Research outputs:

A bibliometric analysis of Wellcome Sanger Institute research publications 2018-2022

Introduction

We use genome technologies to advance understanding of biology and improve health. Working at a scale unachievable for most other research institutions is a central tenet that underpins our research and means that the Institute can tackle some of the most difficult challenges in genomic research.

The 'Research Impact' section provides insight into the breadth and the scope of genomic research conducted at the Sanger Institute. Publication of research articles is the primary way the findings of our research are disseminated to the research community. In order to understand the impact and reach of our research outputs, we analyse a range of citation metrics (i.e., bibliometrics) of these publications over time.

This section provides an analysis of the research articles and reviews published by Sanger Institute scientists.

Background

Sanger Institute research outputs are monitored at an organisational, not individual, level using the following bibliometric indicators:

- Scholarly Output
- Field-Weighted Citation Impact (FWCI)
- Citations per Publication (mean average number of citations)
- Output in Top 5 per cent of Citation Percentiles (field-weighted)
- Output in Top 1 per cent of Citation Percentiles (field-weighted)
- Collaboration (including academic-corporate collaboration)

The choice of indicators was based on our commitment to promote the responsible use of metrics. Several principles underpin this commitment, such as using multiple indicators to analyse different aspects of bibliometric outputs and transparency in how the indicators are calculated and used. As signatories to the San Francisco Declaration on Research Assessment (DORA), the metrics used are not based on impact factor or journal name.

Our bibliometric analyses demonstrate that the Sanger Institute is a world-leading centre of genomics research. They also reveal that the foundational discoveries arising from research at the Sanger Institute have substantial impact within the wider scientific community.

Bibliometrics at a glance

Sanger's publication volume and citation rates are consistently among the highest in comparable fields, despite our comparatively small faculty size. The metrics reflect the Sanger Institute's global influence: research articles are highly cited and frequently referred to in the outputs of research conducted in countries worldwide.

Scholarly output:
Sanger Institute authors have published 565 research articles and reviews in 2022, including 185 articles where Sanger authors are leading the work (as recognised by sole or joint first or last authorship). During the five-year period between 2018 and 2022, 3,005 articles and reviews were published by Sanger Institute authors, including 892 where Sanger authors had a lead role.

Field-Weighted Citation Impact (FWCI):

- Articles and reviews published by Sanger Institute authors are consistently nearly four times more cited than the world average.
- Research articles and reviews published in the past five years (2018-2022) in which Sanger authors are leading the work are 4.64 times more cited than the world average.

Average citations per paper:

- The 3,005 Sanger articles and reviews published in 2018-2022 have accumulated 166,205 citations overall, with a mean of 55.3 citations per paper published during this period.
- First and/or last author Sanger papers have received an average of 65.8 citations per paper for papers published between 2018-2022.
- Sanger Institute papers published in 2022 received, on average, 14.2 citations per paper by September 2023.

Outputs in the top 5 per cent of citation percentiles (field-weighted):

- On average, 23.7 per cent of all Sanger Institute articles and reviews published between 2018-2022 are among the top 5 per cent of the world's most cited publications. Despite expected fluctuations owing to citations accruing over time, consistently between one in every four or five Sanger publications are among the world's top 5 per cent most cited publications.
- On average, 29 per cent of publications published between 2018-2022 in which Sanger authors are leading the research as first and/or last authors, are among the top 5 per cent of the world's most cited publications.
- Of the Sanger research published in 2022, 21.1 per cent were among the world's top 5 per cent most cited. The proportion of publications in 2022 where Sanger authors are leading the research (as first and/or last authors) was 24.9 per cent.

Outputs in the top 1 per cent of citation percentiles (field-weighted):

- 9.8 per cent of Sanger articles and reviews published between 2018 and 2022 – nearly one in every ten publications – are among the top 1 per cent of the world's most cited publications.
- The proportion of articles in the top 1 per cent most cited articles increases where the work is led by Sanger researchers, at 13.3 per cent between 2018 and 2022.

- For Sanger papers published in 2022 alone, 9 per cent were among the top 1 per cent of the world's most cited publications, and 9.7 per cent of papers where Sanger authors were leading the research as first and/or last authors.

