

GTA September 2008

ICH S2R: GUIDANCE ON GENOTOXICITY TESTING AND DATA
INTERPRETATION FOR PHARMACEUTICALS
INTENDED FOR HUMAN USE

Proposed revisions to genotoxicity guidance

Sheila M. Galloway, PhD

Representative of Expert Working Group for ICH S2

Merck Research Laboratories, West Point, PA

Current trends in genetic toxicology – all products, not just drugs

- Recognized need to improve accuracy of genotoxicity testing
- Re-examination of in vitro mammalian cell assays by entire field
 - Ongoing workshops ILSI-HESI IVGT; IWGT, EVCAM, COLIPA
- Concern raised by regulatory scientists (pharmaceuticals) in Europe, Japan
- European regulations (non-drugs):
 - For cosmetics, personal products etc., animal testing is ending
 - Need for more realistic in vitro testing to obviate need for *in vivo* follow-up testing
 - Consideration of concentration limit such as 100 μM
 - High safety margin of topical exposure
 - Physiologically justifiable, since most K_M 's of enzyme reactions are $<100 \mu\text{M}$
- ECVAM review in progress on predictivity of in vitro mammalian cell assays tests at various limits of toxicity, concentration

The revisions to the guidance ...

- Make use of another 10 years of knowledge from testing pharmaceuticals
- Recommend updated test protocols
- Use newer experience e.g.,
 - *in vitro* micronucleus assay
 - rat blood for micronucleus analysis
 - Comet assay for DNA strand breaks
- Make more efficient use of resources
 - e.g., integrating genotoxicity endpoints into existing toxicology studies where feasible
- Provide more options for the test battery
- Increase flexibility and advice on appropriate *in vivo* tests (decrease emphasis on UDS)

Principles agreed by EWG in considering revisions to ICH S2

Make hazard identification more realistic

- We need to reduce our reliance on in vitro assays carried out under somewhat extreme conditions on the principle of *hazard identification*
- Consider tests/protocols that identify potential genotoxic effects under more realistic conditions that provide information more useful to human risk evaluation.

Principles agreed by EWG in considering revisions to ICH S2

Compounds of concern

In designing a test battery, it is important to consider not only predictive value for rodent carcinogens, but prediction of *compounds of concern for human health*.

- In the past, tests or test protocols selected by determining how many rodent carcinogens are identified
- Must take into account information whether
 - primarily genotoxic mechanism(s), ?potential risk to humans
 - non-genotoxic mechanism(s)
 - more likely threshold
 - do not expect short term genotoxicity assays to predict

Principles agreed by EWG in considering revisions to ICH S2

Germ cell mutagens are considered

- Affirming the statement in the original ICH guidance:
 - Battery is primarily aimed at detecting potential genotoxic carcinogens, but in so doing *will protect against potential germ cell mutagens*.
 - Known germ cell mutagens are positive in bone marrow chromosome damage assays*

*Shelby, M.D., *Mutation Research*, 352:159-167, 1996;

Waters, M.D. et al, *Mutation Research*, 341, 109-131, 1994

Principles agreed by EWG in considering revisions to ICH S2

Use the battery approach

- In choosing tests or test conditions appropriate for a battery, consider the battery approach overall
 - Not essential that each individual test detect each carcinogen/compound of concern, provided another test in the required battery effectively detects it.

Principles agreed by EWG in considering revisions to ICH S2

The exception does not drive the rule

- Exceptions should not drive the design of test batteries or protocols.
 - e.g., if certain tests or modified protocols are required to detect nucleoside analogues, cite these as an example where specialized test strategies may be needed; do not change the design of a strategy that effectively detects all other known classes of genotoxins.

What are the arguments for continuing to use mammalian cell assays?

