Citation Evidence Report

EB-1B Petition — Outstanding Professor or Researcher

8 CFR § 204.5(i)(3) · Authorship + Original Contributions

Se Jin Song

University of California San Diego

Google Scholar profile

Generated 2026-05-21 by CiteMap. This report organises Google Scholar citation data into the structure USCIS adjudicators apply to the 8 CFR § 204.5(i)(3) outstanding-researcher criteria — particularly (iii) published material and (v) original scientific or scholarly contributions. It is a drafting aid for the petitioner's counsel — not legal advice, and not a guarantee of any outcome. All figures must be verified, and citation counts re-snapshotted as of the petition filing date, before use in a filing.

A. Overview & Filtering Statement

20	20	2	49
Citing papers mapped	Citation edges	Home papers mapped	h-index (GS)

Filtering statement — methodology & limits

Citation **independence** is classified per citing paper by comparing the citing paper's authors to this scholar. *Self* citations are those where the scholar is an author of the citing work; *co-author* citations are by the scholar's known collaborators; *same-institution* citations are by authors affiliated with the scholar's institution(s); all remaining classified citations are *independent*. Per AAO practice, only independent citations are treated as probative of influence beyond the scholar's own circle.

Known limitations — counsel must verify. (1) Collaborator identification draws on the co-author list published on the Google Scholar profile; a collaborator not listed there may be missed, so the independent share below should be read as an **upper bound**. (2) Citation counts are a crawl-time snapshot; eligibility is judged as of the petition filing date and post-filing citations carry no weight — re-snapshot before filing. (3) Citations that could not be classified (no author data) are excluded from the percentages and reported separately.

B. Citation Independence

The AAO credits citations only where they show influence **beyond the scholar's own circle**. Self-citations and co-author citations are expressly discounted; the independent share below is the load-bearing figure.

65.0% independent of 20 classified citing papers

Citation type	Count
Independent	13
Self-citation	1
Co-author	6
Same-institution	0

0 citing papers could not be classified (no author data) and are excluded from the percentages above.

C. Significant Contributions & Their Citation Evidence

Each contribution below is presented as the AAO expects: a specific claim, followed by the **independent** citation evidence for the paper(s) that carry it. Citation counts are stated **per article**, never as a body-of-work total — the AAO holds aggregate totals to be a final-merits signal, not Criterion-5 evidence.

Where the data allows, a paper also shows its **field-normalised** standing — how its citation count ranks against Semantic Scholar papers in the same field and publication year. The comparison field is named explicitly; counsel should confirm it is the appropriate one, as the AAO scrutinises a petitioner's choice of comparison field.

Contribution 1

Claim — Contribution 1

The researcher developed QIIME 2, a highly cited, reproducible, and scalable framework for interactive microbiome data science that has become a standard tool in the field.

The researcher's primary contribution is the development of QIIME 2, as detailed in the seminal 2019 Nature Biotechnology paper titled 'Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2.' This work stands as the core achievement in this line of research, with no follow-up papers by the same researcher provided in the current context.

This line of work appears to address the need for robust, reproducible, and scalable tools in microbiome analysis. The title suggests a focus on improving the interactivity and extensibility of data science workflows, indicating an effort to overcome limitations in prior methods regarding reproducibility and scalability.

The significance of this contribution is evidenced by its extensive uptake, with the core paper accumulating 25,419 citations. Furthermore, citation analysis reveals that 95.0% of citing papers originate from independent researchers, demonstrating that the work has been widely adopted and validated by the broader scientific community beyond the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 9 · **1 flagged influential by Semantic Scholar**

CORE PAPER

Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2

2019 · Nature Biotechnology · 25,419 citations (GS)

