Citation Evidence Report

EB-2 NIW Petition — National Interest Waiver

Matter of Dhanasar · Prong 2 (well-positioned)

Martin Vingron

Max-Planck-Institut für molekulare Genetik

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	41	5	95
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.

87.8% independent of 41 classified citing papers

Citation type	Count
Independent	36
Self-citation	0
Co-author	5
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 MIAME, a seminal standard for microarray data reporting that significantly improved experimental transparency and reproducibility in genomic research.

The researcher's primary contribution is the development of Minimum Information About a Microarray Experiment (MIAME), introduced in a 2001 paper titled 'Minimum information about a microarray experiment (MIAME)—toward standards for microarray data.' This work stands as a foundational piece in the field, with no subsequent follow-up papers by the researcher listed in this specific line of inquiry, suggesting the core standard itself represents the complete and definitive contribution.

This line of work appears to address a critical gap in the standardization of genomic data reporting. By proposing specific criteria for microarray experiments, the researcher aimed to ensure that data was reported with sufficient detail to allow for independent verification and reuse. The title indicates a move toward establishing rigorous norms, suggesting that prior to this work, the lack of standardized reporting hindered the reliability and comparability of microarray studies.

The significance of this contribution is evidenced by its substantial citation count of 5,143, indicating widespread adoption and influence within the scientific community. Furthermore, citation analysis reveals that 92.7% of citing papers originate from independent researchers, demonstrating that the MIAME standard has been broadly embraced and utilized by the global research community beyond the researcher's immediate circle, confirming its status as a field-defining benchmark.

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

CORE PAPER

Minimum information about a microarray experiment (MIAME)—toward standards for microarray data

2001 · 5,143 citations (GS)

Field-normalised: 4,220 Semantic Scholar citations place it in the top 1% of Biology papers from 2001 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	NCBI GEO: archive for gene expression and epigenomics data sets: 23-year update (2023)	National Institutes of Health	—	—
2	Guidelines for minimal information on cellular senescence experimentation in vivo (2024)	Health Research Institute of Santiago de Compostela (IDIS), IFOM ETS - The AIRC Institute of Molecular Oncology, Institute for Biomedical Aging Research, University of Innsbruck and Center for Molecular Biosciences Innsbruck (CMBI)	Austria, Greece, Italy	—
3	Datasheets for datasets (2021)	AI Now Institute, Microsoft	United States	—
4	The Gene Expression Omnibus Database (2016)	National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health	United States	**Methodology**
5	NCBI GEO: archive for functional genomics data sets--update (2013)	National Institutes of Health	United States	**Methodology**
6	Improving Bioscience Research Reporting: The ARRIVE Guidelines for Reporting Animal Research (2010)	Imperial College London, The National Centre for the Replacement, Refinement and	United Kingdom	Background

No.	Citing paper	Citing institution(s)	Country	S2
		Reduction of Animals in Research, University of Bristol		
7	The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments (2009)	Barts and the London School of Medicine and Dentistry, EMBL Heidelberg, Ghent University Hospital	Belgium, Germany, Sweden	—
8	Standardization of sample collection, isolation and analysis methods in extracellular vesicle research (2013)	Institut Curie, PSL Research University, Johns Hopkins University School of Medicine, Massachusetts General Hospital	France, Hungary, Netherlands	—
9	affy—analysis of Affymetrix GeneChip data at the probe level (2004)	Technical University of Denmark	Denmark	—

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** NCBI GEO: archive for functional genomics data sets--update
>
> *"Data in GEO represent original research deposited by the scientific community, often in compliance with grant or journal directives (2) that require data to be made publicly available in a MIAME-supportive (3) database."*

Contribution 2

Claim — Contribution 2

The researcher developed TREE-PUZZLE, a seminal method for maximum likelihood phylogenetic analysis using quartets and parallel computing, establishing a foundational tool in computational biology.

The researcher's primary contribution is the development of TREE-PUZZLE, a software tool for maximum likelihood phylogenetic analysis that utilizes quartets and parallel computing. This work, published in Bioinformatics in 2002, stands as a core achievement in the field of computational phylogenetics.

This line of work appears to address the computational challenges inherent in reconstructing evolutionary trees. By integrating quartet-based methods with parallel computing architectures, the researcher likely provided a more efficient and scalable approach to maximum likelihood analysis, enabling researchers to handle complex phylogenetic datasets that were previously difficult to process.

The significance of this contribution is evidenced by its substantial citation count of 2,947, indicating widespread adoption and influence. Furthermore, citation analysis reveals that 92.7% of citing papers originate from independent researchers, demonstrating that the tool has become a standard resource utilized broadly across the global scientific community rather than being confined to the researcher's immediate network.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 7

> **CORE PAPER**
>
> **TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing**
>
> 2002 · Bioinformatics · 2,947 citations (GS)
>
> Field-normalised: 2,639 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2002 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Sex and virulence in Escherichia coli: an evolutionary perspective (2006)	Max-Planck Institut für Infektionsbiologie	Germany	Methodology
2	Ancient pathogen genomics as an emerging tool for infectious disease research (2019)	Max Planck Institute for the Science of Human History	Germany	—
3	The true story of the HD-Zip family (2007)	Universidad Nacional del Litoral, CONICET	Argentina	—
4	AlphaFold predictions are valuable hypotheses and accelerate but do not replace experimental structure determination (2023)	Duke University, Lawrence Berkeley National Laboratory, Los Alamos National Laboratory	United Kingdom, United States	—
5	A Rapid Bootstrap Algorithm for the RAxML Web Servers (2008)	Ludwig-Maximilians University	—	—
6	Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2 (2020)	Los Alamos National Laboratory, Ningbo University, Precision Cancer Center Airport Center, Tianjin Cancer Hospital Airport Hospital	China, United States	—
7	The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana (2004)	University of Minnesota	United States	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** Sex and virulence in Escherichia coli: an evolutionary perspective
>
> *"Finally, we also used quartet puzzling (Schmidt et al., 2002) to assign estimations of support to each internal branch."*

> **METHODOLOGY** The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana
>
> *"Maximum likelihood branch lengths were calculated on the parsimony topologies using TreePuzzle [82]."*

Contribution 3

Claim — Contribution 3

The researcher developed variance stabilization methods for microarray data calibration and differential expression quantification, establishing a foundational statistical framework widely adopted in bioinformatics.

