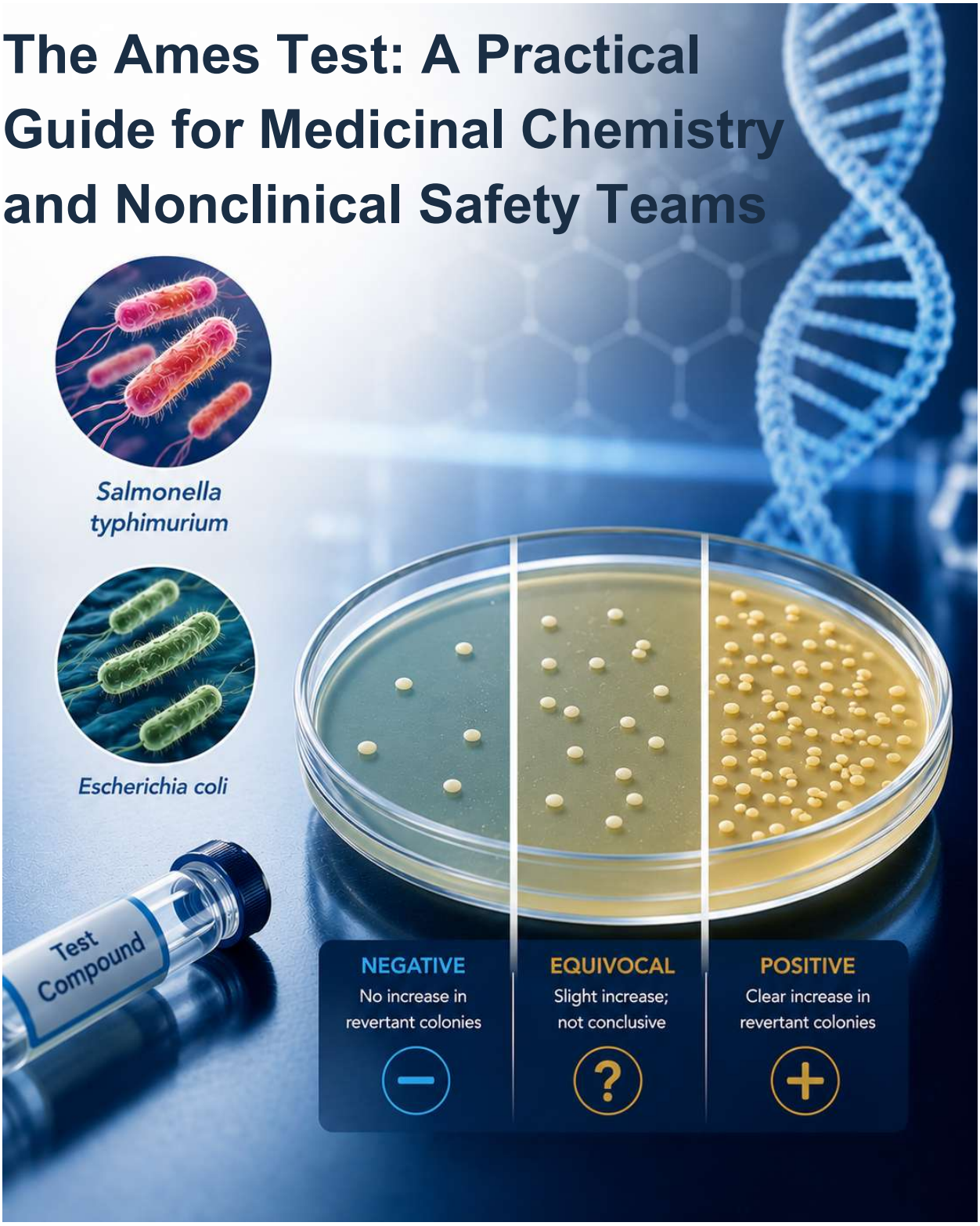
Integrated PharmaConsulting

AUTHOR: Chad Hopkins, PhD

WHITE PAPER TOPIC: Genetic Toxicology | JULY 2026

The Ames Test: A Practical Guide for Medicinal Chemistry and Nonclinical Safety Teams



Executive Summary

The Ames test is an in vitro bacterial reverse mutation assay that can carry substantial consequences for a program. It is the primary in vitro gene mutation assay in the ICH S2(R1) battery, and under ICH M7(R2) it is the confirmatory test that decides whether a structural alert on an impurity is true or false. A positive result often stops the compound's advancement entirely or at best narrows the indications, typically to oncology or other life-threatening diseases.

