

薬食審査発 0920 第 2 号
平成 24 年 9 月 20 日

各都道府県衛生主管部（局）長 殿

厚生労働省医薬食品局審査管理課長

医薬品の遺伝毒性試験及び解釈に関するガイダンスについて

医薬品の製造販売承認に際して添付すべき遺伝毒性に関する資料に関して、日米EU医薬品規制調和国際会議（以下「ICH」という。）における合意に基づき、新たに「医薬品の遺伝毒性試験及び解釈に関するガイダンス」（以下「新ガイダンス」という。）を別添のとおり定めたので、下記事項を御了知の上、貴管内関係者に対する周知方よろしく願いたい。

なお、本通知の適用に伴い、平成 8 年 7 月 2 日付け薬審第 444 号厚生省薬務局審査課長通知「医薬品のための遺伝毒性試験の特定項目に関するガイダンスについて」（以下、「S2A ガイダンス」という。）及び平成 10 年 7 月 9 日付け医薬審第 554 号厚生省医薬安全局審査管理課長通知「「遺伝毒性試験：医薬品の遺伝毒性試験の標準的組合せ」について」（以下、「S2B ガイダンス」という。）は廃止する。

記

1. 背景

優れた医薬品の国際的な研究開発の促進及び患者への迅速な提供を図るため、承認審査資料の国際的な調和を図る必要性が指摘されていることから、このような要請に応えるため、ICH における合意に基づき、近年の科学の進歩及び経験を踏まえて、新ガイダンスが定められたものであること。

2. 新ガイダンスの要点

- (1) S2A ガイダンス及び S2B ガイダンスを一つにまとめたこと。
- (2) 遺伝毒性試験の標準的組合せについて、2 つのオプションを提示したこと（*in vitro* ほ乳類細胞試験を含むものと、含まないもの）。
- (3) *in vitro* ほ乳類細胞試験において、最高濃度の上限を 1 mM 又は 0.5 mg/mL のいずれか低い濃度としたこと。

- (4) *in vitro*ほ乳類細胞試験として *in vitro*小核試験の利用を認めたこと。
- (5) 条件によっては、*in vivo* 遺伝毒性試験を反復投与毒性試験に組み込んでもよいこととしたこと。
- (6) 第2の *in vivo*試験として DNA 傷害性を評価できるコメット試験を推奨することとしたこと。
- (7) 条件によっては、細菌を用いる復帰突然変異試験を1試験だけでもよいこととしたこと。

3. 今後の取扱い

本通知日以降、医薬品の製造販売承認申請に際して添付すべき遺伝毒性に関する資料は、新ガイダンスに基づいて実施されたものであること。

ただし、平成25年3月31日までに平成11年11月1日付け医薬審第1604号厚生省医薬安全局審査管理課長通知「医薬品の遺伝毒性試験に関するガイドラインについて」の別添「遺伝毒性ガイドライン」に基づいて開始又は実施された試験による資料は、当分の間、医薬品の製造販売承認申請に際し添付すべき遺伝毒性試験に関する資料として用いることができること。

医薬品の遺伝毒性試験及び解釈に関するガイダンス

目次

1. 緒言	3
1.1 ガイドラインの目的	3
1.2 背景	3
1.3 ガイドラインの適用範囲	3
1.4 一般原則	3
2. 遺伝毒性試験の標準的組合せ	4
2.1 理論的根拠	4
2.2 標準的組合せの 2 つのオプションの詳細	5
2.3 試験組合せの変更	6
2.3.1 探索的臨床試験	6
2.3.2 細菌に毒性を示す化合物	6
2.3.3 構造的に遺伝毒性が予想される化合物	6
2.3.4 <i>In vivo</i> 試験系の利用の限界	6
2.4 生殖細胞に対する変異原性物質の検出	7
3. <i>In vitro</i> 試験に対する勧告	7
3.1 試験の繰り返し及び解釈	7
3.2 細菌を用いる遺伝子突然変異試験	7
3.2.1 最高用量の選択	7
3.2.2 試験のデザイン及びプロトコール	8
3.3 ほ乳類細胞を用いる試験	8
3.3.1 最高濃度の選択	8
3.3.2 試験デザイン及びプロトコール	9
3.3.3 陽性対照	9
4. <i>In vivo</i> 試験に対する勧告	10
4.1 染色体損傷の検出のための <i>in vivo</i> 試験	10
4.2 その他の <i>in vivo</i> 遺伝毒性試験	10
4.3 <i>In vivo</i> 遺伝毒性試験における用量設定	10
4.3.1 短期投与試験	10
4.3.2 反復投与試験	11
4.3.3 血液又は骨髄毒性のある化合物	11

4.4	<i>In vivo</i> 試験結果が陰性の場合の標的組織での曝露証明	12
4.4.1	<i>In vitro</i> 遺伝毒性試験が陽性の場合（又は実施されていない場合）	12
4.4.2	<i>In vitro</i> 遺伝毒性試験が陰性の場合.....	13
4.5	<i>In vivo</i> 試験における採取時間	13
4.6	観察動物数.....	13
4.7	<i>In vivo</i> 遺伝毒性試験におけるげっ歯類の性の選択.....	13
4.8	投与経路.....	13
4.9	<i>In vivo</i> 試験の陽性対照の使用	13
5.	試験結果の評価及び追加試験に関するガイダンス.....	14
5.1	生物学的妥当性の評価.....	14
5.2	<i>In vitro</i> 試験結果の評価.....	14
5.2.1	細菌を用いる復帰突然変異試験で得られた陽性結果の評価.....	14
5.2.2	ほ乳類細胞を用いる試験で得られた陽性結果の評価.....	15
5.2.3	<i>In vitro</i> の陰性結果の評価.....	15
5.3	<i>In vivo</i> 試験で得られた結果の評価.....	15
5.4	陽性結果に対する追加検討	16
5.4.1	ほ乳類細胞を用いる <i>in vitro</i> 試験結果に対する追加検討.....	16
5.4.2	<i>In vivo</i> 小核試験の陽性結果に対する追加検討.....	17
5.5	がん原性試験で認められた腫瘍発生に関連する追加の遺伝毒性試験	17
6.	注記.....	17
7.	用語の解説.....	24
8.	参考文献.....	26

医薬品の遺伝毒性試験及び解釈に関するガイダンス

1. 緒言

1.1 ガイドラインの目的

このガイダンスは ICH ガイダンス S2A 及び S2B に替わるもので、両者を 1 つにまとめたものである。改訂の目的は、ヒトに対するリスクを予測するための遺伝毒性試験の標準的組合せを最適化すること、及び結果の解釈のためのガイダンスを提供することであり、遺伝子の変化に基づく発がん性のリスク評価の精度向上を最終的な目的としている。本改訂ガイダンスは、国際的に合意された追加試験の基準並びに標準的組合せの *in vitro* 及び *in vivo* 遺伝毒性試験における陽性結果の解釈（妥当性のない[non-relevant]所見の評価を含む。）について述べ、医薬品として開発される製品についてのみ適用される。

1.2 背景

最新の経済開発協力機構（OECD）ガイドラインの勧告及び「遺伝毒性試験に関する国際ワークショップ（IWGT）」の報告書が適切に考慮されてきた。このガイダンスで述べるように、OECD 及び IWGT による勧告とは、いくつかの相違点がある。相違点は本ガイドラインを優先する。このガイダンスに続く注記は他の ICH ガイダンスと併せて適用すべきである。

1.3 ガイドラインの適用範囲

このガイダンスは新規の「低分子」医薬品の試験に関するものであり、生物学的製剤には適用されない。臨床開発段階における遺伝毒性試験実施のタイミングについては、ICH M3(R2)ガイダンスで推奨時期が示されている。

1.4 一般原則

遺伝毒性試験は、種々の機序で遺伝的な傷害を引き起こす物質を検出するために考案された *in vitro* 及び *in vivo* 試験と定義することができる。これらの試験は、DNA 損傷及びその損傷が固定された遺伝的傷害を検出する。DNA に対する遺伝的傷害とは、遺伝子突然変異、より広範な染色体の異常又は組換えのことであり、これらは後世代への遺伝的影響の本質であり、また、がん化における多段階過程の一部を担っていると一般的に考えられている。染色体の数的変化もまた腫瘍発生に関連しているとともに、生殖細胞における数的染色体異常誘発能を持つ可能性があることを示している。このような傷害を検出する試験で陽性となった物質は、ヒトに対する発がん物質や変異原物質である可能性がある。特定の化学物質の曝露とヒトでの発がん性の相関が証明されているが、遺伝性疾患について同様の関係を証明することは困難である。そのため、遺伝毒性試験は主としてがん原性を予測するために用いられてきた。しかしながら、生殖細胞系における突然変異はヒトの疾患と明確に関係していることから、ある物質が後世代への影響を引き起こすことが疑われた場合は、がん原性が疑われたのと同様に重大であると考えられる。また、遺伝毒性試験の結果はがん原性試験結果の解釈に有用である。

2. 遺伝毒性試験の標準的組合せ

2.1 理論的根拠

医薬品の申請のためには遺伝毒性の総合的な評価が要求される。広範なレビューにより、細菌を用いる復帰突然変異（エームス）試験で陽性の化学物質の多くが、げっ歯類での発がん物質であることが示されている。ほ乳類培養細胞を用いる *in vitro* 試験を加えることによって、げっ歯類に対する発がん物質の検出感度が増し、検出される遺伝的傷害の範囲が広がるが、これにより逆に予測の特異性が減少する；すなわち、げっ歯類での発がん性に関連しない陽性結果が増加する。しかしながら、1つの試験でがん原性に関連するすべての遺伝毒性機序を検出できないことから、組合せによる試験の実施は依然として妥当なものと考えられる。

試験の標準的組合せは次のとおりである。

- i. 細菌を用いる復帰突然変異試験での変異原性の評価。この試験は妥当性のある遺伝的变化を検出し、げっ歯類及びヒト遺伝毒性発がん物質の大部分を検出することができるとされている。
- ii. ほ乳類細胞での *in vitro* 及び／又は *in vivo* 遺伝毒性の評価。ここでの遺伝毒性は以下のように評価されるべきである。

In vitro 分裂中期染色体異常試験、*in vitro* 小核試験（注1参照）及び L5178Y 細胞を用いるマウスリンフォーマ *Tk*（チミジンキナーゼ）試験（MLA）の3種類の *in vitro* ほ乳類細胞試験系が広く使用されており、それらの信頼性は十分に保証されていると考えられる。これら3つの試験は同程度の検出力を持つと考えられており、このガイドラインで推奨されているプロトコルを用い、標準的組合せ試験として他の遺伝毒性試験と一緒に行う場合は、染色体傷害の検出に関しては互換性がある。

In vivo で変異原性を示し、*in vitro* で変異原性陰性の化合物が存在すること（注2参照）、及び遺伝毒性の評価に、吸収、分布、代謝及び排泄などの要素を加味した試験法を加えることが望ましいことから、*in vivo* 遺伝毒性試験が標準的組合せに加えられている。この理由により、現時点では、末梢血若しくは骨髓中の赤血球の小核又は骨髓における分裂中期細胞の染色体異常のいずれかが評価の対象として選択される（注3参照）。被験物質で処理した動物の培養リンパ球もまた細胞遺伝学的解析に用いることができるが、その方法はまだ一般的ではない。

分裂中期細胞を用いる *in vitro* 及び *in vivo* の染色体異常試験は、染色体の安定性を変化させる様々な異常を検出することができる。染色分体又は染色体が切断されることで、無動原体染色体断片が生じ、これにより小核が形成される。したがって、染色体異常や小核を検出する試験のいずれも染色体の構造異常の検出に適切であると考えられている。小核はまた、分裂後期における1本以上の染色体の極への移動異常によっても形成されるため、小核試験は染色体の数的異常誘発物質を検出できる可能性がある。MLA はチミジンキナーゼ遺伝子の変異を検出する系であるが、この変異は遺伝子突然変異と染色体異常のいずれの機序によっても発生する。MLA はまた染色体の損失も検出可能であることが知られている。

いくつかの追加の *in vivo* 試験は組合せ試験にも用いることができる。また、*in vitro* 及び *in vivo* 試験結果を評価する際の証拠の重み付け（weight of evidence ; WOE）を得るための追加試験と

しても有用である（下記参照）。評価対象となる試験が十分理にかなない、適正な方法により実施され、かつ、標的組織での曝露が証明されていれば（4.4 項参照）、その *in vivo* 試験（通常 2 試験）の陰性結果は、懸念すべき遺伝毒性リスクがないことを示す十分な証明となる。

2.2 標準的組合せの 2 つのオプションの詳細

標準的組合せに関する以下の 2 つのオプションは同等に適切と判断される（注 4 参照）。

オプション 1

- i. 細菌を用いる復帰突然変異試験
- ii. 染色体傷害を検出するための細胞遺伝学的試験（*in vitro* 分裂中期での染色体異常試験又は *in vitro* 小核試験）又はマウスリンフォーマ *Tk* 試験
- iii. *In vivo* 遺伝毒性試験。一般には、げっ歯類造血細胞での染色体傷害、すなわち小核又は分裂中期細胞の染色体異常を検出する試験

オプション 2

- i. 細菌を用いる復帰突然変異試験
- ii. 2 種類の異なる組織における *in vivo* 遺伝毒性試験。一般的には、げっ歯類造血細胞を用いる小核試験及び 2 つ目の *in vivo* 試験。他に適切な方法がない限り、一般的には肝臓の DNA 鎖切断を検出する試験が勧められる（下記及び 4.2 項、注 12 参照）

オプション 1 は、ICH ガイダンス S2A 及び B に準拠していることもあり、主に歴史的な経験で構成されている。しかしながら、オプション 1 と 2 が同等に受け入れられると考える理由は次のとおりである。*In vitro* ほ乳類培養細胞試験で陽性で、適切な組織で十分な曝露量が得られている適正に実施された 2 種の *in vivo* 試験で明らかな陰性の場合、*in vivo* では遺伝毒性を示さないことの十分な証拠と考えられる（5.4.1.1 項以降参照）。したがって、初めから 2 種の *in vivo* 試験を実施するオプションは、*in vitro* の陽性結果の追加検討と同等である（注 4 参照）。

標準的組合せの両オプションは、*in vivo* の短期又は反復投与どちらの試験方法にも組み入れて使うことができる。反復投与する場合、科学的に正しければ一般毒性試験に遺伝毒性の指標を組み込むことを考慮すべきである。1 つの *in vivo* 試験を独立して短期投与で行う場合には 2 つ以上の指標を組み込むことが望まれる。多くの場合、試験開始前に反復投与毒性試験の投与量が適切かどうかの十分な情報が得られていると考えられるので、短期投与が適切かあるいは反復投与に組み入れるのがよいかの判断に使える。

このガイダンスに従い実施され、評価されたいずれかのオプションの標準的試験の組合せにおいて陰性の結果を示す化合物は、通常、遺伝毒性活性を持たないと考えられ、追加試験は必要ない。標準的組合せ試験で陽性の化合物は、その臨床での使用形態にもよるが、より広範な検討をする必要があるものと考えられる（5 項参照）。

オプション 2 において *in vivo* 評価の第 2 の試験として使用できるものは 4.2 項に示すような試験があり、これらのいくつかは反復投与毒性試験に組み入れることができる。肝臓は曝露及び代謝能の観点から特に最適な組織であると考えられるが、第 2 の組織及び試験法は想定可能な機序、代謝又は曝露情報などの要因を考慮して選択すべきである。

染色体の数的異常に関する情報は *in vitro* でのほ乳類細胞試験及び *in vitro* 又は *in vivo* の小核試験より得ることができる。異数性誘発能を示す指標としては、分裂指数の増加、倍数性及び小核の誘発がある。MLA でも紡錘体阻害剤の検出は可能であるとされている。オプション 2 において推奨される *in vivo* の細胞遺伝学的試験は、染色体の損失（異数体の可能性）を直接検出することができる小核試験が推奨されるが、染色体異常試験は推奨されない。

ここでは、試験の標準的組合せを提示したが、これは他の遺伝毒性試験が不十分又は不適當であることを意味している訳ではない。追加実施した試験結果は、標準的組合せで得られた試験結果をより詳細に調査するために使用することができる（4.2 及び 5 項参照）。必要性が示され、かつ十分に評価されている試験であれば、非げっ歯類を含む代替の試験系も利用可能である。

標準的組合せを構成する試験において、技術的な理由で試験が実施できない場合には、有用性が確認された代替試験を十分な科学的正当性を持つものとして利用することができる。

2.3 試験組合せの変更

下記のような状況下にあつては、標準的組合せを変更することが推奨される場合もある。

2.3.1 探索的臨床試験

特定の探索的臨床試験に関しては、*in vivo* での最高投与量設定に相応の理由があれば、限られた遺伝毒性試験又は異なる判断基準を適用してもよい（ICH M3(R2)ガイダンス参照）。

2.3.2 細菌に毒性を示す化合物

細菌に強い毒性を示す化合物（例えばある種の抗生物質）に関しては、毒性が発現しない低濃度で変異原性が誘発されることもあるため、細胞毒性を示す化合物がほ乳類細胞で試験されるのと同様に、細菌を用いる復帰突然変異（エームス）試験は依然実施すべきである。さらに、このような場合には、*in vitro* のほ乳類培養細胞を用いる試験のいずれか 1 つを合わせて実施すべきである。すなわちオプション 1 を選択する。

