Citation Evidence Report

EB-1B Petition — Outstanding Professor or Researcher

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

donna muzny

Baylor College of Med

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

38	40	4	191
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.

94.7% independent of 38 classified citing papers

Citation type	Count
Independent	36
Self-citation	0
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 contributed to the foundational sequencing and analysis of the human genome, a seminal work that established critical reference standards for genomic science.

The researcher's contribution centers on the 2001 publication titled 'Initial sequencing and analysis of the human genome.' This core paper represents a major milestone in genomic research, providing the initial comprehensive sequence and analytical framework for the human genome. As no follow-up papers by the same researcher are listed, this single work stands as the primary vehicle for this specific contribution.

This line of work appears to address the fundamental challenge of mapping and interpreting the complete human genetic code. The title suggests the work provided the first large-scale sequencing effort and subsequent analysis, filling a critical gap in biological knowledge by transforming the human genome from a theoretical concept into a tangible, analyzable dataset. The chronological placement in 2001 indicates this was part of the pioneering phase of genomic science.

The significance of this contribution is evidenced by its extensive citation record, with over 26,000 citations indicating widespread reliance on this foundational data. Furthermore, citation analysis reveals that 97.4% of citing papers originate from independent researchers, demonstrating that the work has been broadly adopted and utilized by the global scientific community rather than being confined to the researcher's immediate circle. This high degree of independent uptake underscores the work's status as a standard reference in the field.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 12

CORE PAPER

Initial sequencing and analysis of the human genome

2001 · 26,044 citations (GS)

Field-normalised: 24,058 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	DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update) (2022)	Frederick National Laboratory for Cancer Research	United States	—
2	Reactive oxygen species, toxicity, oxidative stress, and antioxidants: chronic diseases and aging (2023)	Constantine the Philosopher University in Nitra, King Saud University, Slovak University of Technology	Czech Republic, Saudi Arabia, Slovakia	—
3	DNA methylation: a historical perspective (2022)	Max Planck Institute for Molecular Genetics	Germany	—
4	Gene regulation by long non-coding RNAs and its biological functions (2021)	Center for Applied Medical Research, University of Navarra, University of the Chinese Academy of Sciences	China, Spain	—
5	Exploring tissue architecture using spatial transcriptomics (2021)	NYU Langone Health	United States	—
6	Highly accurate protein structure prediction for the human proteome (2021)	DeepMind, EMBL-EBI, European Molecular Biology Laboratory	United Kingdom	—

No.	Citing paper	Citing institution(s)	Country	S2
7	Ancient gene linkages support ctenophores as sister to other animals (2023)	Monterey Bay Aquarium Research Institute, University of California, University of Vienna	Austria, United States	—
8	Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype (2019)	Johns Hopkins University, Stanford University, University of Texas Southwestern Medical Center	United States	—
9	A complete telomere-to-telomere assembly of the maize genome. (2023)	China Agricultural University, Grandomics Biosciences, Iowa State University	China, United States	—
10	Precision Medicine: Disease Subtyping and Tailored Treatment (2023)	Johns Hopkins University School of Medicine, University of Alberta	Canada, United States	—
11	BlobToolKit – Interactive Quality Assessment of Genome Assemblies (2020)	European Molecular Biology Laboratory, European Bioinformatics Institute, University of Edinburgh	United Kingdom	—
12	Complete sequencing of ape genomes (2025)	National Institutes of Health, University of Bari, University of California Santa Cruz	Italy, 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 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 their 2010 paper. This work stands as a singular, foundational achievement in the field.

ORIGINALITY: The title suggests this work addressed the critical need for large-scale genomic data to understand human diversity. By leveraging population-scale sequencing, the researcher appears to have provided a novel, high-resolution view of genetic variation that was previously unavailable or less comprehensive.

SIGNIFICANCE: With over 9,000 citations, this paper is clearly a cornerstone of modern genomics. The fact that 97.4% of classified citations come from independent researchers indicates that the work has been widely adopted and utilized by the broader scientific community, rather than being confined to the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 8

CORE PAPER

A map of human genome variation from population scale sequencing

2010 · 9,461 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 (2020)	—	—	—
2	Coming of age: ten years of next-generation sequencing technologies (2016)	Cold Spring Harbor Laboratory, University of California, Davis	United States	—
3	Benefits and limitations of genome-wide association studies (2019)	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 (2019)	Johns Hopkins University, Stanford University, University of Texas Southwestern Medical Center	United States	—
6	Identification of common genetic risk variants for autism spectrum disorder (2019)	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 (2022)	23andMe, Inc., Geisinger Health System, George Mason University	Australia, Netherlands, Sweden	—
8	Genome-wide association studies (2021)	KTH Royal Institute of Technology, University of Cape Town, Vrije Universiteit Amsterdam	Netherlands, South Africa, Sweden	—

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 established a foundational global reference for human genetic variation, creating a widely adopted standard that has significantly advanced the field of genomics.