Adding *in vitro* mammalian cell tests to the Ames test does not improve predictivity for rodent carcinogenesis

Justification same as when battery concept first developed:

- Observation:
 - A few compounds more effectively detected in mammalian cells (e.g., metals)
- Principle:
 - Include a range of genetic endpoints relevant to cancer and germ cell mutations (not seen in bacteria)
 - Drugs have mammalian cell targets (potential for unexpected findings)
- **BUT mammalian cell assay does not have to be *in vitro*.**
- **The combination of the bacterial mutation test and a chromosome damage assay *in vivo* predicts the majority of known human carcinogens**

Principles agreed by EWG in considering revisions to ICH S2

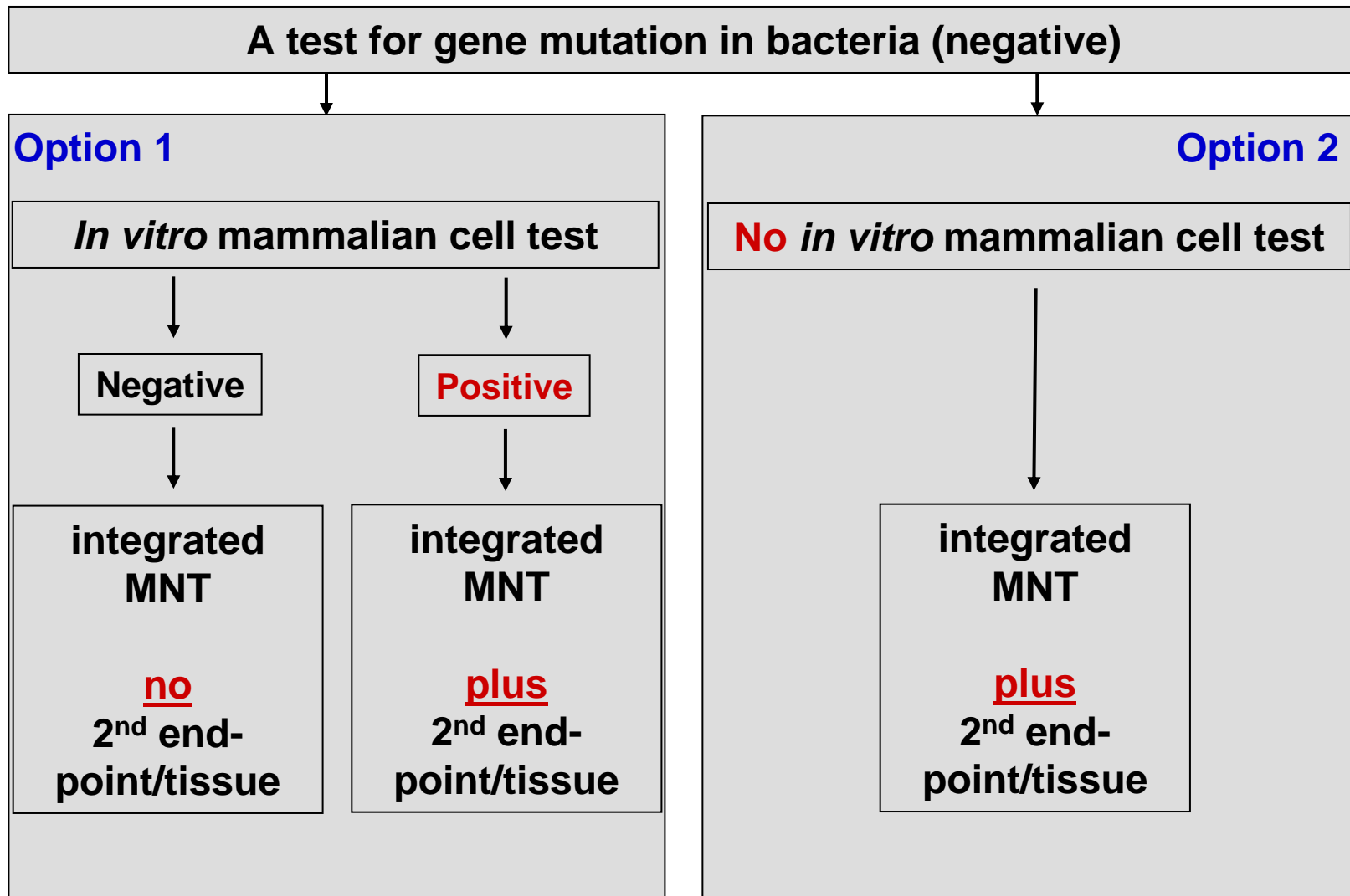
Drugs are different

- OECD guidelines were developed in parallel with the ICH guidances (early 1990s)
- Purposeful attempt to develop methods applicable to genotoxicity testing for all types of chemicals (industrial, agricultural, medical)

