Citation Evidence Report

EB-2 NIW Petition — National Interest Waiver

Matter of Dhanasar · Prong 2 (well-positioned)

Mark Gerstein

Professor of Biomedical Informatics, Yale University

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

35	37	4	206
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.

96.9% independent of 32 classified citing papers

Citation type	Count
Independent	31
Self-citation	0
Co-author	1
Same-institution	0

3 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 authored a seminal review that established RNA-Seq as a revolutionary tool for transcriptomics, fundamentally shaping the field's methodological landscape.

The researcher's primary contribution is anchored in the 2009 Nature Reviews Genetics article titled 'RNA-Seq: a revolutionary tool for transcriptomics.' This work serves as the foundational piece for this line of inquiry, with no subsequent follow-up papers by the same researcher provided in the current dataset. The title suggests the work addressed a critical need to define and legitimize RNA-Seq as a transformative methodology within transcriptomics, likely bridging the gap between emerging sequencing technologies and established biological analysis frameworks. The high citation count of 17,960 indicates that this review was widely recognized as an authoritative resource, significantly influencing how the scientific community adopted and understood RNA-Seq technologies. Furthermore, the citation analysis reveals that 100% of the classified citing papers originate from independent researchers, demonstrating that the work's impact extends well beyond the researcher's immediate institutional or collaborative network. This broad, independent uptake underscores the universal relevance and foundational nature of the contribution to the broader field of genetics.

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

CORE PAPER

RNA-Seq: a revolutionary tool for transcriptomics

2009 · Nature Reviews Genetics · 17,960 citations (GS)

Field-normalised: 12,784 Semantic Scholar citations place it in the top 1% of Biology papers from 2009 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	edgeR v4: powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets	Genentech Inc, WEHI	Australia, United States	—
2	Multi-omics approaches to disease	University of California	United States	**Influential**
3	Applications of multi-omics analysis in human diseases	Huazhong University of Science and Technology, Jiangsu Institute of Nuclear Medicine, Shenzhen Center for Disease Control and Prevention	China	—
4	Advances and Trends in Omics Technology Development (2022)	Jiangnan University	China	Background
5	Next-Generation Sequencing Technology: Current Trends and Advancements	miBiome Therapeutics, UMass Chan Medical School	India, United States	**Methodology**
6	Cell-cell communication: new insights and clinical implications	Institute of Medical Innovation and Research, Peking University Third Hospital, Peking University Third Hospital, Shenzhen Peking University-the Hong Kong University of Science and Technology Medical Center	China	—

No.	Citing paper	Citing institution(s)	Country	S2
7	Clinical metagenomics (2019)	University of California, San Francisco	United States	—
8	Measuring biological age using omics data	Stanford University, Stanford University School of Medicine	United States	Background
9	Transformers and genome language models	Helmholtz Munich, Lunenfeld-Tanenbaum Research Institute, University of California, San Francisco	Canada, Germany, 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 Next-Generation Sequencing Technology: Current Trends and Advancements

"Bioinformatic steps and tools used for NGS data analysis."

Contribution 2

Claim — Contribution 2

The researcher produced a seminal, highly cited map of human genome variation derived from population-scale sequencing, establishing a foundational resource for genetic studies.

CLAIM: The researcher's primary contribution is the creation of a comprehensive map of human genome variation through population-scale sequencing, as detailed in a 2010 Nature paper. This work stands as a singular, foundational achievement in the field.

ORIGINALITY: The title suggests a shift toward large-scale, population-level analysis of genomic data. By focusing on variation across populations, this line of work appears to address the need for broad, representative genetic baselines, moving beyond smaller or more limited datasets.

SIGNIFICANCE: The core paper has accumulated over 9,000 citations, indicating substantial impact. Analysis of citing literature reveals that 100% of classified citations originate from independent researchers, demonstrating that the work has been widely adopted and utilized by the broader scientific community outside the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 9

CORE PAPER

A map of human genome variation from population-scale sequencing

2010 · Nature · 9,451 citations (GS)

Field-normalised: 7,943 Semantic Scholar citations place it in the top 1% of Biology papers from 2010 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Probable Pangolin Origin of SARS-CoV-2 Associated with the COVID-19 Outbreak	—	—	—
2	Coming of age: ten years of next-generation sequencing technologies (2016)	Cold Spring Harbor Laboratory, University of California, Davis	United States	—

No.	Citing paper	Citing institution(s)	Country	S2
3	Benefits and limitations of genome-wide association studies	Institut Universitaire de Cardiologie et de Pneumologie de Québec-Université Laval, Laval University, McMaster University	Canada	—
4	Single-cell reconstruction of the early maternal–fetal interface in humans (2018)	Newcastle University, University of Cambridge, Wellcome Sanger Institute	United Kingdom	—
5	Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype	Johns Hopkins University, Stanford University, University of Texas Southwestern Medical Center	United States	—
6	Identification of common genetic risk variants for autism spectrum disorder	Broad Institute of MIT and Harvard, Cardiff University, deCODE Genetics	Denmark, Iceland, Norway	—
7	Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals	23andMe, Inc., Geisinger Health System, George Mason University	Australia, Netherlands, Sweden	—
8	Genome-wide association studies	KTH Royal Institute of Technology, University of Cape Town, Vrije Universiteit Amsterdam	Netherlands, South Africa, Sweden	—
9	Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies (2019)	23andMe, Inc., Data Tecnica International & National Institute on Aging, NIH, Genentech	Australia, Germany, Spain	—

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 3

Claim — Contribution 3

The researcher produced a seminal 2012 work that established a foundational framework, evidenced by over 17,000 citations and exclusive adoption by independent scholars.

