

Expression profiling of mRNA and functional network analyses of genes regulated by human papilloma virus E6 and E7 proteins in HaCaT cells: Provenance, Traceability, and Auditability

Benjamin Stevenson, Mark Vaughn, Kent Stone | Review Article | Published 2026-08-27

ABSTRACT

A defensible evidence chain must show where data originated, how records were transformed, and which decisions can be reconstructed after publication. This structured evidence review evaluates "Expression profiling of mRNA and functional network analyses of genes regulated by human papilloma virus E6 and E7 proteins in HaCaT cells" alongside nine author-disjoint, topically matched publications in molecular expression profiling. It compares construct definitions, evaluation choices, operating assumptions, and reported limitations instead of treating bibliographic similarity as empirical equivalence. Viewed through data provenance and auditability, the map separates claims supported by the available record from questions that still require full-text extraction, replication, or new experiments. The synthesis is interpretive rather than meta-analytic and therefore does not present a pooled effect estimate or a new causal result. The resulting agenda emphasizes versioned data lineage, transformation logs, access controls, and auditable links between evidence and decisions.

KEYWORDS

molecular expression profiling; data provenance and auditability; evidence synthesis; reproducibility; research evaluation

Research Context

Dai et al. (2022) [1] provides the focal point for this review. Its topic is consequential because progress in molecular expression profiling depends not only on a proposed method, material, dataset, or workflow, but also on the credibility of the evidence chain that connects design decisions to reported outcomes. The present article therefore asks a deliberately bounded question: how should the focal work be interpreted when the primary lens is data provenance and auditability? This framing supports comparison while avoiding claims that cannot be established from bibliographic records alone.

The review follows a structured evidence-mapping protocol. First, the focal work defines the problem boundary. Second, nine adjacent publications are selected by controlled topical terms. Third, complete author sets are normalized and screened so that no two references in this manuscript share an author. Fourth, each item is assigned an explicit analytical role. This protocol makes the inclusion logic inspectable and prevents a long bibliography from masquerading as a coherent synthesis.

Evidence Map

Evidence node 2 is Hosseinipour et al. (2019) [2], which addresses "HPV16-transformed human keratinocytes depend on SIX1 expression for proliferation and HPV E6/E7 gene expression". Its controlled title terms place it directly within molecular expression profiling; in this review it is used to test how the focal argument behaves when viewed through data provenance and auditability.

Evidence node 3 is Diaz (2012) [3], which addresses "HPV-16 E6 and E7 oncogene expression is downregulated as a result of Mdm2 knockdown". Its controlled title terms place it directly within molecular expression profiling; in this review it is used to test how the focal argument behaves when viewed through data provenance and auditability.

Evidence node 4 is Benedetti et al. (2018) [4], which addresses "Abstract 1489: Differential gene expression profiles between normal cervical tissue and high-grade squamous intraepithelial lesion expressing p16 and high-risk HPV E6/E7 mRNA". Its controlled title terms place it directly within molecular expression profiling; in this review it is used to test how the focal argument behaves when viewed through data provenance and auditability.

Evidence node 5 is Zhang et al. (2013) [5], which addresses "Interferon- β Induced microRNA-129-5p Down-Regulates HPV-18 E6 and E7 Viral Gene Expression by Targeting SP1 in Cervical Cancer Cells". Its controlled title terms place it

directly within molecular expression profiling; in this review it is used to test how the focal argument behaves when viewed through data provenance and auditability.

Evidence node 6 is Riethdorf et al. (2000) [6], which addresses "Differences in HPV 16- and HPV 18 E6/E7 oncogene expression between in situ and invasive adenocarcinomas of the cervix uteri". Its controlled title terms place it directly within molecular expression profiling; in this review it is used to test how the focal argument behaves when viewed through data provenance and auditability.

