Mutation Research 726 (2011) 91–97

Contents lists available at ScienceDirect

Mutation Research/Genetic Toxicology and Environmental Mutagenesis

journal homepage: www.elsevier.com/locate/gentox
Community address: www.elsevier.com/locate/mutres

Evidence for hormesis in mutagenicity dose–response relationships

Edward J. Calabrese[a,*], Edward J. Stanek III[b,1], Marc A. Nascarella[c,2]

[a] Department of Public Health, Environmental Health Sciences, University of Massachusetts, Amherst, MA 01003, United States
[b] Department of Public Health, Biostatistics and Epidemiology, Arnold House, University of Massachusetts, Amherst, MA 01003, United States
[c] Gradient, 20 University Road, Cambridge, MA 02138, United States

ARTICLE INFO

Article history:
Received 24 January 2011
Received in revised form 7 April 2011
Accepted 13 April 2011
Available online 21 April 2011

Keywords:
Mutagen
Hormetic
U-shaped
J-shaped
Dose–response
Adaptive response

ABSTRACT

This study assessed the occurrence of hormetic dose responses from three previously published data sets [1–3] with 825 chemicals in three Ames assay tester strains (i.e., TA97, TA98, TA100) with and without the S9 fraction, using a five dose protocol and semi-log dose spacing. Ninety-five (95) (11.5%) chemicals satisfied the multiple *a priori* entry criteria, with a total of 107 assays. Of the assays satisfying the entry criteria, 61 involved TA100, a strain that detects base-pair substitution mutations. 29.5% (18/61) satisfied the statistical evaluative criteria for hormesis, exceeding that predicted by chance by 4.0-fold ($p < 0.001$). The remaining 46 assays involved TA97 and TA98, strains that detect frameshift mutations. Of these 46 assays, the overall responses for the lowest two doses closely approximated the control response (e.g., 101.77% of the control for TA98; 99.20% for TA97). Only 2.2% (1/46) of the assays satisfied the evaluative criteria for hormesis. In conclusion, these data support a hormetic model for TA100, whereas the responses for TA97 and TA98 are consistent with a threshold dose–response model.

1. Introduction

There has long been controversy over the nature of the dose–response for carcinogens. This debate has often centered on whether the data in the low dose zone were best explained by a threshold or linear model. Since data from individual experiments have typically not been sufficient to resolve which model was the most appropriate from a statistical perspective, a public health protectionist philosophy has been adopted by advisory bodies and/or regulatory agencies, leading to the acceptance of linearity at low dose as a matter of policy. This conceptual approach was first adopted over 50 years ago by the Biological Effects of Atomic Radiation (BEAR) I Committee [4] concerning radiation-induced mutation in reproductive cells. Soon after the BEAR I report, the National Council for Radiation Protection and Measurement (NCRPM) [5] generalized the linearity-at-low-dose concept and applied it to somatic effects of mutations induced by ionizing radiation, leading to linearity at low dose modeling for the carcinogenic effects. This perspective came to be widely accepted, generalized to chemical carcinogens [6–8] and eventually integrated within the risk assessment practices of regulatory agencies throughout the world, where it is currently the dominant perspective.

The assumption that the dose of a mutagen is linearly related to response at low dose has been a central theorem underlying the regulatory approach for carcinogens.

Nevertheless, there is an increasing literature over the past decade that a hormesis model better fits response at low doses [9–18]. While the hormesis dose–response model received little attention in the 20th century [19–24], the more recent literature has demonstrated its occurrence, widespread generalizability, reproducibility, mechanistic foundations and frequency [13,14,25–28]. Several large-scale investigations have provided evidence to support claims that the hormetic dose–response model is more common than other dose–response models [29–34]. Given this resurgence of the hormetic hypothesis and its implications, we have investigated the extent to which a hormesis model is applicable in assays which detect base pair and frameshift mutations. In order to do so, we evaluated three previously published datasets which are comprised of assays for 825 chemicals utilizing five bacterial strains tested within the Ames assay, with and without rat or hamster hepatic S9 fraction activation, with a five dose framework, using semi-log dose spacing [1–3].

2. Methods

Mortelmans et al. [1] and Zeiger et al. [2,3] published results of mutagenicity studies with *Salmonella typhimurium* with 825 chemicals as performed by three independent laboratories. Their investigations used a suite of Ames test strains (i.e., TA97, TA98, TA100, TA1535, TA1537) with and without rat and hamster hepatic

* Corresponding author. Tel.: +1 413 545 3164; fax: +1 413 545 4692.
E-mail addresses: edwardc@schoolph.umass.edu (E.J. Calabrese), stanek@schoolph.umass.edu (E.J. Stanek III), mnascarella@gradientcorp.com (M.A. Nascarella).
[1] Tel.: +1 413 545 3812; fax: +1 413 545 1645.
[2] Tel.: +1 617 395 5000; fax: +1 617 395 5001.