Collaboration:
International collaboration is one of the core strengths of the Sanger Institute. With a long tradition of producing research findings as partners in internationally collaborative networks, the proportion of research articles published as a result of international collaboration is consistently high; above 82 per cent of the Sanger Institute's publications between 2018-2022. The proportion of publications authored with industry partners between 2018 and 2022 is also high at 17.8 per cent.

Summary

These data reflect the Sanger Institute's position as a global centre of genomics research. The Institute consistently produces highly cited research, with papers published in 2022 cited 4.71 times more than the global average. A considerable and stable proportion of Sanger primary research articles and reviews are among the world's most cited publications, many of these delivered through international collaboration.

Annexe. Detailed analysis of Sanger research outputs using bibliometric indicators

This Annexe provides a detailed overview of publications and citations, structured using bibliometric indicators. These analyses also take into account publications in which the Sanger Institute scientists take a leading role as indicated by their position as the first or last (or in many cases, both) author position on a research article. The order of the co-authors typically reflects the relative significance of their contributions in the life sciences.

Scholarly Output

Between 2018-2022 the Sanger Institute published 3,005 articles and reviews in peer-reviewed journals, including 892 where Sanger authors had a lead role, either by being the first or last author(s) on the paper. The number of Sanger Institute publications averages over 500 each year, with a peak of 632 publications in 2021 (Figure 1). Publications where Sanger scientists led the work averages over 160 per year, with 186 articles and reviews published in 2022.

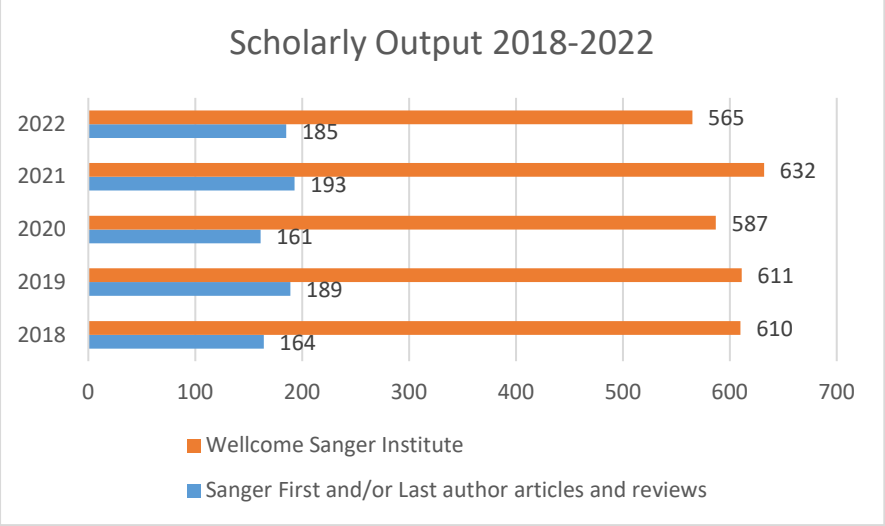

Figure 1: The number of articles and reviews published in academic journals, 2018-2022

Field-Weighted Citation Impact (FWCI)

The Field-Weighted Citation Impact (FWCI) is the ratio of the total citations received by a set of publications and the total expected citations, based on the average citations over the previous three years for all publications of the same age, document type and field. The FWCI greater than 1.00 indicates that a set of publications have been cited more than the world average for similar publications. The field-weighting embedded in this metric accounts for differences in research and citation behaviour across disciplines, removing bias toward older articles which have had longer to accrue citations.

The FWCI is a more robust metric when calculated for publication sets that have over 1,000 publications. The number of publications for individual years are relatively small and therefore subject to fluctuations that are expected within the range, caused by a small number of highly cited publications, so we use the five-year figure as a better indicator of the overall trend.

The FWCI for the Sanger Institute has risen to 3.98 by September 2023 for the current five-year reporting

period (2018-2022). This figure indicates that the articles and reviews published by Sanger Institute authors are 3.98 times more cited than the world average of 1.00 in the same subject field. The FWCI has remained consistently high across the rolling five-year average.

The publications where Sanger scientists lead the work were 4.64 times more cited than the world average by September 2023.

Citations per publication

The 3,005 articles and reviews published 2018-2022 accumulated 166,205 citations overall by September 2023 (Table 1), with an average of 55.3 citations per paper during this period. Older publications have longer to accumulate citations, so the five-year average is a better indicator of the overall trend.