2.3.3 構造的に遺伝毒性が予想される化合物

多くの「警告部分構造」は細菌の変異原性に関して定義されているため、構造的に遺伝毒性が予想される化合物（注 5 参照）は、通常、標準的組合せ試験により検出可能である。少数の化合物クラスは、細菌を用いる遺伝子突然変異試験よりもほ乳類の培養細胞を用いる染色体異常試験で容易に検出できることが知られている。したがって、警告部分構造を有する化合物における標準的組合せにおける陰性結果は、遺伝毒性を有していない十分な保証となると考えられる。しかしながら、ある種の特異的な警告部分構造を有する化合物群に対しては、標準プロトコールを適切に変更することが望ましい（注 5 参照）。追加試験とするかプロトコールの変更とするかの選択は、問題となっている警告部分構造を有する化学物質に関する化学的性質、既知の反応性及び代謝データに基づいて選択する。

2.3.4 *In vivo* 試験系の利用の限界

骨髄、血液又は肝臓を用いる *in vivo* 試験を実施しても有用な情報が得られない化合物がある。トキシコキネティクスやファーマコキネティクスのデータから、全身での体内吸収がなく、標的

臓器に到達できないような化合物がそれに相当する。例として、ある種の造影剤、アルミニウムを主成分とする制酸剤、いくつかの吸入投与剤及び皮膚又は他の局所適用の医薬品があげられる。投与経路を変更しても標的臓器が十分に曝露されず、最も曝露される組織において適切な遺伝毒性試験が実施できない場合には、*in vitro* 試験系のみでの評価が基本的に適切であるかもしれない。汎用されている試験系ではないものの、これらの試験により、曝露部位に対する遺伝毒性評価を適切に実施することは可能である（注 6 参照）。

2.4 生殖細胞に対する変異原性物質の検出

比較検討から、ほとんどの生殖細胞変異原物質は、体細胞を用いる試験で遺伝毒性が検出されることが定性的に示されている。したがって、体細胞を用いる *in vivo* 遺伝毒性試験の陰性結果は生殖細胞に影響がないことを示していると考えられることができる。

3. In vitro 試験に対する勧告

3.1 試験の繰り返し及び解釈

実験結果の再現性は、新しい手法による研究や、予期しない結果が得られた場合に必須の要素である。しかしながら、医薬品のために標準化され汎用されている定型的な遺伝毒性試験では、繰り返しを必要としない場合がある。これらの試験は特性が明確であり、また十分な内部コントロールを有しているので、明らかな陽性又は陰性の場合、試験の繰り返しは通常必要とされない。理想的には、試験結果は明らかな陽性又は明らかな陰性と確定すべきである。しかしながら、時には試験結果が陽性や陰性の判定基準を満たさないことがあり、その場合には「不確か(equivocal)」とせざるを得ない。このような場合には、統計学的手法の適用が解釈に役立つが、適切な生物学的解釈が極めて重要である。「不確か」な場合、再試験において、(i) 明らかな陽性結果を示せば総合して陽性、(ii) 陰性結果を示せば再現性がないため総合して陰性又は (iii) 再び「不確か」を示せば、最終結論も「不確か」となる可能性がある。

3.2 細菌を用いる遺伝子突然変異試験

OECD ガイドライン（1997）及び IWGT 報告書（Gatehouse ら、1994）にプロトコールに関する助言が記述されている。

3.2.1 最高用量の選択

最高用量

溶解性又は菌の生育阻害が問題とならない場合、最高用量は 5000 µg/plate（被験物質が液体の場合、5 µL/plate）とする。

溶解性の限界

菌の生育阻害がなく、また、最高用量が 5000 µg/plate（被験物質が液体の場合、5 µL/plate）以下という条件下では、析出物が変異コロニーの検出を妨げない限り、析出する用量においても測定する。菌の生育阻害が観察されない場合は、析出する最低用量を最高用量とすべきである。も

し、用量相関的に菌の生育阻害又は復帰変異コロニーの増加が認められた場合には、溶解性に関係なく最高用量は以下の基準に基づくべきである。

生育阻害による制限

エームス試験では、生育阻害の現れる用量を最高用量とし、最高濃度 5000 µg/plate を超えない用量とする。生育阻害は復帰変異コロニー数の減少や背景の細菌の生育（background lawn）の透明化（clearing）又は減少によって検出されることがある。

3.2.2 試験のデザイン及びプロトコール

塩基対置換及びフレームシフト突然変異の検出に OECD が推奨する試験菌株のセットは以下のとおりである。

- ネズミチフス菌 (*Salmonella typhimurium*) TA98
- ネズミチフス菌 TA100
- ネズミチフス菌 TA1535
- ネズミチフス菌 TA1537、TA97 又は TA97a
- ネズミチフス菌 TA102、大腸菌 (*Escherichia coli*) WP2uvrA 又は大腸菌 WP2uvrA/pKM101

OECD ガイドライン及び IWGT 報告書との相違は、細菌を用いる復帰突然変異（エームス）試験については、結果が明らかに陰性又は陽性で、代謝活性化系の存在下及び非存在下ですべての試験菌株を含み、最高用量の選択基準を満たす用量範囲並びに適切な陽性及び陰性対照を設置して実施した場合には、医薬品の遺伝毒性試験の経験に基づき、1 試験で十分であると考えられるという点である。また、医薬品の試験ではプレート法及びプレインキュベーション法はともに受け入れ可能である（注 7 参照）。不確か又は弱い陽性結果が得られた場合には、用量レベルの間隔を変更するなどプロトコールを最適化した再試験を実施することが望ましい。

3.3 ほ乳類細胞を用いる試験

OECD ガイドライン（1997）及び IWGT の公表文献（例；Kirsch-Volders ら、2003；Moore ら、2006）にプロトコールに関する助言が記載されている。MLA の結果の解釈についての助言も、総合的評価ファクター（global evaluation factor）の使用も含め記載されている（Moore ら、2006）。ここでは医薬品の試験について推奨されるものと他とのいくつかの違い、特に最高用量の選択について述べる（詳細は下記参照）。

3.3.1 最高濃度の選択

最高濃度

溶媒若しくは培養液中の溶解性又は細胞毒性が問題とならない場合は、最高濃度の上限は 1 mM 又は 0.5 mg/mL のいずれか低い濃度が推奨される（注 8 参照）。

溶解性の限界

難溶又は不溶の場合において細胞毒性が問題とならなければ、最高濃度は培養中で沈殿がみられ、観察を妨げない最も低い濃度とする。沈殿の評価は、肉眼観察又は光学顕微鏡による観察

などで行う。沈殿が最初から継続してみられるのか、又は培養中に新たに生じるのか処理の終了まで観察し、記載する。

細胞毒性

分裂中期の染色体異常又は小核を調べる *in vitro* の細胞遺伝学的試験では、細胞毒性は、細胞増殖抑制が約 50%を超えないようにする（注 9、10 参照）。MLA では、最高用量において細胞毒性が 80～90%（10～20%RTG）になるようにする（注 9 参照）。

3.3.2 試験デザイン及びプロトコール

In vitro の分裂中期細胞における染色体損傷の細胞遺伝学的評価には、陽性及び陰性対照を設けるとともに、代謝活性化系の存在下及び非存在下での試験の実施が必要である。被験物質の処理は 3～6 時間とし、処理開始から約 1.5 正常細胞周期後に標本を作製する。代謝活性化系の存在下及び非存在下の短時間処理の両方で陰性結果又は不確かな結果の場合には、代謝活性化系の非存在下で約 1.5 正常細胞周期の連続処理が必要である。*In vitro* の小核試験にも同じ原則が適用される。ただし、細胞の分裂期を終了させ、次の間期に入らせるため、一般に被験物質の処理開始から 1.5～2 正常細胞周期後に標本を作製する。これらの *in vitro* 細胞遺伝学的試験では、ヌクレオシドアナログ及びニトロソアミンのようなある種の化学物質に対しては、処理時間を長くするか、あるいは採取時間を遅らせるか、又は回復期間を設けるなどのプロトコールの変更によって容易に検出できる場合もある。染色体異常試験では、分裂中期細胞の出現率とともに倍数体（核内倍加を含む）細胞の出現率を記録することによって倍数性の情報を得おくべきである。MLA では、適切な陽性及び陰性対照を設け、代謝活性化系の存在下及び非存在下で試験を実施する。被験物質の処理は 3～4 時間とする。代謝活性化系の存在下及び非存在下の短時間処理の両方において陰性結果あるいは不確かな結果の場合には、代謝活性化系の非存在下で約 24 時間の連続処理を実施すべきである。標準的な MLA では、(i) 主として小さなコロニーを誘発する陽性対照を用い、(ii) 被験物質が陽性を示した場合、陽性対照、溶媒対照及び被験物質で最大突然変異頻度を示した用量を含む少なくとも 1 陽性用量においてコロニーサイズの分類が必要である。

In vitro のほ乳類細胞試験では、上記に概説（例えば異なる処理時間、代謝活性化系の存在下及び非存在下による試験）したように、試験内に再現性をみる要素が組み込まれている。このような試験で明らかに陰性又は陽性の場合には、追加の確認試験は求められない。不確か、又は弱い陽性結果が得られた場合には、処理濃度の間隔を変更するなどプロトコールを最適化した試験の繰り返しが必要となるかもしれない。

3.3.3 陽性対照

同時陽性対照群は重要であるが、遺伝毒性検出のための *in vitro* ほ乳類細胞試験は十分に標準化されているため、（非代謝活性化系の試験と同時に行うのであれば）代謝活性化系の活性確認と試験系の反応性を証明するための代謝活性化系の存在下での陽性対照のみでよい。

4. In vivo 試験に対する勧告

4.1 染色体損傷の検出のための *in vivo* 試験

In vivo での骨髄細胞における染色体異常の分析、又は小核を有する多染性赤血球の解析は、いずれも染色体異常誘発物質の検出に関して適切と見なされる。骨髄細胞の小核試験に使用する動物種としてラット及びマウスはともに適切であると考えられる。小核はまた、マウス末梢血の幼若赤血球（多染性赤血球）又はラット末梢血の新生網赤血球を用いてもよい（注 3 参照）。同様に、他の動物種の幼若赤血球でも、骨髄又は末梢血における染色体異常誘発物質／数的異常誘発物質を検出する試験として適切な感度が示されており使用できる（注 3 参照）。適切に評価されたものであれば、自動解析装置（画像解析及びフローサイトメトリー）を使用することができる（OECD、1997；Hayashi ら、2000、2007）。染色体異常は、被験物質を投与されたげっ歯類から採取培養された末梢リンパ球においても分析が可能である（注 11 参照）。

4.2 その他の *in vivo* 遺伝毒性試験

標準的組合せ 2（オプション 2）において第 2 の試験として示した *in vivo* 試験は、*in vitro* 及び *in vivo* 試験結果の評価における WOE を高めるための追加試験としても使用できる（注 11 及び 12 参照）。*In vitro* で観察された特定の反応や、作用機序に関する情報は *in vivo* 試験を選択する指針となり得るが、染色体異常又は内在性遺伝子における遺伝子突然変異の試験は、大部分の組織においては標準的な方法としては妥当ではない。突然変異はげっ歯類における導入遺伝子によっても検出できるが、特に細胞分裂の頻度が低い組織（注 12 参照）においては突然変異の発現、固定及び蓄積を必要とするため、長期（例えば 28 日間）の投与が必要となる。したがって、第 2 の *in vivo* 試験では多くの場合、代替指標として DNA 傷害性を評価することになる。これまでに発表された多くの文献や、推奨されるプロトコルを考慮すると、DNA 鎖切断を検出する試験である単細胞ゲル電気泳動（「コメット」）試験及びアルカリ溶出試験が推奨され、これに *in vivo* トランスジェニックマウス突然変異試験及び DNA 共有結合試験（これらの試験は多くの組織に適用できる、注 12 参照）が加わる。さらには肝細胞を用いる不定期 DNA 合成（UDS）試験も利用可能である。

4.3 *In vivo* 遺伝毒性試験における用量設定

通常 3 用量を解析する（Hayashi ら、2005）。

4.3.1 短期投与試験

短期投与試験（一般に 1～3 回投与）では、遺伝毒性試験で推奨される最高用量は限界用量の 2000 mg/kg 又は最大耐量とする。最大耐量とは（例えば小核試験[OECD]）、同様の投与方法で、より高用量を投与すると死亡が予測されるような用量として定義される。コメット試験（Hartmann ら、2003）及びトランスジェニック突然変異試験（Heddle ら、2000）についても同様な限界用量設定が推奨される。

用量設定には骨髄赤血球の増殖抑制も考慮に入れるべきである。低用量側の用量は、最高用量から通常公比約 2～3 の間隔で設定する。

4.3.2 反復投与試験

オプション 1 の組合せ

In vivo 遺伝毒性試験を反復投与毒性試験に組み込む場合、反復投与毒性試験の用量が臨床試験の実施を担保するための基準を満たしていれば、通常その用量は適切と判断される。その場合、その用量設定が *in vivo* 小核試験の OECD ガイドラインの基準と異なってもよい。以上の基準は、*in vitro* のほ乳類細胞試験が陰性（あるいは“妥当性のない陽性”、5 項参照）である場合に適用される。

追加試験又はオプション 2 の組合せ

遺伝毒性兆候への追加試験を実施する場合又は *in vitro* のほ乳類細胞試験を行わないオプション 2 を使用する場合には、その試験における最高用量が遺伝毒性の評価に適切か否かを判断するために評価すべき条件がある。下記に示した基準のうち、いずれか 1 つでも満たすのであれば、小核の評価及びその他の遺伝毒性評価を行う試験（通常はラット）の最高用量として適切であると考えられる。

- i. 溶媒中の医薬品の物理化学的特性に基づく投与可能最大量（MFD）（短期投与試験と同じ溶媒を使用した場合）（注 13 参照）。
- ii. 14 日間以上の試験では、耐量の場合には 1000 mg/kg を限界用量とする。
- iii. 曝露がプラトー／飽和に達する場合、あるいは化合物の蓄積が認められる場合は最大可能曝露量。逆に、親化合物の曝露が経時的に大幅に減少（例えば、初期曝露量から 50% 以上減少）する場合は、通常試験は不適切と考えられる（投与開始数日後に採血された血液試料を除く。）。このような現象が片方の性のみにみられる場合、代謝物の高曝露がみられない限り、曝露減少がみられる性を試験終了時に評価対象とすべきでない。
- iv. 短期投与のデータがある場合、その最高用量（最小致死量付近）の 50% 以上の用量（短期投与による小核試験の最高用量に関しては OECD ガイドラインではその上の用量では死亡が予測される用量と記述されている；他の *in vivo* 試験についても同様のガイダンスがある。[例えば Hartmann ら、2003]）。

毒性を伴わない曝露マージン（臨床曝露の複数倍）のみに基づく用量選択は、十分な正当性があるとは考えられない。

4.3.3 血液又は骨髓毒性のある化合物

紡錘体阻害剤のような異数性を誘発する多くの化合物は、骨髓あるいは血液を用いる小核試験において、その作用は、毒性用量に近い狭い用量範囲のみで検出される。この現象は、他の染色体異常誘発物質についても同様である。赤血球系の細胞に強い毒性（例えば、顕著な多染性赤血球（PCE）又は網赤血球の低下）を示す場合には、用量段階は、細胞毒性を示す最高用量以下、約 2 倍を超えない間隔で設定すべきである。適切な用量が反復投与試験に含まれていなければ、異数性誘発物質及び細胞毒性が強い染色体異常誘発物質を検出するための追加試験として下記のものが考えられる。

- i. 投与期間の増加に関連した毒性が顕著に増加した場合は、投与期間初期（投与 3～4 日）

の血液採取が推奨される。例えば、反復投与試験（例えば 28 日間）で血液若しくは骨髓が小核の評価に使用される場合又は網赤血球を評価する場合は、重篤な血液毒性が小核の検出力に影響を及ぼす可能性がある。すなわち、短期投与で十分に小核を誘発する投与量は、反復投与においては毒性が強く発現しすぎる可能性がある（Hamada ら、2001）。このような場合、投与初期のサンプルは、染色体異常誘発物質又は潜在的異数性誘発物質の検出について確証データを提供することができる（注 14、15 参照）。

- ii. *In vitro* のほ乳類細胞を用いる小核試験
- iii. 短期投与による骨髓を用いる小核試験

4.4 *In vivo* 試験結果が陰性の場合の標的組織での曝露証明

In vivo 試験は遺伝毒性を評価する上で重要な役割を担っている。*In vivo* 試験において、標的組織への被験物質の適切な曝露証明が試験結果の意義を左右する。特に、*in vitro* 試験で明白な遺伝毒性が認められたが *in vivo* 試験で陰性の場合又は *in vitro* ほ乳類細胞を用いる試験を実施されていない場合は標的組織の曝露証明が重要である。以下の項で示すように、曝露証明の方法として、標的組織における毒性又はトキシコキネティクスデータがある。