BUT

- Certain attributes of pharmaceutical testing justify specific modifications for drugs
 - e.g., timing of testing; extensive ADME and exposure information from humans
- Some aspects of OECD guidelines need to be updated
- Differences from existing OECD guidelines are pointed out and justified in the revised ICH guidance.

Proposed new recommended test battery



Rationale for revisions to battery

- Two well conducted *in vivo* genotoxicity assays generally outweigh a positive result in an *in vitro* mammalian cell assay, IF
 - in the appropriate tissues
 - appropriate exposure demonstrated
- Already a rather widely accepted concept
 - current ICH S2a and S2b guidance
 - FDA genotoxicity guidance[§]
- The “Ames plus two *in vivo* assay” battery thus treats the test compound as if it were positive *in vitro*
- Need for second tissue affirmed by survey
 - examples of *in vitro* pos, neg *in vivo* in MN, pos in UDS, Comet etc

[§] Recommended Approaches to Integration of Genetic Toxicology Study Results, Jan 2006

Why have options?

- Flexibility
- The “two *in vivo* test” approach may not always be feasible due to
 - Timing in drug development
 - Practical considerations
 - Insufficient exposure in the toxicology study
- Knowledge about the chemical class might indicate an *in vitro* test is preferable

Suitability of repeat-dose toxicity studies for genotoxicity measurement?

Concern about reduced sensitivity

- Top dose typically lower than in acute studies
 - Industry (JPMA, EFPIA, PhRMA) survey to compare exposure in acute versus multi-weeks toxicology studies
- Many genotoxicity endpoints do not reflect cumulative damage
 - Micronucleus (and some DNA damage endpoints) reflect only relatively recent exposure regardless of how long the animals have been treated

**Evaluated published data and set stringent criteria
for selection of top dose**

Criteria for acceptable dose/exposure in subchronic study to justify genotoxicity evaluation

When in vitro mammalian cell assay is positive (Option 1) or not done (Option 2)

Goal is to maximize exposure in this protocol

- Maximum feasible dose
 - Unless this is severely limited in long term treatment of acute by vehicle
- Limit dose of 1000 mg/kg if ≥ 14 days
 - OECD limit dose for micronucleus evaluation
- Exposure:
 - Plateau/saturation in exposure
 - Accumulation
 - Reduced exposure with time would usually disqualify study
- Top dose (or exposure) scored is $\geq 50\%$ of the top dose (exposure) that would be used for acute administration*, if such data are available
 - *Currently described in OECD guidance as the dose above which lethality would be expected

Outcome of industry survey on exposure

- Exposure in 2/4/12 week studies compared with exposure in 1-3 dose acute studies
- In the majority (94/139) of the multi-week rat studies the top dose used would justify use for genotoxicity testing, based on
 - similarity to exposure achieved with acute high-dose protocol*
 - plateau in exposure
 - accumulation of drug with repeat dosing
 - maximum feasible dose
 - limit dose

* Data not often available in EU and US: practical experience in a few companies is that about 50% of studies are usable for integrated MN analysis

What is experience with integrated micronucleus assay?

