COMPARISON OF HISTORICAL NEGATIVE CONTROL DATA FROM THE BACTERIAL REVERSE MUTATION (AMES) TEST: IMPLICATIONS FOR ASSAY ACCEPTABILITY, EVALUATION OF RESULTS, AND UPDATES TO THE OECD TEST GUIDELINE

L.F. Stankowski Jr., J. Callupe and V.Y. Kwok
Charles River Laboratories – Skokie, 8025 Lamon Avenue, Skokie, IL 60077

charles river

1 ABSTRACT

One workgroup at the 2017 International Workshop on Genetic Toxicology reviewed concerns related to the bacterial reverse mutation (Ames) test and OECD TG471[1], which was last updated in 1997. Topics considered included: minimum required strain set; acceptance criteria; laboratory proficiency; and evaluation criteria [including use of historical control data (HCD) and consideration of formal and subjective rules]. Analyses herein focus on HCD for negative controls as related to variability between laboratories, and comparison to published recommendations/values. Different authors have recommended different acceptable ranges for specific strains based upon their experience, and the observed ranges in individual papers (from individual laboratories or ring trials) did not always fall wholly, or even partly, within recommended ranges. While these publications were ~15 – 35 years old, seven laboratories (Labs A-G, including CROs and large Pharma) provided more recent negative HCD. These data showed some significant variations with older published recommendations, and significant variability between some labs for some strains. For example, there was no overlap in 95% control intervals (CI) for: TA98 in Labs F vs. G, TA1537 in Labs C vs. F, or WP2 *uvrA* in Labs E/F vs. C. In some cases, the 95% CI was >4-fold larger between labs [e.g., WP2 *uvrA* pKM101 –S9 in Labs A vs. G]. Such variability suggests: genetic drift or differing environmental conditions between labs; the possible need for a central strain repository; the potential to obtain different outcomes in different labs when using the "2-fold" (or 3-fold) rule to evaluate results; and the need for more rigorous methods to analyze results.

2 INTRODUCTION

The bacterial reverse mutation (Ames) test is likely the most commonly performed genetic toxicology assay for screening and regulatory purposes. And while all of the other commonly used OECD genetic toxicology test guidelines were updated in 2014 – 2016[2], the bacterial reverse mutation (Ames) test guideline (OECD TG471) was not revised. However, one workgroup at the 2017 International Workshop on Genetic Toxicology reviewed concerns related to that assay and OECD TG471, which was last updated in 1997. Topics considered included: minimum required strain set; acceptance criteria; laboratory proficiency; and evaluation criteria [including use of historical control data (HCD) and consideration of formal and subjective rules]. The analyses here focus on HCD for untreated and vehicle (negative) controls as related to variability between laboratories, and comparison to published values and recommendations.

3 MATERIALS AND METHODS

Several historical "methods" papers[3-8], as well as several reporting large scale ring trials[9-13] were reviewed and the negative (untreated and/or vehicle) control data are summarized here. The "methods" papers reported upper and lower values for an acceptable assay based upon the normal ranges observed in the authors' laboratories. For the ring trials, the upper and lower 95% control limits ($\bar{x} \pm 1.96$ SD) were calculated from the negative control data reported for each independent experiment within the paper. Any obvious duplications were removed (i.e., where common controls appeared to have been used for multiple chemicals tested). In addition, a number of laboratories from "Large Pharma" and the contract research organization (CRO) realm kindly provided their historical negative control data that encompassed various, but more recent, time periods. Upper and lower control limits were calculated as above for the published ring trial data.

4 RESULTS AND DISCUSSION

The "methods" papers were published ~15 – 35 years ago and recommended somewhat different acceptable ranges for specific strains based upon the experience of the particular laboratories (Figure 1). While there was some variation in the recommended ranges, there was overlap between them. However, the recommended ranges in the published methods papers did not differentiate between the values expected ±S9.

