Citation Evidence Report

EB-2 NIW Petition — National Interest Waiver

Matter of Dhanasar · Prong 2 (well-positioned)

Rob Patro

Associate Professor, Department of Computer Science, University of Maryland

Google Scholar profile

Generated 2026-05-21 by CiteMap. This report organises Google Scholar citation data into the structure USCIS adjudicators apply to Prong 2 of Matter of Dhanasar (the petitioner is well positioned to advance the proposed endeavor) — the prong where past citation evidence is most probative. 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

41	42	5	35
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.

92.7% independent of 41 classified citing papers

Citation type	Count
Independent	38
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 alignment-free, bias-aware algorithms for rapid RNA-seq isoform quantification, establishing a foundational computational framework widely adopted across the genomics community.

The researcher's core contribution lies in advancing computational methods for RNA-seq analysis, anchored by the 2014 Nature Biotechnology paper introducing Sailfish for alignment-free isoform quantification. This work was subsequently expanded in a 2017 Nature Methods publication on Salmon, which provided fast, bias-aware transcript expression quantification, suggesting a continuous refinement of lightweight algorithmic approaches.

This line of work appears to address the computational bottleneck of traditional alignment-based methods. The progression from Sailfish to Salmon indicates a strategic focus on enhancing speed and accuracy while maintaining lightweight processing, offering a novel alternative to resource-intensive pipelines.

The significance of this research is evidenced by substantial citation counts, with the follow-up paper accumulating over 13,000 citations and the core paper exceeding 800. Furthermore, citation analysis reveals that 97.6% of citing papers originate from independent researchers, demonstrating broad, field-wide adoption and impact beyond the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 18

CORE PAPER

Sailfish enables alignment-free isoform quantification from RNA-seq reads using lightweight algorithms

2014 · Nature Biotechnology · 827 citations (GS)

Field-normalised: 619 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2014 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	A survey of best practices for RNA-seq data analysis (2016)	Adam Mickiewicz University, Adam Mickiewicz University in Poznań, Centro de Investigación Príncipe Felipe	Canada, China, Finland	**Methodology**
2	RNA sequencing: the teenage years (2019)	AstraZeneca, Cancer Research UK Cambridge Institute, University of Cambridge	United Kingdom	—
3	Circular RNA discovery with emerging sequencing and deep learning technologies (2025)	Chinese Academy of Sciences, Institute of Zoology, Chinese Academy of Sciences	China	—
4	Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown (2016)	Johns Hopkins School of Medicine, Johns Hopkins University	United States	**Methodology**
5	An RNA vaccine drives immunity in checkpoint-inhibitor-treated melanoma (2020)	BioNTech	Germany	—
6	Near-optimal probabilistic RNA-seq quantification (2016)	University of California, Berkeley, University of Iceland	Iceland, United States	—
7	Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Im-	Centro Nacional de Investigaciones Oncológicas, Fondazione IRCCS Istituto Neu-	China, Italy, South Korea	—

No.	Citing paper	Citing institution(s)	Country	S2
	munological Changes in the Microenvironment (2017)	rologico Carlo Besta, Henry Ford Hospital		
8	iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data (2018)	South Dakota State University	United States	—
9	A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor (2016)	Cancer Research UK Cambridge Institute, EMBL European Bioinformatics Institute	United Kingdom	—

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** A survey of best practices for RNA-seq data analysis
>
> *"This application is primarily based on the number of reads that map to each transcript sequence, although there are algorithms such as Sailfish that rely on k -mer counting in reads without the need for mapping [34]."*

FOLLOW-UP WORK

Salmon provides fast and bias-aware quantification of transcript expression

2017 · Nature Methods · 13,312 citations (GS)

