Citation Evidence Report

EB-1B Petition — Outstanding Professor or Researcher

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

Sandrine Dudoit

Division of Biostatistics and Department of Statistics, University of California, Berkeley

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

7	7	1	69
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.

100.0% independent of 7 classified citing papers

Citation type	Count
Independent	7
Self-citation	0
Co-author	0
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 established Bioconductor as a foundational open-source framework for computational biology, enabling standardized software development and reproducible research in bioinformatics.

CLAIM: The researcher's primary contribution is the creation of Bioconductor, an open software development platform for computational biology and bioinformatics, as detailed in the seminal 2004 paper published in Genome Biology. This work stands as a singular, foundational achievement in the field.

ORIGINALITY: The title suggests this work addressed the need for structured, open-source software development in bioinformatics. By establishing a dedicated framework, the researcher appears to have provided a standardized infrastructure that facilitated collaborative and reproducible computational research, distinguishing it from ad-hoc scripting approaches prevalent at the time.

SIGNIFICANCE: With nearly 15,000 citations, the paper is highly influential. Analysis of citing literature indicates that 100% of classified citations originate from independent researchers, demonstrating broad adoption across the global scientific community rather than self-citation or institutional clustering. This widespread independent uptake underscores the framework's critical role in advancing the field.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 7

CORE PAPER

Bioconductor: open software development for computational biology and bioinformatics

2004 · Genome Biology · 14,977 citations (GS)

Field-normalised: 12,833 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2004 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	limma powers differential expression analyses for RNA-sequencing and microarray studies (2015)	Harvard University, Murdoch Childrens Research Institute, The Walter and Eliza Hall Institute of Medical Research	Australia, Switzerland, United States	—
2	edgeR v4: powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets (2025)	Genentech Inc, WEHI	Australia, United States	—
3	Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis (2019)	McGill University, University of Alberta	Canada	—
4	Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data (2018)	North Carolina State University, Stanford University, Stanford University School of Medicine	United States	—
5	CZ CELLxGENE Discover: a single-cell data platform for scalable exploration, analysis and modeling of aggregated data (2025)	Chan Zuckerberg Initiative, Stanford University School of Medicine, Wellcome Sanger Institute	—	—
6	Maftools: efficient and comprehensive analysis of somatic variants in cancer (2018)	Cancer Science Institute of Singapore, National University of Singapore, Cedars-Sinai Med-	Germany, Singapore, United States	—

No.	Citing paper	Citing institution(s)	Country	S2
		ical Center, German Cancer Research Center (DKFZ)		
7	Welcome to the Tidyverse (2019)	RStudio	—	—

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
Stanford University School of Medicine	United States	—	2
McGill University	Canada	SCImago #168 · THE =41 · QS 27	1
National University of Singapore	Singapore	SCImago #59 · THE 17 · QS 8	1
Murdoch Childrens Research Institute	Australia	—	1
The Walter and Eliza Hall Institute of Medical Research	Australia	SCImago #580	1
Cedars-Sinai Medical Center	United States	SCImago #705	1
North Carolina State University	United States	SCImago #484 · THE 301–350 · QS =272	1
Wellcome Sanger Institute	UK	SCImago #204	1
University of Alberta	Canada	SCImago #262 · THE 119 · QS =94	1
Chan Zuckerberg Initiative	United States	—	1
Harvard University	United States	SCImago #4 · THE =5 · QS 5	1
WEHI	Australia	—	1
Genentech Inc	United States	SCImago #185	1
University of Zurich	Switzerland	SCImago #313 · QS 100	1
Stanford University	United States	SCImago #18 · THE =5 · QS 3	1

Geographic distribution of citing authors

Country	Citing papers
United States	4
Australia	2
Canada	1
Germany	1
Singapore	1
Switzerland	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).

Disclaimer

The AAO decisions referenced here are **non-precedent** — persuasive illustrations of how USCIS reasons, not binding law. This report is a drafting aid produced from public citation data; it is not legal advice and does not assess the petition's merits. All analysis must be reviewed by qualified immigration counsel.

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	Bioconductor: open software development for computational biology and bioinformatics	7	8 CFR 204.5(i)(3) — Outstanding Researcher