This white paper, prepared by Integrated Pharma Consulting, covers how the assay works, how to stage strain selection between lead optimization and the definitive GLP study, what the published evidence says about redundancy in the five-strain battery, how to read the three possible outcomes (including the equivocal one) and what follow-ups an Ames-positive result actually requires.

KEY INSIGHTS

- Triage with two strains (TA98 + TA100) for lead optimization, then the full five-strain battery for the GLP study.
- Equivocal can go either way. Repeat with refined doses or method; if still unresolved, treat as positive.
- Verify the positive before starting any follow-up study. Impurities, reproducibility, bacteria-specific mechanisms, and free histidine/tryptophan in the sample can cause false positives.
- Follow-up depends on what was tested. API, metabolite, and impurity have different tracks.
- Match the endpoint to the finding. An Ames positive reports small-scale gene mutation. The follow-up must measure mutation, not the large-scale chromosomal damage that micronucleus and comet detect. These are not interchangeable.

1. How the Assay Works

Each *Salmonella* tester strain carries a mutation in the histidine operon that leaves it unable to synthesize histidine (his^-). Plated on histidine-free agar, it cannot grow. A *reverse* mutation, a second mutation that restores the gene's function, returns the ability to make histidine (his^+), and only those revertant cells go on to form colonies. More colonies equal more mutations.

Expose the strains to a range of doses, with and without metabolic activation, and a dose-related rise in revertant colonies above the historical control range indicates mutagenicity. *E. coli* WP2 uvrA works the same way, but on tryptophan rather than histidine (Figure 1).

2. Strain Strategy: Triage First, Then the Definitive Battery

Strain selection depends on the question being asked. During lead optimization, two strains are generally sufficient, whereas for definitive studies, the full battery is required.

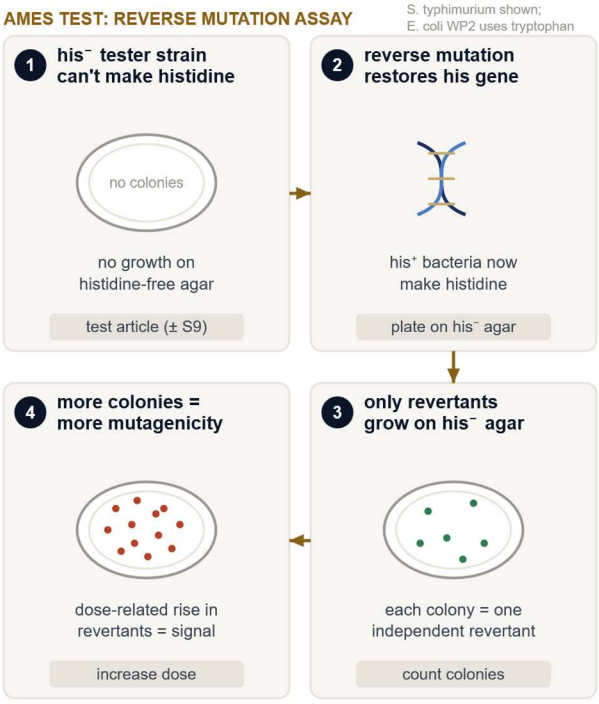


Figure 1. 1) A his^- tester strain cannot grow on histidine-free agar; 2) Reverse mutation restores the gene; 3) Each resulting colony represents one independent mutational event; 4) A dose-related increase in revertants indicates mutagenicity

2.1 Step 1 · Triage (Non-GLP)

TA98 + TA100: the mini-Ames. Two strains detect the large majority of bacterial mutagens while consuming less compound, time, and money. That makes them appropriate for lead-optimization where the goal is to rank and deselect chemotypes. Sponsors wanting broader early coverage can add a third strain, *E. coli* WP2 uvrA or TA102, to pick up oxidative and cross-linking mutagens at A·T sites before committing to the definitive study.

A single-strain screen is a false economy. The two mutation classes the mini-Ames assay detects are mechanistically distinct, and one strain cannot see both.

