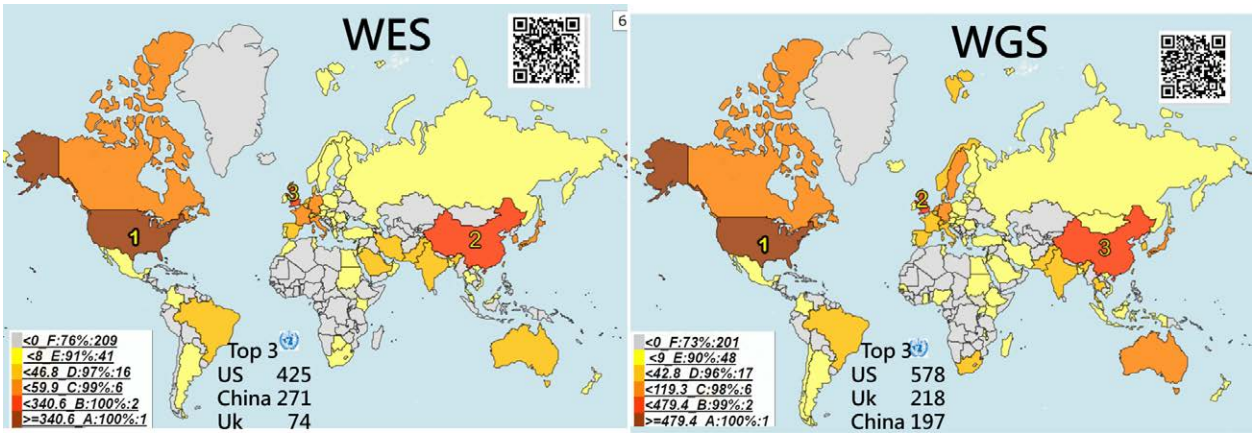

Figure 2. The most productive countries on WES/WGS since 2012 (Note. The darker the area is, the greater the number of publications). WES = whole exome sequencing, WGS = whole genome sequencing.

Readers are invited to scan the QR code in Figure 2 to examine the details about the information on the choropleth maps.

3.2. Enhanced TBGs showing the top 5 entities regarding publications on WES/WGS

The 5 countries with the most publications on WES/WGS are shown in Figure 3, including the stacked bar charts on the top and the traditional TBGs on the bottom. Viewing the bottom panel in Figure 3, we can see additional information about the hot spots (denoted by the black bar). The BS (represented by the bar width), and the trend determined by the last 4 years and colored by the trend types[19,20] for increasing, ready to rise, decreasing, and ready to decline using green, yellow, red, and light green colors, respectively, to denote.

The enhanced TBGs are presented in Figures 4–6 for journals, institutes, and MeSH terms, respectively. We can see that

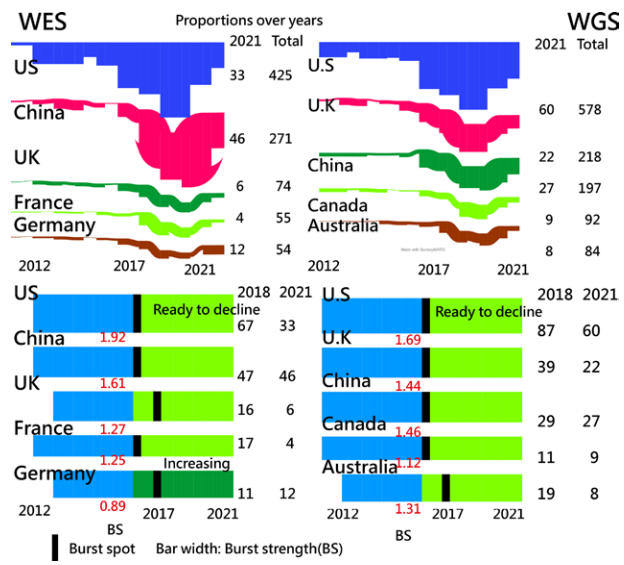

Figure 3. Top 5 productive countries in WES and WGS since 2012 (Note. Comparison of the Top 5 countries in publications between the topics of WES and WGS made in this diagram. In the bottom panel, the TBG tells us that the black bars represent the hot spots. The last block in color indicates the trend determined by the last 4 time points. The burst strength is determined by the accumulative number of cases at the beginning of the hot spot). TBG= temporal bar graphs, WES = whole exome sequencing, WGS = whole genome sequencing.

the most productive entities (either journals or institutes) have more proportions since 2016, with the feature of traditional stacked bar chart, distinct hot spots (with black areas), BS shown on the right-side hand, and trends with colors at the far right side.