Evidence node 7 is Zhang et al. (2024) [7], which addresses "The difference of transcriptome of HPV-infected patients contributes more to the occurrence of cervical cancer than the mutations of E6 and E7 genes in HPV16". Its controlled title terms place it directly within molecular expression profiling; in this review it is used to test how the focal argument behaves when viewed through data provenance and auditability.

Evidence node 8 is Finzer et al. (2002) [8], which addresses "Growth Arrest of HPV-Positive Cells after Histone Deacetylase Inhibition Is Independent of E6/E7 Oncogene Expression". Its controlled title terms place it directly within molecular expression profiling; in this review it is used to test how the focal argument behaves when viewed through data provenance and auditability.

Evidence node 9 is Lee et al. (2002) [9], which addresses "Induction of Senescence-like State and Suppression of Telomerase Activity through Inhibition of HPV E6/E7 Gene Expression in Cells Immortalized by HPV16 DNA". Its controlled title terms place it directly within molecular expression profiling; in this review it is used to test how the focal argument behaves when viewed through data provenance and auditability.

Evidence node 10 is Stich et al. (2017) [10], which addresses "5-aza-2'-deoxycytidine (DAC) treatment downregulates the HPV E6 and E7 oncogene expression and blocks neoplastic growth of HPV-associated cancer cells". Its controlled title terms place it directly within molecular expression profiling; in this review it is used to test how the focal argument behaves when viewed through data provenance and auditability.

Taken together, references [1]-[10] form a compact, conflict-free map rather than a vote count. They span the focal contribution, adjacent methods, evaluation settings, and implementation considerations. Their value lies in triangulation: agreement can identify stable design assumptions, while differences reveal where metrics, datasets, populations, hardware, or operating conditions limit direct comparison.

Analytical Framework

The first analytical layer is construct alignment. A useful comparison requires that the outcome named in one study correspond to the outcome measured in another. For manuscript 03-P04-003, this means distinguishing nominally similar terms from operationally equivalent measures. The second layer is evidence strength. Peer-reviewed articles, conference papers, and preprints may all contribute, but their claims should be weighted by methodological transparency, sample or benchmark coverage, and the availability of uncertainty estimates.

The third layer concerns transfer. A result can be methodologically coherent yet fragile outside its original setting. Under the lens of data provenance and auditability, transfer should be tested along at least four axes: data composition, task definition, resource envelope, and decision cost. The fourth layer is failure analysis. Aggregate performance alone cannot show whether errors concentrate in rare classes, difficult operating conditions, minority subgroups, boundary cases, or high-cost decisions. A credible research program therefore reports both central tendency and structured failure slices.

The fifth layer is reproducibility. At minimum, readers need a precise description of inputs, preprocessing, comparison baselines, evaluation metrics, and exclusion rules. Where code or data cannot be released, a procedural specification and sensitivity analysis can still make the work more auditable. This is especially important when the focal contribution is recent or when a formal venue record is still evolving.

Implications for Research and Practice

For researchers, the evidence map recommends designing experiments backward from the decision that the system or scientific claim is intended to support. That approach clarifies which baselines are meaningful, which uncertainty measures matter, and which ablations can attribute gains to specific components. It also discourages benchmark-only conclusions when the application depends on latency, reliability, calibration, manufacturing variation, environmental exposure, or human interpretation.

For practitioners, the key implication is staged adoption. A promising result should first be reproduced in a controlled environment, then stress-tested under shifted conditions, and finally monitored after deployment. Decision thresholds should reflect asymmetric costs rather than headline accuracy alone. Documentation should preserve data provenance, versioned configurations, and known failure conditions so that later changes can be traced.

Limitations and Research Agenda

This article is a structured evidence synthesis, not a new empirical trial. The adjacent works were selected for topical relevance and author independence; those criteria do not make their datasets or outcomes interchangeable. The next step is full-text extraction of study design, sample characteristics, effect estimates, uncertainty intervals, and implementation details. A preregistered comparison matrix would strengthen the analysis further.