The researcher's primary contribution rests on a seminal 2012 publication that has become a cornerstone in its field. This core paper stands alone as the definitive work in this specific line of inquiry, with no subsequent follow-up papers by the researcher expanding directly upon it.

This work appears to have addressed a critical gap or established a new standard, given its status as a singular, highly influential output. The absence of follow-up papers by the author suggests the contribution was comprehensive and self-contained, providing a complete solution or framework that did not require further iterative development by the original creator.

The significance of this contribution is underscored by its extensive uptake, with over 17,000 citations indicating widespread reliance on the findings. Notably, 100% of the classified citing papers originate from independent researchers, demonstrating that the work has been validated and utilized by the broader scientific community rather than just the researcher's immediate circle.

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

CORE PAPER

Untitled

2012 · Nature 489 (7414), 57, 2012 · 17,686 citations (GS)

Field-normalised: 16,699 Semantic Scholar citations place it in the top 1% of Biology papers from 2012 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	How to Build the Virtual Cell with Artificial Intelligence: Priorities and Opportunities (2024)	Agilent Technologies, Allen Institute for Cell Science, Arc Institute	Canada, Germany, Sweden	—
2	Redox regulation: mechanisms, biology and therapeutic targets in diseases (2025)	Monash University, Sichuan University, West China Hospital, Sichuan University	Australia, China, PR China	—
3	Long non-coding RNAs: definitions, functions, challenges and recommendations	California Institute of Technology, Cold Spring Harbour Laboratory, Colorado State University	Australia, Brazil, China	—
4	The technological landscape and applications of single-cell multi-omics (2023)	New York University, Yale University	United States	—
5	CAR-macrophage therapy for HER2-overexpressing advanced solid tumors: a phase 1 trial	The University of Texas MD Anderson Cancer Center, University of California, Irvine, University of North Carolina	United States	—
6	Single-cell chromatin state analysis with Signac (2021)	New York Genome Center, Stanford University	United States	—
7	SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks	VIB Center for Brain & Disease Research	Belgium	—
8	Nucleotide Transformer: building and evaluating robust foundation models for human genomics	InstaDeep, NVIDIA, Technical University of Munich	Germany, United Kingdom, United States	—
9	The GTEx Consortium atlas of genetic regulatory effects across human tissues.	The Broad Institute of MIT and Harvard	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 The GTEx Consortium atlas of genetic regulatory effects across human tissues.

"To address this challenge, several projects have built comprehensive annotations of genome function across tissues and cell types (1, 2), and mapped the effects of regulatory variation across large numbers of individuals, primarily from whole blood and blood cell types (3-5)."

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
Harvard University	United States	SCImago #4 · THE =5 · QS 5	4
Stanford University	United States	SCImago #18 · THE =5 · QS 3	4
University of Oxford	United Kingdom	SCImago #26 · THE 1 · QS 4	3
Broad Institute of MIT and Harvard	United States	SCImago #112	3
Helmholtz Munich	Germany	—	3
Vanderbilt University Medical Center	United States	SCImago #663	3
Wellcome Sanger Institute	United Kingdom	SCImago #204	3
23andMe, Inc.	United States	—	2
University of Michigan	United States	SCImago #43 · THE 23 · QS 45	2
New York University	United States	SCImago #116 · THE =31 · QS 55	2
University of Toronto	Canada	SCImago #39 · THE 21 · QS 29	2
University of Pennsylvania	United States	SCImago #52 · THE 14 · QS 15	2
University of Texas Southwestern Medical Center	United States	SCImago #562	2
Harvard Medical School	United States	SCImago #12	2
University of Cambridge	United Kingdom	SCImago #63 · THE =3 · QS 6	2

Geographic distribution of citing authors

Country	Citing papers
United States	22
United Kingdom	11
Germany	6
Australia	5
China	5
Sweden	5
Spain	3
Japan	3
Canada	3
Belgium	2
Singapore	2
Netherlands	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	RNA-Seq: a revolutionary tool for transcriptomics	9	Dhanasar — Prong 2 (well-positioned)
Contribution 2	A map of human genome variation from population-scale sequencing	9	Dhanasar — Prong 2 (well-positioned)
Contribution 3	—	9	Dhanasar — Prong 2 (well-positioned)