The top 3 MeSH terms are {genetics, methods, diagnosis} and {genetics, methods, genome human} for WES and WGS, respectively. The top 5 terms each year are shown in Figure 6 using the Sankey technique, which is rarely seen in bibliographical studies.

3.3. hTs for entity comparisons in IRAs on the Sankey

The hTs for the top 3 entities are shown in Figure 7. One look is worth a thousand words and quite a few numbers, too.[44] The RAs of 6 article entities with their hT-indices are succinctly and simultaneously displayed on a single Sankey diagram that was never launched in bibliographical studies. Until now, such a picture – the Sankey – we perceived is worth an effort by piecing them into a drawing program,[45] as we demonstrated in Supplemental Digital Content 2, http://links.lww.com/MD/H379.

Additionally, the top 3 elements in each entity are distinctly different from the productive ones in Figures 3–6, particularly in journals (e.g., PLos One and Sci Rep vs E engl J Med and Nature, except Genet Med in WES). We confirmed that only hT (or other bibliometric indices) could be the IRA proxy in lieu of publications, as reported in many bibliographical studies.[24–28]

4. Discussion

We observed that the 3 countries (the US, China, and the UK) occupied most articles in the WES/WGS fields since 2012, the 3 entities (i.e., top 5 journals, research institutes, and MeSH terms) were demonstrated on the enhanced TBGs, the top 2 MeSH terms were genetics and methods in WES and WGS, and the RAs of 6 article entities with their hT-indices were succinctly and simultaneously displayed on a single Sankey diagram that stipulated that one look is worth a thousand words.

4.1. Additional information

Technological advances in gene sequencing have enabled the rapid reading of any part of the genome at an affordable price.[5,6] NICUs are a key target for the implementation of genomic tools.[7,8] Some studies focused on the use of these WES/WGS

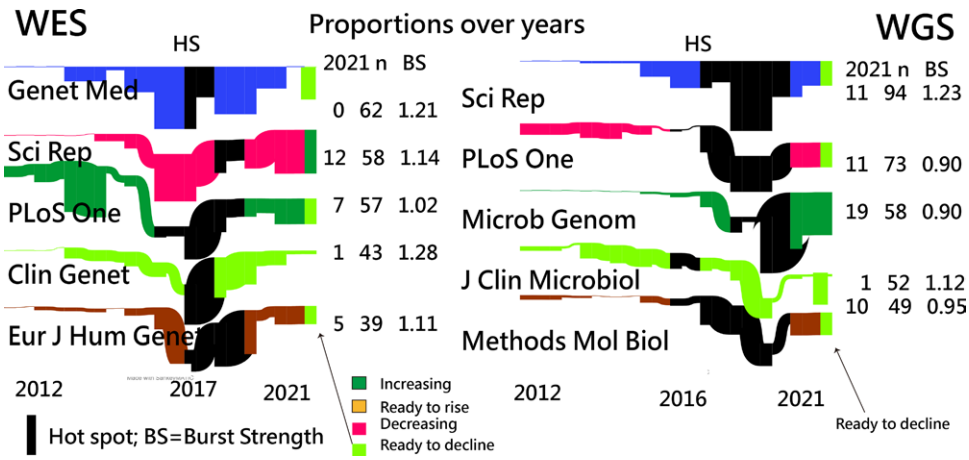

Figure 4. Top 5 productive journals in WES and WGS since 2012 (Note. Comparison of Top 5 journals in publications between topics of WES and WGS made in this diagram. The black bars represent the hot spots in the period. The last block in color indicates the trend determined by the last 4 time points. The burst strength is determined by the accumulative number of cases at the beginning of the hot spot. The length of hot spots is the number of observed cases not less than the number of cases at the hot spot). WES = whole exome sequencing, WGS = whole genome sequencing.