1383-5718/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.mrgentox.2011.04.006

S9. While essentially all chemicals were tested in strains TA98, TA100, TA1535 and TA1537, only five of the 270 chemicals were tested in strain TA97 by Mortelmans et al. [1], but the strain was included in the assessments by Zeiger et al. [2,3]. In the present evaluation, only data for tester strains TA97, TA98 and TA100 are considered. Tester strains TA1535 and TA1538 are excluded due to very low background control colony counts (about 6–18 colonies per plate), resulting in high control group variability. Among the tester strains used, TA97 had a control revertant count (colonies per plate) of about 100–130, whereas TA98 and TA100 had approximately 20–35 and 90–150, respectively.

In general, an assay consists of five doses assessed per chemical with three plates per dose along with concurrent solvent (i.e., water, DMSO, 95% ethanol or acetone) controls. Available dose–response data for each assay corresponded to the average numbers of revertant colonies per plate, based on the mean of the three plates at each dose, and the standard error of the mean (SEM). Although the assays were replicated, only the average response per dose for the final replication was typically published for the individual chemicals (Zeiger, personal communication). When the results of the replicate assay did not agree or the replication was equivocal/weak, data from such assays were published along with the final "replication".

An examination of the replicated assays indicated that they were not generally designed to be exact replications. For example, in the cases of allyl isocyanate (AI) and dimethyl hydrogen phosphite (DHP), the replications did not include the lower doses (i.e., the lowest dose for AI and the two lowest doses for DHP). For methoxychlor and methdilazine, the replicates were not tested at the same percentage of S9 fraction (10% S9 in one replicate but 30% S9 in the other). In the case of methyl phenidate, the replication did not include the same tester strain (i.e., TA97) [1]. Due to the lack of exact replication, each assay was treated as an independent evaluation. Judgments were made by the original authors as to whether the agent exhibited evidence of mutagenicity. In their papers [1–3], an agent was deemed to cause a mutagenic response if responses were dose related, causing a reproducible increase in the number of revertants above background, even if the increase was less than two-fold.

2.1. Entry criteria

The present paper assessed the frequency of hormetic dose responses for mutagenic endpoints within the Ames assay, using the data sets of Mortelmans et al. [1] and Zeiger et al. [2,3]. To assess the frequency of hormesis, *a priori* entry and evaluative criteria were employed as described below (Fig. 1).

2.1.1. Entry Criteria #1: response at dose #5 is ≥110% of control and entry Criteria #2: response at dose #4 is <110% of control

The initial entry criteria (Criteria #1 and #2) involved the selection of dose responses which had a local Benchmark Dose $(BMD)_{10}$ [31]. The local BMD_{10} is defined as a dose between dose #4 and dose #5 such that the response at dose #4 is below 110% and that at dose #5 is ≥110% of the control. With particular reference to the data we describe, this means that there are >10% more mutations (i.e., increase in the number of revertants) at Dose 5 relative to the control. This ensures that a response at dose #5 would equal or exceed 110% of the control rate and thereby provide potential evidence of a mutagenic response. This tentative conclusion is further strengthened by statistical evaluation (see Criteria #5). These criteria resulted in assays having three doses below the lower bound (LB) of the local BMD_{10}.

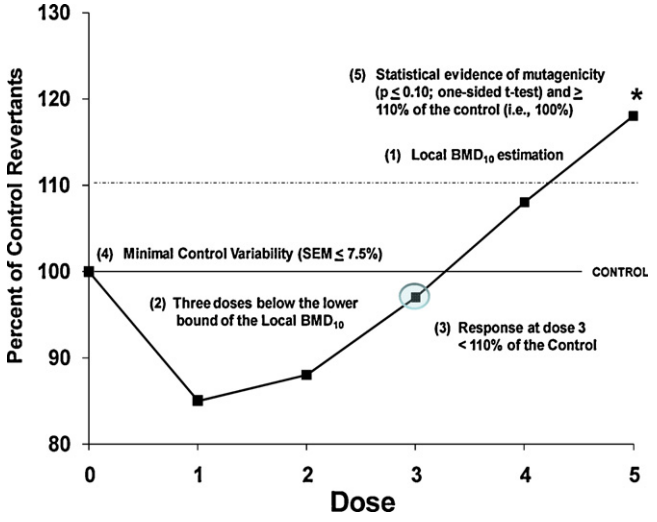

Fig. 1. Dose–response entry criteria. (1) The evaluative strategy centered on assessing the responses of the two lowest doses (doses #1 and #2) below the lower bound (LB) dose of the BMD_{10}. Dose #5 was required to satisfy two minimum criteria statistic evidence of mutagenicity ($p \leq 0.1$), one-side *t*-test and have a response ≥110% of the control group (100%). (2) The LB dose bracketing the BMD_{10} (i.e., dose #4) and the third lowest dose (i.e., dose #3) were not used to assess hormesis. Responses for the doses #3 and #4 were required to be <110%. (3) In an effort to minimize variability, a criterion also required that the control group of each dose–response display an SEM of <7.5%. (4) A dose–response demonstrated evidence of hormesis if one or two mean responses for doses #1 and #2 were significantly less than the control response (two-way *t*-test, $p < 0.10$).

2.1.2. Entry Criteria #3: response to dose #3 is <110% of the control

In order to enhance the likelihood of a more stable and accurate estimate of the toxicological threshold, dose responses were eliminated if the response of the third highest dose (i.e. dose #3) was ≥110% of the control.