- Based on *mouse* data with a few potent mutagens, [e.g., MacGregor et al 1990 and Heddle et al 1991] considered
 - repeat dose protocol sensitivity \cong acute protocol, provided
 - top dose was the same proportion of the MTD in the relevant protocol
- IWGT [Hayashi et al 1994] concluded MN assay in repeat dose study was acceptable as long as toxicity demonstrated
- Concern raised by *rat* 28 day micronucleus studies
 - well known clastogens
 - 4 compounds such as dimethylhydrazine and MMC, detectable after acute dosing, were negative at 28 days because MTD was so much lower.
- Demonstrates potential difficulty of getting the dose right in toxicology studies
- Guideline endorses recommendation of the Japanese CSGMT [Hamada et al 1991]
 - Sample blood at 4 days, and bone marrow + blood at 28 days

Use of rat blood for MN evaluation

- Rats selectively remove micronucleated erythrocytes from peripheral blood due to spleen function
- Collaborative studies now show rat blood can be used if
 - Stain to identify most newly formed reticulocytes
 - Increase sample size scored:
 - manual evaluation: ≤ 8000 PCE/animal (depends on background level)
 - flow cytometry: 20,000 PCE/animal
- Concerns about difficulty of detecting aneugens in routinely used protocols; rat strain differences
 - Comparison of aneugens in two strains in two labs
 - No substantial Sprague-Dawley-Wistar rat difference
 - Aneugens are detectable in rat blood, but difficult to be sure of exactly correct dose selection for blood or bone marrow
- Conclusion: rat blood is acceptable
 - Caveats about toxicity, dose selection & sampling times in guidance

Choice of *in vivo* assays: 2nd endpoint/tissue

- Remove emphasis on the *in vivo* UDS assay as best validated *in vivo* assay (current ICH guideline)
- Options considered as acceptable
 - DNA strand break assays (Comet, alkaline elution)
 - transgenic mouse mutation assays
 - DNA binding assays
 - UDS assay
- Liver typically the preferred tissue but choice should be based on factors such as
 - type of effect seen *in vitro* (if relevant) [*all info, not just genetox*]
 - any knowledge of the potential mechanism
 - any knowledge of the metabolism *in vivo* and
 - of the exposed tissues thought to be relevant

Need to develop more experience with integrated Comet assay

- Practical issues including
 - Sampling time in relation to last dose
 - Freezing tissues
- Sensitivity: Potential false negatives
 - Same issue as MN assay
 - 2/4 week etc MTD may be too low
- Specificity: strand breaks associated with toxicity
 - So far, acutely toxic compounds have not given false positives in Comets, e.g. galactosamine
 - What about prolonged treatment, proliferation, oxidative damage etc?

EU, US, Japan labs beginning collaborative studies to look at all above issues before finalization of guidance

In vitro mammalian cell genotoxicity assays

The good side:

- These are sensitive assays, that do their job, i.e., detect “bad actors”, DNA damaging chemicals and some potential inducers of aneuploidy
 - **Note: Typically these are also detectable *in vivo***

The problem:

- The results of the *in vitro* chromosome aberration and mouse lymphoma cell mutation (MLA) assays as applied to date are not predictive of rodent carcinogenicity, and by extension, not accurate indicators of potential human risk
 - Kirkland et al, 2005 [Mutation Research 584: 1-256]
 - Matthews et al , 2006 [Reg. Tox Pharmacol 44: 97 – 110]
 - Snyder & Green, 2001 [Mutat Res. 488:151-169] re PDR

Reinforce toxicity limits for in vitro mammalian cell assays

- Chromosome damage assays
 - Not necessary to exceed about 50% cytotoxicity for
 - Structural aberrations
 - Micronuclei (in vitro)
 - Recommend cell growth measurement, not just cell counts
- Published data show frequency of non-relevant positive results increases with toxicity, even before approaching 50% growth reduction
- Clear DNA damaging clastogens detected without much toxicity based on growth/counts at time of sampling
- Limit of 50% cytotoxicity was agreed by IWGTP expert group in early 1990's
 - “>” 50% in guidelines was put in to avoid repeated attempts to obtain exactly 50%, but has led to continued use of excessively toxic doses

Toxicity limits in MLA

- Consensus (IWGT workshops; Mitchell et al Genetox report 1997)
 - Caution is needed in interpreting results when increases are seen only below 20% RTG
 - A result would not be considered positive if the increase in mutant fraction occurred only at $\leq 10\%$ RTG.
- Few mutagenic carcinogens detected only with toxicity $>80\%$
- ICHS2(R1): Use IWGT recommendation and reinforce “about” 80% cytotoxicity (reduction in RTG)

Re-examine 10 mM limit

- What are “relevant” levels of exposure?
 - Human drug levels
- ICPEMC review
 - Re-examine data used in setting the 10 mM limit
- Frequency of “relevant” positives
 - Industry survey

What was basis for 10 mM top concentration (OECD) for *in vitro* mammalian cell assays?