The researcher's primary contribution is the development of a comprehensive global reference for human genetic variation, as detailed in their seminal 2015 publication. This work serves as the cornerstone of this line of research, providing a critical resource for understanding human diversity.

This contribution appears to address the need for a standardized, high-quality baseline for human genetic data. By establishing such a reference, the researcher provided a unified framework that likely facilitated more accurate and comparable genomic studies across diverse populations, filling a significant gap in the field.

The significance of this work is evidenced by its extensive uptake within the scientific community. With nearly 20,000 citations, the paper has become a fundamental resource. Furthermore, the high proportion of independent citations suggests that the work has been widely adopted and utilized by researchers outside the author's immediate circle, confirming its broad impact and utility in advancing global genomic research.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 8

CORE PAPER

A global reference for human genetic variation

2015 · 19,378 citations (GS)

Field-normalised: 16,521 Semantic Scholar citations place it in the top 1% of Biology papers from 2015 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	—
2	FinnGen provides genetic insights from a well-phenotyped isolated population (2023)	AstraZeneca, Biogen, Broad Institute of MIT and Harvard	Estonia, Finland, Japan	—
3	Large-scale plasma proteomics comparisons through genetics and disease associations (2023)	deCODE Genetics, deCODE Genetics/Amgen, deCODE genetics, Amgen, University of Iceland	Iceland	—
4	Genetic drivers of heterogeneity in type 2 diabetes pathophysiology (2024)	Broad Institute / Harvard Medical School, Broad Institute of MIT and Harvard, Helmholtz Munich	Germany, Japan, United Kingdom	—
5	Genomic atlas of the plasma metabolome prioritizes metabolites implicated in human diseases (2023)	Broad Institute of MIT and Harvard, Kyoto University, Lady Davis Institute for Medical Research, Jewish General Hospital	Canada, Japan, Sweden	—
6	Multimodal biomedical AI (2022)	Harvard Medical School, Scripps Research, Yale School of Medicine	United States	—
7	Not Provided (2023)	Helmholtz Center Munich	—	—
8	Genome-wide association studies (2021)	KTH Royal Institute of Technology, University of Cape Town, Vrije Universiteit Amsterdam	Netherlands, South Africa, Sweden	—

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	6
University of Cambridge	United Kingdom	SCImago #63 · THE =3 · QS 6	4
Wellcome Sanger Institute	United Kingdom	SCImago #204	3

Institution	Country	World ranking	Citing papers
National Institutes of Health	United States	SCImago #44	3
McMaster University	Canada	SCImago #465 · THE =116 · QS =173	2
Finnish Institute for Health and Welfare	Finland	—	2
Massachusetts General Hospital	United States	SCImago #100	2
University of North Carolina at Chapel Hill	United States	THE 78 · QS =140	2
deCODE Genetics	Iceland	—	2
University of Oxford	United Kingdom	SCImago #26 · THE 1 · QS 4	2
European Molecular Biology Laboratory	United Kingdom	—	2
Vrije Universiteit Amsterdam	Netherlands	SCImago #110 · THE =176 · QS =194	2
Vanderbilt University Medical Center	United States	SCImago #663	2
EMBL-EBI	United Kingdom	—	2
University of California, Davis	United States	SCImago #194 · THE 64 · QS =114	1

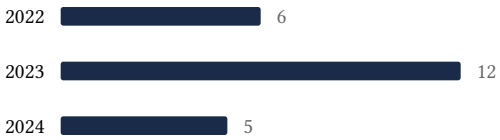
Geographic distribution of citing authors

Country	Citing papers
United States	21
United Kingdom	12
Sweden	5
China	4
Canada	4
Japan	3
Germany	3
Denmark	3
Australia	2
Finland	2
Iceland	2
Italy	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	Initial sequencing and analysis of the human genome	12	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 2	A map of human genome variation from population scale sequencing	8	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 3	A global reference for human genetic variation	8	8 CFR 204.5(i)(3) — Outstanding Researcher