Data were compiled from an older series of published ring trials and are reported separately ±S9 (Figure 1). Seven laboratories also provided more recent unpublished historical negative control data (Labs A-G, including large Pharma and CROs; Figure 2). All of these recent large Pharma and CRO data were reported separately for cultures treated ±S9, and some laboratories specified the type(s) of treatments, as indicated. For example, Laboratory C reported results separately for plate incorporation vs. liquid pre-incubation treatments, Laboratories D and F reported only plate incorporation results, and the remaining laboratories reported combined plate and pre-incubation treatments. In addition, Lab G changed their historical data compilation software at the beginning of 2017 and they provided two sets of historical control data.

The 95% control intervals from the multi-laboratory ring trials, as well as those from the more recent large Pharma/CRO historical control data, did not always fall wholly – or even partly – within recommended ranges in those earlier published methods papers (Figure 1). Some of the more recent ring trial and large Pharma/CRO data sets showed very large variations between the data sets, as well as relative to the older published recommendations, for some of the tester strains (Figure 2). For example, there was no overlap in 95% control intervals for: TA1537 between Labs C vs. F; TA98 between Labs F vs. G; and WP2 *uvrA* between Labs E/F vs. C. Likewise, the 95% control intervals were significantly larger in some laboratories than others: 2.6- to 3.2-fold in strain TA1535 ±S9 for Labs C/E vs. G; 2.5- to 2.6-fold in strain TA1537 ±S9 for Labs E vs. F; and 4.1-fold in strain WP2 *uvrA* (pKM101) –S9 for Labs A vs. G.

Figure 1. HCD, Literature Values

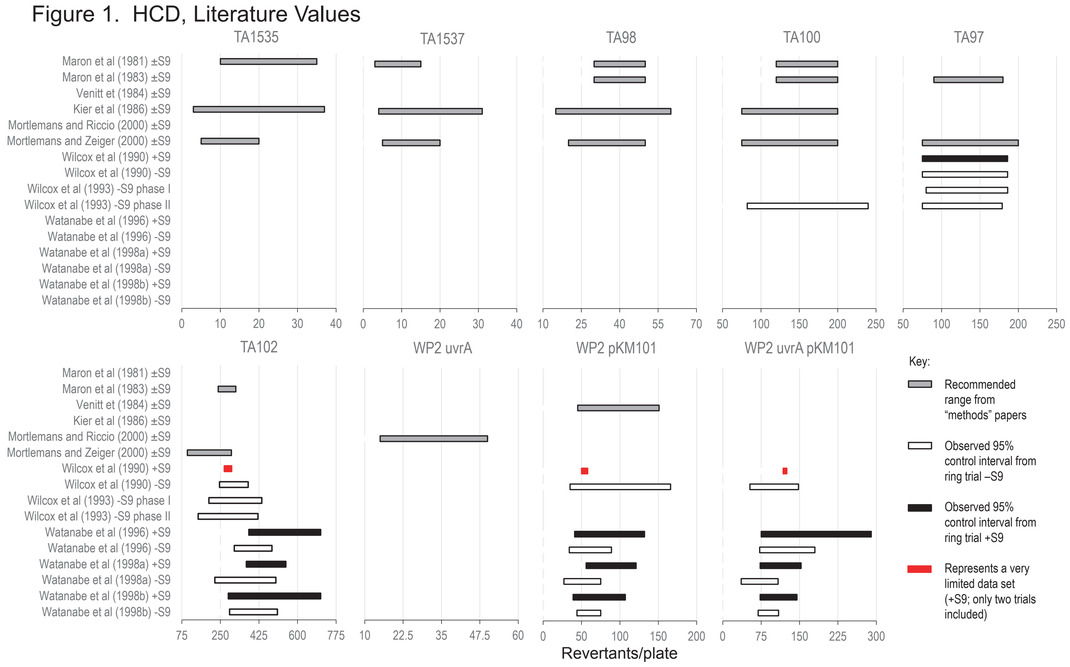
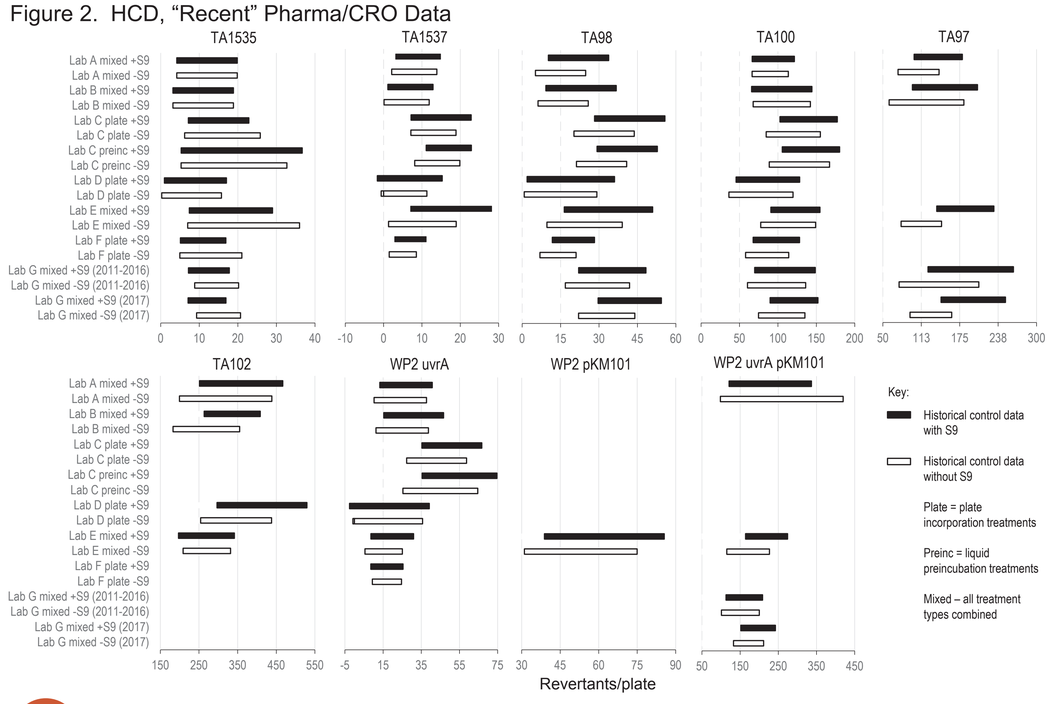