4.4.1 *In vitro* 遺伝毒性試験が陽性の場合（又は実施されていない場合）

In vivo の曝露評価は、遺伝毒性試験と同じ動物種、系統及び投与経路を用いて、最高用量又は適切な用量で行われるべきである。遺伝毒性が一般毒性試験に組み込まれて評価される場合には、曝露の情報は毒性学的評価と共有できる。

In vivo の曝露証明は次のいずれかによって行われる。

- i. 細胞毒性
 - a. 細胞遺伝学的試験：小核試験では、試験で使用した用量及び採取時間で評価した組織（骨髓あるいは血液）における全赤血球に対する幼若赤血球の割合の有意な変化。染色体異常試験では、分裂指数の有意な減少
 - b. 他の *in vivo* 遺伝毒性試験：肝臓あるいは評価した組織における毒性。例えば、病理組織学的評価又は血液生化学的毒性指標など
- ii. 曝露
 - a. 被験物質又はその関連物質の血中又は血漿中濃度の測定。骨髓は血液がよく灌流する組織であり、被験物質又はその関連物質の血中又は血漿中濃度は、骨髓濃度と通常同等である。また肝臓は、投与経路に関係なく全身曝露によって曝露されると予想される。
 - b. 標的組織における被験物質若しくはその関連物質濃度の直接測定又はオートラジオグラフィによる組織曝露評価

全身曝露が、予測される臨床曝露と同等あるいは低い場合には以下の方法が求められる。

- i. 異なる投与経路の利用
- ii. より高い曝露が得られる異なる動物種の利用
- iii. 異なる組織又は試験法の利用（2.3.4 項「*In vivo* 試験系の利用の限界」参照）

被験物質に適切に曝露されていない場合（例えば標的組織への取り込みが非常に低い場合）には、通常の *in vivo* 遺伝毒性試験はほとんど意味をなさないものと考えられる。

4.4.2 *In vitro* 遺伝毒性試験が陰性の場合

In vitro 試験で遺伝毒性が認められなかった場合、上記の方法を用いて *in vivo*（全身）の曝露を評価する。また、他の目的で実施されたげっ歯類における標準的な吸収、分布、代謝、排泄（ADME）試験の結果から評価してもよい。

4.5 *In vivo* 試験における採取時間

In vivo 小核試験、染色体異常試験及び UDS 試験における採取時間は、OECD ガイドライン（1997）に従うべきである。

小核試験が反復投与試験に組み込まれる場合には、血液あるいは骨髄の採取は最終投与の翌日に行う（上述の追加血液採取の推奨を参照）。

その他の遺伝毒性試験では、採取時間は試験に合わせて選択する；例えば DNA 損傷/DNA 鎖切断の判定は、通常最終投与後数時間（例えば 2～6 時間）に行われる。単回投与する場合は、投与後数時間と投与後 24 時間の 2 ポイントで採取する。

原則として、最高用量又は曝露が適切であれば、どのような投与期間の試験でも受け入れられる。

4.6 観察動物数

観察動物数は、小核試験（OECD）又は他の遺伝毒性試験で現在推奨されている動物数を設定する。一般的にはすべての投与動物を観察する必要はない。遺伝毒性評価用の動物は毒性試験に用いた動物から無作為に抽出されるべきである。

4.7 *In vivo* 遺伝毒性試験におけるげっ歯類の性の選択

男性又は女性にのみに用いられる医薬品の試験を行う場合には、適切な性を用いて試験を実施する。短期投与の *in vivo* 試験では、一般的には片側の性のみで実施する。短期投与試験で両性の使用が考慮されるのは、使用する動物種で、毒性、代謝又は曝露（C_{max} 又は AUC）において毒性学的に意味のある性差が示されている場合のみである。その他の場合は、雄のみの使用が適切である。遺伝毒性試験を雌雄動物の反復投与毒性試験に組み入れる場合には、標本は両性から採取するが、毒性又は代謝において性差を示す十分な根拠がなければ、片側の性の観察のみでよい。投与量は「適切な投与量の基準（4.3.2 及び 4.3.3 項）」に従う。

その他の *in vivo* 遺伝毒性試験についても同じ原則が適用される。

4.8 投与経路

投与経路は通常、経口、静脈内又は皮下などの予定臨床投与経路とするが、局所用剤のような場合には全身曝露を得るために、投与経路を変更してもよい（2.3.4 項参照）。

4.9 *In vivo* 試験の陽性対照の使用

In vivo 試験では、試験施設が試験を行うのに十分な能力がある場合には、定期的に陽性対照

物質の動物に対する反応が確認されていれば、試験ごとに同時陽性対照群を置く必要はないと考えられる（注 16 参照）。

5. 試験結果の評価及び追加試験に関するガイダンス

In vitro 試験系では、げっ歯類のがん原性予測に対して偽陰性及び偽陽性の結果を与えることが試験結果の比較から明らかにされている。遺伝毒性試験の組合せ（*in vitro* 及び *in vivo* 試験）は、大部分の既知ヒト発がん物質でみられるように、直接的に遺伝的傷害を引き起こし作用すると考えられている発がん物質を検出する。したがって、これらの試験の組合せでは非遺伝毒性発がん物質は検出できない。*In vitro* の代謝活性化系に限界があるように、試験条件によっては *in vitro* 試験は偽陰性の結果をもたらすことがある。組合せ試験は、遺伝毒性を示す化合物が偽陰性となる危険性を減少させるようにデザインされているが、ある試験で陽性となった化合物が必ずしもヒトに対して遺伝毒性／発がん性を持つことを意味するものではない。

陽性の *in vitro* データは、医薬品が化学物質の特性として遺伝毒性を持つことを示しているが、多くの場合、これら *in vitro* の陽性結果の生物学的意義は適切な *in vivo* 試験で検証される必要がある。さらに、ある濃度以上でのみ作用を現す間接的遺伝毒性の機序が知られており、このような作用機序を持つ医薬品については、安全レベル（閾値）を設定することが可能であるとされている（5.2 項、Müller と Kasper, 2000 ; Scott ら, 1991 ; Thybaud ら, 2007）。

5.1 生物学的妥当性の評価

試験が適切な用量間隔、適切な毒性レベルで実施された場合、以下のように考えることができる。

In vitro あるいは *in vivo* で、遺伝毒性の明らかな増加が認められるが、その程度が弱い場合はまずその再現性及び生物学的な意義について評価すべきである。生物学的に意味が乏しいと判断される例を以下に示す。

- i. 陰性又は溶媒対照の値と比較して統計学的に有意であるが、試験施設での適切な背景データの統計信頼区間の範囲内にある軽度の増加
- ii. 再現性のない弱い反応／不確かな反応

上記のいずれかの条件があてはまり、WOE から遺伝毒性がないと考えられれば、陰性又は生物学的に妥当性がない所見と判断され、追加試験の必要はない。

5.2 *In vitro* 試験結果の評価

特に細菌を用いる変異原性試験では、陽性結果が不純物に起因していないかを判断するため、被験物質の純度を考慮すべきである。

5.2.1 細菌を用いる復帰突然変異試験で得られた陽性結果の評価

エームス試験で得られた陽性結果は、DNA との反応性があることを示しているため、適切なリスクベネフィット解析により担保されない限り、*in vivo* での変異原性又は発がん性を評価するための広範囲の追加試験が、患者に投与する際の予測される潜在的风险を評価するために必要である。変異体ではない疑似的なコロニー増加の例が知られている。これらはアミノ酸の混入

によって引き起こされる（ネズミチフス菌の試験系ではヒスチジン、大腸菌の試験系ではトリプトファン）。そのため、細菌を用いる復帰突然変異試験は分解しやすいペプチドの試験には適さない。また、例えば細菌のニトロ還元酵素による活性化のように、細菌の特異的な代謝が関与する陽性反応も存在し、ヒトでの遺伝毒性とは無関係の場合もある。

5.2.2 ほ乳類細胞を用いる試験で得られた陽性結果の評価

陽性結果が得られた場合の WOE による評価及び追加試験の推奨に関しては IWGT の報告書の中で議論されている（例えば Thybaud ら、2007）。さらに、論文などで *in vitro* 試験での陽性結果の妥当性を疑わせるようないくつかの条件についても報告されている。したがって、いかなる *in vitro* 試験での陽性結果についても、以下に示すような WOE を考慮して評価されるべきである。以下の項目はすべてを網羅したものではないが、判定を下すための一助となろう。

- i. *In vivo* では起こりえない条件（pH；浸透圧；析出物）
1 mM までは浸透圧の増加を考慮する必要はない。また、被験物質が pH を変化させる場合は、被験物質処理時の pH を無処理群の正常 pH に調整することが推奨される。
- ii. 強い毒性が発現する濃度のみでの作用
MLA において RTG が 80%以上低下した場合
In vitro 細胞遺伝学的試験において細胞増殖が 50%以上抑制された場合

上記の条件があてはまり、WOE から遺伝毒性の可能性がないと判断できれば、標準的組合せ（オプション 1）に従う。このような場合には、1 つの *in vivo* 試験の実施で十分であると考えられる。

5.2.3 *In vitro* の陰性結果の評価

In vitro 試験で陰性でも、次のような場合には追加試験を考慮すべきである（ここにあげた例はすべてを網羅したものではないが、判断を下すための一助となろう。）：化合物の構造やその既知の代謝経路から考えて、標準的な *in vitro* 代謝活性化法（例えば、げっ歯類肝 S9）が不適切と考えられる場合。化合物の構造や既知の活性から考えて、他の方法／試験系が適切と考えられる場合。

5.3 *In vivo* 試験で得られた結果の評価

In vivo 試験には、吸収、分布及び排泄という *in vitro* 試験にはない要素が勘案されているという特徴があり、したがって、ヒトへの適用に重要な意味を持つ。さらに、薬物代謝は *in vitro* で通常使用されている系と比べると、*in vivo* の系の方がより生物学的に妥当性がある。*In vivo* と *in vitro* の結果が一致しない場合には両者の相違についてケース・バイ・ケースで判断され、説明がなされるべきである（例えば、代謝の差、*in vivo* での効率的な排泄など）。

In vivo 試験においても偽陽性（misleading positive）結果を引き起こす可能性がある。例えば、

- i. 遺伝毒性物質の投与がない場合でも小核の増加がみられることがあり、これは造血障害によることが知られている（Tweats ら、2007, I）。
- ii. DNA 付加体のデータは、既知の内因性付加体の背景レベルを考慮に入れて解釈すべきである。

- iii. 毒性に関連した間接的な作用が、DNA 鎖切断の結果に影響を及ぼすことがある（例えば、アルカリ溶出試験及びコメット試験において）。

このように、遺伝毒性データを評価する際にはすべての毒性学的及び血液学的所見を考慮することが重要である（注 15 参照）。毒性学的変化に関連する間接的な作用には安全域があり、これらが臨床で発現するとは考えにくい。

5.4 陽性結果に対する追加検討

5.4.1 ほ乳類細胞を用いる *in vitro* 試験結果に対する追加検討

以下の論議は細菌を用いる復帰突然変異試験は陰性であることを前提としてなされている。

5.4.1.1 作用機序/*in vivo* の追加検討

ほ乳類細胞を用いる *in vitro* 試験での陽性結果に妥当性がないことを示す WOE が不十分である場合、追加の *in vitro* 試験（下記 i）又は2 種類の適切な *in vivo* 試験を実施することを推奨する（下記 ii）。これらは、ほ乳類細胞試験陽性に対して推奨される追加検討として実験的根拠を与えるものとなるであろう。

- i. 陽性結果が妥当性を欠くものであることを示す作用機序に関する情報は、しばしば *in vitro* から得られる。例えば、染色体異常を誘発したり、また MLA で陽性を示したりしても DNA 損傷性がない化合物であるとする証拠もある（例：エームス試験に加え、他の突然変異/DNA 傷害試験で陰性；化学構造に関する考察）。また、*in vivo* では妥当とされない又は閾値が想定される間接的な作用機序の証拠（例：DNA 合成阻害、高濃度のみで産生される活性酸素種）があげられる（Galloway ら、1998；Scott ら、1991；Muller と Kasper、2000）。同様のケースは、*in vitro* 小核試験における陽性結果の追加検討についても言える。染色体の損失又は異数性の機序が考えられ、染色体の損失を示す動原体染色実験（注 17 参照）が証拠としてあげられる。倍数性は *in vitro* の染色体異常試験ではよくみられる所見である。異数性誘発物質は倍数性を誘発するが、倍数性だけでは異数性誘発の証拠としては不十分である。単に細胞周期の遅延によるのかもしれない。それは、通常、細胞毒性の増加も伴う。*In vitro* 試験において構造的な染色体異常はないが倍数性がみられる場合には、適切に曝露されたことが確認された *in vivo* 小核試験の陰性結果により、異数性誘発能を有さないことが十分に担保される。

上記の機序に関する情報及び WOE が、陽性結果が妥当性を欠くことの裏づけとなるのであれば、適切な曝露情報を有する 1 つの *in vivo* 試験の陰性により、遺伝毒性がないことを示すことができる。これに用いられるのは、一般に細胞遺伝学的試験であり、染色体の損失を追加検討する場合は *in vivo* 小核試験が推奨される。

妥当性がないと判断しうる十分な WOE がない場合、又は作用機序に関する情報がない場合には、2 種類の *in vivo* 試験が要求される。この場合、適切な遺伝毒性の指標及び組織（通常 2 つの異なる組織）で行う。さらに、*in vivo* モデルで十分な曝露が得られることが重要である。

又は

ii. 曝露証明を伴い異なる組織を用いての適切な *in vivo* 試験を 2 種類実施する。

以上を要約すると、試験が的確になされ、曝露が確認された適切な *in vivo* 試験の陰性結果は (4.4.1 項参照)、遺伝毒性を示さない十分な証拠となる。

5.4.1.2 S9 活性化系存在下での *in vitro* 試験の陽性結果に対する追加検討

S9 活性化系の存在下のみで陽性結果がみられた場合には、代謝活性化がその要因であり、代謝活性化以外の条件（例えば、非代謝活性化系のインキュベーションにおける 10%以上の血清と比較した、S9 mix 存在下での低濃度の血清又は無血清）が関与していないことを確認する。ここでの追加試験の目的は *in vitro* での結果の *in vivo* 条件に対する妥当性を確認することであり、通常、肝臓を用いる *in vivo* 試験が対象となる（注 18 参照）。

5.4.2 *In vivo* 小核試験の陽性結果に対する追加検討

In vivo で小核の増加がみられる場合には、非遺伝毒性作用が原因あるいは関与している可能性を判断するため、すべての毒性試験データを考察する（注 15 参照）。造血障害あるいは生理的攪乱（低体温、高体温）のような非特異的作用が疑われる場合には、*in vitro* 染色体異常試験がより適切かもしれない。小核の「真の」増加が懸念される場合には、その増加が染色体の損失によるものか、又は染色体の切断に起因するかを立証する必要がある（注 17 参照）。例えば紡錘体阻害剤による異数性の誘発は、非線形の用量相関反応性を示すという証拠がある。したがって、染色体の損失には閾値があり、それより低い曝露では染色体損失は生じず、臨床での曝露と比較して適切な安全域があると判断することは可能であろう。

結論として、化合物の遺伝毒性誘発能の評価は、得られた所見全体を考慮に入れ、*in vitro* 及び *in vivo* 両試験の本質的な意義及び限界を認識すべきである。

5.5 がん原性試験で認められた腫瘍発生に関連する追加の遺伝毒性試験

標準的組合せ試験では陰性であるが、がん原性試験で腫瘍の発生頻度の増加を示し、不十分な証拠ではあるが非遺伝毒性の機序が示された化合物に関しては、適切な試験系を用いた追加試験の実施が望まれる。作用機序の理解を助けるための補助的な試験として、*in vitro* 試験における代謝活性化の条件の変更や、腫瘍が誘発された標的臓器における遺伝子損傷を指標とする *in vivo* 試験、例えばコメット試験あるいはアルカリ溶出試験のような DNA 鎖切断試験、肝 UDS 試験、DNA 共有結合性（例えば、³²P ポストラベル法）、導入遺伝子における突然変異誘発又は腫瘍関連遺伝子の遺伝的変異の分子レベルでの解析が含まれる（Kasper ら、2007）。

6. 注記

1. *In vitro* の小核試験は国際共同研究（Kirsch-Volders ら、2003）において幅広く評価され、ECVAM（Corvi ら、2008）によって有効性が検証され、OECD ガイドライン 487（2010）として策定された。
2. 標準的組合せの *in vitro* 試験では陰性若しくは弱陽性又は相反する結果しか得られないが、骨髄の染色体損傷試験で明らかに陽性となる遺伝毒性発がん物質が少数ながら存在する。

プロカルバジン、ヒドロキノン、ウレタン、ベンゼンなどがこの分類に入る。企業の調査から、その他のいくつかの例が Tweats ら (2007,II) によって紹介されている。