2.2 Step 2 · Definitive (GLP)

The full five-strain battery. TA1535, TA1537, TA98, TA100, plus *E. coli* WP2 uvrA or TA102, run with and without S9. This is the study that supports IND submission, and OECD TG 471 (2020) requires all five. Triage results do not substitute for the full battery.

2.3 Metabolic Activation (± S9)

Bacteria do not metabolize xenobiotics in the same manner as mammals, therefore a compound whose mutagenic species is a metabolite would go undetected. The assay compensates with S9: a post-mitochondrial liver fraction from rats treated with Aroclor, a polychlorinated biphenyl mixture that induces P450. Running each strain both with and without S9 allows a single assay to interrogate the parent and its metabolites simultaneously.

> **CAUTION**
> It is important to keep in mind that human-specific metabolites will be missed in the presence of rat S9. Inversely, a rat-specific metabolite could trigger a false positive. It's advisable to compare results against a cross-species metabolite ID study.

3. What Each Strain Detects

The value of the battery is easier to see from the chemistry side. Each strain is tuned to a mutational mechanism, and each mechanism has recognizable structural signatures (Table 1). Reviewing the strain-specific positive can provide insight into the offending structural feature for future SAR.

4. How Much Do the Extra Strains Add?

If TA98 and TA100 catch most mutagens, it is fair to ask what the other three are doing. The International Workshops on Genotoxicity Testing (IWGT) addressed this directly.

Williams and colleagues (2019) reported that **TA98 and TA100 alone detect 93% of TG 471 bacterial mutagens**, rising to **99%** when paired with an in vitro clastogenicity assay such as chromosomal aberration or the mouse lymphoma assay (MLA). They found that removing TA1535, TA1537, TA102, and WP2 uvrA would miss only **2 mutagens out of the more than 1,000** they started with, but with two critical conditions: it holds only after the reported positives are re-examined statistically, *and* only once responses in other common in vitro genotoxicity assays are taken into account. If a broader panel is desired, the authors suggest TA100 over TA1535, TA97 over TA1537, and WP2 uvrA pKM101 over both WP2 uvrA and TA102.

The mechanism behind the pattern is noteworthy. The pKM101 plasmid carries *mucAB*, which drives error-prone SOS repair and raises sensitivity. TA100 is essentially TA1535 plus pKM101, which is precisely why it out-detects TA1535. The same logic makes WP2 uvrA pKM101 a better performer than either WP2 uvrA or TA102.

> **CAUTION**
> Williams et al. did propose that TA1535, TA1537, TA102, and WP2 uvrA could be removed from TG 471 with little loss of sensitivity. That recommendation has not been adopted: TG 471 as revised in 2020 still calls for at least five strains. The evidence is a sound basis for designing a focused triage; not a basis for reducing the definitive battery.

However, there are chemotype-specific blind spots worth mentioning. Among the mutagenic oximes in the database they searched, those picked up by TA1535 but not TA100 account for more than 40%, and for N- and O-allyl compounds the figure is above a third. On their own account, more than a third of the mutagens in those classes would be missed if TA1535 were not in the battery. It's important to note that if these types of structures are being screened, the standard mini-Ames may need to be modified. Meanwhile, low-background strains combined with the 2–3× fold-increase rule can trigger weak positives that a trend test will not confirm, which is why many reported strain-specific positives were recategorized under Williams et al.'s statistical re-analysis. Sensitivity and specificity can move in opposite directions, and the full battery is a compromise between the two.

5. Reading the Result

Three outcomes are possible, each evaluated with and without S9 (Figure 2). The positive control validates the assay in each, including the negative, which only provides an interpretable negative result if the assay demonstrably worked.

The outcome that causes the most confusion is **equivocal; not quite positive, not quite negative.** A slight rise that does not clearly fall below the historical negative control's upper range, or lacks a clear dose-response, should be repeated with refined doses or an adjusted method. If it stays unresolved, weigh the evidence and manage the compound as a positive. Treating an unresolved equivocal as a negative risks the program discovering a problem later in development.

Table 1. What each strain detects, and the structural features that tend to trigger it.