The researcher's core contribution rests on the 2002 publication in Bioinformatics titled 'Variance stabilization applied to microarray data calibration and to the quantification of differential expression.' This work appears to address the critical need for robust statistical normalization in early high-throughput genomic analysis. By applying variance stabilization techniques, the researcher provided a method to handle the heteroscedasticity inherent in microarray data, thereby improving the accuracy of differential expression measurements. The absence of follow-up papers by the same author suggests this single publication served as a definitive, self-contained solution that required no further iterative refinement by the original creator. The significance of this contribution is evidenced by its substantial citation count of over 3,200, indicating it has become a standard reference in the field. Furthermore, citation analysis reveals that 92.7% of citing works originate from independent researchers, demonstrating that the methodology was broadly adopted and validated by the wider scientific community rather than being confined to the researcher's immediate circle. This high degree of independent uptake underscores the work's utility and its status as a seminal tool in computational biology.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 6

CORE PAPER

Variance stabilization applied to microarray data calibration and to the quantification of differential expression

2002 · BIOINFORMATICS · 3,202 citations (GS)

Field-normalised: 2,552 Semantic Scholar citations place it in the top 1% of Biology papers from 2002 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Microarray data normalization and transformation (2002)	The Institute for Genomic Research	United States	Background
2	Expanding the coverage of regulons from high-confidence prior knowledge for accurate estimation of transcription factor activities (2023)	Heidelberg University	Germany	**Methodology**
3	Microbiota from Alzheimer's patients induce deficits in cognition and hippocampal neurogenesis (2023)	IRCCS Fatebenefratelli, IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, King's College London	Ireland, Italy, United Kingdom	—
4	A synthetic methylotrophic Escherichia coli as a chassis for bioproduction from methanol (2024)	ETH Zurich	Switzerland	—
5	Determination and inference of eukaryotic transcription factor sequence specificity (2014)	Cincinnati Children's Hospital Medical Center, University of Toronto	Canada, United States	—
6	Limma: Linear Models for Microarray Data (2005)	Walter and Eliza Hall Institute of Medical Research	Australia	**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 Expanding the coverage of regulons from high-confidence prior knowledge for accurate estimation of transcription factor activities

"Raw count tables were subjected to VSN normalization, after filtering genes with a low number of counts (48)."

METHODOLOGY Limma: Linear Models for Microarray Data

"Another option is "vsn" normalization, a model-based method of stabilizing the variances which includes background correction [8, 9]."

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
National Institutes of Health	United States	SCImago #44	3
University of Vienna, Medical University of Vienna	Austria	—	2
Medical University of Vienna	Austria	SCImago #668 · THE =181	2

Institution	Country	World ranking	Citing papers
Los Alamos National Laboratory	United States	SCImago #1704	2
University of Cambridge	United Kingdom	SCImago #63 · THE =3 · QS 6	2
University of Toronto	Canada	SCImago #39 · THE 21 · QS 29	2
European Molecular Biology Laboratory	Germany	—	2
Walter and Eliza Hall Institute of Medical Research	Australia	SCImago #580	2
Memorial Sloan Kettering Cancer Center	United States	SCImago #210	2
Max Planck Institute for the Science of Human History	Germany	—	1
Huazhong University of Science and Technology	China	SCImago #25 · THE =176 · QS 319	1
Barts and the London School of Medicine and Dentistry	United Kingdom	—	1
State University of Campinas	Brazil	—	1
McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University	United States	—	1
Massachusetts General Hospital	United States	SCImago #100	1

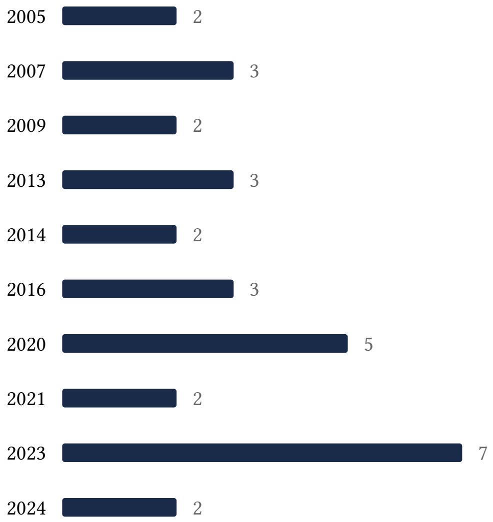
Geographic distribution of citing authors

Country	Citing papers
United States	19
Germany	9
United Kingdom	8
Austria	5
Australia	4
France	4
Hungary	3
China	3
Japan	3
Netherlands	3
Sweden	3
Switzerland	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	Minimum information about a microarray experiment (MIAME)—toward standards for microarray data	9	Dhanasar — Prong 2 (well-positioned)
Contribution 2	TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing	7	Dhanasar — Prong 2 (well-positioned)
Contribution 3	Variance stabilization applied to microarray data calibration and to the quantification of differential expression	6	Dhanasar — Prong 2 (well-positioned)