2.1.3. Entry Criteria #4: minimize control group variation via SEM limit

The fourth entry criteria required that the control group of each dose–response display a standard error of the mean (SEM) of ≤7.5% to minimize variability. We used this criterion to select assays with higher power to detect differences in response at low doses.

2.1.4. Entry Criteria #5: statistical evidence of mutagenicity at dose #5

Entry criteria #5 required that the *p*-value for a one-sided test of the null hypothesis that response at dose #5 is less than or equal to 110% of control be rejected with a *p* value of ≤0.10.

2.1.5. Simulation study validation of entry criteria and tester strain hyper-poisson control group distribution assumption

Detailed simulation studies assessed the impact of each *a priori* entry criteria specifically or in combination. These simulations used several different strategies. One involved an assumption of a normal distribution of control counts and employed an estimate of a standard deviation to generate simulated individual "experiments" to assess the role of bias in selection criteria. A second approach involved the use of data simulated directly from repeated measures or control samples [35]. This second approach was designed to account for the possible hyper-poisson distribution that has been proposed for control samples, thereby providing a sensitivity analysis to the normality assumption. The results from both simulation strategies revealed no evidence that bias was introduced into the assessment of hormetic responses at low doses.

2.2. Evaluative strategy

2.2.1. Evaluative criteria for low potency mutagens

The evaluative strategy centered on assessing the responses of the two lowest doses (doses #1 and #2) below the lower bound (LB) dose of the BMD_{10}. The lower bounding dose bracketing the BMD_{10} (i.e., dose #4) and the third lowest dose (i.e., dose #3) were not used to assess the possible occurrence of hormesis. An assay was considered to provide evidence of hormesis if a two sided test of equality of response to control was rejected at $\alpha = 0.10$ at dose #1 and/or dose #2. A similar assessment was also made using $\alpha = 0.05$.

We classified assays with a statistically significant result as having response below the control (which we refer to as hormesis) or response above the control. We based the classification on doses where the response was statistically significant. In no case was there a conflict in the direction of significance.

2.2.2. Type I error—false positive estimation

Associated with any hypothesis test is the type I error, equal to the probability of rejecting the null hypothesis when the null is true. Such rejected tests are false positive results. Since two statistical tests were conducted for each assay, the probability that one (or more) test is statistically significant by chance is given by $1-(1-\alpha)^2$, i.e., 0.19 when $\alpha = 0.10$ and 0.0975 when $\alpha = 0.05$. We use these false positive rates to predict the number of false positive assays, which we expect to be evenly distributed above and below the control mean.

2.2.3. Additional statistical criteria

We were particularly interested in whether or not the response was above or below control response when the difference was considered to be statistically significant. To answer this question, we tested the null hypothesis that the proportion of statistically significant assays above and below control was equal using a binomial test.

2.2.4. Entry/evaluative criteria for chemicals with higher mutation production

Entry criteria: Criteria #1 – The response of dose 5 must be ≥150% relative to the concurrent control of 100% and statistically significantly greater than the control ($p < 0.1$). Criteria #2 – The dose–response must display a monotonic decrease in response from dose 5 to dose 4 to dose 3. Criteria #3 – The response of dose 3 must be ≤110% compared to the control of 100%. The evaluative criteria are the same as for the mutagens described in Section 2.2.1.

3. Results

3.1. Simulation of the Impact of the a priori entry criteria

We conducted a simulation study of the *a priori* entry criteria to examine whether the criteria themselves may lead to bias when assessing evidence for hormesis in the assays. The simulation was conducted by randomly generating a response for a chemical, assuming that the true response follows a normal distribution with a mean given by the observed control mean, and a standard deviation given by the observed control standard deviation. Three responses were generated for control and at each of five doses, and the average response calculated for each dose. In the simulation, we assumed that there was no dose (i.e., treatment) effect, so that the only difference between average responses at different doses was a result of the control standard deviation. In this manner, data were simulated for an assay, where we refer to one simulation as a trial. This process of simulating data was repeated for each of the 106 chemicals with control data.

Since there was no difference in mean response between any doses, the simulated data will result in identical expected responses at each dose, and no hormesis. Our goal in the simulation was to implement the entry criteria, and evaluate whether, after implementation, the expected response for assays that were judged to be eligible was equal to the control response. We interpret a deviation from this expected response as a bias induced by the entry criteria. The results of this analysis are summarized in Table 1. At each stage of the entry criteria, we determined the number of trials per chemical so that approximately 50,000 assays would be eligible for evaluation.

The impact of adding the entry criteria is given by reviewing consecutive rows in Table 1. The last five columns of the first row present the average response as a percent of control at each dose. Notice that when averaging over the 50,350 simulated assays, all responses are very close to 100%, indicating essentially no bias. The first selection criteria required response at dose 5 to be greater than 110% of control. Since no dose effect was included in the simulation, the requirement selected 5.8% of the assays, and resulted in an average response of 107.8% of control for dose 1 and dose 2, even though there was no dose (i.e., treatment) effect. Since the expected response at these doses is 100%, the simulation results reveal a 7.8% positive bias consequent to using this criterion. The addition of other entry criteria modify the estimated bias, but illustrate that after adding all five criteria, only 0.3% of the simulated assays would remain, with a positive bias (i.e., a bias against showing hormesis) at doses 1 and 2 of 6.07%.