- Improvement over no limit, previously
- 10 mM known to avoid increased osmolality
 - NaCl, sucrose etc. shown to cause aberrations, mutations
- A few clastogens active in mM range, e.g. EMS, DMN
 - **Detectable in battery approach or by SAR**
- *In vivo* clastogens shown to be detected *in vitro* at ≤ 10 mM (ICPEMC review*)
 - Rechecked these for the 1 mM limit
 - All *in vivo* clastogens reviewed were pos *in vitro* at 1 mM or below
 - (except benzene in hamster cells; it was detectable at <1 mM in HL)
 - OR were positive in Ames test
 - **Detectable in battery approach**

*Scott, D., S. M. Galloway, R. R. Marshall, M. Ishidate Jr., D. Brusick, J. Ashby, B. C. Myhr, ICPEMC Task Group 9," *Mutation Research*, 257:147-204, 1991.

In vitro mammalian cell assays: is 1 mM limit adequate to detect compounds of concern?

- Are there compounds not detected by rest of battery (Ames and *in vivo* cytogenetics) that are compounds of concern?
 - *in vivo* DNA damaging compounds?
 - rodent carcinogens with established genotoxic mechanism?
- What are *in vitro* conditions required to detect them?
 - Cytotoxicity
 - Concentration
- Compounds from ICPEMC review compiled by Ishidate;
 - 16, Ames and *in vivo* cytogenetics negative
 - *In vitro* positive at > 1 mM
 - Do not appear to be compounds of concern

Human exposures to pharmaceuticals

- Peak exposure generally $< 10\text{-}50\ \mu\text{M}$
 - Exceptions : Antibiotics, antivirals, anti-tumor drugs
- The K_m of most biochemical processes is $<100\ \mu\text{M}$
 - metabolic activation/inactivation
 - (in tests with S-9 etc metabolic activation, enzymes are generally saturated at higher concentrations)
 - general cellular defence, cellular transport/turnover
- In vitro testing done with 0 – 15% serum so there is more free drug than *in vivo* (many drugs highly protein bound)
- ➔ In most situations, a top in vitro concentration of $100\ \mu\text{M}$ would give meaningful data, the in vivo relevance of which can easily be assessed
 - ➔ A top concentration of 1mM would capture low potency drugs and other high dose drugs including cases of extensive tissue accumulation

Industry survey: Case studies of convincing positive results found in vitro, in mammalian cell assays

- To see if these tests are important in the battery, and if so, could a new concentration limit be justified
- Emphasis was on compounds that convinced the investigators they **might be a potential risk to humans**
- Emphasized chemicals negative in Ames test
- Definition of “relevant” positive; examples
 - High potency (e.g., strong response, low concentration)
 - Lack of toxicity
 - Knowledge of mechanism,
 - Other evidence for DNA damage from non-routine tests,
 - In vivo positive genotoxicity data

In vitro assays of pharmaceuticals can be made more relevant by limiting top concentration

- Conclusions from survey:
 - There are examples of *in vitro* positive results seen only above 1 mM
 - In many cases, these were considered “non relevant” based on
 - weight of evidence from other genotoxicity assays *in vitro* and *in vivo*
 - mechanistic studies including DNA binding, use of antioxidants etc
- Proposed 1 mM limit maintains hazard identification function
 - Higher than clinical exposure
 - Higher than typical exposure in toxicology studies and acute *in vivo* genotox studies

Conclusions

- It is time to move forward in realistic risk evaluation
 - In principle, genotox results at high, toxic doses are questionable, in the same way we question the relevance of high-dose toxicities in animals to lower, pharmacological doses
- We have known about the liabilities of *in vitro* mammalian cell assays for 25 years!
- Time to move on to “how to do appropriate follow-up to a positive Ames test”, or we’ll be in the same place another 25 years from now