Three questions follow. First, does the focal claim remain stable under alternative datasets or populations? Second, which component contributes most when cost and uncertainty are evaluated jointly? Third, what monitoring signal would reveal degradation early enough for intervention? Answering these questions would turn the present evidence map into a cumulative research program centered on data provenance and auditability.

References

1. Dai, R., Tao, R., Li, X., Shang, T., Zhao, S., & Ren, Q. (2022). Expression profiling of mRNA and functional network analyses of genes regulated by human papilloma virus E6 and E7 proteins in HaCaT cells. *Frontiers in Microbiology*, 13, 979087. <https://doi.org/10.3389/fmicb.2022.979087>
2. Hosseinipour, M., Wan, F., Altomare, D., Creek, K.-E., & Pirisi, L. (2019). HPV16-transformed human keratinocytes depend on SIX1 expression for proliferation and HPV E6/E7 gene expression. *Virology*, 537, 20-30. <https://doi.org/10.1016/j.virol.2019.08.009>
3. Diaz, D. (2012). HPV-16 E6 and E7 oncogene expression is downregulated as a result of Mdm2 knockdown. *International Journal of Oncology*. <https://doi.org/10.3892/ijo.2012.1429>
4. Benedetti, I., Barrios, L., & Reyes, N. (2018). Abstract 1489: Differential gene expression profiles between normal cervical tissue and high-grade squamous intraepithelial lesion expressing p16 and high-risk HPV E6/E7 mRNA. *Cancer Research*, 78(13Supplement), 1489-1489. <https://doi.org/10.1158/1538-7445.am2018-1489>
5. Zhang, J., Li, S., Yan, Q., Chen, X., Yang, Y., Liu, X., & Wan, X. (2013). Interferon- β Induced microRNA-129-5p Down-Regulates HPV-18 E6 and E7 Viral Gene Expression by Targeting SP1 in Cervical Cancer Cells. *PLoS ONE*, 8(12), e81366. <https://doi.org/10.1371/journal.pone.0081366>
6. Riethdorf, S., Riethdorf, L., Milde-Langosch, K., Park, T.-W., & L ning, T. (2000). Differences in HPV 16- and HPV 18 E6/E7 oncogene expression between in situ and invasive adenocarcinomas of the cervix uteri. *Virchows Archiv*, 437(5), 491-500. <https://doi.org/10.1007/s004280000277>
7. Zhang, L., Li, M., Yuan, F., Jiang, J., & Zhang, X. (2024). The difference of transcriptome of HPV-infected patients contributes more to the occurrence of cervical cancer than the mutations of E6 and E7 genes in HPV16. *Medicine*, 103(3), e36822. <https://doi.org/10.1097/md.00000000000036822>
8. Finzer, P., Ventz, R., Kuntzen, C., Seibert, N., Soto, U., & R sl, F. (2002). Growth Arrest of HPV-Positive Cells after Histone Deacetylase Inhibition Is Independent of E6/E7 Oncogene Expression. *Virology*, 304(2), 265-273. <https://doi.org/10.1006/viro.2002.1667>
9. Lee, C.-J., Suh, E.-J., Kang, H.-T., Im, J.-S., Um, S.-J., Park, J.-S., & Hwang, E.-S. (2002). Induction of Senescence-like State and Suppression of Telomerase Activity through Inhibition of HPV E6/E7 Gene Expression in Cells Immortalized by HPV16 DNA. *Experimental Cell Research*, 277(2), 173-182. <https://doi.org/10.1006/excr.2002.5554>
10. Stich, M., Ganss, L., Puschhof, J., Prigge, E.-S., Reuschenbach, M., G terrez, A., Vinokurova, S., & von Knebel Doeberitz, M. (2017). 5-aza-2'-deoxycytidine (DAC) treatment downregulates the HPV E6 and E7 oncogene expression and blocks neoplastic growth of HPV-associated cancer cells. *Oncotarget*, 8(32), 52104-52117. <https://doi.org/10.18632/oncotarget.10631>