These estimates of bias assume there is no effect of the chemical. When there is a chemical treatment effect, the selection bias associated with the entry criteria is reduced. We examine the impact of a chemical effect on the selection bias by adding a dose-5 chemical effect, incrementing the effect from 0% to 10% (see last 4 rows of Table 1). If there is a true response of 105% of control at dose 5, the bias is reduced to 1.1% of control. When the true response is 110% of control at dose 5, the bias is reduced to 0.58% of control. The actual chemical effect is not known for the 106 chemicals tested. If the true response at dose 5 is greater than 110% of control, we expect that a small positive bias (i.e., a bias against showing hormesis) to occur at doses 1 and 2, with the magnitude of the bias equal to approximately 0.5% of control response. Similar results were obtained when control data were based on TA100 from Margolin et al. [35].

3.2. Data analysis

Entry Criteria #1 eliminated approximately 80% of the assays. This was due to the chemicals either demonstrating very high

Table 1
Summary of simulation results evaluating possible selection bias for entry criteria.

Eligibility criteria	Additional criteria	Eligible assays	Percent eligible assays	Total assays	Assays per chemical	True eligible assay control mean	Simulated mean over all eligible assays					
							Control mean	Dose 1 (% of control)	Dose 2 (% of control)	Dose 3 (% of control)	Dose 4 (% of control)	Dose 5 (% of control)
None	None	50,350	100%	50,350	475	114.14	114.18	100.18	100.14	100.14	100.18	100.17
1	Dose 5 > 110%	51,857	5.8%	901,000	8500	111.15	103.39	107.82	107.81	107.83	107.81	114.58
1,2	Dose 4 < 110%	49,992	3.7%	1,356,800	12,800	112.31	106.24	105.85	105.87	105.85	103.05	113.68
1,2,3	Dose 3 < 110%	50,166	2.7%	1,855,000	17,500	112.43	107.36	104.83	104.76	102.67	102.66	113.25
1,2,3,4	SEM < 7.5%	51,214	2.2%	2,332,000	22,000	114.24	109.12	104.77	104.8	102.79	102.81	113.06
1,2,3,4,5	Dose 5 versus cntl $p < 0.1$	50,972	0.3%	19,822,000	187,000	113.59	107.23	106.07	106.07	103.37	103.35	116.94
1,2,3,4,5	Dose 5: 0%	2660	0.3%	1,060,000	10,000	112.43	106.11	105.93	105.98	103.36	103.33	116.94
1,2,3,4,5	Dose 5: 5%	113,295	10.7%	1,060,000	10,000	35.58	35.86	101.13	101.14	100.73	100.74	120.69
1,2,3,4,5	Dose 5: 7.5%	174,402	16.5%	1,060,000	10,000	41.46	40.82	100.58	100.58	100.21	100.21	126.33
1,2,3,4,5	Dose 5: 10%	219,949	20.7%	1,060,000	10,000	50.81	50.02	100.58	100.58	100.19	100.19	131.02

Table 2
The frequency of assays (dose responses) in which dose #1 and/or dose #2 were statistically significantly different from the control.

p-Value	Significant dose #1 or #2 effect	Response below control	Response above control	Expected below (above) control by chance	Net response below control	% net hormesis
TA100 (detects based pair substitution mutations) (n = 61 assays)						
0.10	28	24	4	6	18	29.5
0.05	17	16	1	3	13	21.3
TA97/TA98 combined (detects frameshift mutations) (n = 46 assays)						
0.10	12	7	5	6	1	2.2
0.05	8	4	4	3	1	2.2

Table 3
Mutation responses by tester strain for all assays satisfying the entry criteria.

	Control	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5
TA100						
Mean (N = 61)	100.00	95.79	98.46	99.14	101.63	118.05
StDev	8.16	9.68	8.42	5.89	5.28	18.20
SEM	1.04	1.23	1.07	0.75	0.67	2.33
TA98						
Mean (N = 22)	100.00	100.55	100.22	92.48	96.37	126.21
StDev	2.08	23.17	14.85	13.19	9.76	14.75
SEM	0.44	4.94	3.16	2.81	2.08	3.14
TA97						
Mean (N = 24)	100.00	102.30	99.07	101.15	104.33	117.93
StDev	11.49	8.56	9.87	5.94	4.00	10.91
SEM	2.34	1.75	2.01	1.21	0.81	2.23
TA98/TA97 combined						
Mean (N = 46)	100.00	101.77	99.20	96.67	100.44	121.40
StDev	6.97	16.87	12.25	10.87	8.29	12.81
SEM	1.02	2.48	1.80	1.60	1.22	1.88

mutagenic responses or no-evidence of mutagenicity. In each case, a local BDM_{10} is not derivable. In cases of assays with a derivable local BMD_{10} only those assays with three doses below the BMD_{10} were eligible for further evaluation. This resulted in the further elimination of chemicals that displayed moderate to higher mutagen potency. Approximately one-third of the assays satisfying Criterion #1 also satisfied Criterion #2, resulting in 6.5% of the initial assays. Application of the remaining entry criteria (Criteria #3–5) resulted in the final selection of 107 assays assessed from 95 of the initial 825 chemicals satisfying the entry criteria. This is approximately 0.74% of the initial 14,500 dose responses and 11.5% of the 825 chemicals.