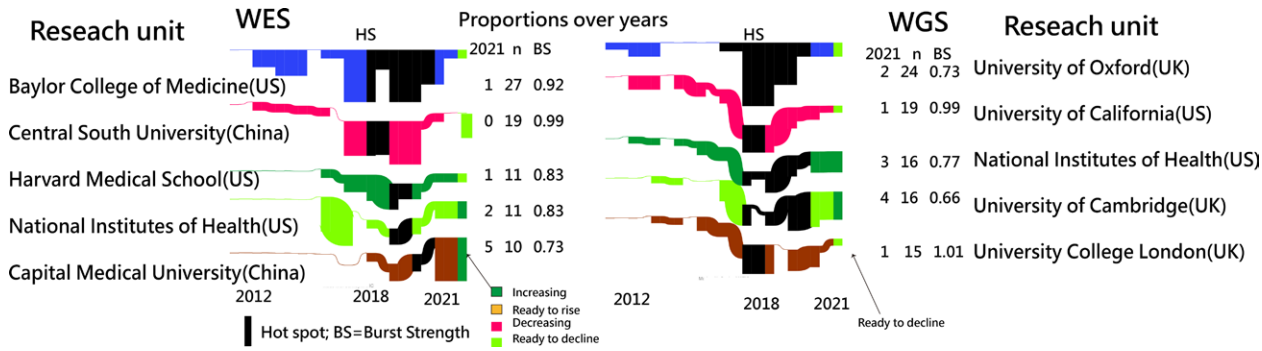

Figure 5. Top 5 productive research institutes in WES and WGS since 2012 (Note. Comparison of Top 5 institutes in publications between topics of WES and WGS made in this diagram. The black bars represent the hot spots in the period. The last block in color indicates the trend determined by the last 4 time points. The burst strength is determined by the accumulative number of cases at the beginning of the hot spot. The length of hot spots is the number of observed cases not less than the number of cases at the hot spot). WES = whole exome sequencing, WGS = whole genome sequencing.

technologies and conducted in NICU settings[46–57] have yielded highly promising results. Further advances are required to make precision, personalized, and predictive Medicine into reality.[1]

Although integration and interpretation of the data produced by genomic sequencing is a key obstacle to the incorporation of these strategies into clinical routine practice,[58,59] phenotype-based filtering and prioritization could greatly facilitate the interpretation of genetic variants detected by genome sequencing.[10,60]

There were 2 major objectives achieved: we applied an enhanced TBG to interpret article entity evolution in WES and WGS, and we could understand the IRAs compared between entities using the hT indices shown on the Sankey. Two studies[61,62] have applied the Sankey method to bibliometric analysis. However, no information on how to draw the Sankey is given to readers who may wish to utilize it in future relevant studies.

A demonstration and introduction to the 3 Sankey diagrams (e.g., Figs. 5–7) can be found at the following link[43] and Supplemental Digital Content 2, http://links.lww.com/MD/H379. The hT-index (or other bibliometric indicators) are suitable for measuring RAs due to the fact that they take both publications and citations into account. We used the hT index to calculate the RA for entities and demonstrate the Sankey method used in this study as the only index that takes into account all publications (see Eqs. 1 and 2).[32,33]

Several articles have been published in the literature related to WES/WGS. In spite of this, none of the studies conducted a bibliometric analysis of the topic. The top 3 countries (the US, the UK, and Germany) have higher hTs, which are reflected in the IRAs of their institutions, for example, Baylor College of Medicine (US), University of Oxford (UK), and Harvard Medical School (US).

4.2. Implications and changes

New genomic sequencing techniques have shown considerable promise in neonatology,[1] and they will likely increase diagnostic rates and reduce diagnosis times. The use of WGS for diagnosis of rare disorders in ill neonates in NICUs has been raised as an ethical concern, but it has the potential to augment or modify the care provided to this vulnerable population of patients.[10] In addition to the knowledge provided from this study regarding the most influential entities in articles on WES/WGS, there are several features that could be improved in the future.