3. 原則として、造血細胞の小核は、骨髄ではいかなる動物種においても評価可能であり、また、血液では全身を循環している小核含有赤血球が脾臓で取り除かれない動物種において評価可能と考えられる。マウスの場合、血液では多染性赤血球を用いて小核の計測が可能であり、約 4 週間以上の継続投与をした場合には成熟（正染性）赤血球も使用可能である。ラットの場合、小核を有する赤血球は速やかに血液中から取り除かれるが、一連の染色体異常誘発物質及び異数性誘発物質による小核の誘発がラットの網赤血球を用いて検出可能なことが確認されている (Wakata ら、1998 ; Hamada ら、2001)。ラットの血液は新しく生成された網赤血球を確実に観察でき (Hayashi ら、2007 ; MacGregor ら、2006)、骨髄よりも低い血液での小核頻度を検出できる適当な統計学的な感度を与えられるだけの十分な観察細胞数があれば、小核試験に利用できる (Kissling ら、2007)。骨髄又は血液で、自動計測又はマニュアル計測のいずれの方法を選択しても、各試験施設は動物間のばらつきを下回るレベルに測定誤差を維持できるだけの最小限の観察細胞数を決定する必要がある。

イヌ又はアカゲザルを用いた小核誘発の検討が行われており、現在利用可能となっている (Harper ら、2007 ; Hotchkiss ら、2008)。このような代替動物種が有用とされる一例として、げっ歯類では十分に認められないがイヌやサルでは生成されるようなヒト代謝物の評価があげられる。

4. 試験の組合せに示す 2 つのオプションは同様に適切であるが、それぞれの被験物質特有の性質により、一方がより適切となる場合がある。例えば、実験動物における全身曝露が、臨床で予想される曝露と同等かそれ以下の場合には、*in vitro* 試験を選択すべきである：オプション 1 (2.3.4 及び 4.4.1 参照)。一方、オプション 2 は肝臓での試験も含んでおり、肝臓で短寿命の活性代謝物が生成されると予想される場合に推奨される。
5. ある種の警告部分構造を有する化学物質は、発がん性や変異原性とある程度関連があると考えられている。警告部分構造の例として、アルキル化求電子中心、不安定型エポキシド、芳香族アミン類、アゾ構造、N-ニトロソ基及び芳香族ニトロ基があげられる (Ashby と Paton、1994)。特別な警告部分構造を有するある種の化学物質では、プロトコールの特別な変更／追加による試験が遺伝毒性の検出に重要であることが明らかにされている（例えばアゾ基含有化学物質、配糖体、活性化にニトロ還元を必要とするニトロイミダゾールのような化学物質、代謝活性化に別種のげっ歯類の S9 を必要とするフェナセチンのような化学物質）。
6. 皮膚及び結腸における *in vivo* 小核試験が開発されている (Hayashi ら、2007)。また、これらの組織では DNA 損傷試験も用いることが可能である。
7. プレート法又はプレインキュベーション法で検出感度に差のある場合があるが、量的な

違いのみで、結果が逆転することはない (Gatehouse ら、1994)。両プロトコールで試験した製薬企業の経験では、2 つの試験法で異なる結果は得られておらず、また、IWGT の報告書 (Gatehouse ら、1994) ではプレインキュベーション法でより容易に検出できるとされる化学物質は一般に医薬品ではなく、また、肝臓を用いる *in vivo* 遺伝毒性試験で陽性を示すものであった。これらには短鎖の脂肪族ニトロソアミン；二価金属；アルデヒド (例えばホルムアルデヒド、クロトンアルデヒド)；アゾ色素 (例えばバタージェロー)；ピロリジジンアルカロイド；アリル化合物 (イソチオシアン酸アリル、塩化アリル) 及びニトロ (例えば芳香族、脂肪族) 化合物が含まれる。

8. *In vitro* のほ乳類培養細胞試験における最大濃度を 1 mM とする理論的根拠は以下のとおりである。試験の標準的組合せはエームス試験と *in vivo* 試験が含まれている。この標準的組合せは、個別の試験結果のみに頼らずに遺伝毒性発がん物質を検出することに最適化されている。エームス試験又は *in vivo* 遺伝毒性試験では検出されないが、*in vitro* のほ乳類細胞試験の 1 mM 以上でのみ検出されるような化学物質 (DNA を損傷する発がん物質) が存在する可能性は低い。さらに、上限の 1 mM は、既知医薬品の組織中の濃度を含めた臨床曝露より高く (Goodman と Gilman、2001)、また、*in vivo* の非臨床試験で一般的に達する量より高いことからハザード同定の色合いが強い。ヌクレオシドアナログ及びいくつかの抗生物質のように、ある種の薬物は極めて高い臨床曝露を必要とすることが知られている。既存薬物との強さの比較は興味のあるところであり、1 mM の上限以上での評価が必要であったとしても、ヒトの安全性を最終的に決定するのは *in vivo* 試験である。極めて低い分子量 (200 以下のような) の医薬品の場合には、より高い試験濃度を考慮すべきである。

9. ある種の遺伝毒性発がん物質は、ある程度の細胞毒性を引き起こす濃度で試験しない限り、*in vitro* 遺伝毒性試験では検出されないが、DNA に直接損傷を与えるような物質は一般に中等度の毒性レベルで検出可能である (Greenwood ら、2004)。細胞毒性が強くなるに従って、試験化合物又はその代謝物による直接的な DNA 損傷以外の作用により、遺伝毒性ではなく細胞毒性に関連した「陽性」結果をもたらすことがある。このような非 DNA 損傷性によって、二次的に起こる間接的 DNA 損傷の誘発は、ある閾値濃度以上で引き起こされることが多い。このような細胞機能障害が薬理学的に妥当性のあるような低濃度で引き起こされるとは考えられない。

染色体異常誘発能の弱い既知発がん物質であっても、細胞遺伝学的試験においては、50%より低い細胞増殖抑制濃度で陽性結果を示す。一方、DNA 損傷、変異原性又はがん原性を有しない化学物質でも、細胞毒性が認められるような濃度で染色体切断を誘発することがある。*In vitro* 細胞遺伝学的試験 (染色体異常試験及び *in vitro* 小核試験) において、約 50%の増殖抑制を上限とすることは適切と考えられる。

株化細胞を用いた細胞遺伝学的試験では、細胞を計数するだけでは毒性を過小評価するおそれのあることが知られているため、時間経過における細胞集団増殖の測定値 (培

養中の細胞数の変化を対照と比較して測定する。例えば細胞集団倍加[PD；注 10 参照]と呼ばれる方法)が、細胞毒性の指標として有用であることが示されている。リンパ球の培養では、約 50%を超えない分裂抑制が十分と考えられる。このためには、分裂中期像の異常をみる試験の場合は分裂指数 (MI) を、*in vitro* 小核試験では細胞分裂阻止に基づく指標を用いることができる。さらに *in vitro* 小核試験については、小核は細胞分裂に続く間期で計数されることから、細胞周期全体が回っているのを確認することが重要である。このためには、細胞分裂はさせないが核分裂には影響しないサイトカラシン B の使用が考えられ、二核細胞における小核を計数すればよいことになる (リンパ球を用いる場合に推奨される。)。株化細胞では、上に記した時間経過による細胞集団倍加 (PD) など、細胞増殖を他の方法で証明することもできる (Kirsch-Volders ら、2003)。

MLA にはソフトアガー法及びマイクロウェル法があるが、ともに最高濃度を相対総増殖率 (RTG) が 20%に近い (10~20%) 濃度に設定することで適切な感受性が得られる (Moore ら、2002)。公表データを現在の基準で再調査すると、RTG が 20%未満の濃度でのみ MLA 結果が陽性であったげっ歯類発がん物質は極めて少数であり、この範疇の化学物質については遺伝毒性発がん物質である信頼に足る証拠がない。20%以下の RTG でのみ突然変異の増加がみられる場合には結果の解釈を慎重に行う必要があり、RTG が 10%以下の濃度のみで突然変異の誘発が認められる場合、陽性とは判断できない。

結論として、細胞増殖/生存細胞率の減少が、細胞遺伝学的試験では 50%、MLA では 80%に達するかそれ以上で得られた陽性結果の解釈には注意が必要である。このような細胞毒性/コロニー形成率 (clonal survival) レベルで処理された細胞を評価することは、感度を高め、不適切な陽性結果をもたらす危険性を増加させる。遺伝毒性試験の組合せは、強い細胞毒性発現用量を用いた単一の *in vitro* ほ乳類細胞試験に頼らなくても、適切な感度を保証できるようデザインされている。

適切な毒性の範囲を得るため、広い範囲の濃度を用いた用量設定のための予備試験は有用であるが、遺伝毒性試験ではしばしば極めて狭い間隔 (公比 2 以下) で数段階の濃度を用いることが重要である。標準的な設定数以上の用量で試験が実施されることもありうるが、すべての処理濃度群について評価する必要はない。正確な 50%の増殖抑制あるいは 80%の RTG 抑制を示す濃度で試験するために、何度も試験を繰り返すことを意図するものではない。

10. *In vitro* の細胞遺伝学的試験において、細胞毒性を評価するためには、細胞の計数では細胞毒性を過小に見積もることがあるため細胞毒性評価には相対細胞増殖率が適している (Greenwood ら、2004)。50%増殖抑制レベルを算定するために細胞集団倍加 (用語の解説を参照) を用いると、DNA に直接損傷を与える物質は確実に陽性となる一方で、変異原性あるいは発がん性を有さない物質の陽性頻度が減少することが示されている。
11. 骨髄の分裂中期細胞が使用されるのと同様に、被験物質を 1 回以上投与された試験動物から採取して培養したリンパ球の分裂中期における染色体異常を調べるのが有用な場

合がある。循環しているリンパ球は複製しないため、遺伝毒性作用に複製が必要な薬剤（例えばいくつかのヌクレオシドアナログ）では、通常のリンパ球細胞を用いた場合には検出されないと考えられる。ある種のリンパ球の寿命は長いため、原理的に修復されない DNA 損傷の蓄積が起こる可能性があり、これらの細胞が *in vitro* で分裂を誘導された時に染色体異常が増加する可能性がある。*In vivo* のリンパ球試験は染色体異常誘発能を調べるための追加試験として有用となりうるが、一般には、造血細胞の小核試験に加え、肝臓のような他の組織を用いた方が、より多くの情報が得られる。すなわち薬物及び代謝物の曝露がしばしば肝臓でより高いためである。

12. 試験の組合せに第 2 の *in vivo* 試験を含めるのは、薬物やその代謝物に十分に曝露される組織を使用することにより遺伝毒性がないことを保証するためである；遺伝毒性を有すると判断される発がん物質のいくつかは肝臓を用いる試験では陽性結果を示すが、骨髓を用いる *in vivo* 細胞遺伝学的試験では陰性を示すことがある。これらの例は、適切な代謝活性化能の欠如又は活性中間体が骨髓の血液系細胞には到達していないことを反映したものであろう。

DNA 鎖切断試験、DNA 付加体試験及び導入遺伝子の突然変異試験は多くの組織に適用できるという利点がある。それらすべての *in vivo* 試験について、国際的に合意されたプロトコールはまだないが、UDS 試験に加え、DNA 鎖切断試験（コメット試験及びアルカリ溶出試験）、DNA 付加体（共有結合）測定及びげっ歯類を用いたトランスジェニック遺伝子突然変異試験に関する多数の公表データや、推奨されるプロトコールがある。MLA で陽性を示し、主に大きなコロニーを誘発するが、*in vitro* 分裂中期細胞試験で染色体切断を示さない化合物に対しては、トランスジェニックマウス遺伝子突然変異試験のような *in vivo* 突然変異試験が DNA 鎖切断試験よりも優先度が高いと考えられる。UDS 試験は、大きな DNA 付加体を誘発、あるいは細胞を用いる復帰突然変異（エームス）試験陽性の化合物に対して有用と考えられる。

細胞毒性は DNA 鎖切断を誘発するため、DNA 鎖切断試験の結果から混乱を招かぬよう注意深い細胞毒性評価が必要である。細胞毒性の問題は、*in vitro* アルカリ溶出試験では十分にその特性が評価されているが（Storer ら、1996）、コメット試験では完全には評価されていない。原則として、DNA 鎖切断試験を反復投与毒性試験で行う場合は適切な用量設定と採取時間が重要である。

成熟動物の肝臓は分裂が盛んな組織ではないことから、第 2 の試験としては非細胞遺伝学的指標が汎用される。しかし、部分肝切除や幼若ラット（Hayashi ら、2007）などの利用により肝細胞の分裂が認められれば、肝臓を用いた小核試験は実施可能であり、既知の遺伝毒性物質を検出できる。

13. メチルセルロース水溶液に代表されるような水溶性溶媒では通常問題ないが、Tween 80 のような溶媒では反復投与可能な容量は単回投与の 30 分の 1 程度である。
14. 毒性試験において、曝露の測定などの目的で追加の採血が計画されている場合には注意

を要する。そうした失血が小核試験結果に影響を与えることがある。すなわち、失血により刺激された赤血球の新生が小核を有する赤血球を増加させることがある。

15. 小核の増加は、遺伝毒性物質を投与しなくても赤血球生成障害（再生性貧血；髄外造血など）、ストレス、低体温及び高体温により生ずる（Tweats らによる総説、2007,I）。血液の中では、脾臓の機能変化が血液からの小核含有赤血球細胞の除去に影響を及ぼし、循環する小核含有赤血球の増加をきたすと考えられる。

16. 短期試験あるいは反復投与毒性試験における陽性対照：

小核（及びその他の細胞遺伝学的）試験における陽性対照の目的は、標本観察者が小核の増加を間違いなく検出できることを証明することである。これには、陽性対照を短期投与した少数の動物（片性で可）を用いる定期的な試験（数ヵ月毎）から採取したサンプルを使用すればよい。マニュアル計測では、このようなスライドを各試験で観察するコード化したスライドに加えることが可能である。陽性対照のスライドは、その染色特性や小核頻度によって標本観察者に容易に識別されてはならない。自動計測では、各試験で適切な品質保証対照サンプルを使用する。

他の *in vivo* 遺伝毒性試験における陽性対照の目的は、選択した動物種、組織及びプロトコールによる試験が DNA 損傷／変異原性の増加を間違いなく検出できることを証明することである。試験施設が、複数回の独立した試験において適切な陽性対照物質を常に検出可能であることを証明すれば、一般的には陽性対照と被験物質の実験条件も変わらないことを十分に示している。しかしながら、現在のところ、コメット試験では同時の陽性対照をおくことが望ましいと考えられる。

17. 小核の誘発が主に染色体損失によるものか、又は染色体切断によるものかを証明するには、動原体の存在を確認するために、*in vitro* 又は *in vivo* での小核の染色が有効である。例えば、動原体部位の DNA 塩基配列プローブを用いる *in situ* 蛍光ハイブリダイゼーション（FISH）又はキネトコア蛋白への標識抗体を使用する。誘発された小核の大部分が動原体陽性であれば、染色体損失が示唆される（コルヒチン及びビンブラスチンのような強力なチューブリン重合阻害物質でも 100%のキネトコア陽性小核を生ずることはなく、大体 70～80%程度であるが、リスク評価の際はまず異数性誘発物質として認められている。）。代替法として、分裂中期の構造異常を調べる *in vitro* 又は *in vivo* 試験がある；陰性であれば、小核の誘発は染色体損失に関連することを意味する。

18. 標準的な方法で誘導された S9 mix は、ヒト S9 よりも高い活性化能を有し、特定の補因子が供給されない限り第 2 相の代謝（解毒）能を欠いている。また、*in vitro* では高濃度の被験物質の場合、非特異的な活性化が生ずる可能性がある（Kirkland ら、2007）。ヒト S9 又は他のヒトに関連した代謝活性化系を用いた遺伝毒性試験は有用である。前臨床試験で用いた動物種（非誘導のマイクロゾーム、肝細胞又は *in vivo* 組織を含む。）又はヒト試料における既知の代謝物プロファイルと、遺伝毒性試験の代謝物プロファイルを解析し、

比較する事も試験結果の関連性を判断するのに有用であり (Ku ら、2007)、その追加試験では通常、肝臓を用いた *in vivo* 試験に焦点を当てる。S9 存在下の *in vitro* 試験で陽性結果を示す化合物が *in vivo* では遺伝毒性を誘発しない可能性があり、それは、*in vitro* で生成される代謝物が *in vivo* では生成されないか、生成されても極めて少量の場合又は代謝的に解毒されるか若しくは速やかに排泄されるためであり、このような場合、*in vivo* ではリスクがないことを示唆している。

7. 用語の解説

アルカリ溶出試験 (*Alkaline elution assay*) : DNA 鎖切断試験を参照。

異数性 (*Aneuploidy*) : 細胞あるいは生物種に固有な染色体基本数からずれていること。

塩基対置換 (*Base substitution*) : 塩基配列において1つ以上の塩基が他の塩基と置き換わっていること。これにより元のものとは異なるたん白質が合成される可能性がある。

細胞増殖 (*Cell proliferation*) : 細胞分裂して娘細胞をつくる能力。

動原体/キネトコア (*Centromere/kinetochore*) : 姉妹染色分体の接着及び娘染色体の極への移動と娘核への封入を確実にする紡錘糸の付着に重要な染色体の構造。