Figure 2. HCD, "Recent" Pharma/CRO Data



5 REFERENCES

[1]OECD (1997) Test No. 471: Bacterial Reverse Mutation Test, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris.
[2]OECD (2017) Overview of the set of OECD Genetic Toxicology Test Guidelines and updates performed in 2014-2015, Series on Testing & Assessment, No. 238 (2nd edition), OECD Publishing, Paris.
[3]Maron and Ames (1981) https://doi.org/10.1016/0165-1218(81)90025-2
[4]Maron and Ames (1983) https://doi.org/10.1016/0165-1161(83)90010-9
[5]Venitt, S., Crofton-Sleigh, D., Forster, R., Bacterial mutation assays using reverse mutation, In: S. Venitt , J. M. Parry (Eds.), Mutagenicity Testing: a Practical Approach, IRL Press, Oxford,1984, pp. 45–97.
[6]Kier et al (1986) https://doi.org/10.1016/0165-1110(86)90002-3
[7]Mortlemans and Riccio (2000) https://doi.org/10.1016/S0027-5107(00)00076-2
[8]Mortlemans and Zeiger (2000) https://doi.org/10.1016/S0027-5107(00)00064-6
[9]Wilcox et al (1990) https://doi.org/10.1093/mutage/5.3.285
[10]Wilcox et al (1993) https://doi.org/10.1093/mutage/8.2.93
[11]Watanabe et al (1996) https://doi.org/10.1016/S0165-1161(96)90249-6
[12]Watanabe et al (1998a) https://doi.org/10.1016/S1383-5718(97)00155-1
[13]Watanabe et al (1998b) https://doi.org/10.1016/S0165-1161(96)90249-6

6 CONCLUSIONS

- Large differences were observed between some of the data sets and relative to published recommendations
- The observed variability may be due to genetic drift or differing environmental conditions between labs
- These results suggest: the possible need for a central strain repository; the potential to obtain different outcomes in different laboratories when using the "2-fold" (or 3-fold) rule to evaluate results; and the need for more rigorous methods to analyze results