The first and/or last author papers received an average of 65.8 citations per paper for papers published between 2018 and 2022 (Figure 2). Publications where Sanger scientists lead the work receive more citations per publication. First and/or last author papers published in 2022 alone have received 12.3 citations per paper by September 2023, similar to the Sanger papers overall and one of the highest citation rates among similar organisations.

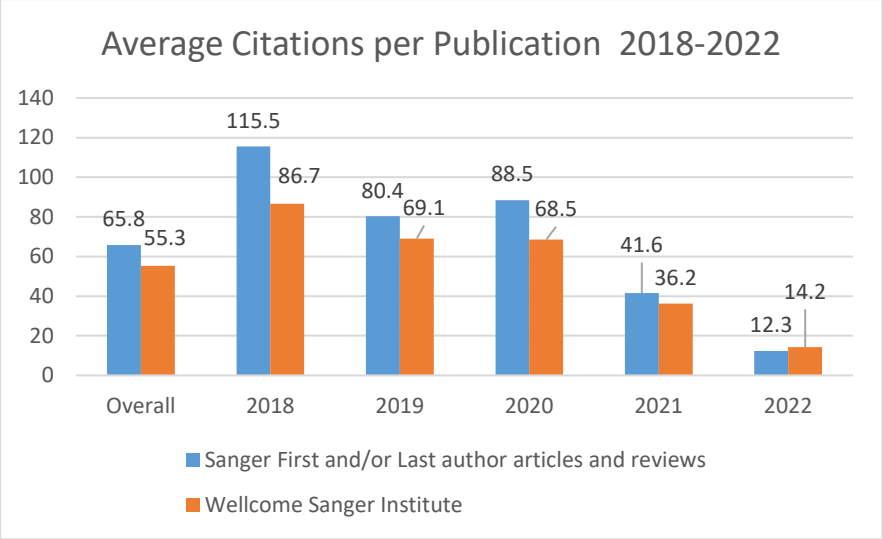

Figure 2: The change in the mean average citations per publication in academic journals, 2018-2022.

Top 5 per cent most cited papers

On average, 23.7 per cent of all Sanger Institute articles and reviews published between 2018 and 2022 are among the 5 per cent of the world's most cited publications. The metric used here is field-weighted and accounts for the variation in citation rates across multiple research fields. The thresholds for the top-cited percentiles are revised up or down as new citations are accrued and change regularly. Nevertheless, around 1 in 4 or 1 in 5 of Sanger Institute publications are consistently among the world's top 5 per cent most cited publications. In June 2023, 21.1 per cent of Sanger publications were among the world's top 5 per cent most cited.

The proportion of publications where Sanger authors are leading the research (as first and/or last authors) was 29 per cent on average between 2018 and 2022, and 24.9 per cent of those published in 2022 alone (Figure 3). A particularly strong year was 2018: over one in three articles published in 2018 that Sanger were leading on were among the top 5 per cent most cited publications globally.

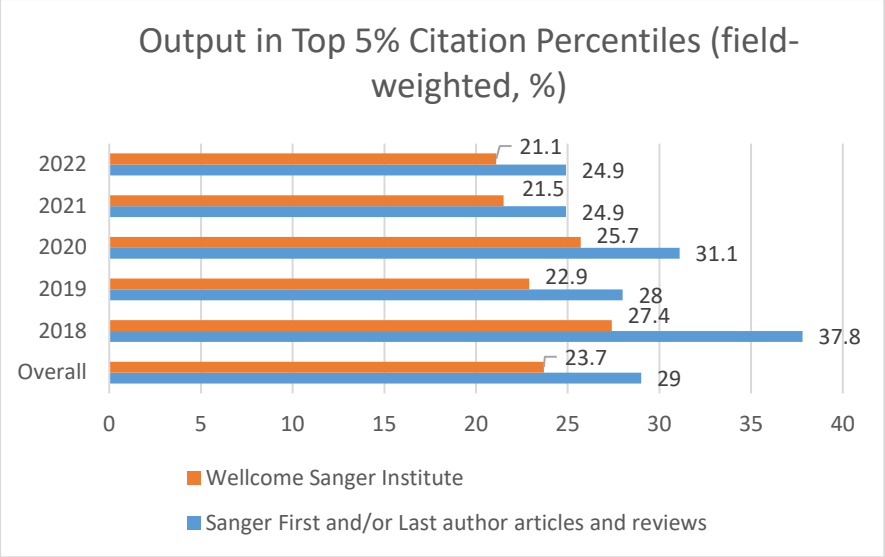

Figure 3: The change in the Outputs in Top 5 per cent of Citation Percentiles, 2018-2022.