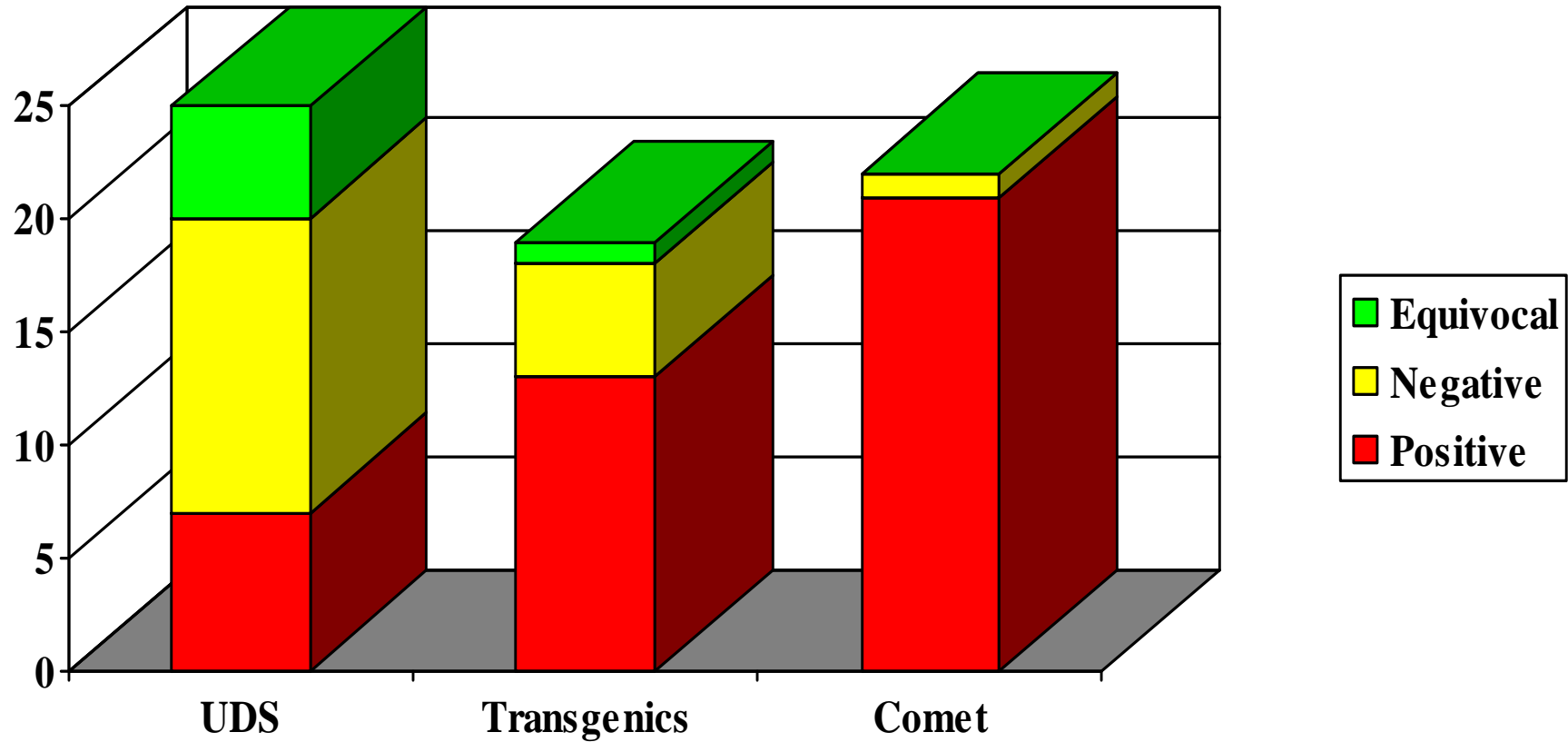
Back-ups

Example of an *in vitro* positive result considered “non-relevant” *in vivo*

- Dopamine Agonist
 - CHO cell aberration induction at 1.4 mM, only with S-9
 - Cytotoxicity by cell counts was 44% at least effective concentration
- *In vitro* alkaline elution in rat hepatocytes
 - positive at toxic doses
- Tested catechol metabolite of drug: same results
- Dopamine had same effects (more potent)
 - Suppressed by oxygen radical scavengers, SOD or anaerobic incubation
- V79 *hprt*: negative
- *In vivo* assays negative:
 - Mouse bone marrow chromosome aberrations
 - DNA strand breaks by alkaline elution in liver
- Catechol metabolite undetectable *in vivo*