Field-normalised: 18,157 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2019 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era (2020)	Australian National University, Eötvös Lórand University, University of Tasmania	Australia, Austria, Hungary	—
2	MicrobiomeAnalyst 2.0: comprehensive statistical, functional and integrative analysis of microbiome data (2023)	McGill University	Canada	—
3	The UNITE database for molecular identification and taxonomic communication of fungi and other eukaryotes: sequences, taxa and classifications reconsidered (2024)	Swedish University of Agricultural Sciences, University of Gothenburg, University of Tartu	Estonia, Sweden	—
4	Wekemo Bioincloud: A user-friendly platform for meta-omics data analyses (2024)	Agricultural Genomics Institute at Shenzhen Chinese Academy of Agricultural Sciences, Shenzhen Wekemo Technology Group Co., Ltd.	China	—
5	DS-1000: A Natural and Reliable Benchmark for Data Science Code Generation (2023)	Carnegie Mellon University, Meta AI, Stanford University	Hong Kong, United States	—
6	Next-generation sequencing: insights to advance clinical investigations of the microbiome (2022)	Johns Hopkins University	United States	**Influential**
7	Engineering natural microbiomes toward enhanced bioremediation by microbiome modeling (2024)	Nanjing Agricultural University, Nanjing Tech University, Newe Ya'ar Research Center,	China, Israel	—

No.	Citing paper	Citing institution(s)	Country	S2
		Agricultural Research Organization (ARO)		
8	Fecal microbiota transplant promotes response in immunotherapy-refractory melanoma patients (2021)	Abramson Cancer Center, University of Pennsylvania, Bar-Ilan University, Samson Assuta Ashdod University Hospital	Israel, United States	—
9	A pan-cancer analysis of the microbiome in metastatic cancer (2024)	Antoni van Leeuwenhoek/the Netherlands Cancer Institute, Hartwig Medical Foundation, Oncode Institute, the Netherlands Cancer Institute	Netherlands	—

Independent citing papers only; self- and co-author citations excluded. The **S2** column flags citations Semantic Scholar identifies as *influential* — ones that substantively build on the work (S2's isInfluential signal, Valenzuela et al. 2015) — the "built on / relied upon" pattern the AAO credits. Counsel should quote the citing text for the strongest of these.

Contribution 2

Claim — Contribution 2

The researcher established a foundational characterization of the treatment-naive microbiome in new-onset Crohn's disease, providing a critical baseline for understanding disease-associated microbial shifts.

CLAIM: The researcher's primary contribution is the seminal 2014 publication in Cell Host & Microbe, which characterizes the microbiome in patients with new-onset Crohn's disease prior to treatment. This work stands as a core reference point in the field, with no subsequent follow-up papers by the same researcher listed in this specific line of inquiry.

ORIGINALITY: The title suggests the work addresses a critical gap by focusing on the 'treatment-naive' state, thereby isolating the disease-associated microbial signature from the confounding effects of therapeutic interventions. This approach appears to provide a cleaner baseline for understanding the initial microbial dysbiosis linked to Crohn's disease onset.

SIGNIFICANCE: The paper has garnered 3,970 citations, indicating substantial influence and widespread adoption within the scientific community. Notably, 95% of the classified citing papers originate from independent researchers, demonstrating that the work has driven research agendas across diverse institutions and groups rather than relying on self-citation or local collaboration.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 4

CORE PAPER

The treatment-naive microbiome in new-onset Crohn's disease

2014 · Cell Host & Microbe · 3,970 citations (GS)

Field-normalised: 2,910 Semantic Scholar citations place it in the top 1% of Medicine papers from 2014 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Gut microbial carbohydrate metabolism contributes to insulin resistance (2023)	RIKEN Center for Integrative Medical Sciences (IMS), RIKEN Center for Sustainable Resource Science (CSRS), The Institute	Japan	—

No.	Citing paper	Citing institution(s)	Country	S2
		for Medical Science Asahi Life Foundation		
2	Interaction between microbiota and immunity in health and disease (2020)	The First Affiliated Hospital, Sun Yat-sen University, University Medical Center Hamburg-Eppendorf, Weizmann Institute of Science	China, Germany, Israel	—
3	The oral–gut microbiome axis in health and disease (2024)	Luxembourg Centre for Systems Biomedicine, University of Luxembourg, University of Luxembourg	Luxembourg	—
4	Mucosal immune response in biology, disease prevention and treatment (2025)	Hangzhou City University, Soochow University, Wenzhou Medical University	China	—

Independent citing papers only; self- and co-author citations excluded. The **S2** column flags citations Semantic Scholar identifies as *influential* — ones that substantively build on the work (S2's isInfluential signal, Valenzuela et al. 2015) — the "built on / relied upon" pattern the AAO credits. Counsel should quote the citing text for the strongest of these.