Mutation class	Strains	What it means	Example Structural Features
Base-pair substitution	TA100 · TA1535	One base swapped for another (G·C→A·T), giving a single miscoded codon.	Direct electrophiles and alkylators: epoxides, aziridines, sulfonate esters, alkyl halides, Michael acceptors.
Frameshift	TA98 · TA1537	A base pair inserted or deleted, causing the reading frame to shift and everything downstream is read incorrectly.	Flat, planar polyaromatics that intercalate the helix and cause slippage: fused aromatics, acridines, nitroaromatics.
Base-pair substitution Oxidative & cross-linking	WP2 uvrA · TA102	ROS base damage and strand breaks, or covalent bridges tying the strands together.	Redox-active or bifunctional groups: quinones, catechols, nitroaromatics, hydrazines, cross-linkers.

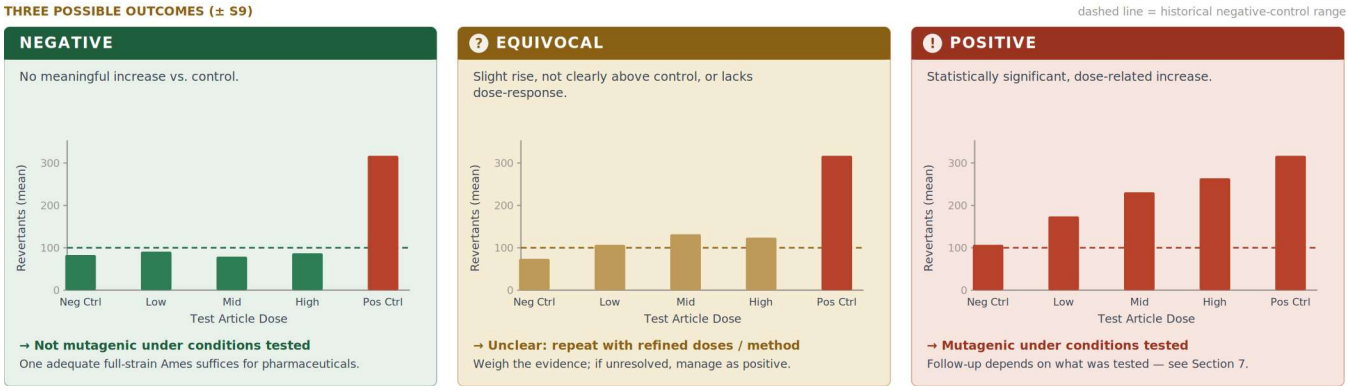


Figure 2. The three outcomes based on revertant counts across a dose range. The dashed line marks the historical negative-control upper range. The positive control (red in all three panels) validates the assay regardless of the test article's result.

6. A Positive Result: What Next?

6.1 First, Confirm the Positive

Before running additional follow-up studies, establish that the positive is real. A common reason for a false positive is impurities, residual solvents, or residual elements. Test using high-purity material (i.e. >99%) and confirm absence of residual solvents and elements that may trigger a false positive.

Additionally, the FDA's 2024 draft guidance and Levy et al. (2019) provide recommendations for evaluating bacterial mutagenicity data that can lower concern without any new in vivo work: (Q)SAR analysis and read-across to existing carcinogenicity data; a demonstrably bacteria-specific mechanism such as bacterial nitro reduction; an artifact from free histidine or tryptophan in the sample; or rat-specific metabolism with no human relevance. (Zeller et al. propose azides as another classic bacteria-specific mechanism). Each of these is worth evaluation since a true positive and human-relevant Ames result typically means termination of the compound for most indications as continued development can become costly and burdensome for the Sponsor.

EXCLUDED COHORTS OF CONCERN

! N-nitroso, polyaromatic hydrocarbon, aflatoxin-like, and alkyl-azoxy structures are outside the follow-up scheme. There are no FIH healthy-subject trials appropriate for these cohorts. The FDA draft does not apply to biologics, advanced-cancer therapeutics, and mutagenic impurities.

6.2 Then, Ask What Was Actually Tested

The next question to ask is whether the positive is on the drug substance (i.e. API), on a metabolite, or on a process impurity. Depending on the answer, there are three different frameworks that govern the path forward (Figure 3).

If the Ames result is determined to be a true positive on the API, this often terminates advancement of the compound, particularly if it falls outside of oncology (ICH S9), due to the burdens required to de-risk. If the Sponsor chooses to continue development, the API route aligns with ICH S2(R1) and the FDA draft: an in vitro mammalian gene mutation assay (MLA or hypoxanthine-guanine phosphoribosyl transferase (HPRT)) first, then in vivo gene mutation work (TGR under OECD 488, and/or *Pig-a* under OECD 470) if the in vitro results were negative. Colony size is important for the MLA, and careful interpretation is required. Large colonies indicate a gene-mutation event rather than a clastogenic one (e.g. small colonies). Positive or equivocal results in these follow-up assays would preclude dosing in healthy subjects.