3.3. An estimate of hormesis frequency

Table 2 provides a summary of the number of assays satisfying entry criteria and the number of assays where response from dose #1 and or dose #2 differed from control, and the number of assays where the difference was above or below the control response using the evaluative statistical ($p \leq 0.10$ or 0.05) criteria. The results are reported as tester strain specific. Responses statistically significantly less than the control response (hormesis) were more commonly reported in TA100, which detects based pair substitution mutations. There was no net evidence of chemical effects at low doses, after accounting for type I error, for TA97 and TA98, both of which detect frameshift mutations. Table 2 indicates that after accounting for type I error, 29.58% of the 61 (18/61) assays satisfying entry criteria for TA100 had response statistically significantly less than control. Among the assays for TA100 where differences were significant between low doses and control (at either 0.05 or 0.01), a null hypothesis that the proportion of assays with low dose–response below the control was equal to the proportion of assays with low dose–response above the control was rejected ($p < 0.0001$). The ratio of below to ≥control (100%) value for doses #1 and #2 is 1.65 to 1 ($p < 0.01$). This 1.65 to 1 value represents the average of the below to ≥control for dose #1 (1.90 to 1) and dose #2 (1.34 to 1). Tables 3 and 4 provide the specific dose–response values for the dose responses represented in Tables 2 and 3 for $p < 0.10$ and $p < 0.05$ evaluative criteria, respectively.

3.4. Chemicals displaying hormesis

The chemicals that satisfied the evaluative criteria for hormesis are given in Table 5 along with the TA strains in which the hormetic response occurred. The hormetic responses were distributed principally with agents requiring bioactivation using the S9 fraction of either rat or hamster liver.

Table 4
Mutation responses for TA100 for all assays satisfying both entry and evaluative criteria (i.e., hormesis) (i.e., $p \leq 0.1$ and $p \leq 0.05$).

	Control (100)	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5
TA100 – 0.10						
Average (N = 24)	100.00	89.30	93.11	97.26	101.02	121.86
StDev		7.65	8.28	5.92	5.16	14.14
SEM		1.56	1.69	1.21	1.05	1.89
TA100 – 0.05						
Average (N = 16)	100.00	90.71	93.93	97.88	102.28	122.89
StDev		8.12	8.71	5.95	4.62	16.61
SEM		2.03	2.17	1.48	1.15	4.15

Table 5
Chemicals from Tables 2 and 3 with evidence of mutagen response that satisfied evaluative hormesis criteria with TA100.

Chemical agent	WO S9	W S9	Hamster (H) Rat (R)	(0.10)	(0.05)	Reference
Mercuric chloride		X	H	X	X	[3]
Mercuric chloride		X	R	X		[3]
Triphenylamine		X	H	X	X	[3]
Triphenyl phosphite		X	H	X	X	[3]
Allyl isothiocyanate		X	H	X		[1]
Bis(2-methoxyethyl)ether		X	R	X		[1]
Dimethoxane	X			X	X	[1]
Dimethyl hydrogenphosphite		X	R	X	X	[1]
Hydrochlorothiazide	X			X		[1]
Malonaldehyde, NA salt		X	R	X	X	[1]
Maltol		X	H	X	X	[1]
Methoxychlor		X	H	X		[1]
6-Amino-4-chloro-1-phenol-2-sulfonic acid		X	R	X	X	[2]
p-Aminophenol		X	H	X	X	[2]
Chlorinated trisodium phosphate		X	R	X	X	[2]
N-(3-chloroallyl) hexaminium chloride		X	R	X		[2]
N,N′-di-sec-butyl-p-phenylenediamine		X	R	X		[2]
Direct violet 32 (CI 24105)		X	H	X		[2]
Alpha-naphthyl isothiocyanate		X	R	X	X	[2]
4-Nitrophthalimide		X	H	X	X	[2]
p-Phenylenediamine 2HCl		X	R	X	X	[2]
Pigment green 7 (CI 74260)		X	R		X	[2]
Succinonitrile		X	H	X	X	[2]
Tetraethyl dithiopyrophosphate		X	R	X	X	[2]
Zirconocene dichloride		X	R	X		[2]
Total				24	16	

Table 6
Hormetic dose responses for TA100 tester strain for all dose responses satisfying the entry criteria for chemicals with an entry criteria for mutagenicity in dose 5 of ≥150%.

Higher mutagen potency – TA100 (n = 28 assays)						
p-Value	Significant dose #1 or #2 effect	Response below control	Response above control	Expected below (above) control by chance	Net response below control	% net hormesis
≤0.1	28	8	1	3	5	17.9 (5/28)

3.5. Chemicals displaying higher mutation rates than controls at high doses

An attempt was made to assess the frequency of hormetic dose responses for chemicals displaying a higher mutation yield from the same three data sets. These criteria had the strength of identifying stronger mutation responses but were less effective in confidently estimating the toxic threshold. Of the 28 TA100 dose responses satisfying the entry requirements, seven displayed evidence of hormesis (Tables 6 and 7). These findings were generally consistent with that observed with the lower mutation production criteria. Only nine dose responses satisfied the entry criteria for TA97 and TA98 combined, precluding an adequate comparison with TA100.

