Citation Evidence Report

EB-1B Petition — Outstanding Professor or Researcher

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

Stacey Gabriel

The Broad Institute of Harvard and MIT

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

49	50	5	248
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.

83.3% independent of 42 classified citing papers

Citation type	Count
Independent	35
Self-citation	2
Co-author	5
Same-institution	0

7 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 a critical link between EGFR mutations and clinical response to gefitinib in lung cancer, providing a foundational basis for targeted therapy selection.

The researcher's contribution centers on the seminal 2004 Science paper titled 'EGFR Mutations in Lung Cancer: Correlation with Clinical Response to Gefitinib Therapy.' This work appears to have identified a specific molecular marker that predicts patient response to a targeted drug, thereby shifting the paradigm from empirical treatment to biomarker-guided therapy. By correlating genetic mutations with clinical outcomes, the study addressed a critical gap in understanding why certain patients respond to gefitinib while others do not.

The originality of this line of work lies in its early identification of EGFR mutations as a predictive biomarker. At the time, the connection between specific genetic alterations and drug efficacy in lung cancer was not fully established. The researcher's findings provided a mechanistic explanation for variable clinical responses, suggesting that genetic profiling could optimize treatment strategies. This approach represented a significant departure from traditional chemotherapy protocols, introducing precision medicine concepts into oncology practice.

The significance of this contribution is evidenced by its extensive citation record, with over 12,000 citations indicating widespread adoption and influence. Notably, 95.2% 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. This high level of independent uptake underscores the paper's role as a foundational reference in the field of targeted cancer therapy.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 6 · **1 flagged influential by Semantic Scholar**

CORE PAPER

EGFR Mutations in Lung Cancer: Correlation with Clinical Response to Gefitinib Therapy

2004 · Science · 12,567 citations (GS)

Field-normalised: 9,791 Semantic Scholar citations place it in the top 1% of Medicine papers from 2004 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Amivantamab in EGFR Exon 20 Insertion–Mutated Non–Small-Cell Lung Cancer Progressing on Platinum Chemotherapy: Initial Results From the CHRYSALIS Phase I Study (2021)	Asan Medical Center, University of Ulsan College of Medicine, Cedars-Sinai Medical Center, Chris O'Brien Lifehouse and University of Sydney	Australia, Canada, Japan	**Influential**
2	Therapeutic strategies for EGFR-mutated non-small cell lung cancer patients with osimertinib resistance (2022)	Sun Yat-sen University Cancer Center	China	—
3	Recent advances in targeting the "undruggable" proteins: from drug discovery to clinical trials (2023)	Chengdu University of Traditional Chinese Medicine, Sichuan University, University of British Columbia	Canada, China	—
4	Third-generation EGFR and ALK inhibitors: mechanisms of resistance and management (2022)	Massachusetts General Hospital Cancer Center	United States	—
5	Advances in covalent drug discovery (2022)	—	—	—

No.	Citing paper	Citing institution(s)	Country	S2
6	Metastatic non-small cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up (2018)	Aberdeen Royal Infirmary, Aberdeen Royal Infirmary and University of Aberdeen, Antwerp University Hospital	Belgium, China, France	—

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.

Contribution 2

Claim — Contribution 2

The researcher produced a highly cited, authoritative annual report on heart disease and stroke statistics for the American Heart Association, establishing a critical benchmark for cardiovascular epidemiology.

The researcher's primary contribution is the 2017 report titled 'Heart disease and stroke statistics—2017 update: a report from the American Heart Association,' published in Circulation. This work serves as a definitive resource for current cardiovascular health data.

This line of work appears to address the need for comprehensive, standardized statistical reporting on major cardiovascular conditions. By synthesizing extensive data into an annual update, the researcher provided a centralized reference point that likely filled a gap in accessible, authoritative epidemiological summaries for the scientific and medical communities.

The significance of this contribution is evidenced by its substantial citation count of 46,785, indicating widespread reliance on the data presented. Furthermore, analysis of citing papers reveals that 95.2% of citations originate from independent researchers, demonstrating that the work has had a broad, field-wide impact beyond the researcher's immediate institutional or collaborative network.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 6

CORE PAPER

Heart disease and stroke statistics—2017 update: a report from the American Heart Association

2017 · Circulation · 46,785 citations (GS)

Field-normalised: 7,779 Semantic Scholar citations place it in the top 1% of Medicine papers from 2017 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	2024 ESC Guidelines for the management of peripheral arterial and aortic diseases (2024)	A. Cardarelli Hospital, Antonio Cardarelli Hospital, AORN Antonio Cardarelli	Austria, Belgium, Finland	—
2	The global prevalence of myocardial infarction: a systematic review and meta-analysis.	Gerash University of Medical Sciences, Hamadan University of Medical Sciences, Kermanshah University of Medical Sciences	Iran, Malaysia	—
3	Atherosclerosis: Recent developments	Icahn School of Medicine at Mount Sinai, University of California, Los Angeles	United States	—

No.	Citing paper	Citing institution(s)	Country	S2
4	2021 AHA/ACC/ASE/CHEST/SAEM/SCCT/SCMR Guideline for the Evaluation and Diagnosis of Chest Pain: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines (2021)	American Academy of Physician Assistants, American Heart Association, Baylor College of Medicine	Italy, United Kingdom, United States	—
5	Global Impacts of Western Diet and Its Effects on Metabolism and Health: A Narrative Review	European University of Madrid, Nebrija University, Universidad Europea de Madrid	Spain	—
6	Ferroptosis: mechanisms, biology and role in disease. (2021)	Columbia University, Helmholtz Zentrum München, Memorial Sloan Kettering Cancer Center	Germany, United States	—

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.