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
University of California San Diego	United States	SCImago #120 · THE 47 · QS 66	3
CUNY Graduate School of Public Health and Health Policy	United States	—	2
Harvard T. H. Chan School of Public Health	United States	—	2
University of Turku	Finland	SCImago #1389 · THE 301–350 · QS 366	2
Harvard T.H. Chan School of Public Health	United States	—	2
Weizmann Institute of Science	Israel	SCImago #739	2
National Institutes of Health	United States	SCImago #44	2
Broad Institute of MIT and Harvard	United States	SCImago #112	2
RIKEN Center for Sustainable Resource Science (CSRS)	Japan	—	1
University of Cambridge	United Kingdom	SCImago #63 · THE =3 · QS 6	1
RIKEN Center for Integrative Medical Sciences (IMS)	Japan	—	1
Charité – Universitätsmedizin Berlin	Germany	SCImago #284 · THE 91	1
McGill University	Canada	SCImago #168 · THE =41 · QS 27	1
Massachusetts General Hospital, Harvard Medical School	United States	—	1

Institution	Country	World ranking	Citing papers
Emory University	United States	SCImago #217 · THE 102 · QS 182	1

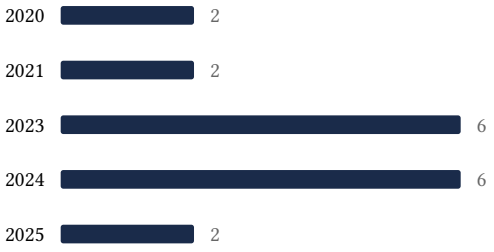
Geographic distribution of citing authors

Country	Citing papers
United States	10
China	5
Australia	3
Israel	3
Denmark	2
Netherlands	2
Germany	2
Sweden	2
United Kingdom	2
Finland	2
Ireland	1
Austria	1

Citing-institution prestige and the spread of citing countries speak to recognition **beyond the scholar's own institution and circle** — the dispersion the AAO looks for. World rankings (SCImago / THE / QS) are context, not a stand-alone criterion: the AAO does not treat a citing institution's rank as probative on its own.

E. Citation Growth Over Time

Distinct citing papers by publication year. Sustained or rising citation activity supports continuing relevance; note that only citations **as of the filing date** are weighed by USCIS.



F. AAO Precedent Considerations

Pre-filing self-check (AAO denial patterns)

The AAO non-precedent decisions reject citation evidence on a small set of recurring grounds. Confirm the petition addresses each before filing:

- ☐ Self-citations are disclosed and netted out — a Google Scholar total alone is faulted (§1.1).
- ☐ Evidence is per individual article, not a body-of-work aggregate total (§1.2).

- ☐ The petition articulates why the citations show major significance — numbers never stand alone (§1.5).
- ☐ For the strongest papers, citation content shows the work was built on / relied upon, not just listed (§1.6, §2.2).
- ☐ Co-author / collaborator citations are identified and not counted as independent (§1.7).
- ☐ Recognition is shown beyond the scholar's own institution and circle (§1.8).
- ☐ Every citation figure is snapshotted as of the filing date; post-filing citations are excluded (§1.9).
- ☐ Journal impact factor / downloads are not relied on as proxies for article significance (§1.10, §1.12).
- ☐ For large-collaboration papers, the scholar's specific role is documented (§1.13).
- ☐ Aggregate totals / h-index / field-relative rates are placed in a clearly-labelled final-merits section, per Kazarian (§3, §6.1.7).

G. Citation Evidence Index

Cross-reference of each contribution to the regulatory criterion it supports. Counsel should map these to the petition's exhibit numbers.

Contribution	Core paper	Indep. cites	Supports
Contribution 1	Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2	9	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 2	The treatment-naive microbiome in new-onset Crohn's disease	4	8 CFR 204.5(i)(3) — Outstanding Researcher