4. Discussion

This is the first retrospective assessment of the frequency of responses at low doses consistent with hormesis studies of mutagenicity employing Ames tester strains used in large-scale screening assays to detect base pair and frameshift mutations. The findings indicate that the TA100 strain, which detects base pair substitution mutations, showed an excess of statistically significant low dose responses (i.e., responses for doses #1 and #2) that were below the control response. Thus, statistically significant deviations from control at low doses (i.e., doses #1 and #2) were 4–5 times more likely to have responses below the control (i.e., in the hormesis direction) than expected. These results are not consistent with what would be expected if mutagens are assumed to act linearly at low doses [7,36] (see Freese [37] and the discussion commentaries offered at the end of his paper). In contrast, TA97 and TA98, which detect frameshift mutations, did not displaying evidence of response different from control at low doses.

While it was uncertain that hormetic effects would be discerned in these assays, due to the elimination of error-free DNA nucleotide excision repair (deletion of the *uvrB* gene) and the inclusion of other genetic alterations that enhance the occurrence of mutagenic responses in the Ames assay, dose responses satisfying the evaluative criteria for hormesis in the TA100 strain were observed with high frequency after satisfying the entry criteria. This suggests that other adaptive/repair mechanisms are effective in this tester strain at low doses of mutagen exposures. It is well known that a broad array of mechanisms exist within bacteria strains, including *S. typhimurium*, that may reduce the occurrence of chemically-induced mutations, including polyamine induction [38,39], DNA repair methyltransferases [40], and glutathione transferase enzyme systems [41,42]. Further research will be needed to clarify the basis

Table 7
Mean dose responses for mutagens from Table 6 satisfying entry and evaluative criteria.

	Control (100)	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5
TA100						
Mean (N = 8)		86.21	98.14	98.75	128.48	167.09
StDev		4.47	2.91	6.48	10.78	11.60

of the dose-dependent transitions resulting in the expression of hormetic dose responses as reported in the present paper.

The trend toward a higher proportion of responses below, rather than above, the control for dose #1 versus dose #2 was first observed in the report of Calabrese and Baldwin [17,30]. This earlier observation raised the possibility that the dose most immediately below an estimated toxicological threshold may at times display a low level of toxicity due to lack of precision in the estimation of the threshold response. This phenomenon was called "residual toxicity". Since responses at doses #1 and #2 were used to calculate the presence or absence of the hormetic response, the residual toxicity phenomenon, if operational here, may have led to an underestimate of the frequency of the hormetic response by several percentage points.

The quantitative features of the hormetic dose responses suggested in the present study of mutagenic endpoints are consistent with that reported in the hormesis database [27,28] for a broad range of biological models, endpoints and chemicals from a large number of highly diverse chemical classes. For example, the average decrease of revertant colonies was approximately 10–12%. This decrease was typically occurring within a dose range 1/10–1/30 of the estimated threshold dose. The use of only two responses below the BMD_{10} zone limited an assessment of the width of the hormetic zone, that is, the range of the decreased incidence of the mutation response.

Despite the large number of initial bioassays (i.e., 14,500), fewer than one percent satisfied the multiple entry criteria for the evaluation of possible hormetic responses. The low percentage of assays satisfying the *a priori* entry criteria illustrate the challenges in assessing hormesis, especially retrospectively, when the studies were not specifically designed to test this hypothesis. The failure to satisfy entry criteria was principally due to the screening nature of the studies, in which a sizeable proportion of test agents were either relatively potent mutagens or non-mutagens. As noted above, both types of agents would have failed to satisfy the entry criteria. In addition, all agents requiring bioactivation would have failed the entry criteria for those assays tested without the S9 fraction. Thus, we believe that the occurrence of a low proportion of assays satisfying the entry criteria has no scientific relationship to the hormesis hypothesis.

The data set employed had both strengths and limitations. Strengths were that the biological models (i.e., bacterial strains) and test methods employed are extremely well known, broadly validated and extensively applied in toxicological studies and in hazard assessments. The data set was generated by genotoxicity experts from multiple laboratories, integrated, with the findings published in the peer-reviewed literature and in a journal specializing in mutagenesis. The study tested each compound with and without rat or hamster S9. The study also included at least five doses for each dose–response, making it potentially attractive for the assessment of hormesis, and required a replication experiment. The data set provided SEM and sample size information for each dose–response, thereby allowing hypothesis testing to be conducted. This permitted statistical criteria/analyses to guide decisions concerning the occurrence of mutation at high doses and hormetic effects at lower doses.

Some limitations, such as low colony counts, were minimized by the strict percentage SEM criteria, with those dose responses satisfying such criteria still being subject to the hypothesis testing statistical criteria. An important limitation is that the study was not designed to evaluate hormesis. More extensive testing is necessary in order to establish conclusively a hormetic dose–response for these specific agents.