Mutagenic metabolites are typically captured in the Ames assay when performed in the presence of Aroclor-induced rat S9, but be aware of rat- or human-specific or disproportionate metabolites that may trigger a false positive or negative. An Ames-positive metabolite present at >10% of total drug-related exposure is generally evaluated much as the API would be under ICH M3(R2). Because an Ames positive correlates strongly with a positive two-year rodent assay, Robison et al. (2021) and the FDA draft guidance handle an Ames-positive metabolite at ≤10% of total drug-related exposure on a **case-by-case basis**, using an approach comparable to ICH M7(R2) for low-level mutagenic impurities. Note that the threshold of toxicological concern (TTC) itself cannot be applied to a metabolite, since tissue exposure cannot be quantified precisely enough or controlled. Zeller et al. (2020) outline an “Aware-Avoid-Assess” paradigm with instructive case examples.

The impurity route follows ICH M7(R2) and is briefly discussed by Robison et al. (2021), but notably sits outside the FDA draft's scope entirely. The draft explicitly excludes DNA-reactive impurities. The Ames test can further categorize a (Q)SAR alert into Class 2 (i.e. control ≤TTC) or Class 5 (i.e. treat as non-mutagenic). Note that nitrosamines are handled separately and are the subject of continued discussions.

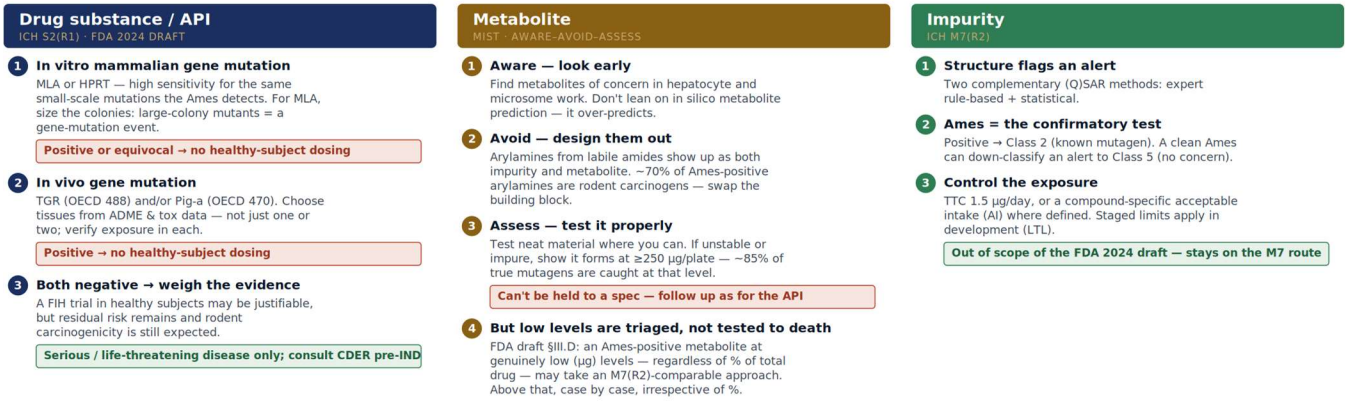


Figure 3. Follow-up depends on what was tested: three routes under three frameworks.

7. Managing Divergence

There is an open disagreement in the literature about whether the micronucleus and comet assays are appropriate follow-ups to an Ames positive. It is worth noting because it is critical to answer the appropriate question without performing unnecessary studies. It is also important to remember that genotoxic does not equal mutagenic. While all mutagens are genotoxic, not all genotoxic substances are mutagenic.

Remember that mutagenic refers to the ability to induce gene mutations (base-pair substitutions and frameshifts), and genotoxic is broader and includes clastogenic and aneugenic agents, which underlies the following.

The disagreement is not about which assay is better *per se*, but about **whether the endpoint matches the finding** (Figure 4). Robison and colleagues (2021), a group of FDA authors from CDER, NCTR, and CDRH, presented the argument that is restated in the FDA draft guidance.