First, the hT-index with decimal places can be used to complement the original h-index in order to improve the discrimination power for identifying the RAs and rankings of a given group.[14] We suggested using the hT index to present the IRAs in the Sankey and in future bibliographical studies.

The second feature is that the Sankey method was presented as a way of highlighting all-in-one entities in bibliometrics, and has proven to be viable and feasible.

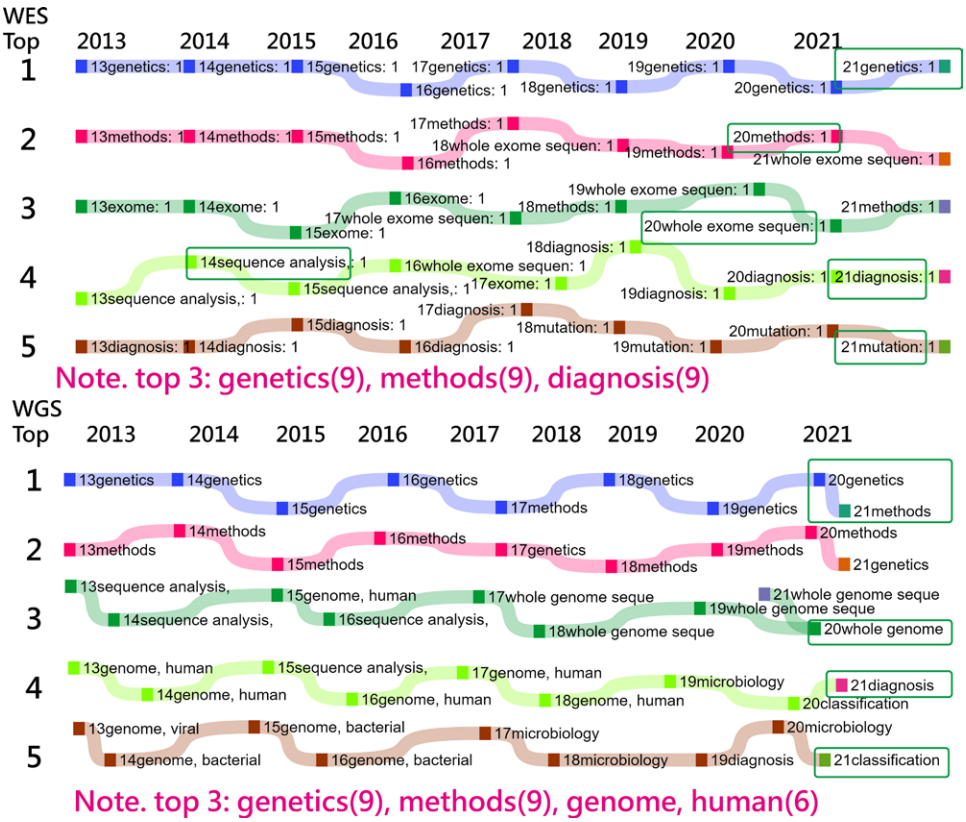

Figure 6. Top 3 observed MeSH terms in WES and WGS since 2012 (Note. Showing names on the diagram, we defined all elements column by column with the flow from 2013 to 2022 for MeSH terms with most frequently observed in years. Terms in 2012 were absent). MeSH = medical subject heading, WES = whole exome sequencing, WGS = whole genome sequencing.