Top 1 per cent most cited papers

For the five-year reporting period between 2018-2022, on average 9.8 per cent of Sanger publications – nearly 1 in every ten publications - were among the top 1 per cent of the world's most cited publications (Figure 4) indicates that the foundational knowledge produced at the Sanger Institute is influencing further research.

The proportion of Sanger-led articles that are among the world's top 1 per cent most cited papers was higher at 13.3 per cent in the five-year period between 2018-2022. The proportion of Sanger-led articles among the top 1 per cent most cited publications released in 2022 was 9.7 per cent by September 2023, reflecting that the more recent publications have not had long to accrue citations.

Regular monitoring shows that the Sanger Institute's publications enjoy a stable position of being consistently highly cited, with a substantial proportion of papers among the top 5 per cent and top 1 per cent most cited publications.

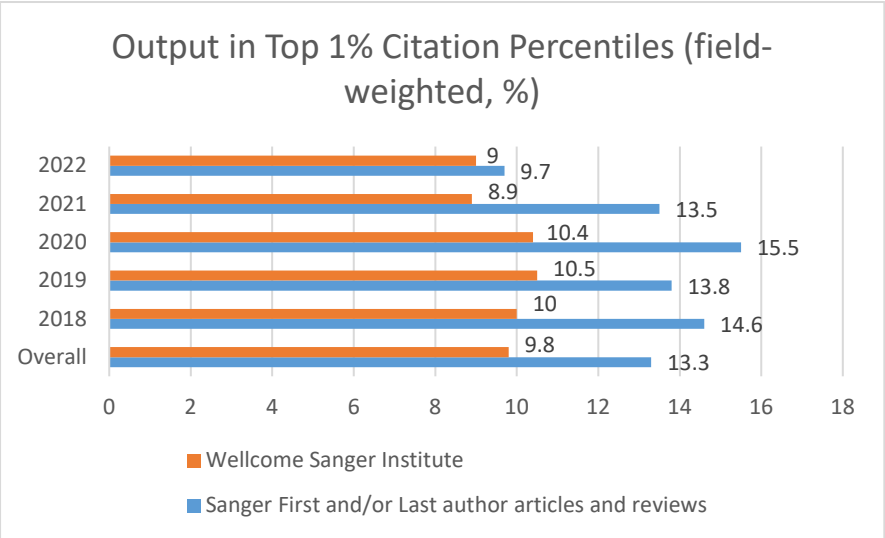

Figure 4: The change in the Outputs in Top 1% Citation Percentiles, 2018-2022.

Collaboration indicators: international collaboration and academic-industry collaboration

The foundational insights we gain from conducting research at scale help accelerate our understanding of human biology in both health and disease. However, answering some of the biggest and long-standing questions in biology requires international collaboration to harness our collective capabilities and take on translational challenges. Collaborative approach is at the heart of the Sanger Institute's strategy. Through initiating new partnerships, convening globally and taking a leading role in the global scientific community we are able to leverage our capabilities for large-scale data generation to tackle complex challenges and enrich the global scientific community.

The publication analysis indicate that the Sanger Institute collaborates widely, with an average of 82.4 per cent of publications between 2018-2022 arising from collaborations, and 82.7 per cent of those published in 2022 alone (Figure 5). In the five years between 2018-2022, Sanger scientists led the work in 77.2 per cent of publications carried out in collaboration, as indicated by the first and/or last author positions on the research articles.

Of the Sanger Institute's publications 17.8 per cent were authored with industry partners between 2018-2022 (Figure 6). The proportion of those led by Sanger Institute scientists was 15.1 per cent of articles between 2018 and 2022.