There are two primary endpoints discussed. **Small-scale** damage is base-pair substitutions and frameshifts from small insertions or deletions (Ames test). **Large-scale** damage is chromosomal effects initiated by a DNA strand break or interaction with cellular targets other than DNA: clastogenicity and aneugenicity, respectively. TGR assays and Pig-a detect small-scale damage, and micronucleus and comet detect large-scale damage.

FDA 2024 DRAFT

Micronucleus and comet measure large-scale chromosomal damage, not the small-scale gene mutation an Ames positive flags. The draft directs sponsors to MLA or HPRT, then TGR and/or Pig-a instead. Keep in mind this is still in draft form as of publication of this white paper, and therefore any recommendations could change based on final guidance.

WHY THE TWO LOOK INTERCHANGEABLE — AND AREN'T

Genotoxic substances can cause both small- and large-scale damage. Surveys of the published literature therefore show an apparent parity between the two assay classes. Robison et al. argue that parity is a selection artifact of which compounds get studied, not evidence that the assays are equivalent. A compound that causes only small-scale damage would be missed by a micronucleus / comet package.

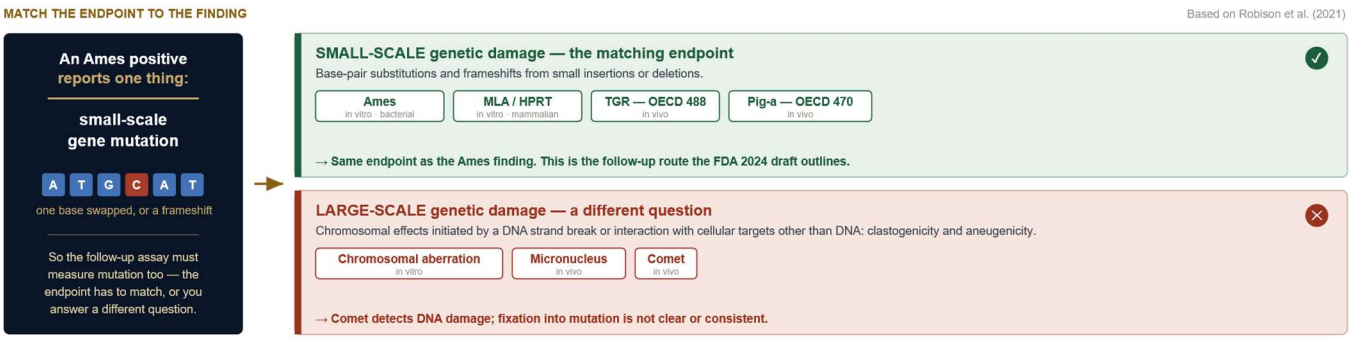


Figure 4. The argument underneath the FDA position. Genetic damage comes in two forms and the assays that detect them are not necessarily interchangeable.

8. Closing Summary

The Ames test rewards programs that treat it as a design tool rather than a checkbox. Five points are worth carrying forward:

1. **Skip the single-strain screen.** Triage with TA98 + TA100, then run the full five-strain battery for the GLP study. Two strains for a screen; five for a submission.
2. **An equivocal Ames is not a negative.** Repeat with refined doses or method. If it stays unresolved, manage it as a positive.
3. **Qualify the positive before you spend on follow-ups.** Impurity issues, reproducibility, bacteria-specific mechanisms, and histidine/tryptophan artifacts may resolve many calls.
4. **Follow-up depends on what was tested.** API, metabolite, and impurity are three different tracks. Low-level metabolites may be triaged through an ICH M7 approach.
5. **Match the endpoint to the finding.** An Ames positive is small-scale gene mutation, and therefore the follow-up must measure mutation. That is why the guidance diverges on micronucleus / comet and why inclusion of relevant questions in a pre-IND meeting can be useful for alignment.

Used strategically and interpreted carefully, the assay does more than act as a gatekeeper: it informs a medicinal chemistry team which structural feature to change. Leverage the results as a design tool.

References

1. OECD. *Test No. 471: Bacterial Reverse Mutation Test*, OECD Guidelines for the Testing of Chemicals, Section 4. OECD Publishing, Paris, 2020. Adopted 26 June 2020. https://doi.org/10.1787/9789264071247-en

2. ICH Harmonised Tripartite Guideline. *Guidance on Genotoxicity Testing and Data Interpretation for Pharmaceuticals Intended for Human Use S2(R1)*. International Conference on Harmonisation, Step 4, November 2011; issued as FDA guidance for industry, June 2012.