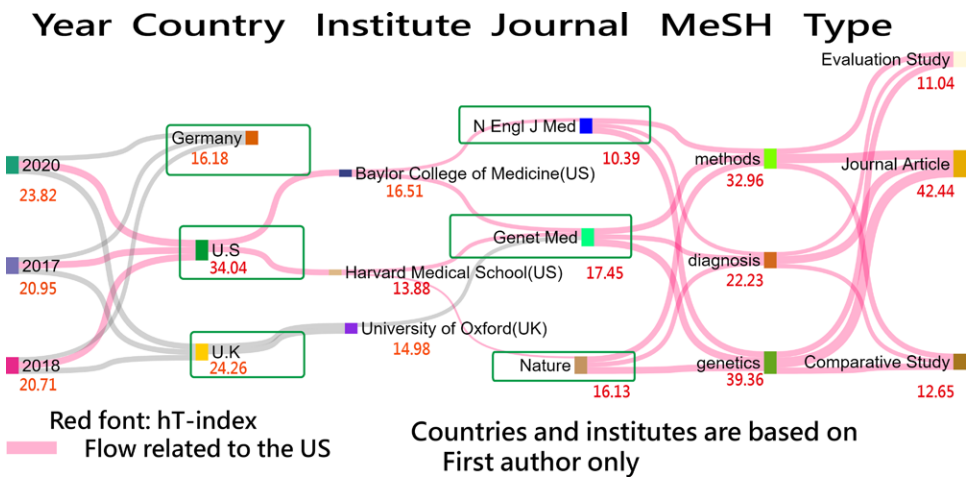

Figure 7. Top 3 article entities with hT indices in WES and WGS since 2012 (note. Red curves linked to all entities related to the US; Germany has nothing to link due to no any institute related to Germany in these 3 research institutes; the highest hT is onto the journal articles due to the number of articles more than other entities). WES = whole exome sequencing, WGS = whole genome sequencing.

As a third feature, readers can access teaching material at the link[43] and Supplemental Digital Content 2, http://links.lww.com/MD/H379 which will assist them in replicating the relevant studies in the future.

Fourth, some scholars pointed out that *while Sankey diagrams are better known, alluvial plots are generally a good deal easier to generate*.[63] After viewing the teaching materials[44] and using the software[42] to draw the Sankey diagrams, it is anticipated that more research will apply the Sankey (or alluvial) diagrams to visualize (or dance) their data in the future.

Additionally, citations have been replaced with RCRs.[37] The reason for this is that the RCRs do not depend on the age of the article: older articles earn more citations than younger articles. Based on normalized years and disciplines, RCRs are considered a fair indicator for comparing article impacts.

There is a dedicated software program that can solve the problem of calculating the hT-index even though it is more complex than the h-index in computation. Here is a link[39] that provides the programming codes to understand how the hT-index is calculated within a second.

4.3. Limitations and suggestions

Several issues should be considered thoroughly in further studies. The first concern is that the software to draw the Sankey is not unique and inevitable. Other software[64] easily draws the Sankey online. However, we have not yet had any experience in comparison thus far.

Second, the dashboards (e.g., the choropleth maps shown in Fig. 2) are displayed on Google Maps. Since Google Maps requires a paid project key for use of the cloud platform, these installments are not free of charge. Therefore, it is difficult for other authors to replicate the use within a short period of time.

Third, the calculation of the hT-index based on the summation of weights in the Ferrers tableau (i.e., all the cited papers in the list) requires a substantial amount of time. By using advanced hardware, the time-consuming job is now trivial, approximately equivalent to computing other bibliometric indices using dedicated software.

Finally, even though most WES/WGS articles were retrieved from PubMed, the results were very different when articles were retrieved from other databases (e.g., Google Scholar, Scopus, Web of Science). It will be necessary to extract articles from more bibliometric databases in future studies.

5. Conclusion

The use of the Sankey diagram to display network characteristics of articles related to WES/WGS led to a breakthrough. With an all-in-one feature involving frequently used entities with hT indices on a Sankey diagram, article space could be reduced, and readers could focus on the fewer influential entities that are important to bibliometrics. In future studies, the Sankey diagram can be applied to bibliometrics on other topics in addition to WES/WGS.

Acknowledgments

We thank Enago (www.enago.tw) for English language revision.

Author contributions

MJ developed the study. TW and KW analyzed the data. F.J.L monitored the process of this study and helped respond to the reviewers' advice and comments. MJ and TW drafted the manuscript, and all authors provided critical revisions for important intellectual content. The study was supervised by F.J.L. All authors have read and agreed to the published version of the manuscript.
Conceptualization: Meng-Ju Li.
Data curation: Kuang-Wen Liao.
Investigation: Feng-Jie Lai.
Methodology: Tsair-Wei Chien.

