Citation Evidence Report

EB-1B Petition — Outstanding Professor or Researcher

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

Avi Srivastava

Assistant Professor, The Wistar Institute

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

15	15	2	19
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.

80.0% independent of 15 classified citing papers

Citation type	Count
Independent	12
Self-citation	1
Co-author	2
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 Bioconda, a sustainable and comprehensive software distribution system for the life sciences, establishing a critical infrastructure for reproducible computational biology.

CLAIM: The researcher's primary contribution is the creation of Bioconda, a software distribution platform designed to support the life sciences community. This work is anchored in the 2018 publication in Nature Methods, which serves as the foundational reference for this line of inquiry.

ORIGINALITY: The titles suggest this work addresses the need for sustainable and comprehensive software management in biological research. By focusing on distribution mechanisms, the researcher appears to have tackled the challenge of ensuring reliable access to computational tools, a gap that likely hindered reproducibility and efficiency in the field prior to this intervention.

SIGNIFICANCE: The core paper has accumulated 1755 citations, indicating substantial uptake by the scientific community. Furthermore, analysis of citing literature reveals that 86.7% of citations originate from independent researchers, demonstrating that the work has been widely adopted and valued by peers outside the researcher's immediate circle, underscoring its broad impact on the field.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 5 · **1 flagged influential by Semantic Scholar**

CORE PAPER

Bioconda: sustainable and comprehensive software distribution for the life sciences

2018 · Nature Methods · 1,755 citations (GS)

Field-normalised: 1,405 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2018 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Twelve years of SAMtools and BCFtools (2021)	Dana-Farber Cancer Institute, Harvard Medical School, EMBL-EBI, European Molecular Biology Laboratory	United Kingdom, United States	Methodology
2	A practical guide to amplicon and metagenomic analysis of microbiome data (2020)	Children's Hospital, Zhejiang University School of Medicine, China Academy of Chinese Medical Sciences, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences	China	—
3	BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes (2021)	University of Geneva	Switzerland	Background
4	JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles (2022)	Aix Marseille Univ, BC Children's Hospital Research Institute, University of British Columbia, Centre for Molecular Medicine Norway	Belgium, Canada, Denmark	—

No.	Citing paper	Citing institution(s)	Country	S2
		(NCMM), Nordic EMBL Partnership, University of Oslo		
5	The PRIDE database and related tools and resources in 2019: improving support for quantification data (2019)	European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Max Planck Institute for Biochemistry, Medical University of Vienna	Austria, Germany, United Kingdom	Methodology

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

Citing-text excerpts — how the field used this work

METHODOLOGY Twelve years of SAMtools and BCFtools

"Both packages have been installed over a million times via Bioconda."

METHODOLOGY The PRIDE database and related tools and resources in 2019: improving support for quantification data

"The main rationale is to make possible the use of that software infrastructures (using the EMBL-EBI cloud as the starting point), so that in the future the pipelines can be used by the community in the cloud using software container technologies (39,40)."

Contribution 2

Claim — Contribution 2

The researcher developed an integrated analytical framework for multimodal single-cell data, establishing a foundational methodology widely adopted across the field.

The researcher's primary contribution is the development of an integrated analysis framework for multimodal single-cell data, as detailed in their 2021 Cell paper. This work stands as a seminal core contribution, with no follow-up papers by the same researcher listed in this specific line of inquiry, suggesting the original publication itself constitutes the complete methodological advance.

This line of work appears to address the challenge of synthesizing diverse single-cell data types into a unified analytical model. The title indicates a focus on integration, suggesting the researcher provided a novel approach to handling the complexity of multimodal datasets, a gap that likely hindered comprehensive biological interpretation prior to this publication.

The significance of this contribution is evidenced by its substantial citation count of 15,339, indicating widespread adoption and influence. Furthermore, the high proportion of independent citations, with 86.7% originating from researchers outside the scholar's immediate circle, underscores the work's broad impact and acceptance across the global scientific community.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 7 · **1 flagged influential by Semantic Scholar**

CORE PAPER

Integrated analysis of multimodal single-cell data

2021 · Cell · 15,339 citations (GS)

Field-normalised: 11,458 Semantic Scholar citations place it in the top 1% of Biology papers from 2021 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	TBtools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining (2023)	Henan University, Hunan Agricultural University, Institute of Tropical Bioscience and Biotechnology, Chinese Academy of Tropical Agricultural Sciences	China	—
2	Best practices for single-cell analysis across modalities (2023)	Helmholtz Center Munich, German Research Center for Environmental Health, Helmholtz Munich, Technical University of Munich	Germany	Methodology
3	Not Provided (2023)	Helmholtz Center Munich	—	—
4	Deterministic reprogramming of neutrophils within tumors (2024)	Agency for Science, Technology and Research, A*STAR, Centro Nacional de Investigaciones Cardiovasculares Carlos III	Australia, China, France	—
5	Persistent complement dysregulation with signs of thromboinflammation in active Long Covid (2024)	Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, ETH Zurich, ETH Zurich & Swiss Institute of Bioinformatics (SIB)	Sweden, Switzerland, United Kingdom	—
6	The dawn of spatial omics (2023)	Cancer Research UK (CRUK) Cambridge Institute, University of Cambridge	United Kingdom	Background
7	cGAS–STING drives ageing-related inflammation and neurodegeneration (2023)	Lausanne University Hospital, Swiss Federal Institute of Technology Lausanne (EPFL), The Netherlands Cancer Institute	Germany, Netherlands, Switzerland	—

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

Citing-text excerpts — how the field used this work

METHODOLOGY Best practices for single-cell analysis across modalities

> *"A recent benchmark 44 recommends using the built-in cell cycle labelling and correction functions in Scanpy 4 or Seurat 45 as a baseline, which compare the mean expression values to a reference signature."*

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
EMBL-EBI	United Kingdom	—	2

Institution	Country	World ranking	Citing papers
New York University	United States	SCImago #116 · THE =31 · QS 55	2
Centre for Molecular Medicine Norway (NCMM), Nordic EMBL Partnership, University of Oslo	Norway	—	1
BC Children's Hospital Research Institute, University of British Columbia	Canada	—	1
MRC London Institute of Medical Sciences	United Kingdom	—	1
Swiss Institute of Bioinformatics (SIB)	Switzerland	SCImago #523	1
University of Copenhagen Globe Institute	Denmark	—	1
Western University	Canada	THE 201–250 · QS 151	1
University of Mainz	Germany	—	1
SIB Swiss Institute of Bioinformatics / ELIXIR Switzerland	Switzerland	—	1
Earle A Chiles Research Institute	United States	—	1
European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI)	United Kingdom	—	1
University of Tuebingen	Germany	—	1
Max Planck Institute for Biochemistry	Germany	—	1
Technical University of Munich	Germany	SCImago #187 · THE 27 · QS =22	1

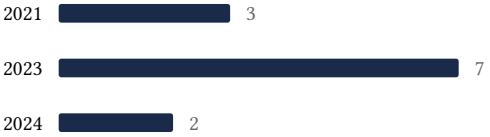
Geographic distribution of citing authors

Country	Citing papers
United States	6
United Kingdom	6
Switzerland	4
Germany	4
China	3
Canada	2
France	2
Denmark	2
Sweden	1
Italy	1
Austria	1
Belgium	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	Bioconda: sustainable and comprehensive software distribution for the life sciences	5	8 CFR 204.5(i)(3) — Outstanding Researcher

Contribution	Core paper	Indep. cites	Supports
Contribution 2	Integrated analysis of multimodal single-cell data	7	8 CFR 204.5(i)(3) — Outstanding Researcher