染色体 (構造) 異常誘発物質 (*Clastogen*) : 染色体の構造的切断を引き起こす物質で、通常光学顕微鏡で検出可能。

コロニー形成率 (*Cloning efficiency*) : 1 個の細胞がクローンを形成する割合。通常少量の細胞を適切な培養条件下で播種した後に測定する。

コメット試験 (*Comet assay*) : DNA 鎖切断試験を参照。

培養飽和密度 (*Culture confluency*) : 目視検査による培養における細胞密度の飽和状態。

細胞遺伝学的評価 (*Cytogenetic evaluation*) : 有糸分裂及び減数分裂における染色体構造の光学顕微鏡による解析、あるいは小核の解析。

DNA 付加体 (*DNA adduct*) : 化学物質と DNA が共有結合した生成物。

DNA 修復 (*DNA repair*) : DNA 損傷後の本来の DNA 配列への再構成。

DNA 鎖切断 (*DNA strand breaks*) : DNA の単鎖あるいは二本鎖切断。

DNA 鎖切断試験 (*DNA strand break assay*) : アルカリ処理により、特定の型の DNA 損傷が一本鎖切断に転換される。これらはフィルターを通過する移動率を測定するアルカリ溶出試験、単細胞ゲル電気泳動試験又はコメット試験 (スライドガラス上に単層ゲルを重ねし、そこに包埋された細胞に電流をあて、電気泳動する方法。DNA の短い断片が核の外側に移動し、「彗星の尾」の様な像を呈し、DNA の移動した程度は、細胞を染色することにより顕微鏡下で計測する。) で検出可能である。

フレームシフト突然変異 (*Frameshift mutation*) : 遺伝子の塩基配列に1個又は2個の塩基が付加 (挿入) 又は欠失した突然変異 (遺伝コードの変化)。これにより、元のものとは異なる、又は短い蛋白質が合成される可能性がある。

遺伝子突然変異 (*Gene mutation*) : 単一の遺伝子又はその調節遺伝子の配列内に生じた恒久的な変化。変化としては点突然変異、挿入、欠失などがある。

遺伝的指標 (*Genetic endpoint*) : 対象とする遺伝的变化の型又はそのクラス (例えば、遺伝子突然変異、染色体異常、DNA 鎖切断、DNA 修復、DNA 付加体の生成など)

遺伝毒性 (*Genotoxicity*) : 誘発の機序に関係なく、遺伝物質に生じた有害な変化の総称。

小核 (*Micronucleus*) : 細胞質中の核 DNA の粒子: 染色体全体又は染色体断片を含んでいる。

分裂指数 (*Mitotic index*) : 染色体標本 (スライド) において、細胞分裂していない (間期) 細胞を含む全細胞中に占める各段階の分裂細胞の割合。

染色体の数的変化 (*Numerical chromosome changes*) : 固有の一倍体あるいは二倍体の染色

体数から染色体の数が増加していること；株細胞では、染色体数のモード値から外れていること。

プラスミド (*Plasmid*) : バクテリアの通常の遺伝子とは別の微小遺伝子。プラスミドは宿主の遺伝子に挿入されるか、染色体外の微小遺伝子として存在する。

点突然変異 (*Point mutations*) : 遺伝コードの変化のことで、通常は単一の DNA 塩基対に限定される。

多染性赤血球 (*Polychromatic erythrocyte*) : 細胞分化の途中にあるリボゾームを持つ未成熟な赤血球で、成熟した正染色性赤血球 (リボゾームを欠く) とは RNA の特異染色により容易に判別ができる。

倍数性 (*Polyplidy*) : 染色体のモード値の数的な異常、半数体数のおよその倍数。核内倍加は、分裂中期に染色体の対が「二重染色体」として結合している状態を意味する。

細胞集団倍加あるいは培養増殖 (*Population doubling or culture growth*) : これはいくつかの方法で算出されうる；適切な計算式の一例を示す：集団倍加＝処理開始時の細胞数 (初期値) (X_0) に対する最終時の細胞数 (N) の割合を対数化し、2 の対数で割った値。 $PD = [\log(N \div X_0)] \div \log 2$

組換え (*Recombination*) : DNA 切断とそれに続く均衡あるいは不均衡な再結合。

RTG (相対総増殖率) (*RTG [relative total growth]*) : この細胞毒性の計算値は、被験物質処理後の細胞増加率 (処理開始から処理後 2 日目までの細胞損失及び細胞増殖に基づく計算値) と処理後 2 日の細胞生存率の積により求められる。

単細胞ゲル電気泳動試験 (*Single cell gel electrophoresis assay*) : コメット試験。DNA 切断試験を参照。

生存 (率) (変異原性試験における) (*Survival [in the context of mutagenicity testing]*) : 死細胞を含む全細胞に占める生存細胞の割合で、通常、ある期間の処理後、細胞染色法あるいはコロニー形成法により求める。

導入遺伝子 (*Transgene*) : 体細胞や生殖系列細胞の宿主遺伝子に組み込まれた外来の遺伝子。

不定期 DNA 合成試験 (UDS) (*Unscheduled DNA synthesis [UDS]*) : DNA 損傷によって誘発される S 期以外の DNA 合成。通常 DNA 除去修復と関連している。

8. 参考文献

Ashby, J., D. Paton, "The influence of chemical structure on the extent and sites of carcinogenesis for 522 rodent carcinogens and 55 different human carcinogen exposures," *Mutation Research* 286:3-74, 1994.

Corvi, R., S. Albertini, T. Hartung, S. Hoffmann, D. Maurici, S. Pfuhler, J. van Benthem, P. Vanparys, "ECVAM Retrospective Validation of the *in vitro* micronucleus test (MNT)", *Mutagenesis*, submitted 2007.

.
Gatehouse, D., S. Haworth, T. Cebula, E. Gocke, L. Kier, T. Matsushima, C. Melcion, T. Nohmi, T. Ohta, S. Venitt, E. Zeiger, "Report from the working group on bacterial mutation assays: International Workshop on Standardisation of Genotoxicity Test Procedures," *Mutation Research* 312: 217-233, 1994

Goodman & Gilman "The Pharmacological Basis of Therapeutics". J. G. Hardman, L E. Limbird, A. G. Gilman (Eds.), McGraw-Hill Professional, New York; 10th edition (August 13, 2001).

Greenwood, S.K., R.B. Hill, J.T. Sun, M.J. Armstrong, T.E. Johnson, J.P. Gara, S.M. Galloway, "Population Doubling: A simple and more accurate estimation of cell growth suppression in the *in vitro* assay for chromosomal aberrations that reduces irrelevant positive results," *Environmental and Molecular Mutagenesis*, 43:36-44, 2004.

Hamada S., S. Sutou, T. Morita, A. Wakata, S. Asanami, S. Hosoya, et al., "Evaluation of the rodent micronucleus assay by a 28-day treatment protocol: summary of the 13th collaborative study by the collaborative study group for the micronucleus test (CSGMT)/Environmental Mutagen Society of Japan (JEMS)–Mammalian Mutagenicity Study Group (MMS)," *Environmental and Molecular Mutagenesis*, 37:93-110, 2001.

Hartmann A., E. Agurell, C. Beevers, S. Brendler-Schwaab, B. Burlinson, P. Clay, A. Collins, A. Smith, G. Speit, V. Thybaud, R.R. Tice, "Recommendations for conducting the *in vivo* alkaline Comet assay", *Mutagenesis* 18:45-51, 2003.

Hayashi, M., J.T. MacGregor, D.G. Gatehouse, I. Adler, D.H. Blakey, S.D. Dertinger, G. Krishna, T. Morita, A. Russo, S. Sutou, "In vivo Rodent Erythrocyte Micronucleus Assay. II. Some Aspects of Protocol Design Including Repeated Treatments, Integration With Toxicity

Testing, and Automated Scoring," *Environmental and Molecular Mutagenesis*, 35:234-252, 2000.

Hayashi, M., J.T. MacGregor, D.G. Gatehouse, D.H. Blakey, S.D. Dertinger, L. Abramsson-Zetterberg, G. Krishna, T. Morita, A. Russo, N. Asano, H. Suzuki, W. Ohyama, D. Gibson, "In vivo erythrocyte micronucleus assay. III. Validation and regulatory acceptance of automated scoring and the use of rat peripheral blood reticulocytes, with discussion of non-hematopoietic target cells and a single dose-level limit test," *Mutation Research*, 627:10-30, 2007.

Heddle, J.A., S. Dean, T. Nohmi, M. Boerrigter, D. Casciano, G..R Douglas, B.W. Glickman, N.J. Gorelick, J.C. Mirsalis, H-J Martus, T.R. Skopek, V. Thybaud, K.R. Tindall, N. Yajima, "In vivo transgenic mutation assays", *Environmental and Molecular Mutagenesis*, 35:253-259, 2000.

Hotchkiss CE, Bishop ME, Dertinger SD, Slikker W, Moore MM, MacGregor JT. Flow cytometric analysis of micronuclei in peripheral blood reticulocytes IV: an index of chromosomal damage in the rhesus monkey (macaca mulatta). *Toxicol Sci* 2008 ;102:352-8.

Kasper, P., Y. Uno, R. Mauthe, N. Asano, G. Douglas, E. Matthews, M. Moore, L. Müller, M. Nakajima, T. Singer, G. Speit, "Follow-up testing of rodent carcinogens not positive in the standard genotoxicity testing battery: IWGT workgroup report," *Mutation Research*, 627:106-116, 2007.

Kenelly, J.C., R. Waters, J. Ashby, P.A. Lefevre, B. Burlinson, D.J. Benford, S.W. Dean, I deG. Mitchell, "In vivo rat liver UDS assay. In: Supplementary Mutagenicity Tests, UKEMS Recommended Procedures, ed Kirkland D.J. and Fox M., Cambridge University Press, pp 52-77, 1993

Kirkland, D.J., S. Pfuhler, D. Tweats, M. Aardema, R. Corvi, F. Darroudi, A. Elhajouji, H. Glatt, P. Hastwell, M. Hayashi, P. Kasper, S. Kirchner, A. Lynch, D. Marzin, D. Maurici, J-R. Meunier, L. Müller, G. Nohynek, J. Parry, E. Parry, V. Thybaud, R. Tice, J. van Benthem, P. Vanparys, P. White, "How to reduce false positive results when undertaking *in vitro* genotoxicity testing and thus avoid unnecessary follow-up animal tests: Report of the ECVAM workshop," *Mutation Research*, 628:31-55, 2007.

Kirsch-Volders, M., T. Sofuni, M. Aardema, S. Albertini, D. Eastmond, M. Fenech, M. Ishidate, S. Kirchner, E. Lorge, T. Morita, H. Norppa, J. Surralles, A. Vanhauwaert, A. Wakata, "Report from the *in vitro* micronucleus assay working group", *Mutation Research*, 540:153-163, 2003.

Kissling, GE., S.D. Dertinger, M. Hayashi, J.T. MacGregor., "Sensitivity of the erythrocyte micronucleus assay: Dependence on number of cells scored and inter-animal variability", *Mutation Research* 634:235–240, 2007.

Ku, W.W., A. Bigger, G. Brambilla, H. Glatt, E. Gocke, P.J. Guzzie, A. Hakura, M. Honma, H-J. Martus, R.S. Obach, S. Roberts, "Strategy for genotoxicity testing-Metabolic considerations," *Mutation Research*, 627:59-77, 2007.

MacGregor J.T., M.E. Bishop, J.P. McNamee, M. Hayashi, N. Asano, A. Wakata, M. Makajima, J. Saito, A. Aidoo, M.M. Moore, S.D. Dertinger, "Flow Cytometric analysis of micronuclei in peripheral blood reticulocytes: II. An efficient method of monitoring chromosomal damage in the rat", *Toxicological Sciences*, 94:92-107, 2006.

Moore MM, Honma M, Clements J, Harrington-Brock K, Awogi T, Bolcsfoldi G et al. Mouse lymphoma thymidine kinase locus gene mutation assay: follow-up international workshop on genotoxicity test procedures–New Orleans, Louisiana, April 2000. *Environ Mol Mutagen* 2002;40:292-9.

Moore MM, Honma M, Clements J, Bolcsfoldi G, Burlinson B, Cifone M et al. Mouse lymphoma thymidine kinase gene mutation assay: follow-up meeting of the international workshop on genotoxicity testing–Aberdeen, Scotland, 2003–Assay acceptance criteria, positive controls, and data evaluation. *Environ Mol Mutagen* 2006;47:1-5.

Müller, L, P. Kasper, "Human biological relevance and the use of threshold arguments in regulatory genotoxicity assessment: Experience with pharmaceuticals." *Mutation Research*, 464: 9-34, 2000.

OECD Guidelines for Genetic Toxicology (1997) , www.oecd.org/dataoecd.

Scott, D., S.M. Galloway, R.R. Marshall, M. Ishidate Jr., D. Brusick, J. Ashby, B.C. Myhr, "Genotoxicity under extreme culture conditions. A report from ICPEMC Task Group 9", *Mutation Research*, 257:147-204, 1991.

Storer, R.D., T.W. McKelvey, A.R. Kraynak, M.C. Elia, J.E. Barnum, L.S. Harmon, W.W. Nichols, J.G. DeLuca, "Revalidation of the *in vitro* alkaline elution/rat hepatocyte assay for DNA damage: improved criteria for assessment of cytotoxicity and genotoxicity and results for 81 compounds," *Mutation Research*, 368:59-101, 1996.

Suzuki, H., N. Ikeda, K. Kobayashi, Y. Terashima, Y. Shimada, T. Suzuki, T. Hagiwara, S.Hatakeyama, K.Nagaoka, J. Yoshida, Y. Saito, J. Tanaka, M. Hayashi, "Evaluation of liver and peripheral blood micronucleus assays with 9 chemicals using young rats A study by the Collaborative Study Group for the Micronucleus Test (CSGMT)/Japanese Environmental Mutagen Society (JEMS) - Mammalian Mutagenicity Study Group (MMS)," *Mutation Research*, 583: 133-145, 2005

Thybaud, V., M. Aardema, J. Clements, K. Dearfield, S. Galloway, M. Hayashi, D. Jacobson-Kram, D. Kirkland, J. T. MacGregor, D. Marzin, W. Ohyama, M. Schuler, H. Suzuki, E. Zeiger, "Strategy for genotoxicity testing: Hazard identification and risk assessment in relation to *in vitro* testing," *Mutation Research*, 627:41-58, 2007.

Tweats, D. J., D. Blakey, R. H. Heflich, A. Jacobs, S. D. Jacobsen, T. Morita, T. Nohmi, M. R. O'Donovan, Y. F. Sasaki, T. Sofuni, R. Tice, "Report of the IWGT working group on strategies and interpretation of regulatory *in vivo* tests. I. Increases in micronucleated bone marrow cells in rodents that do not indicate genotoxic hazards," *Mutation Research*, 627:78-91, 2007, I.

Tweats, D. J., D. Blakey, R. H. Heflich, A. Jacobs, S. D. Jacobsen, T. Morita, T. Nohmi, M. R. O'Donovan, Y. F. Sasaki, T. Sofuni, R. Tice, "Report of the IWGT working group on strategy/interpretation of regulatory *in vivo* tests. II. Identification of *in vivo*-only positive compounds in the bone marrow micronucleus test," *Mutation Research*, 627:92-105, 2007, II.

Wakata, A., Y. Miyamae, S. Sato, T. Suzuki, T. Morita, N. Asano, T. Awogi, K. Kondo, M. Hayashi, "Evaluation of the Rat Micronucleus Test with Bone Marrow and Peripheral Blood: Summary of the 9th Collaborative Study by CSGMT/JEMS • MMS," *Environmental and Molecular Mutagenesis*, 32:84-100, 1998.

INTERNATIONAL CONFERENCE ON HARMONISATION OF TECHNICAL
REQUIREMENTS FOR REGISTRATION OF PHARMACEUTICALS FOR HUMAN USE

ICH HARMONISED TRIPARTITE GUIDELINE

**GUIDANCE ON GENOTOXICITY TESTING AND
DATA INTERPRETATION FOR
PHARMACEUTICALS INTENDED FOR HUMAN USE**

S2(R1)

Current *Step 4* version
dated 9 November 2011

This Guideline has been developed by the appropriate ICH Expert Working Group and has been subject to consultation by the regulatory parties, in accordance with the ICH Process. At Step 4 of the Process the final draft is recommended for adoption to the regulatory bodies of the European Union, Japan and USA.