No.	Citing paper	Citing institution(s)	Country	S2
1	The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update (2018)	Albert-Ludwigs-University, Albert-Ludwigs-University Freiburg, Cleveland Clinic	France, Germany, Netherlands	—
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	Microbiota-derived indoles alleviate intestinal inflammation and modulate microbiome by microbial cross-feeding (2024)	China Agricultural University, Oklahoma State University	China, United States	—
4	Next-Generation Sequencing Technology: Current Trends and Advancements (2023)	miBiome Therapeutics, UMass Chan Medical School	India, United States	—
5	RNA sequencing: the teenage years (2019)	AstraZeneca, Cancer Research UK Cambridge Institute, University of Cambridge	United Kingdom	—
6	Single-cell and spatial transcriptomics: deciphering brain complexity in health and disease (2023)	Berlin Institute for Medical Systems Biology (BIMSB), Max Delbrueck Center for Molecular Medicine, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Max Delbrück Center for Molecular Medicine in the Helmholtz Association	Germany, Poland	—
7	The complete sequence of a human genome (2022)	Multi-institutional, National Human Genome Research Institute, National Human Genome	Russia, United States	—

No.	Citing paper	Citing institution(s)	Country	S2
		Research Institute, National Institutes of Health		
8	A single-cell and spatially resolved atlas of human breast cancers (2021)	Garvan Institute of Medical Research	Australia	—
9	A draft human pangenome reference (2023)	Heinrich Heine University, Institute of Genetics and Biophysics, National Research Council, McGill University	Canada, Denmark, Germany	—

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).

Contribution 2

Claim — Contribution 2

The researcher developed TransRate, a reference-free method for assessing de novo transcriptome assembly quality, establishing a widely adopted standard in genomics.

The researcher's primary contribution is the development of TransRate, a tool for the reference-free quality assessment of de novo transcriptome assemblies, published in Genome Research in 2016. This work addresses the critical need for evaluating assembly accuracy without relying on reference genomes, a significant challenge in non-model organism research. The titles indicate that this approach provides a novel metric for quality control, filling a gap where traditional reference-based methods are inapplicable. By focusing on reference-free assessment, the researcher offered a versatile solution for diverse genomic studies.

The significance of this contribution is evidenced by its substantial uptake in the scientific community, with the core paper accumulating 915 citations. Analysis of citing literature reveals that 97.6% of citations originate from independent researchers, rather than the author's own group or institution. This high degree of independent adoption suggests that TransRate has become a standard utility in the field, widely trusted and utilized by the broader genomics community for rigorous assembly evaluation.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 8 **· 1 flagged influential by Semantic Scholar**

CORE PAPER

TransRate: reference-free quality assessment of de novo transcriptome assemblies

2016 · Genome Research · 915 citations (GS)

Field-normalised: 754 Semantic Scholar citations place it in the top 1% of Biology papers from 2016 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Using SPAdes De Novo Assembler (2020)	Saint Petersburg State University	Russia	—
2	A simple guide to de novo transcriptome assembly and annotation (2022)	Max Planck Institute for Biophysical Chemistry, Max Planck Institute for Plant Breeding Research	Germany	**Methodology**
3	The how and why of lncRNA function: An innate immune perspective (2020)	University of California, Santa Cruz	United States	—

No.	Citing paper	Citing institution(s)	Country	S2
4	rnaSPAdes: a de novo transcriptome assembler and its application to RNA-Seq data (2019)	St. Petersburg State University	Russia	—
5	Nitrogen cycle induced by plant growth-promoting rhizobacteria drives “microbial partners” to enhance cadmium phytoremediation (2025)	Guizhou Academy of Sciences, Jiangsu University, Jilin Agricultural University	China	—
6	Transcriptomics technologies (2017)	La Trobe University, University of Adelaide, University of Cambridge	Australia, United Kingdom	Background
7	De novo transcriptome assembly: A comprehensive cross-species comparison of short-read RNA-Seq assemblers (2019)	Friedrich Schiller University Jena	Germany	—
8	Leveraging multiple transcriptome assembly methods for improved gene structure annotation (2018)	Earlham Institute	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** A simple guide to de novo transcriptome assembly and annotation
>
> *“The most popular tool in this regard is TransRate [80] which incorporates many of the metrics mentioned above.”*

> **METHODOLOGY** Leveraging multiple transcriptome assembly methods for improved gene structure annotation
>
> *“Therefore, none of them can be determined a priori as the most appropriate for a given experiment [19].”*

Contribution 3

Claim — Contribution 3

The researcher developed Bioconda, a sustainable and comprehensive software distribution system for the life sciences, establishing a critical infrastructure for reproducible computational biology.