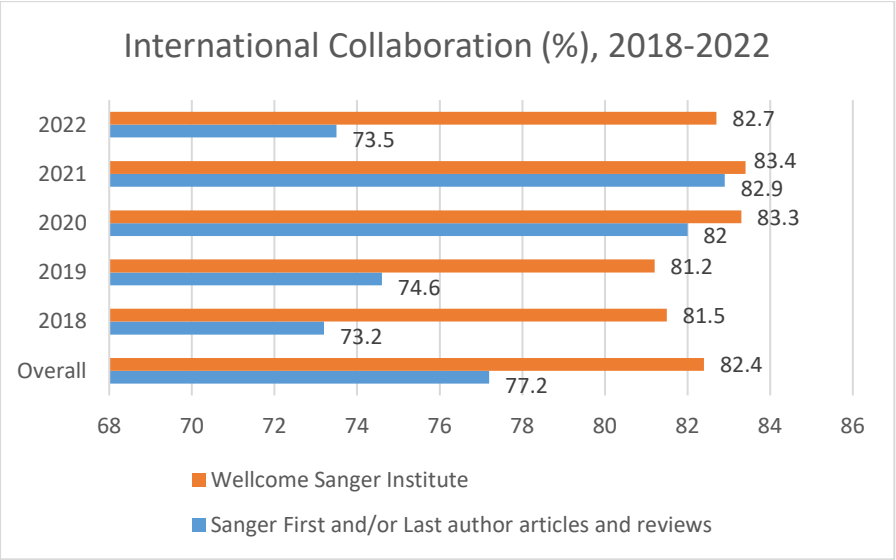

Figure 5: The proportion of articles and reviews published as a result of international collaboration, 2018-2022.

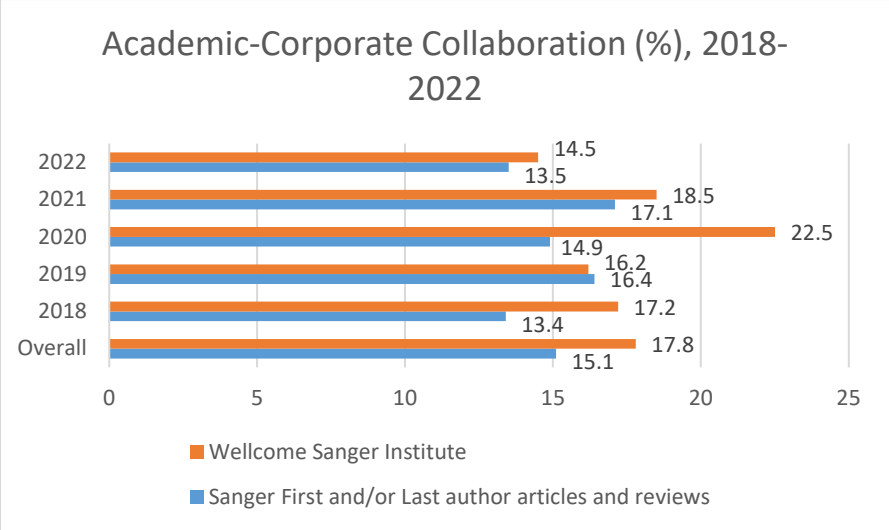

Figure 6: The proportion of articles and reviews published as a result of academic-corporate collaboration, 2018-2022.

The research articles and reviews published in the five years between 2018-2022 have been cited in further 166,205 research publications (Table 1), reflecting the Sanger Institute's influence on further genomics research.

Table 1: Total citation count, 2018-2022

Citation Count	Overall	2018	2019	2020	2021	2022
Sanger First and/or Last author articles and reviews	58,686	18,946	15,198	14,243	8,027	2,272
Wellcome Sanger Institute	166,205	52,870	42,208	40,223	22,887	8,017

Since 1996 - the first year for which the data are available – the Sanger Institute's articles and reviews have been cited over a million times (1,388,662 in September 2021), with an overall citation rate of 140 citations per publication.

Conclusion

The bibliometric analyses shows that the research articles produced by the Sanger Institute are among the most cited publications globally, underscoring the profound impact of this foundational knowledge on subsequent studies.

Despite a relatively modest faculty size, the Sanger Institute consistently ranks among the top in terms of publication volume and citation rates, and has a longstanding tradition of contributing to research within international collaborative networks.

Collectively, this data firmly establishes the Sanger Institute as a global hub for genomics research. The Institute's authors consistently produce highly cited research, with a significant and stable proportion of the primary research articles and reviews ranking among the world's most cited publications, many of which result from international collaborative efforts. These citation indicators highlight the extensive influence of Sanger Institute publications in advancing further research on a global scale.