In vivo results with Ames +ve carcinogens that are negative in bone marrow MN (micronucleus) assay. Literature survey by Kirkland

Numbers of chemicals
With each *in vivo* result



Note many Comet data are from 1 lab, and protocol not yet internationally validated

Additional guidance on use of subchronic study for micronucleus assessment when in vitro mammalian cell assay is not done (Option 2)

Dose selection and strategy to increase ability to detect aneugens

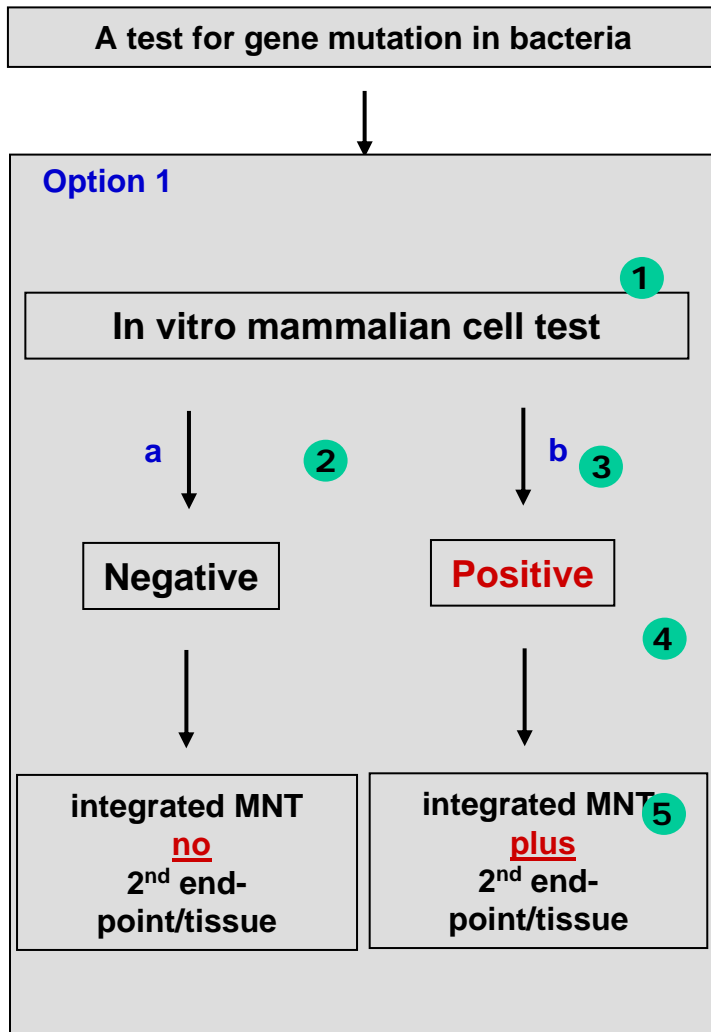
- Since marked hematotoxicity is often seen following exposure to aneugens, when rat blood or bone marrow is used for micronucleus measurement in a multiweek study (e.g. 28 days), and marked hematotoxicity is seen, it is recommended to sample blood both on day 2 to 4 of dosing and after the last dose to provide assurance that chromosome loss is detected (potential aneugens). (day 2 – 4 sample is only scored if required)

Additional guidance on use of subchronic study for micronucleus assessment when *in vitro* mammalian cell assay is not done (Option 2)