References

[1] de Castro MJ, González-Vioque E, Barbosa-Gouveia S, et al. Rapid phenotype-driven gene sequencing with the NeoSeq panel: a diagnostic tool for critically ill newborns with suspected genetic disease. J Clin Med. 2020;9:2362.
[2] Lee H, Deignan JL, Dorrani N, et al. Clinical exome sequencing for genetic identification of rare Mendelian disorders. JAMA. 2014;312:1880–7.
[3] Wojcik MH, Schwartz TS, Thiele KE, et al. Infant mortality: the contribution of genetic disorders. J Perinatol. 2019;39:1611–9.
[4] Wojcik MH, Schwartz TS, Yamin I, et al. Genetic disorders and mortality in infancy and early childhood: delayed diagnoses and missed opportunities. Genet Med. 2018;20:1396–404.
[5] Reuter JA, Spacek DV, Snyder MP. High-throughput sequencing technologies. Mol Cell. 2015;58:586–97.
[6] Gaff CL, Winship IM, Hansen DF, et al. Preparing for genomic medicine: a real world demonstration of health system change. NPJ Genom Med. 2017;2:16.
[7] Holm IA, Agrawal PB, Ceyhan-Birsoy O, et al. The BabySeq project: Implementing genomic sequencing in newborns. BMC Pediatr. 2018;18:225.
[8] Friedman JM, Bombard Y, Cornel MC, et al. Genome-wide sequencing in acutely ill infants: Genomic Medicine's critical application? Genet Med. 2019;21:498–504.
[9] Rafehi H, Szmulewicz DJ, Bennett MF, et al. Bioinformatics-based identification of expanded repeats: a nonreference intronic pentamer expansion in RFC1 causes CANVAS. Am J Hum Genet. 2019;105:151–65.
[10] Smith LD, Willig LK, Kingsmore SF. Whole-exome sequencing and whole-genome sequencing in critically ill neonates suspected to have single-gene disorders. Cold Spring Harb Perspect Med. 2015;6:a023168.
[11] Manickam K, McClain MR, Demmer LA, et al. Exome and genome sequencing for pediatric patients with congenital anomalies or intellectual disability: an evidence-based clinical guideline of the American College of Medical Genetics and Genomics (ACMG). Genet Med. 2021;23:2029–37.
[12] Pubmed. Over 1325 article on WES/WGS published in Pubmed since 2016. Available at https://pubmed.ncbi.nlm.nih.gov/?term=%28Whole-Exome+Sequencing%5BMeSH+Major+Topic%5D%29+AND+%28Whole-Genome+Sequencing%5BMeSH+Major+Topic%5D%29&sort=date. [Access date April 4, 2022].
[13] Block JH, Fisch C. Eight tips and questions for your bibliographic study in business and management research. Manag Rev Q. 2020.
[14] Wu JW, Chien TW, Tsai YC, et al. Using the forest plot to compare citation achievements in bibliographic and meta-analysis studies since 2011 using data on PubMed Central: a retrospective study. Med (Baltimore). 2022; in print.
[15] Barnett A, Doubleday Z. The growth of acronyms in the scientific literature. Elife. 2020;9:e60080.
[16] Barnett A, Doubleday Z. Demonstrating the ascendancy of COVID-19 research using acronyms. Scientometrics. 2021;126:6127–30.
[17] Kuo SC, Chien TW, Chou W. Questions to the article: demonstrating the ascendancy of COVID-19 research using acronyms. Scientometrics. 2021;126:8761–4.
[18] Shen L, Xiong B, Li W, Lan F, Evans R, Zhang W. Visualizing collaboration characteristics and topic burst on international mobile health research: bibliometric analysis. JMIR Mhealth Uhealth. 2018;6:e135.
[19] Chow JC, Chien TW, Chou W. Suggestions to the article: demonstrating the ascendancy of COVID-19 research using acronyms. Scientometrics. 2022;127:2897–9.
[20] Ho SY, Chien TW, Shao Y, et al. Visualizing the features of inflection point shown on a temporal bar graph using the data of COVID-19 pandemic. Medicine (Baltim). 2022;101:e28749.
[21] Wang C, Lv T, Deng X. Bibliometric and visualized analysis of China's smart grid research 2008-2018. Front Res Metr Anal. 2020;5:551147.
[22] Lamer A, Laurent G, Pelayo S, El Amrani M, Chazard E, Marcilly R. Exploring patient path through sankey iagram: a proof of concept. Stud Health Technol Inform. 2020;270:218–22.
[23] Yu B, Silva CT. VisFlow - Web-based visualization framework for tabular data with a subset flow model. IEEE Trans Vis Comput Graph. 2017;23:251–60.
[24] Yeh YT, Chien TW, Kan WC, et al. The use of hx-index to compare research achievements for ophthalmology authors in Mainland China, Hong Kong, and Taiwan since 2010. Medicine (Baltim). 2021;100:e24868.
[25] Kung SC, Chien TW, Yeh YT, et al. Using the bootstrapping method to verify whether hospital physicians have different h-indices regarding individual research achievement: a bibliometric analysis. Medicine (Baltim). 2020;99:e21552.
[26] Lin CH, Chou PH, Chou W, Chien T-W. Using the Kano model to display the most cited authors and affiliated countries in schizophrenia research. Schizophr Res. 2020;216:422–8.
[27] Hsieh WT, Chien TW, Kuo SC, et al. Whether productive authors using the national health insurance database also achieve higher individual research metrics: a bibliometric study. Medicine (Baltim). 2020;99:e18631.
[28] Chien TW, Wang HY, Chang Y, Kan W-C. Using Google Maps to display the pattern of coauthor collaborations on the topic of schizophrenia: a systematic review between 1937 and 2017. Schizophr Res. 2019;204:206–13.