S2(R1)
Document History

Code	History	Date
------	---------	------

S2A:
**Guidance on Specific Aspects of Regulatory Genotoxicity Tests
for Pharmaceuticals**

S2A	Approval by the Steering Committee under <i>Step 2</i> and release for public consultation.	10 March 1994
S2A	Approval by the Steering Committee under <i>Step 4</i> and recommendation for adoption to the three ICH regulatory bodies.	19 July 1995

S2B:
**Genotoxicity: A Standard Battery for Genotoxicity
Testing of Pharmaceuticals**

S2B	Approval by the Steering Committee under <i>Step 2</i> and release for public consultation.	2 October 1996
S2B	Approval by the Steering Committee under <i>Step 4</i> and recommendation for adoption to the three ICH regulatory bodies.	16 July 1997

S2(R1):
**Revision of the S2A and S2B Guidelines
which have been merged as part of the revision**

S2(R1)	Approval by the Steering Committee of S2(R1) under <i>Step 2</i> and release for public consultation.	6 March 2008
--------	---	--------------

Current *Step 4* version

S2(R1)	Approval by the Steering Committee of S2(R1) under <i>Step 4</i> and recommendation for adoption to the three ICH regulatory bodies.	9 November 2011
--------	--	-----------------

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

ICH Harmonised Tripartite Guideline

Having reached *Step 4* of the ICH Process at the ICH Steering Committee meeting on 9 November 2011, this Guideline is recommended for adoption to the three regulatory parties to ICH

TABLE OF CONTENTS

1. INTRODUCTION.....	1
1.1 OBJECTIVES OF THE GUIDELINE.....	1
1.2. Background.....	1
1.3. Scope of the Guideline	1
1.4 General Principles.....	1
2 THE STANDARD TEST BATTERY FOR GENOTOXICITY.....	2
2.1 Rationale.....	2
2.2 Description of the Two Options for the Standard Battery.....	2
2.3 Modifications to the Test Battery.....	4
<i>2.3.1 Exploratory Clinical Studies.....</i>	<i>4</i>
<i>2.3.2 Testing Compounds that are Toxic to Bacteria</i>	<i>4</i>
<i>2.3.3 Compounds Bearing Structural Alerts for Genotoxic Activity.....</i>	<i>4</i>
<i>2.3.4 Limitations to the Use of In Vivo Tests.....</i>	<i>4</i>
2.4 Detection of Germ Cell Mutagens	4
3. RECOMMENDATIONS FOR IN VITRO TESTS	5
3.1 Test Repetition and Interpretation	5
3.2 Recommended Protocol for the Bacterial Mutation Assay	5
<i>3.2.1 Selection of Top Dose Level</i>	<i>5</i>
<i>3.2.2 Study Design/Test Protocol</i>	<i>5</i>
3.3 Recommended Protocols for the Mammalian Cell Assays	6
<i>3.3.1 Selection of Top Concentration</i>	<i>6</i>
<i>3.3.2 Study Design/Test Protocols.....</i>	<i>6</i>
<i>3.3.3 Positive Controls</i>	<i>7</i>
4. RECOMMENDATIONS FOR IN VIVO TESTS	7
4.1 Tests for the Detection of Chromosome Damage <i>In Vivo</i>	7
4.2 Other <i>In Vivo</i> Genotoxicity Tests	7
4.3 Dose Selection for <i>In Vivo</i> Assays	8

4.3.1	<i>Short-Term Studies</i>	8
4.3.2	<i>Multiple Administration Studies</i>	8
4.3.3	<i>Testing Compounds that are Toxic for Blood or Bone Marrow</i>	9
4.4	Demonstration of Target Tissue Exposure for Negative <i>In Vivo</i> Test Results	9
4.4.1	<i>When an In Vitro Genotoxicity Test is Positive (or not done)</i>	9
4.4.2	<i>When In Vitro Genotoxicity Tests are Negative</i>	10
4.5	Sampling Times for <i>In Vivo</i> Assays	10
4.6	Number of Animals Analyzed.....	10
4.7	Use of Male/Female Rodents in <i>In Vivo</i> Genotoxicity Tests	10
4.8	Route of Administration	11
4.9	Use of Positive Controls for <i>In Vivo</i> Studies.....	11
5.	GUIDANCE ON EVALUATION OF TEST RESULTS AND ON FOLLOW-UP TEST STRATEGIES.....	11
5.1	Assessment of Biological Relevance	11
5.2	Evaluation of Results Obtained in <i>In Vitro</i> Tests	12
5.2.1	<i>Evaluation of Positive Results Obtained In Vitro in a Bacterial Mutation Assay</i>	12
5.2.2	<i>Evaluation of Positive Results Obtained In Vitro in Mammalian Cell Assays</i>	12
5.2.3	<i>Evaluation of In Vitro Negative Results</i>	12
5.3	Evaluation of Results Obtained from <i>In Vivo</i> Tests.....	13
5.4	Follow-up Strategies for Positive Results	13
5.4.1	<i>Follow-up to Findings in In Vitro in Mammalian Cell Tests</i>	13
5.4.2	<i>Follow-up to a Positive In Vivo Micronucleus Assay</i>	14
5.5	Follow-up Genotoxicity Testing in Relation to Tumor Findings in a Carcinogenicity Bioassay	14
6.	NOTES.....	15
7.	GLOSSARY.....	20
8.	REFERENCES	23

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

1. INTRODUCTION

1.1 Objectives of the Guideline

This guidance replaces and combines the ICH S2A and S2B Guidelines. The purpose of the revision is to optimize the standard genetic toxicology battery for prediction of potential human risks, and to provide guidance on interpretation of results, with the ultimate goal of improving risk characterization for carcinogenic effects that have their basis in changes in the genetic material. The revised guidance describes internationally agreed upon standards for follow-up testing and interpretation of positive results *in vitro* and *in vivo* in the standard genetic toxicology battery, including assessment of non-relevant findings. This guidance is intended to apply only to products being developed as human pharmaceuticals.

1.2. Background

The recommendations from the latest Organization for Economic Co-operation and Development (OECD) guidelines and the reports from the International Workshops on Genotoxicity Testing (IWGT) have been considered where relevant. In certain cases, there are differences from the OECD or IWGT recommendations, which are noted in the text. The following notes for guidance should be applied in conjunction with other ICH guidances.

1.3. Scope of the Guideline

The focus of this guidance is testing of new “small molecule” drug substances, and the guidance does not apply to biologics. Advice on the timing of the studies relative to clinical development is provided in the ICH M3 (R2) guidance.

1.4 General Principles

Genotoxicity tests can be defined as *in vitro* and *in vivo* tests designed to detect compounds that induce genetic damage by various mechanisms. These tests enable hazard identification with respect to damage to DNA and its fixation. Fixation of damage to DNA in the form of gene mutations, larger scale chromosomal damage or recombination is generally considered to be essential for heritable effects and in the multi-step process of malignancy, a complex process in which genetic changes might possibly play only a part. Numerical chromosome changes have also been associated with tumorigenesis and can indicate a potential for aneuploidy in germ cells. Compounds that are positive in tests that detect such kinds of damage have the potential to be human carcinogens and/or mutagens. Because the relationship between exposure to particular chemicals and carcinogenesis is established for humans, whilst a similar relationship has been difficult to prove for heritable diseases, genotoxicity tests have been used mainly for the prediction of carcinogenicity. Nevertheless, because germ line mutations are clearly associated with human disease, the suspicion that a compound might induce heritable effects is considered to be just as serious as the suspicion that a compound might induce cancer. In addition, the outcome of genotoxicity tests can be valuable for the interpretation of carcinogenicity studies.

2 THE STANDARD TEST BATTERY FOR GENOTOXICITY

2.1 Rationale

Registration of pharmaceuticals requires a comprehensive assessment of their genotoxic potential. Extensive reviews have shown that many compounds that are mutagenic in the bacterial reverse mutation (Ames) test are rodent carcinogens. Addition of *in vitro* mammalian tests increases sensitivity for detection of rodent carcinogens and broadens the spectrum of genetic events detected, but also decreases the specificity of prediction; i.e., increases the incidence of positive results that do not correlate with rodent carcinogenicity. Nevertheless, a battery approach is still reasonable because no single test is capable of detecting all genotoxic mechanisms relevant in tumorigenesis.

The general features of a standard test battery are as follows:

- i. Assessment of mutagenicity in a bacterial reverse gene mutation test. This test has been shown to detect relevant genetic changes and the majority of genotoxic rodent and human carcinogens.
- ii. Genotoxicity should also be evaluated in mammalian cells *in vitro* and/or *in vivo* as follows.

Several *in vitro* mammalian cell systems are widely used and can be considered sufficiently validated: The *in vitro* metaphase chromosome aberration assay, the *in vitro* micronucleus assay (Note 1) and the mouse lymphoma L5178Y cell *Tk* (thymidine kinase) gene mutation assay (MLA). These three assays are currently considered equally appropriate and therefore interchangeable for measurement of chromosomal damage when used together with other genotoxicity tests in a standard battery for testing of pharmaceuticals, if the test protocols recommended in this Guideline are used.

In vivo test(s) are included in the test battery because some agents are mutagenic *in vivo* but not *in vitro* (Note 2) and because it is desirable to include assays that account for such factors as absorption, distribution, metabolism and excretion. The choice of an analysis either of micronuclei in erythrocytes (in blood or bone marrow), or of chromosome aberrations in metaphase cells in bone marrow, is currently included for this reason (Note 3). Lymphocytes cultured from treated animals can also be used for cytogenetic analysis, although experience with such analyses is less widespread.

In vitro and *in vivo* tests that measure chromosomal aberrations in metaphase cells can detect a wide spectrum of changes in chromosomal integrity. Breakage of chromatids or chromosomes can result in micronucleus formation if an acentric fragment is produced; therefore assays that detect either chromosomal aberrations or micronuclei are considered appropriate for detecting clastogens. Micronuclei can also result from lagging of one or more whole chromosome(s) at anaphase and thus micronucleus tests have the potential to detect some aneuploidy inducers. The MLA detects mutations in the *Tk* gene that result from both gene mutations and chromosome damage. There is some evidence that MLA can also detect chromosome loss.

There are several additional *in vivo* assays that can be used in the battery or as follow-up tests to develop weight of evidence in assessing results of *in vitro* or *in vivo* assays (see below). Negative results in appropriate *in vivo* assays (usually two), with adequate justification for the endpoints measured, and demonstration of exposure (see Section 4.4) are generally considered sufficient to demonstrate absence of significant genotoxic risk.

2.2 Description of the Two Options for the Standard Battery

The following two options for the standard battery are considered equally suitable (Note 4):

Option 1

- i. A test for gene mutation in bacteria.
- ii. A cytogenetic test for chromosomal damage (the *in vitro* metaphase chromosome aberration test or *in vitro* micronucleus test), or an *in vitro* mouse lymphoma *Tk* gene mutation assay.
- iii. An *in vivo* test for genotoxicity, generally a test for chromosomal damage using rodent hematopoietic cells, either for micronuclei or for chromosomal aberrations in metaphase cells.

Option 2

- i. A test for gene mutation in bacteria.
- ii. An *in vivo* assessment of genotoxicity with two different tissues, usually an assay for micronuclei using rodent hematopoietic cells and a second *in vivo* assay. Typically this would be a DNA strand breakage assay in liver, unless otherwise justified (see below; also Section 4.2 and Note 12).

There is more historical experience with Option 1, partly because it is based on S2A and B. Nevertheless, the reasoning behind considering Options 1 and 2 equally acceptable is as follows: When a positive result occurs in an *in vitro* mammalian cell assay, clearly negative results in two well conducted *in vivo* assays, in appropriate tissues and with demonstrated adequate exposure, are considered sufficient evidence for lack of genotoxic potential *in vivo* (see Section 5.4.1.1 below). Thus a test strategy in which two *in vivo* assays are conducted is the same strategy that would be used to follow up a positive result *in vitro* (Note 4).

Under both standard battery options, either acute or repeat dose study designs *in vivo* can be used. In case of repeated administrations, attempts should be made to incorporate the genotoxicity endpoints into toxicity studies, if scientifically justified. When more than one endpoint is evaluated *in vivo* it is preferable that they are incorporated into a single study. Often sufficient information on the likely suitability of the doses for the repeat-dose toxicology study is available before the study begins and can be used to determine whether an acute or an integrated test will be suitable.

For compounds that give negative results, the completion of either option of the standard test battery, performed and evaluated in accordance with current recommendations, will usually provide sufficient assurance of the absence of genotoxic activity and no additional tests are warranted. Compounds that give positive results in the standard test battery might, depending on their therapeutic use, need to be tested more extensively (see Section 5).

There are several *in vivo* assays that can be used as the second part of the *in vivo* assessment under Option 2 (see Section 4.2), some of which can be integrated into repeat-dose toxicology studies. The liver is typically the preferred tissue because of exposure and metabolizing capacity, but choice of *in vivo* tissue and assay should be based on factors such as any knowledge of the potential mechanism, of the metabolism *in vivo*, or of the exposed tissues thought to be relevant.

Information on numerical changes can be derived from the mammalian cell assays *in vitro* and from the micronucleus assays *in vitro* or *in vivo*. Elements of the standard protocols that can indicate such potential are elevations in the mitotic index, polyploidy induction and micronucleus evaluation. There is also experimental evidence that spindle poisons can be detected in MLA. The preferred *in vivo* cytogenetic test under Option 2 is the micronucleus assay, not a chromosome aberration assay, to include more

direct capability for detection of chromosome loss (potential for aneuploidy).

The suggested standard set of tests does not imply that other genotoxicity tests are generally considered inadequate or inappropriate. Additional tests can be used for further investigation of genotoxicity test results obtained in the standard battery (see Sections 4.2 and 5). Alternative species, including non-rodents, can also be used if indicated, and if sufficiently validated.

Under conditions in which one or more tests in the standard battery cannot be employed for technical reasons, alternative validated tests can serve as substitutes provided sufficient scientific justification is given.

2.3 Modifications to the Test Battery

The following sections describe situations where modification of the standard test battery might be advisable.

2.3.1 Exploratory Clinical Studies

For certain exploratory clinical studies, fewer genotoxicity assays or different criteria for justification of the maximum dose *in vivo* might apply (see ICH M3(R2) guidance).

2.3.2 Testing Compounds that are Toxic to Bacteria

In cases where compounds are highly toxic to bacteria (e.g., some antibiotics), the bacterial reverse mutation (Ames) test should still be carried out, just as cytotoxic compounds are tested in mammalian cells, because mutagenicity can occur at lower, less toxic concentrations. In such cases, any one of the *in vitro* mammalian cell assays should also be done, i.e., Option 1 should be followed.

2.3.3 Compounds Bearing Structural Alerts for Genotoxic Activity

Structurally alerting compounds (Note 5) are usually detectable in the standard test battery since the majority of “structural alerts” are defined in relation to bacterial mutagenicity. A few chemical classes are known to be more easily detected in mammalian cell chromosome damage assays than bacterial mutation assays. Thus negative results in either test battery with a compound that has a structural alert is usually considered sufficient assurance of a lack of genotoxicity. However, for compounds bearing certain specific structural alerts, modification to standard protocols can be appropriate (Note 5). The choice of additional test(s) or protocol modification(s) depends on the chemical nature, the known reactivity and any metabolism data on the structurally alerting compound in question.

2.3.4 Limitations to the Use of In Vivo Tests

There are compounds for which many *in vivo* tests (typically in bone marrow, blood or liver) do not provide additional useful information. These include compounds for which data on toxicokinetics or pharmacokinetics indicate that they are not systemically absorbed and therefore are not available to the target tissues. Examples of such compounds are some radioimaging agents, aluminum based antacids, some compounds given by inhalation, and some dermally or other topically applied pharmaceuticals. In cases where a modification of the route of administration does not provide sufficient target tissue exposure, and no suitable genotoxicity assay is available in the most exposed tissue, it might be appropriate to base the evaluation only on *in vitro* testing. In some cases evaluation of genotoxic effects at the site of contact can be warranted, although such assays have not yet been widely used (Note 6).

2.4 Detection of Germ Cell Mutagens

Results of comparative studies have shown that, in a qualitative sense, most germ cell

mutagens are likely to be detected as genotoxic in somatic cell tests so that negative results of *in vivo* somatic cell genotoxicity tests generally indicate the absence of germ cell effects.

3. RECOMMENDATIONS FOR *IN VITRO* TESTS

3.1 Test Repetition and Interpretation

Reproducibility of experimental results is an essential component of research involving novel methods or unexpected findings; however, the routine testing of drugs with standard, widely used genotoxicity tests often does not call for replication. These tests are sufficiently well characterized and have sufficient internal controls that repetition of a clearly positive or negative assay is not usually warranted. Ideally it should be possible to declare test results clearly negative or clearly positive. However, test results sometimes do not fit the predetermined criteria for a positive or negative call and therefore are declared “equivocal”. The application of statistical methods can aid in data interpretation; however, adequate biological interpretation is of critical importance. An equivocal test that is repeated might result in (i) a clearly positive outcome, and thus an overall positive result; (ii) a negative outcome, so that the result is not reproducible and overall negative, or (iii) another equivocal result, with a final conclusion that remains equivocal.

3.2 Recommended Protocol for the Bacterial Mutation Assay

Advice on the protocols is given in the OECD guideline (1997) and the IWGT report (Gatehouse et al., 1994).

3.2.1 Selection of Top Dose Level

Maximum dose level

The maximum dose level recommended is 5000 µg/plate (or 5 µL/plate for liquid test substance) when not limited by solubility or cytotoxicity.

Limit of solubility

For bacterial cultures, precipitating doses are scored provided precipitate does not interfere with scoring, toxicity is not limiting, and the top concentration does not exceed 5000 µg/plate (or 5 µL/plate for liquid test substance). If no cytotoxicity is observed, then the lowest precipitating dose should be used as the top dose scored. If dose related cytotoxicity or mutagenicity is noted, irrespective of solubility, the top dose scored should be based on cytotoxicity as described below.

Limit of cytotoxicity

In the Ames test, the doses scored should show evidence of significant toxicity, but without exceeding a top dose of 5000 µg/plate. Toxicity might be detected by a reduction in the number of revertants, and/or clearing or diminution of the background lawn.