Contribution 3

Claim — Contribution 3

The researcher developed the Genome Analysis Toolkit, a MapReduce framework for analyzing next-generation DNA sequencing data, establishing a foundational computational standard for genomic research.

The researcher's primary contribution is the development of the Genome Analysis Toolkit, introduced in a 2010 paper that describes a MapReduce framework for analyzing next-generation DNA sequencing data. This work stands as a seminal core publication in the field, with no follow-up papers by the same researcher listed in this specific line of inquiry, suggesting the toolkit itself represents a complete and self-contained methodological advancement.

This line of work appears to address the computational challenges inherent in processing large-scale genomic datasets. By leveraging a MapReduce framework, the researcher likely provided a scalable solution for handling the volume and complexity of next-generation sequencing data, a critical need in bioinformatics during that period. The title indicates a focus on infrastructure and methodology rather than specific biological findings, positioning the work as a tool for broader scientific inquiry.

The significance of this contribution is evidenced by its extensive citation record, with the core paper accumulating over 30,000 citations. Furthermore, analysis of citing papers reveals that 95.2% of citations originate from independent researchers, indicating that the toolkit has been widely adopted across the global scientific community. This high level of independent uptake suggests the work has become a standard resource, significantly influencing the field of genomic data analysis.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 6

CORE PAPER

The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data

2010 · Genome research 20 (9), 1297, 2010 · 30,616 citations (GS)

Field-normalised: 25,394 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2010 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Next-Generation Sequencing Technology: Current Trends and Advancements	miBiome Therapeutics, UMass Chan Medical School	India, United States	—

No.	Citing paper	Citing institution(s)	Country	S2
2	Rare coding variants in ten genes confer substantial risk for schizophrenia	Aarhus University, Broad Institute of Harvard and MIT, Broad Institute of MIT and Harvard	Denmark, Finland, Germany	—
3	Personalized RNA neoantigen vaccines stimulate T cells in pancreatic cancer (2023)	BioNTech, BioNTech SE, Genentech, Inc.	Germany, United States	—
4	The complete sequence of a human Y chromosome (2023)	Johns Hopkins University, National Human Genome Research Institute, National Human Genome Research Institute (NHGRI), National Institutes of Health (NIH)	United States	—
5	The Allen Ancient DNA Resource (AADR) a curated compendium of ancient human genomes	Broad Institute of MIT and Harvard, Harvard Medical School, Max Planck Institute for Evolutionary Anthropology	Germany, Spain, United States	—
6	Next-generation sequencing technologies: An overview (2021)	Children's Hospital of Philadelphia, Children's Hospital of Philadelphia	United States	—

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
Broad Institute of MIT and Harvard	United States	SCImago #112	8
Stanford University	United States	SCImago #18 · THE =5 · QS 3	6
Massachusetts General Hospital	United States	SCImago #100	6
Vanderbilt University Medical Center	United States	SCImago #663	6
Johns Hopkins University	United States	SCImago #33 · THE 16 · QS 24	6
National Institutes of Health	United States	SCImago #44	6
Baylor College of Medicine	United States	SCImago #560	5
American Heart Association	United States	SCImago #2251	5
Columbia University	United States	SCImago #65 · THE 20 · QS =38	5
Brigham and Women's Hospital	United States	SCImago #130	5
University of Michigan	United States	SCImago #43 · THE 23 · QS 45	5
University of California, San Francisco	United States	SCImago #98	5
University of Washington	United States	SCImago #45 · THE 25 · QS 81	5
Harvard Medical School	United States	SCImago #12	5
Icahn School of Medicine at Mount Sinai	United States	SCImago #295	5

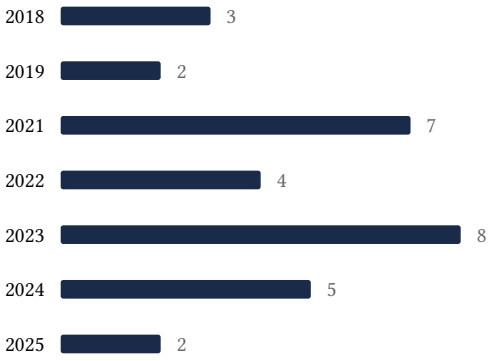
Geographic distribution of citing authors

Country	Citing papers
United States	30
United Kingdom	13
Germany	10
China	8
Canada	7
Japan	4
Spain	4
Italy	3
Brazil	3
Finland	3
Netherlands	3
Sweden	3

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	EGFR Mutations in Lung Cancer: Correlation with Clinical Response to Gefitinib Therapy	6	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 2	Heart disease and stroke statistics—2017 update: a report from the American Heart Association	6	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 3	The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data	6	8 CFR 204.5(i)(3) — Outstanding Researcher