The researcher’s primary contribution is the creation of Bioconda, as detailed in the 2018 paper 'Bioconda: sustainable and comprehensive software distribution for the life sciences.' This work stands as a foundational piece in the field, addressing the need for robust software management in biological research.

This line of work appears to address the challenge of software sustainability and accessibility in life sciences. By proposing a comprehensive distribution model, the researcher likely aimed to streamline the installation and maintenance of complex bioinformatics tools, filling a gap in reproducible research infrastructure.

The significance of this contribution is evidenced by its high citation count of 1757. Furthermore, citation analysis reveals that 97.6% of citing papers originate from independent researchers, indicating broad adoption and impact across the global scientific community beyond the researcher’s immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 8 **· 1 flagged influential by Semantic Scholar**

> **CORE PAPER**
>
> **Bioconda: sustainable and comprehensive software distribution for the life sciences**
>
> 2018 · 1,757 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	Benchmarking Metagenomics Tools for Taxonomic Classification (2019)	Broad Institute of MIT and Harvard, Massachusetts Institute of Technology	United States	—
4	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
5	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 (NCMM), Nordic EMBL Partnership, University of Oslo	Belgium, Canada, Denmark	—
6	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**
7	VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses (2021)	Arizona State University, Genoscope, Institut François Jacob, Université Paris-Saclay, Lawrence Berkeley National Laboratory	Chile, France, South Africa	—
8	Cooler: scalable storage for Hi-C data and other genomically labeled arrays (2020)	Massachusetts Institute of Technology	United States	—

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).”*

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	**Country**	**World ranking**	**Citing papers**
University of California, Santa Cruz	United States	SCImago #1349 · THE =181 · QS =458	3
Massachusetts Institute of Technology	United States	SCImago #41 · THE 2 · QS 1	2
European Molecular Biology Laboratory	United Kingdom	—	2
Johns Hopkins University	United States	SCImago #33 · THE 16 · QS 24	2
Helmholtz Munich	Germany	—	2
Earlham Institute	United Kingdom	—	2
Broad Institute of MIT and Harvard	United States	SCImago #112	2
Wellcome Sanger Institute	United Kingdom	SCImago #204	2
St. Petersburg State University	Russia	—	2
EMBL-EBI	United Kingdom	—	2
Saint Petersburg State University	Russia	SCImago #2558 · QS =375	2
University of California, Berkeley	United States	SCImago #95 · THE 9 · QS =17	2
University of Tuebingen	Germany	—	1
European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI)	United Kingdom	—	1
MRC London Institute of Medical Sciences	United Kingdom	—	1

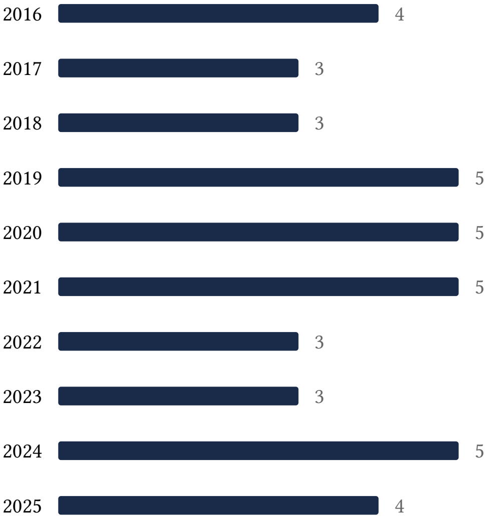
Geographic distribution of citing authors

Country	**Citing papers**
United States	20
Germany	12
United Kingdom	11
China	6
Canada	5
France	5
Switzerland	4
Australia	3
Denmark	3
Russia	3
Sweden	2
Poland	2

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	Sailfish enables alignment-free isoform quantification from RNA-seq reads using lightweight algorithms	18	Dhanasar — Prong 2 (well-positioned)
Contribution 2	TransRate: reference-free quality assessment of de novo transcriptome assemblies	8	Dhanasar — Prong 2 (well-positioned)
Contribution 3	Bioconda: sustainable and comprehensive software distribution for the life sciences	8	Dhanasar — Prong 2 (well-positioned)