We report results based on a data analysis using multiple two-sided *t*-test for the evaluative criteria, but other statistical criteria were investigated. Analyses were conducted using an ANOVA in which the SEM variance was pooled across the control and three doses below the LB of the BMD_{10} under assumptions of equal variances, or assuming unequal variances, using Welch's test and using multiple Dunnett's *t*-test. The results were consistent with the analyses reported here.

Aspects of the present study may limit its generalizability. These include the fact that the final number of chemicals (95) satisfying the *a priori* entry criteria was a low percentage (11.5%) of the original total. This circumstance was principally due to the fact that possible below-threshold responses were not a consideration in the construction of the original study design and often led to dose responses failing to pass the entry criteria for the present paper. We note that other analyses could have been conducted to investigate low-dose–response. One possibility was to alter the entry criteria to allow a larger number of eligible assays. We did not systematically examine such changes.

Of the 61 dose responses satisfying the entry criteria for TA100, the BMD_{10} values varied from the lowest category range of 10–33 μg/plate to 3300–10,000 μg/plate, indicating a broad mutagenic potency range of dose responses satisfying the entry criteria. Furthermore, dose responses satisfying the evaluative criteria also occurred across a broad potency range. For example, of the 37 dose responses with a dose range of 100–10,000 μg/plate 11 (29.7%) satisfied the evaluative criteria. Of the 13 dose responses with the most potent mutagens (with dose–response range starting at either 0.33 μg or 3.3 μg per plate) 6 (46.1%) satisfied the evaluative criteria. Thus, these data suggest that the hormetic response can occur independent of agent potency, further enhancing the generalizability of the findings.

Conflict of interest statement

None.

Acknowledgments

Thanks are given to Dr. Errol Zeiger for reading a draft manuscript and offering helpful suggestions. Likewise we appreciate the assistance and critiques offered by Professor George Hoffmann, Biology Department, College of the Holy Cross, Worcester, MA. We would like to acknowledge the helpful written comments provided by Dr. Kenneth Crump to our poster presented at the 2010 annual meeting of the Society of Toxicology.

Effort sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant number FA9550-07-1-0248. The U.S. Government is authorized to reproduce and distribute for governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsement, either expressed or implied, of the Air Force Office of Scientific Research or the U.S. Government.

References

[1] K. Mortelmans, S. Haworth, T. Lawlor, W. Speck, B. Tainer, E. Zeiger, *Salmonella* mutagenicity tests. II. Results from the testing of 270 chemicals, Environ. Mutagen. 8 (1986) 1–119.
[2] E. Zeiger, B. Anderson, S. Haworth, T. Lawlor, K. Mortelmans, *Salmonella* mutagenicity tests. IV. Results from the testing of 300 chemicals, Environ. Mol. Mutagen. 11 (1988) 1–158.
[3] E. Zeiger, B. Anderson, S. Haworth, T. Lawlor, K. Mortelmans, W. Speck, *Salmonella* mutagenicity tests. III. Results from the testing of 255 chemicals, Environ. Mutagen. 9 (1987) 1–110.
[4] BEAR I Committee Report—National Academy of Sciences, The Biological Effects of Atomic Radiation: A Report to the Public, NAS/NRC, Washington, DC, 1956.
[5] National Committee on Radiation Protection and Measurements (NCRPM), Somatic radiation dose for the general population, Science 131 (1960) 482–486.