3. ICH Harmonised Guideline. *Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk M7(R2)*. International Council for Harmonisation, 2023; issued as FDA guidance for industry, July 2023.

4. ICH Harmonised Tripartite Guideline. *Nonclinical Evaluation for Anticancer Pharmaceuticals S9*. International Conference on Harmonisation, Step 4, October 2009; issued as FDA guidance for industry, March 2010.

5. ICH Harmonised Tripartite Guideline. *Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals M3(R2)*. International Conference on Harmonisation, Step 4, June 2009; issued as FDA guidance for industry, January 2010.

6. U.S. Food and Drug Administration, Center for Drug Evaluation and Research. *Recommended Follow-Up Testing for an Ames-Positive Drug (Active Ingredient) or Metabolite To Support First-in-Human Clinical Trials With Healthy Subjects*. Guidance for Industry and Review Staff, DRAFT, November 2024. Docket FDA-2024-D-4540. Nonbinding; draft status verified July 2026.

7. Williams, R. V.; DeMarini, D. M.; Stankowski, L. F., Jr.; Escobar, P. A.; Zeiger, E.; Howe, J.; Elespuru, R.; Cross, K. P. Are All Bacterial Strains Required by OECD Mutagenicity Test Guideline TG471 Needed? *Mutat. Res. Genet. Toxicol. Environ. Mutagen.* **2019**, *848*, 503081. https://doi.org/10.1016/j.mrgentox.2019.503081

8. Levy, D. D.; Zeiger, E.; Escobar, P. A.; Hakura, A.; van der Leede, B.-J. M.; Kato, M.; Moore, M. M.; Sugiyama, K.-I. Recommended Criteria for the Evaluation of Bacterial Mutagenicity Data (Ames Test). *Mutat. Res. Genet. Toxicol. Environ. Mutagen.* **2019**, *848*, 403074. https://doi.org/10.1016/j.mrgentox.2019.07.004

9. Zeller, A.; Brigo, A.; Brink, A.; Guerard, M.; Lang, D.; Muster, W.; Runge, F.; Sutter, A.; Vock, E.; Wichard, J.; Schadt, S. Genotoxicity Assessment of Drug Metabolites in the Context of MIST and Beyond. *Chem. Res. Toxicol.* **2020**, *33* (1), 10–19. https://doi.org/10.1021/acs.chemrestox.9b00348

10. Robison, T. W.; Heflich, R. H.; Manjanatha, M. G.; Elespuru, R.; Atrakchi, A.; Mei, N.; Ding, W. Appropriate In Vivo Follow-Up Assays to an In Vitro Bacterial Reverse Mutation (Ames) Test Positive Investigational Drug Candidate (Active Pharmaceutical Ingredient), Drug-Related Metabolite, or Drug-Related Impurity. *Mutat. Res. Genet. Toxicol. Environ. Mutagen.* **2021**, *868–869*, 503386. https://doi.org/10.1016/j.mrgentox.2021.503386

11. OECD. *Test No. 488: Transgenic Rodent Somatic and Germ Cell Gene Mutation Assays*, OECD Guidelines for the Testing of Chemicals, Section 4. OECD Publishing, Paris, 2022. https://doi.org/10.1787/9789264203907-en

12. OECD. *Test No. 470: Mammalian Erythrocyte Pig-a Gene Mutation Assay*, OECD Guidelines for the Testing of Chemicals, Section 4. OECD Publishing, Paris, 2022. https://doi.org/10.1787/4faea90e-en

If your program is approaching development candidate nomination, working through an Ames-positive finding, or building a genotoxicity strategy ahead of a pre-IND meeting, reach out to Integrated Pharma Consulting at info@integratedpharmaconsulting.com. We have experience designing genotoxicity and impurity qualification packages across small molecule programs, from healthy-volunteer FIH through accelerated oncology development.

Integrated PharmaConsulting
integratedpharmaconsulting.com

About Integrated Pharma Consulting, LLC

Founded to give early-stage biotech companies and academic spinouts access to senior preclinical & CMC expertise typically found only inside larger pharmaceutical organizations. With over 15 years leading preclinical R&D at Zentalis and Kalyra Pharmaceuticals, IPC provides strategic depth across the full translational continuum from discovery science to regulatory submission.