Citation Evidence Report

EB-1A Petition — Original Contributions of Major Significance

8 CFR § 204.5(h)(3)(v) · Criterion 5

Charles Lee, PhD, DSc, FACMG, FRSC

The Jackson Laboratory for Genomic Medicine

Google Scholar profile

Generated 2026-05-21 by CiteMap. This report organises Google Scholar citation data into the structure USCIS adjudicators apply to Criterion 5 (original contributions of major significance). 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

12	15	3	103
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.

100.0% independent of 12 classified citing papers

Citation type	Count
Independent	12
Self-citation	0
Co-author	0
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 produced a seminal, highly cited study mapping human genome variation through population-scale sequencing, establishing a foundational resource for genetic research.

CLAIM: The researcher's primary contribution is the publication of a seminal paper in Nature (2010) that maps human genome variation using population-scale sequencing. This work stands as a singular, high-impact achievement in the field.

ORIGINALITY: The title indicates 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 comprehensive reference maps of human genetic diversity, moving beyond smaller-scale studies.

SIGNIFICANCE: With over 9,000 citations, the paper is clearly highly influential. Analysis of citing literature reveals that 100% of sampled citations come from independent researchers, suggesting the work has been widely adopted and utilized by the broader scientific community rather than just the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 5

CORE PAPER

A map of human genome variation from population scale sequencing

2010 · Nature · 9,448 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	Benefits and limitations of genome-wide association studies	Institut Universitaire de Cardiologie et de Pneumologie de Québec-Université Laval, Laval University, McMaster University	Canada	—
3	Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype	Johns Hopkins University, Stanford University, University of Texas Southwestern Medical Center	United States	—
4	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	—
5	Genome-wide association studies	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 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 established a foundational global reference for human genetic variation, a seminal contribution that has become a standard resource in the field.

CLAIM: The researcher's primary contribution is the creation of a comprehensive global reference for human genetic variation, detailed in a 2015 Nature paper. This work stands as a singular, foundational achievement in the field.

ORIGINALITY: The title suggests the work addressed a critical need for a standardized, worldwide dataset of human genetic diversity. By providing a unified reference, the research appears to have resolved fragmentation in prior studies, offering a new baseline for understanding human genomic variation.

SIGNIFICANCE: With over 19,000 citations, the paper is highly influential. Analysis of citing literature indicates that 100% of sampled citations come from independent researchers, demonstrating broad, field-wide adoption and validating the work's status as an essential, unbiased resource for the global scientific community.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 5

CORE PAPER

A global reference for human genetic variation

2015 · Nature · 19,394 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	FinnGen provides genetic insights from a well-phenotyped isolated population	AstraZeneca, Biogen, Broad Institute of MIT and Harvard	Estonia, Finland, Japan	—
2	A genomic mutational constraint map using variation in 76,156 human genomes	Broad Institute, Broad Institute; Massachusetts General Hospital, Broad Institute of MIT and Harvard	United States	—
3	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	—
4	Multimodal biomedical AI	Harvard Medical School, Scripps Research, Yale School of Medicine	United States	Background
5	Genome-wide association studies	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 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, highly cited integrated map of genetic variation from over 1,000 human genomes, establishing a foundational resource for understanding human genomic diversity.

The researcher's contribution centers on the publication of a seminal core paper titled 'An integrated map of genetic variation from 1,092 human genomes' in 2012. This work stands alone as the primary output in this specific line of inquiry, with no follow-up papers by the same researcher building directly upon it within the provided dataset. The title suggests the creation of a comprehensive, unified resource that synthesizes genetic data from a large cohort, likely aiming to provide a detailed reference for human genomic diversity.

This line of work appears to address the need for a consolidated and high-resolution view of human genetic variation. By integrating data from more than a thousand genomes, the research likely moved beyond fragmented or smaller-scale studies to offer a broader, more robust framework for analyzing genetic differences. The absence of follow-up papers by the researcher indicates that this single publication served as a definitive, standalone resource rather than the start of a prolonged, iterative series of incremental updates by the same author.

The significance of this contribution is underscored by its substantial citation count of 8,858, indicating widespread adoption and influence within the scientific community. Furthermore, analysis of citing papers reveals that 100% of the classified citations originate from independent researchers, excluding the scholar, co-authors, or same-institution colleagues. This high degree of independent citation suggests that the work has been broadly recognized and utilized by the wider field as a critical reference point, validating its impact beyond the researcher's immediate network.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 5

CORE PAPER

An integrated map of genetic variation from 1,092 human genomes

2012 · 8,858 citations (GS)

Field-normalised: 7,980 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	CADD v1.7: using protein language models, regulatory CNNs and other nucleotide-level scores to improve genome-wide variant predictions (2024)	Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Berlin Institute of Health (BIH) at Charité – Universitätsmedizin Berlin, University Hospital Schleswig-Holstein, University of Lübeck	Germany	—
2	Empowering biomedical discovery with AI agents (2024)	Harvard Medical School, Harvard University, Massachusetts Institute of Technology	United States	—
3	Applications of single-cell RNA sequencing in drug discovery and development (2023)	AbbVie Inc., Boehringer Ingelheim Pharmaceuticals Inc., Bristol Myers Squibb	Belgium, France, United Kingdom	—
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	Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype	Johns Hopkins University, Stanford University, University of Texas Southwestern Medical Center	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).

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
Broad Institute of MIT and Harvard	United States	SCImago #112	3
Vrije Universiteit Amsterdam	Netherlands	SCImago #110 · THE =176 · QS =194	2
Harvard University	United States	SCImago #4 · THE =5 · QS 5	2
University of Oxford	United Kingdom	SCImago #26 · THE 1 · QS 4	2
Harvard Medical School	United States	SCImago #12	2
Kuopio University Hospital	Finland	SCImago #4388	1
University of Oulu and Oulu University Hospital	Finland	—	1
Massachusetts General Hospital	United States	SCImago #100	1
University of Cambridge	United Kingdom	SCImago #63 · THE =3 · QS 6	1
University of Pennsylvania	United States	SCImago #52 · THE 14 · QS 15	1
University of Cape Town	South Africa	SCImago #1052 · THE =164 · QS 150	1
AstraZeneca	Sweden	SCImago #244	1
University of Eastern Finland	Finland	SCImago #1834 · THE 401–500 · QS =604	1
University of Tartu	Estonia	SCImago #1820 · THE 301–350 · QS =362	1
Orton Orthopaedic Hospital	Finland	—	1

Geographic distribution of citing authors

Country	Citing papers
United States	8
United Kingdom	5
Sweden	3
Germany	2
Japan	2
Netherlands	2
Australia	1
Canada	1
Belgium	1
South Africa	1
Estonia	1
Finland	1

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	A map of human genome variation from population scale sequencing	5	8 CFR 204.5(h)(3)(v) — Criterion 5

Contribution	Core paper	Indep. cites	Supports
Contribution 2	A global reference for human genetic variation	5	8 CFR 204.5(h)(3)(v) — Criterion 5
Contribution 3	An integrated map of genetic variation from 1,092 human genomes	5	8 CFR 204.5(h)(3)(v) — Criterion 5