[29] Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl Acad Sci USA. 2005;102:16569–72.
[30] Fenner T, Harris M, Levene M, Bar-Ilan J. A novel bibliometric index with a simple geometric interpretation. PLoS One. 2018;13:e0200098.
[31] Egghe L. Theory and practice of the g-index. Scientometrics. 2006;69:131–52.
[32] Anderson TR, Hankin RKS, Killworth PD. Beyond the Durfee square: enhancing the h-index to score total publication output. Scientometrics. 2008;76:577–88.
[33] Hua PH, Wan JK, Wu JH. A perfect hirsch-type index? Experiences using the tapered h-Index (hT). Chin J Sci Tech Periodicals. 2010;21:33–7.
[34] Chien TW, Chang Y, Wang HY. Understanding the productive author who published papers in Medicine using National Health Insurance Database: a systematic review and meta-analysis. Medicine (Baltim). 2018;97:e9967.
[35] Liu MY, Chou W, Chien TW, et al. Evaluating the research domain and achievement for a productive researcher who published 114 sole-author articles: a bibliometric analysis. Medicine (Baltim). 2020;99:e20334.
[36] NIH. ICite tool. Available at: https://icite.od.nih.gov/analysis. [Access date January 7, 2022].
[37] Hutchins BI, Yuan X, Anderson JM, Santangelo GM. Relative citation ratio (RCR): a new metric that uses citation rates to measure influence at the article level. PLoS Biol. 2016;14:e1002541.
[38] Yie KY, Chien TW, Chen CH, et al. Suitability of h- and x-indices for evaluating authors' individual research achievements in a given short period of years: a bibliometric analysis. Medicine (Baltim). 2021;100:e25016.
[39] Chien TW. The computation of hT-index on the website. Available at: http://www.healthup.org.tw/kpiall/hTindexurl.asp. [Access date March 3, 2022].
[40] Kan WC, Chou W, Chien TW, Yeh Y-T, Chou P-H. The most-cited authors who published papers n JMIR mHealth and uHealth using the authorship-weighted scheme: bibliometric analysis. JMIR Mhealth Uhealth. 2020;8:e11567.
[41] Yie KY, Chien TW, Yeh YT, Chou W, Su SB. Using social network analysis to identify spatiotemporal spread patterns of COVID-19 around the World: online dashboard development. Int J Environ Res Public Health. 2021;18:2461.
[42] SankeyNATUC. Software for drawing the sankey diagram. Available at: https://sankeymatic.com/build/. [Access date January 22, 2022].
[43] Chien TW. How to draw the Sankey diagram. Available at: https://youtu.be/GybTApNt1M0. [Access date March 22, 2022].
[44] Hepting DH. One look is worth a thousand words. Available at: http://www2.cs.uregina.ca/~hepting/projects/pictures-worth/1921-dec-08.html. [Access date March 24, 2022].
[45] Doka G. Sankey diagrams in excel. Available at: https://www.doka.ch/sankey24.htm. [Access date March 24, 2022].
[46] Daoud H, Luco SM, Li R, et al. Next-generation sequencing for diagnosis of rare diseases in the neonatal intensive care unit. Can Med Assoc J. 2016;188:E254–60.
[47] Meng L, Pammi M, Saronwala A, et al. Use of exome sequencing for infants in intensive care units: ascertainment of severe single-gene disorders and effect on medical management. JAMA Pediatr. 2017;171:e173438.