[6] National Academy of Sciences (NAS) Safe Drinking Water Committee, Drinking Water and Health, National Academy of Sciences, Washington, DC, 1977.
[7] E.J. Calabrese, The road to linearity: why linearity at low doses became the basis for carcinogen risk assessment, Arch. Toxicol. 83 (2009) 203–225.
[8] E.J. Calabrese, Getting the dose–response wrong: why hormesis became marginalized and the threshold model accepted, Arch. Toxicol. 83 (2009) 227–247.
[9] J.M. Cuttler, M. Pollycove, Nuclear energy and health: and the benefits of low-dose radiation hormesis, Dose Response 7 (2009) 52–89.
[10] E.J. Calabrese, Paradigm lost, paradigm found: the re-emergence of hormesis as a fundamental dose response model in the toxicological sciences, Environ. Pollut. 138 (2005) 379–412.
[11] E.J. Calabrese, Threshold – dose–response model – RIP: 1911 to 2006, BioEssays 29 (2007) 686–688.
[12] E.J. Calabrese, Hormesis: why it is important to toxicology and toxicologists, Environ. Toxicol. Chem. 27 (2008) 1451–1474.
[13] B.R. Scott, It's time for a new low-dose-radiation risk assessment paradigm—one that acknowledges hormesis, Dose Response 6 (2008) 333–351.
[14] B.R. Scott, Low-dose-radiation stimulated chemical and biological protection against lung cancer, Dose Response 6 (2008) 299–318.
[15] J.L. Redpath, Radiation-induced neoplastic transformation in vitro, hormesis and risk assessment, Environ. Mol. Mutagen. 47 (2006) 403–1403.
[16] J.L. Redpath, Nonlinear response for neoplastic transformation following low doses of low let radiation, Nonlinear Biol. Toxicol. Med. 3 (2005) 113–124.
[17] E.J. Calabrese, L.A. Baldwin, Toxicology rethinks its central belief, Nature 421 (2003) 691–692.
[18] E.J. Calabrese, L.A. Baldwin, Hormesis: U-shaped dose responses and their centrality in toxicology, Trends Pharm. Sci. 22 (2001) 285–291.
[19] E.J. Calabrese, L.A. Baldwin, Chemical hormesis: its historical foundations as a biological hypothesis, Hum. Exp. Toxicol. 19 (2000) 2–31.
[20] E.J. Calabrese, L.A. Baldwin, The marginalization of hormesis, Hum. Exp. Toxicol. 19 (2000) 32–40.
[21] E.J. Calabrese, L.A. Baldwin, Radiation hormesis: its historical foundations as a biological hypothesis, Hum. Exp. Toxicol. 19 (2000) 41–75.
[22] E.J. Calabrese, L.A. Baldwin, Radiation hormesis: the demise of a legitimate hypothesis, Hum. Exp. Toxicol. 19 (2000) 76–84.
[23] E.J. Calabrese, L.A. Baldwin, Tales of two similar hypotheses: the rise and fall of chemical and radiation hormesis, Hum. Exp. Toxicol. 19 (2000) 85–97.
[24] E.J. Calabrese, Historical blunders: how toxicology got the dose–response relationship half right, Cell. Mol. Biol. 51 (2005) 643–654.
[25] G.R. Hoffmann, A perspective on the scientific, philosophical, and policy dimensions of hormesis, Dose Response 7 (2009) 1–51.
[26] T.G Son, S. Camandola, M.P. Mattson, Hormetic dietary phytochemicals, Neuromol. Med. 10 (2008) 236–246.
[27] E.J. Calabrese, R.B. Blain, Hormesis and plant biology, Environ. Pollut. 157 (2009) 42–48.
[28] E.J. Calabrese, R. Blain, The occurrence of hormetic dose responses in the toxicological literature, the hormesis database: an overview, Toxicol. Appl. Pharmacol. 202 (2005) 289–301.
[29] E.J. Calabrese, L.A. Baldwin, The frequency of U-shaped dose-responses in the toxicological literature, Toxicol. Sci. 62 (2001) 330–338.
[30] E.J. Calabrese, L.A. Baldwin, The hormetic dose response model is more common than the threshold model in toxicology, Toxicol. Sci. 71 (2003) 246–250.
[31] E.J. Calabrese, J.W. Staudenmayer, E.J. Stanek, Drug development and hormesis: changing conceptual understanding of the dose response creates new challenges and opportunities for more effective drugs, Curr. Opin. Drug Discov. Dev. 9 (2006) 117–123.
[32] E.J. Calabrese, J.W. Staudenmayer, E.J. Stanek, G.R. Hoffmann, Hormesis outperforms threshold model in NCI anti-tumor drug screening data, Toxicol. Sci. 94 (2006) 368–378.
[33] E.J. Calabrese, G.R. Hoffmann, E.J. Stanek, M.A. Nascarella, Hormesis in high-throughput screening of antibacterial compounds in *E. coli*, Hum. Exp. Toxicol. 29 (2010) 667–677.
[34] E.J. Calabrese, E.J. Stanek III, M.A. Nascarella, G.R. Hoffmann, Hormesis predicts low-dose responses better than threshold models, Int. J. Toxicol. 27 (2008) 369–378.
[35] B.H. Margolin, N. Kaplan, E. Zeiger, Statistical analysis of the Ames Salmonella/Microsome Test, Proc. Natl. Acad. Sci. U.S.A. 78 (1981) 3779–3783.
[36] J.W. Drake, S. Abrahamson, J.F. Crow, A. Hollaender, S. Lederberg, M.S. Legator, J.V. Neel, M.W. Shaw, E.E. Sutton, R.C. von Borstel, S. Zimmering, F.J. de Serres, Environmental mutagenic hazards, Science 187 (1975) 503–514.
[37] E. Freese, Thresholds in toxic, teratogenic, mutagenic, and carcinogenic effects, Environ. Health Perspect. 6 (1973) 171–178.
[38] S.P. Pillai, D.M. Shankel, Effects of antimutagens on development of drug/antibiotic resistance in microorganisms, Mut. Res. 402 (1998) 139–150.
[39] S.P. Pillai, D.M. Shankel, Polyamines and their potential to be antimutagens, Mut. Res. 377 (1997) 217–224.
[40] M. Yamada, K. Matsui, T. Sofuni, T. Nohmi, New tester strains of *Salmonella typhimurium* lacking O^6-methylguanine DNA methyltransferases and highly sensitive to mutagenic alkylating agents, Mut. Res. 381 (1997) 15–24.
[41] L.I. Vorobjeva, E.Y. Khodjaev, T.A. Cherdinceva, The study of induced antimutagenesis of propionic acid bacteria, J. Microbiol. Methods 24 (1996) 249–258.
[42] L.I. Vorobjeva, E.Y. Khodjaev, T.A. Cherdinceva, Antimutagenic and reactivative activities of dairy propionibacteria, LAIT 75 (1995) 473–487.