3.2.2 Study Design/Test Protocol

The recommended set of bacterial strains (OECD) includes those that detect base substitution and frameshift mutations as follows:

- *Salmonella typhimurium* TA98;
- *Salmonella typhimurium* TA100;
- *Salmonella typhimurium* TA1535;
- either *Salmonella typhimurium* TA1537 or TA97 or TA97a;

- and either *Salmonella typhimurium* TA102 or *Escherichia coli* WP2 *uvrA* or *Escherichia coli* WP2 *uvrA* (pKM101).

One difference from the OECD and IWGT recommendations is that, based on experience with testing pharmaceuticals, a single bacterial mutation (Ames) test is considered sufficient when it is clearly negative or positive, and carried out with a fully adequate protocol including all strains with and without metabolic activation, a suitable dose range that fulfills criteria for top dose selection, and appropriate positive and negative controls. Also, for testing pharmaceuticals, either the plate incorporation or the pre-incubation method is considered appropriate for this single experiment (Note 7). Equivocal or weak positive results might indicate that it would be appropriate to repeat the test, possibly with a modified protocol such as appropriate spacing of dose levels.

3.3 Recommended Protocols for the Mammalian Cell Assays

Advice on the protocols is given in the OECD guidelines (1997) and the IWGT publications (e.g., Kirsch-Volders et al., 2003; Moore et al., 2006). Advice on interpretation of MLA results is also given (Moore et al., 2006), including use of a global evaluation factor. Several differences from these recommendations are noted here for testing pharmaceuticals, notably for selection of the top concentration. (See details below.)

3.3.1 Selection of Top Concentration

Maximum concentration

The maximum top concentration recommended is 1 mM or 0.5 mg/ml, whichever is lower, when not limited by solubility in solvent or culture medium or by cytotoxicity (Note 8).

Limit of solubility

When solubility is limiting, the maximum concentration, if not limited by cytotoxicity, should be the lowest concentration at which minimal precipitate is visible in cultures, provided there is no interference with scoring. Evaluation of precipitation can be done by naked eye or by methods such as light microscopy, noting precipitate that persists or appears during culture (by the end of treatment).

Cytotoxicity

For *in vitro* cytogenetic assays for metaphase chromosome aberrations or for micronuclei, cytotoxicity should not exceed a reduction of about 50% in cell growth (Notes 9 and 10). For the MLA, at the top dose there should be 80-90% cytotoxicity as measured by an RTG between 20-10% (Note 9).

3.3.2 Study Design/Test Protocols

For the cytogenetic evaluation of chromosomal damage in metaphase cells *in vitro*, the test protocol should include the conduct of tests with and without metabolic activation, with appropriate positive and negative controls. Treatment with the test articles should be for 3 to 6 hours with a sampling time approximately 1.5 normal cell cycles from the beginning of the treatment. A continuous treatment without metabolic activation up to the sampling time of approximately 1.5 normal cell cycles should be conducted in case of negative or equivocal results for both short treatments, with and without metabolic activation. The same principles apply to the *in vitro* micronucleus assay, except that the sampling time is typically 1.5 to 2 normal cell cycles from the beginning of treatment to allow cells to complete mitosis and enter the next interphase. For both *in vitro* cytogenetic assays, there might be a need to modify the protocol for certain types of chemicals that could be more readily detected by longer treatment, delayed sampling times or recovery periods, e.g., some nucleoside analogues and some

nitrosamines. In the metaphase aberration assay, information on the ploidy status should be obtained by recording the incidence of polyploid (including endoreduplicated) metaphases as a percentage of the number of metaphase cells. For MLA, the test protocol should include the conduct of tests with and without metabolic activation, with appropriate positive and negative controls, where the treatment with the test article is for 3 to 4 hours. A continuous treatment without metabolic activation for approximately 24 hours should be conducted in case of a negative or equivocal result for both short treatments, with and without metabolic activation. A standard MLA should include (i) the incorporation of positive controls that induce mainly small colonies, and (ii) colony sizing for positive controls, solvent controls and at least one positive test compound concentration (should any exist), including the culture that gave the greatest mutant frequency.

For mammalian cell assays *in vitro*, built-in confirmatory elements, such as those outlined above (e.g., different treatment lengths, tests with and without metabolic activation), should be used. Following such testing, further confirmatory testing in the case of clearly negative or positive test results is not usually warranted. Equivocal or weak positive results might call for repeating tests, possibly with a modified protocol such as appropriate spacing of the test concentrations.

3.3.3 Positive Controls

Concurrent positive controls are important, but *in vitro* mammalian cell tests for genetic toxicity are sufficiently standardized that use of positive controls can generally be confined to a positive control with metabolic activation (when it is done concurrently with the non-activated test) to demonstrate the activity of the metabolic activation system and the responsiveness of the test system.

4. RECOMMENDATIONS FOR *IN VIVO* TESTS

4.1 Tests for the Detection of Chromosome Damage *In Vivo*

Either the analysis of chromosomal aberrations or the measurement of micronucleated polychromatic erythrocytes in bone marrow cells *in vivo* is considered appropriate for the detection of clastogens. Both rats and mice are considered appropriate for use in the bone marrow micronucleus test. Micronuclei can also be measured in immature (e.g., polychromatic) erythrocytes in peripheral blood in the mouse, or in the newly formed reticulocytes in rat blood (Note 3). Likewise, immature erythrocytes can be used from any other species which has shown an adequate sensitivity to detect clastogens/aneuploidy inducers in bone marrow or peripheral blood (Note 3). Systems for automated analysis (image analysis and flow cytometry) can be used if appropriately validated (OECD, 1997; Hayashi et al., 2000; 2007). Chromosomal aberrations can also be analyzed in peripheral lymphocytes cultured from treated rodents (Note 11).

4.2 Other *In Vivo* Genotoxicity Tests

The same *in vivo* tests described as the second test in the standard battery (Option 2) can be used as follow-up tests to develop weight of evidence in assessing results of *in vitro* or *in vivo* assays (Notes 11 and 12). While the type of effect seen *in vitro* and any knowledge of the mechanism can help guide the choice of *in vivo* assay, investigation of chromosomal aberrations or of gene mutations in endogenous genes is not feasible with standard methods in most tissues. Although mutation can be measured in transgenes in rodents, this entails prolonged treatment (e.g., 28 days) to allow for mutation expression, fixation and accumulation, especially in tissues with little cell division (Note 12). Thus the second *in vivo* assay will often evaluate a DNA damage endpoint as a surrogate. Assays with the most published experience and advice on protocols include

the DNA strand break assays such as the single cell gel electrophoresis (“Comet”) assay and alkaline elution assay, the *in vivo* transgenic mouse mutation assays and DNA covalent binding assays, (all of which may be applied in many tissues, Note 12), and the liver unscheduled DNA synthesis (UDS) assay.

4.3 Dose Selection for *In Vivo* Assays

Typically three dose levels are analyzed (Hayashi et al., 2005).

4.3.1 Short-Term Studies

For short-term (usually 1 to 3 administrations) studies, the top dose recommended for genotoxicity assays is a limit dose of 2000 mg/kg, if this is tolerated, or a maximum tolerated dose defined (for example for the micronucleus assay (OECD)) as the dose producing signs of toxicity such that higher dose levels, based on the same dosing regimen, would be expected to produce lethality. Similar recommendations have been made for the Comet assay (Hartmann et al., 2003) and transgenic mutation assay (Heddle et al., 2000). Suppression of bone marrow red blood cell production should also be taken into account in dose selection. Lower doses are generally spaced at approximately two to three fold intervals below this.

4.3.2 Multiple Administration Studies

Option 1 Battery

When the *in vivo* genotoxicity test is integrated into a multiple administration toxicology study, the doses are generally considered appropriate when the toxicology study meets the criteria for an adequate study to support human clinical trials; this can differ from dose selection criteria in the OECD guideline for the *in vivo* micronucleus assay. This applies when the *in vitro* mammalian cell test is negative (or “non-relevant positive”; see Section 5).

Follow-up studies or Option 2 Battery

When carrying out follow-up studies to address any indication of genotoxicity, or when using Option 2 with no *in vitro* mammalian cell assay, several factors should be evaluated to determine whether the top dose is appropriate for genotoxicity evaluation. Any one of the criteria listed below is considered sufficient to demonstrate that the top dose in a toxicology study (typically in rats) is appropriate for micronucleus analysis and for other genotoxicity evaluation:

- i. Maximum Feasible Dose (MFD) based on physico-chemical properties of the drug in the vehicle (provided the MFD in that vehicle is similar to that achievable with acute administration; Note 13).
- ii. Limit dose of 1000 mg/kg for studies of 14 days or longer, if this is tolerated.
- iii. Maximal possible exposure demonstrated either by reaching a plateau/saturation in exposure or by compound accumulation. In contrast, substantial reduction in exposure to parent drug with time (e.g., $\geq 50\%$ reduction from initial exposure) can disqualify the study (unless a blood sample taken in the first few days is available). If this is seen in one sex, generally the sex with reduced exposure would not be scored at the end of the study, unless there is enhanced exposure to a metabolite of interest.
- iv. Top dose is $\geq 50\%$ of the top dose that would be used for acute administration, i.e., close to the minimum lethal dose, if such acute data are available for other reasons. (The top dose for acute administration micronucleus tests is currently described in

OECD guidance as the dose above which lethality would be expected; similar guidance is given (e.g., Hartmann et al., 2003) for other *in vivo* assays.)

Selection of a top dose based only on an exposure margin (multiple over clinical exposure) without toxicity is not considered sufficient justification.

4.3.3 Testing Compounds that are Toxic for Blood or Bone Marrow

Many compounds that induce aneuploidy, such as potent spindle poisons, are detectable in *in vivo* micronucleus assays in bone marrow or blood only within a narrow range of doses approaching toxic doses. This is also true for some clastogens. If toxicological data indicate severe toxicity to the red blood cell lineage (e.g., marked suppression of Polychromatic Erythrocytes (PCEs) or reticulocytes), doses scored should be spaced not more than about 2 fold below the top, cytotoxic dose. If suitable doses are not included in a multi-week study, additional data that could contribute to the detection of aneugens and some toxic clastogens could be derived from any one of the following:

- i. Early blood sampling (at 3-4 days) is advisable when there are marked increases in toxicity with increasing treatment time. For example, when blood or bone marrow is used for micronucleus measurement in a multiweek study (e.g., 28 days), and reticulocytes are scored, marked hematotoxicity can affect the ability to detect micronuclei; i.e., a dose that induces detectable increases in micronuclei after acute treatment might be too toxic to analyze after multiple treatments (Hamada et al., 2001). The early sample can be used to provide assurance that clastogens and potential aneugens are detected (but see Notes 14 and 15).
- ii. An *in vitro* mammalian cell micronucleus assay.
- iii. An acute bone marrow micronucleus assay.

4.4 Demonstration of Target Tissue Exposure for Negative *In Vivo* Test Results

In vivo tests have an important role in genotoxicity test strategies. The value of *in vivo* results is directly related to the demonstration of adequate exposure of the target tissue to the test compound. This is especially true for negative *in vivo* test results when *in vitro* test(s) have shown convincing evidence of genotoxicity, or when no *in vitro* mammalian cell assay is used. Evidence of adequate exposure could include toxicity in the tissue in question, or toxicokinetic data as described in the following section.

4.4.1 When an *In Vitro* Genotoxicity Test is Positive (or not done)

Assessments of *in vivo* exposure should be made at the top dose or other relevant doses using the same species, strain and dosing route used in the genotoxicity assay. When genotoxicity is measured in toxicology assays, exposure information is generally available as part of the toxicology assessment.

Demonstration of *in vivo* exposure should be made by any of the following measurements:

- i. Cytotoxicity:
 - a. For cytogenetic assays: By obtaining a significant change in the proportion of immature erythrocytes among total erythrocytes in the tissue used (bone marrow or blood) at the doses and sampling times used in the micronucleus test or by measuring a significant reduction in mitotic index for the chromosomal aberration assay.
 - b. For other *in vivo* genotoxicity assays: Toxicity in the liver or tissue being assessed, e.g., by histopathological evaluation or blood biochemistry toxicity

indicators.

ii. Exposure:

- a. Measurement of drug related material either in blood or plasma. The bone marrow is a well perfused tissue and levels of drug related materials in blood or plasma are generally similar to those observed in bone marrow. The liver is expected to be exposed for drugs with systemic exposure regardless of the route of administration.
- b. Direct measurement of drug-related material in target tissue, or autoradiographic assessment of tissue exposure.

If systemic exposure is similar to or lower than expected clinical exposure, alternative strategies might be called for such as:

- i. Use of a different route of administration;
- ii. Use of a different species with higher exposure;
- iii. Use of a different tissue or assay (see Section 2.3.4, "Limitations to the use of standard *in vivo* tests").

When adequate exposure cannot be achieved (e.g., with compounds showing very poor target tissue availability) conventional *in vivo* genotoxicity tests have little value.

4.4.2 When *In Vitro* Genotoxicity Tests are Negative

If *in vitro* tests do not show genotoxic potential, *in vivo* (systemic) exposure can be assessed by any of the methods above, or can be assumed from the results of standard Absorption, Distribution, Metabolism and Excretion (ADME) studies in rodents done for other purposes.

4.5 Sampling Times for *In Vivo* Assays

Selection of the sampling time in the *in vivo* Micronucleus (MN), chromosomal aberration and UDS test should follow OECD (1997).

When micronucleus analysis is integrated into multi-week studies, sampling of blood or bone marrow can be done the day after the final administration (see recommendation for additional blood sampling time above).

For other genotoxicity assays, sampling time should be selected as appropriate for the endpoint measured; for example, DNA damage/strand break measurements are usually made a few (e.g., 2-6) hours after the last administration for the multiple daily administration. In the case of single administration, two sampling times should be used: a few hours and 24 hours after the treatment.

In principle, studies of any length can be considered appropriate, provided the top dose/exposure is adequate.

4.6 Number of Animals Analyzed

The number of animals analyzed is determined by current recommendations for the micronucleus assay (OECD) or other genotoxicity assays and generally does not include all the animals treated for a toxicology study. Animals used for genotoxicity analyses should be randomly selected from the group used for the toxicology study.

4.7 Use of Male/Female Rodents in *In Vivo* Genotoxicity Tests

If sex-specific drugs are to be tested, then the assay can be done in the appropriate sex. *In vivo* tests with the acute protocol can generally be carried out in only one sex. For

acute tests, both sexes should be considered only if any existing toxicity, metabolism or exposure (C_{max} or AUC) data indicate a toxicologically meaningful sex difference in the species being used. Otherwise, the use of males alone is considered appropriate for acute genotoxicity tests. When the genotoxicity test is integrated into a repeat-dose toxicology study in two sexes, samples can be collected from both sexes, but a single sex can be scored if there is no substantial sex difference evident in toxicity/metabolism. The dose levels for the sex(es) scored should meet the criteria for appropriate dose levels (see Sections 4.3.2 and 4.3.3).

Similar principles can be applied for other established *in vivo* genotoxicity tests.

4.8 Route of Administration

The route of administration is generally the expected clinical route, e.g., oral, intravenous or subcutaneous, but can be modified if appropriate in order to obtain systemic exposure, e.g., for topically applied compounds (see Section 2.3.4).

4.9 Use of Positive Controls for *In Vivo* Studies

For *in vivo* studies, it is considered sufficient to treat animals with a positive control only periodically, and not concurrently with every assay, after a laboratory has established competence in the use of the assay (Note 16).

5. GUIDANCE ON EVALUATION OF TEST RESULTS AND ON FOLLOW-UP TEST STRATEGIES

Comparative trials have shown conclusively that each *in vitro* test system generates both false negative and false positive results in relation to predicting rodent carcinogenicity. Genotoxicity test batteries (of *in vitro* and *in vivo* tests) detect carcinogens that are thought to act primarily via a mechanism involving direct genetic damage, such as the majority of known human carcinogens. Therefore, these batteries are not expected to detect non-genotoxic carcinogens. Experimental conditions, such as the limited capability of the *in vitro* metabolic activation systems, can lead to false negative results in *in vitro* tests. The test battery approach is designed to reduce the risk of false negative results for compounds with genotoxic potential. On the other hand a positive result in any assay for genotoxicity does not always mean that the test compound poses a genotoxic/carcinogenic hazard to humans.

Although positive *in vitro* data could indicate intrinsic genotoxic properties of a drug, appropriate *in vivo* data determine the biological significance of these *in vitro* signals in most cases. Also, because there are several indirect mechanisms of genotoxicity that operate only above certain concentrations, it is possible to establish a safe level (threshold) for classes of drugs with evidence for such mechanisms (see 5.2. below, Müller and Kasper, 2000; Scott et al., 1991; Thybaud et al., 2007).

5.1 Assessment of Biological Relevance

The recommendations below assume that the test has been conducted using appropriate spacing of doses, levels of toxicity etc.

Small increases in apparent genotoxicity *in vitro* or *in vivo* should first be assessed for reproducibility and biological significance. Examples of results that are not considered biologically meaningful include:

- i. Small increases that are statistically significant compared with the negative or solvent control values but are within the confidence intervals of the appropriate historical control values for the testing facility.