[48] Van Diemen CC, Kerstjens-Frederikse WS, Bergman KA, et al. targeted genomics in critically ill newborns. Pediatrie. 2017;140:e20162854.
[49] Petrikin JE, Cakici JA, Clark MM, et al. The NSIGHT1-randomized controlled trial: rapid whole-genome sequencing for accelerated etiologic diagnosis in critically ill infants. NPJ Genom Med. 2018;3:6.
[50] Stark Z, Tan TY, Chong B, et al. A prospective evaluation of whole-exome sequencing as a first-tier molecular test in infants with suspected monogenic disorders. Genet Med. 2016;18:1090–6.
[51] Brunelli L, Jenkins SM, Gudgeon JM, et al. Targeted gene panel sequencing for the rapid diagnosis of acutely ill infants. Mol Genet Genom Med. 2019;7:e00796.
[52] Ceyhan-Birsoy O, Murry JB, Machini K, et al. Interpretation of genomic sequencing results in healthy and ill newborns: results from the BabySeq project. Am J Hum Genet. 2019;104:76–93.
[53] Elliott AM, du Souich C, Lehman A, et al. RAPIDOMICS: rapid genome-wide sequencing in a neonatal intensive care unit-successes and challenges. Eur J Pediatr. 2019;178:1207–18.
[54] French CE, Disease NB, Delon I, et al. Whole genome sequencing reveals that genetic conditions are frequent in intensively ill children. Intensive Care Med. 2019;45:627–36.
[55] Kingsmore SF, Cakici JA, Clark MM, et al. A randomized, controlled trial of the analytic and diagnostic performance of singleton and trio, rapid genome and exome sequencing in ill infants. Am J Hum Genet. 2019;105:719–33.
[56] Wang H, Lu Y, Dong X, et al. Optimized trio genome sequencing (OTGS) as a first-tier genetic test in critically ill infants: practice in China. Hum Genet. 2020;139:473–82.
[57] Gubbels CS, VanNoy GE, Madden JA, et al. Prospective, phenotype-driven selection of critically ill neonates for rapid exome sequencing is associated with high diagnostic yield. Genet Med. 2020;22:736–44.
[58] Ross LF, Clayton EW. Ethical issues in newborn sequencing research: the case study of BabySeq. Pediatrics. 2019;144:e20191031.
[59] Yang H, Robinson PN, Wang K. Phenolyzer: phenotype-based prioritization of candidate genes for human diseases. Nat Methods. 2015;12:841–3.
[60] Wise AL, Manolio TA, Mensah GA, et al. Genomic medicine for undiagnosed diseases. Lancet 2019;394:533–40.
[61] Liu PC, Lu Y, Lin HH, et al. Classification and citation analysis of the 100 top-cited articles on adult spinal deformity since 2011: a bibliometric analysis. J Chin Med Assoc. 2022;85:401–8.
[62] Kuo YC, Chien TW, Kuo SC, et al. Predicting article citations using data of 100 top-cited publications in the journal Medicine since 2011: a bibliometric analysis. Medicine (Baltim). 2020;99:e22885.
[63] Peterson B. Alluvial plots vs Sankey diagrams. Available at: http://www.datasmith.org/2020/05/02/alluvial-plots-vs-sankey-diagrams/. [Access date March 20, 2022].
[64] Marystone M. 10 Best websites to create a Sankey chart. Available at: https://www.online-tech-tips.com/cool-websites/10-best-websites-to-create-a-sankey-chart/. [Access date April 3, 2022].