Assessment of suitability of dose levels

- Compounds that induce aneuploidy, such as spindle poisons, are typically detectable in *in vivo* micronucleus assays in bone marrow or blood over a narrow range of doses approaching toxic doses.
- If toxicological data indicate severe toxicity to red blood cell lineage (e.g. marked suppression of PCEs or reticulocytes), doses scored should be selected to aid in detection of potential aneugens:
 - dose spacing should not be more than 2 to 3 fold below the top, cytotoxic dose
- If suitable doses are not included in a multiweek study, additional data may be required to ensure detection of aneugens; these could be derived from any one of the following:
 - 2 -4 day blood sampling from the multiweek study before substantial hematotoxicity developed
 - an *in vitro* mammalian cell micronucleus assay
 - [metaphase chromosome aberration assay and MLA are not appropriate – this is now a follow up assay for ensuring detection of aneugens]
 - An acute bone marrow micronucleus assay

Proposed new recommended test battery



- 1 *in vitro* test for chromosomal aberrations, or *in vitro* assay for micronuclei, or *in vitro* mouse lymphoma cell *tk* mutation assay
- 2 including initially positive results considered non-relevant based on weight of evidence
- 3 positive and considered relevant (or insufficient evidence to establish that it is not relevant)
- 4 *in vivo* MNT integrated into a general toxicological study only if the top dose is appropriate and 2nd end-point/tissue (integrated where possible) also with an acceptable top dose.
- 5 if top dose is not appropriate:
2 "acute" genotoxicity assays

Requirements 4 + 5 apply also to OPTION 2
If toxicology study not suitable for integrated genotox,
Can revert to option 1 and do *in vitro* assay

Predictivity of short term tests for carcinogenicity

- Kirkland et al [Mutation Research 584 (2005) 1-256]
- Gold database : genotoxicity results available for
 - 553 rodent carcinogens; 177 non-carcinogens
- Re-evaluated data with up to date criteria as far as possible
- Sensitivity (Positive results with carcinogens)
 - Ames or mammalian cell assays: 60-80%; increased when tests were combined (>90%)
- Specificity (negative results with non-carcinogens)
 - Ames test >70%
 - **Mammalian cell tests only 35-45%**

What if the chemicals in the Gold database are not representative of e.g., new pharmaceuticals?

- Matthews et al 2006 (Reg. Tox Pharmacol 44, 97 – 110) reached similar conclusions with a wider database and using conservative weight of evidence criteria to classify chemical carcinogens
 - MLA and CAbs in vitro are not predictive of rodent carcinogenicity
 - Roughly equal numbers of false positives and true positives
- 15/76 = 20% of non-carcinogenic pharmaceuticals in PDR are +ve in CAbs (Snyder & Green, 2001, Mutat Res. 488,151-169)
 - How many compounds were dropped from development because of +ve CA or MLA results that would also have been non-carcinogenic?

17% (of 313) marketed pharmaceuticals have a C_{max} of more than ~ 10 µg/ml (or ~50 µM)

- 19 Antibiotics; 6 Antitumor; 4 Antiviral
- Others: 24

e.g.

- Acetaminophen: 20 µg/ml (~130 µM)
 - Acetylsalicylic acid: 24 µg/ml (~130 µM)
 - Clofibrate: 109 µg/ml (450 µM)
 - Ibuprofen: 61 µg/ml (~300 µM)
 - Theophylline: 14 µg/ml (~60 µM)
- 11 pharmaceuticals reach plasma exposure ≥ 100 µg/ml
 - 6 antibiotics, 3 antitumor, lithium (1-2 mM); clofibrate (109 µg/ml)

1 mM is typically 300 – 500 µg/ml

What about tissue accumulation of drugs?

- Tissue levels may be higher than in plasma
 - especially if high lipophilicity and/or long half-life (often connected)
 - e.g. fluoxetine (Prozac®)
 - cationic amphiphilic drug for depression
 - elimination half life 1-3 days
 - active metabolite has elimination half life 4 -16 days)
 - Accumulates in the brain to $\sim 10\mu\text{g/ml}$ ($\sim 35\mu\text{M}$)
 - brain/plasma ratio 20:1
- Tissue levels for drugs may be 10-20 times higher than plasma levels but high plasma levels plus 10-20 fold accumulation in tissue is not known for pharmaceuticals