- ii. Weak/equivocal responses that are not reproducible.

If either of the above conditions applies, the weight of evidence indicates a lack of genotoxic potential, the test is considered negative or the findings not biologically relevant, and no further testing is called for.

5.2 Evaluation of Results Obtained in *In Vitro* Tests

In evaluating positive results, especially for the microbial mutagenicity test, the purity of the test compound should be considered, to determine whether the positive result could be attributable to a contaminant.

5.2.1 Evaluation of Positive Results Obtained *In Vitro* in a Bacterial Mutation Assay

Since positive results in the Ames test are thought to indicate DNA reactivity, extensive follow-up testing to assess the *in vivo* mutagenic and carcinogenic potential would be warranted to assess the potential risk for treatment of patients, unless justified by appropriate risk-benefit analysis.

There are some well characterized examples of artifactual increases in colonies that are not truly revertants. These can occur due to contamination with amino acids (i.e., providing histidine for *Salmonella typhimurium* strains or tryptophan for *Escherichia coli* strains), so that the bacterial reversion assay is not suitable for testing a peptide that is likely to degrade. Certain cases exist where positive results in bacterial mutation assays might be shown not to indicate genotoxic potential *in vivo* in humans, for example when bacterial-specific metabolism occurs, such as activation by bacterial nitroreductases.

5.2.2 Evaluation of Positive Results Obtained *In Vitro* in Mammalian Cell Assays

Recommendations for assessing weight of evidence and follow-up testing for positive genotoxicity results are discussed in IWGT reports (e.g., Thybaud et al., 2007). In addition, the scientific literature gives a number of conditions that can lead to a positive *in vitro* result of questionable relevance. Therefore, any *in vitro* positive test result should be evaluated based on an assessment of the weight of evidence as indicated below. This list is not exhaustive, but is given as an aid to decision-making.

- i. The conditions do not occur *in vivo* (pH; osmolality; precipitates).

(Note that the 1 mM limit avoids increases in osmolality, and that if the test compound alters pH it is advisable to adjust pH to the normal pH of untreated cultures at the time of treatment).

- ii. The effect occurs only at the most toxic concentrations.

In the MLA increases at $\geq 80\%$ reduction in RTG.

For *in vitro* cytogenetics assays when growth is suppressed by $\geq 50\%$.

If any of the above conditions apply the weight of evidence indicates a lack of genotoxic potential; the standard battery (Option 1) can be followed. Thus, a single *in vivo* test is considered sufficient.

5.2.3 Evaluation of *In Vitro* Negative Results

For *in vitro* negative results further testing should be considered in special cases, such as (the examples given are not exhaustive, but are given as an aid to decision-making): The structure or known metabolism of the compound indicates that standard techniques for *in vitro* metabolic activation (e.g., rodent liver S9) might be inadequate; the structure or

known activity of the compound indicates that the use of other test methods/systems might be appropriate.

5.3 Evaluation of Results Obtained from *In Vivo* Tests

In vivo tests have the advantage of taking into account absorption, distribution and excretion, which are not factors in *in vitro* tests, but are potentially relevant to human use. In addition metabolism is likely to be more relevant *in vivo* compared to the systems normally used *in vitro*. If the *in vivo* and *in vitro* results do not agree, then the difference should be considered/explained on a case-by-case basis, e.g., a difference in metabolism; rapid and efficient excretion of a compound *in vivo*.

In vivo genotoxicity tests also have the potential to give misleading positive results that do not indicate true genotoxicity. As examples:

- i. Increases in micronuclei can occur without administration of any genotoxic agent, due to disturbance in erythropoiesis (Tweats et al., 2007, I).
- ii. DNA adduct data should be interpreted in the light of the known background level of endogenous adducts.
- iii. Indirect, toxicity-related effects could influence the results of the DNA strand break assays (e.g., alkaline elution and Comet assays).

Thus it is important to take into account all the toxicological and hematological findings when evaluating the genotoxicity data (Note 15). Indirect effects related to toxicological changes could have a safety margin and might not be clinically relevant.

5.4 Follow-up Strategies for Positive Results

5.4.1 Follow-up to Findings in *In Vitro* in Mammalian Cell Tests

The following discussion assumes negative results in the Ames bacterial mutation assay.

5.4.1.1 Mechanistic/*In Vivo* Follow-up

When there is insufficient weight of evidence to indicate lack of relevance, recommended follow-up for positive mammalian cell assays would be to provide experimental evidence, either by additional *in vitro* studies (i, below) **or** by carrying out two appropriate *in vivo* assays (ii, below), as follows:

- i. Mechanistic information that contributes to a weight of evidence for a lack of relevant genotoxicity is often generated *in vitro*, for example evidence that a test compound that induces chromosome aberrations or mutations in the MLA is not a DNA damaging agent (e.g., other negative mutation/DNA damage tests in addition to the Ames test; structural considerations), or evidence for an indirect mechanism that might not be relevant *in vivo* or might have a threshold (e.g., inhibition of DNA synthesis, reactive oxygen species produced only at high concentrations) (Galloway et al., 1998; Scott et al., 1991; Müller and Kasper, 2000). Similar studies can be used to follow up a positive result in the *in vitro* micronucleus assay, or in this case evidence can include a known mechanism that indicates chromosome loss/aneuploidy, or centromere staining experiments (Note 17) that indicate chromosome loss. Polyploidy is a common finding in chromosome aberration assays *in vitro*. While aneugens can induce polyploidy, polyploidy alone does not indicate aneugenic potential and can simply indicate cell cycle perturbation; it is also commonly associated with increasing cytotoxicity. If polyploidy, but no structural chromosome breakage, is seen in an *in vitro* assay, generally a negative *in vivo* micronucleus assay with assurance of appropriate exposure would provide

sufficient assurance of lack of potential for aneuploidy induction.

If the above mechanistic information and weight of evidence supports the lack of relevant genotoxicity, only a single *in vivo* test with appropriate evidence of exposure is called for in order to establish the lack of genotoxic activity. This is typically a cytogenetic assay, and the micronucleus assay *in vivo* is called for when following up potential for chromosome loss.

If there is not sufficient weight of evidence or mechanistic information to rule out relevant genotoxic potential, two *in vivo* tests are generally called for, with appropriate endpoints and in appropriate tissues (usually two different tissues), and with an emphasis on obtaining sufficient exposure in the *in vivo* models.

Or

- ii. Two appropriate *in vivo* assays are done, usually with different tissues, and with supporting demonstration of exposure.

In summary, negative results in appropriate *in vivo* assays, with adequate justification for the endpoints measured and demonstration of exposure (see Section 4.4.1) are considered sufficient to demonstrate absence of significant genotoxic risk.

5.4.1.2 Follow-up to an In Vitro Positive Result That is Dependent upon S9 Activation

When positive results are seen only in the presence of the S9 activation system, it should first be verified that metabolic activation is responsible and not some other difference in conditions (e.g., low or no serum in the S9 mix, compared with $\geq 10\%$ serum in the non-activated incubations). The follow-up strategy is then aimed at determining the relevance of the results *in vitro* to conditions *in vivo*, and will generally focus on *in vivo* studies in liver (Note 18).

5.4.2 Follow-up to a Positive In Vivo Micronucleus Assay

If there is an increase in micronuclei *in vivo*, all the toxicological data should be evaluated to determine whether a non-genotoxic effect could be the cause or a contributing factor (Note 15). If non-specific effects of disturbed erythropoiesis or physiology (such as hypo/hyperthermia) are suspected, an *in vivo* assay for chromosome aberrations might be more appropriate. If a “real” increase is suspected, strategies should be used to demonstrate whether the increase is due to chromosome loss or chromosome breakage (Note 17). There is evidence that aneuploidy induction, e.g., with spindle poisons, follows a non-linear dose response. Thus, it might be possible to determine that there is a threshold exposure below which chromosome loss is not expected and to determine whether an appropriate safety margin exists compared with clinical exposure.

In conclusion, the assessment of the genotoxic potential of a compound should take into account the totality of the findings and acknowledge the intrinsic values and limitations of both *in vitro* and *in vivo* tests.

5.5 Follow-up Genotoxicity Testing in Relation to Tumor Findings in a Carcinogenicity Bioassay

Additional genotoxicity testing in appropriate models can be conducted for compounds that were negative in the standard test battery but which have shown increases in tumors in carcinogenicity bioassay(s) with insufficient evidence to establish a non-genotoxic mechanism. To help understand the mode of action, additional testing can include modified conditions for metabolic activation in *in vitro* tests or can include *in vivo* tests measuring genetic damage in target organs of tumor induction, such as DNA

strand break assays (e.g., comet or alkaline elution assays), liver UDS test, DNA covalent binding (e.g., by ^{32}P -postlabelling), mutation induction in transgenes, or molecular characterization of genetic changes in tumor-related genes (Kasper et al., 2007).

6. NOTES

1. The *in vitro* micronucleus assay has been widely evaluated in international collaborative studies (Kirsch-Volders et al., 2003), is validated by ECVAM (Corvi et al., 2008), and is the subject of an OECD guideline 487 (2010).
2. There is a small but significant number of genotoxic carcinogens that are reliably detected by the bone marrow tests for chromosomal damage but have yielded negative/weak/conflicting results in the *in vitro* tests outlined in the standard battery options. Carcinogens such as procarbazine, hydroquinone, urethane and benzene fall into this category. Some other examples from a survey of companies are described by Tweats et al., 2007, II.
3. In principle, micronuclei in hematopoietic cells can be evaluated in bone marrow from any species, and in blood from species that do not filter out circulating micronucleated erythrocytes in the spleen. In laboratory mice, micronuclei can be measured in polychromatic erythrocytes in blood, and mature (normochromatic) erythrocytes can be used when mice are treated continuously for about 4 weeks or more. Although rats rapidly remove micronucleated erythrocytes from the circulation, it has been established that micronucleus induction by a range of clastogens and aneugens can be detected in rat blood reticulocytes (Wakata et al., 1998; Hamada et al., 2001). Rat blood can be used for micronucleus analysis provided methods are used to ensure analysis of the newly formed reticulocytes (Hayashi et al., 2007; MacGregor et al., 2006), and the sample size is sufficiently large to provide appropriate statistical sensitivity given the lower micronucleus levels in rat blood than in bone marrow (Kissling et al., 2007). Whichever method is chosen, bone marrow or blood, automated or manual analysis, each laboratory should determine the appropriate minimum sample size to ensure that scoring error is maintained below the level of animal-to-animal variation.

Some experience is now available for micronucleus induction in the dog and rhesus monkey (Harper et al., 2007; Hotchkiss et al., 2008). One example where such alternative species might be useful would be in evaluation of a human metabolite that was not sufficiently represented in rodents but was formed in the dog or monkey.

4. While the two options in the battery are equally suitable, specific knowledge about an individual test compound can indicate that one option is preferable. For example, if systemic exposure in animal models is equal to or less than anticipated clinical exposure, *in vitro* assays should be employed: Option 1 (see also Sections 2.3.4 and 4.4.1). On the other hand Option 2, including a test in liver, is recommended in cases where short-lived reactive metabolites are expected to be generated in the liver.
5. Certain structurally alerting molecular entities are recognized as being causally related to the carcinogenic and/or mutagenic potential of chemicals. Examples of structural alerts include alkylating electrophilic centers, unstable epoxides, aromatic amines, azo-structures, *N*-nitroso groups, and aromatic nitro-groups (Ashby and Paton, 1994). For some classes of compounds with specific structural alerts, it is established that specific protocol modifications/additional tests are important for optimum detection of genotoxicity (e.g., molecules containing an azo-group, glycosides, compounds such as nitroimidazoles requiring nitroreduction for activation,

compounds such as phenacetin requiring a different rodent S9 for metabolic activation).

6. There is some experience with *in vivo* assays for micronucleus induction in skin and colon (Hayashi et al., 2007), and DNA damage assays in these tissues can also be an appropriate substitute.
7. A few chemicals are more easily detectable either with plate-incorporation or with pre-incubation methods, though differences are typically quantitative rather than qualitative (Gatehouse et al., 1994). Experience in the pharmaceutical industry where drugs have been tested in both protocols has not resulted in different results for the two methods, and, in the IWGT report (Gatehouse et al., 1994), the examples of chemical classes listed as more easily detectable in the pre-incubation protocol are generally not pharmaceuticals and are positive in *in vivo* genotoxicity tests in liver. These include short chain aliphatic nitrosamines; divalent metals; aldehydes (e.g., formaldehyde, crotonaldehyde); azo dyes (e.g., butter yellow); pyrrolizidine alkaloids; allyl compounds (allylisothiocyanate, allyl chloride), and nitro (aromatic, aliphatic) compounds.
8. The rationale for a maximum concentration of 1 mM for *in vitro* mammalian cell assays includes the following: The test battery includes the Ames test and an *in vivo* assay. This battery optimizes the detection of genotoxic carcinogens without relying on any individual assay alone. There is a very low likelihood of compounds of concern (DNA damaging carcinogens) that are not detected in Ames test or *in vivo* genotoxicity assay, but are detectable in an *in vitro* mammalian assay only above 1 mM. Second, a limit of 1 mM maintains the element of hazard identification, being higher than clinical exposures to known pharmaceuticals, including those that concentrate in tissues (Goodman & Gilman, 2001), and is also higher than the levels generally achievable in preclinical studies *in vivo*. Certain drugs are known to require quite high clinical exposures for therapeutic effect, e.g., nucleoside analogs and some antibiotics. While comparison of potency with existing drugs can be of interest to sponsors, perhaps even above the 1 mM limit, it is ultimately the *in vivo* tests that determine relevance for human safety. For pharmaceuticals with unusually low molecular weight (e.g., less than 200) higher test concentrations should be considered.
9. Although some genotoxic carcinogens are not detectable in *in vitro* genotoxicity assays unless the concentrations tested induce some degree of cytotoxicity, DNA damaging agents are generally detectable with only moderate levels of toxicity (Greenwood et al., 2004). As cytotoxicity increases, mechanisms other than direct DNA damage by a compound or its metabolites can lead to 'positive' results that are related to cytotoxicity and not genotoxicity. Such indirect induction of DNA damage secondary to damage to non-DNA targets is more likely to occur above a certain concentration threshold. The disruption of cellular processes is not expected to occur at lower, pharmacologically relevant concentrations.

In cytogenetic assays, even weak clastogens that are known to be carcinogens are positive without exceeding a 50% reduction in cell counts. On the other hand, compounds that are not DNA damaging, mutagenic or carcinogenic can induce chromosome breakage at toxic concentrations. For both *in vitro* cytogenetic assays, the chromosome aberration assay and the *in vitro* micronucleus assay, a limit of about 50% growth reduction is considered appropriate.

For cytogenetic assays in cell lines, measurement of cell population growth over time (by measuring the change in cell number during culture relative to control, e.g., by

the method referred to as Population Doubling (PD; Note 10), has been shown to be a useful measure of cytotoxicity, as it is known that cell numbers can underestimate toxicity. For lymphocyte cultures, an inhibition of proliferation not exceeding about 50% is considered sufficient; this can be measured by Mitotic Index (MI) for metaphase aberration assays and by an index based on cytokinesis block for *in vitro* micronucleus assays. In addition, for the *in vitro* micronucleus assay, since micronuclei are scored in the interphase subsequent to a mitotic division, it is important to verify that cells have progressed through the cell cycle. This can be done by use of cytochalasin B to allow nuclear division but not cell division, so that micronuclei can be scored in binucleate cells (the preferred method for lymphocytes). For cell lines, other methods to demonstrate cell proliferation, including cell population growth over time (PD) as described above, can be used (Kirsch-Volders et al., 2003).

For MLA, appropriate sensitivity is achieved by limiting the top concentration to one with close to 20% Relative Total Growth (RTG) (10-20%) both for soft agar and for microwell methods (Moore et al., 2002). Reviews of published data using the current criteria found very few chemicals that were positive in MLA only at concentrations with less than 20% RTG and that were rodent carcinogens, and convincing evidence of genotoxic carcinogenesis for this category is lacking. The consensus is that caution is appropriate in interpreting results when increases in mutation are seen only below 20% RTG, and a result would not be considered positive if the increase in mutant fraction occurred only at $\leq 10\%$ RTG.

In conclusion, caution is appropriate in interpreting positive results obtained as reduction in growth/survival approaches or exceeds 50% for cytogenetics assays or 80% for MLA. It is acknowledged that the evaluation of cells treated at these levels of cytotoxicity/clonal survival can result in greater sensitivity but bears an increased risk of non-relevant positive results. The battery approach for genotoxicity is designed to ensure appropriate sensitivity without relying on single *in vitro* mammalian cell tests at high cytotoxicity.

To obtain an appropriate toxicity range, a preliminary range-finding assay over a broad range of concentrations is useful, but in the genotoxicity assay it is often critical to use multiple concentrations that are spaced quite closely (less than two-fold dilutions). Extra concentrations can be tested but not all concentrations need be evaluated for genotoxicity. It is not intended that multiple experiments be carried out to reach exactly 50% reduction in growth, for example, or exactly 80% reduction in RTG.

10. For *in vitro* cytogenetic assays it is appropriate to use a measure of relative cell growth to assess toxicity, because cell counts can underestimate toxicity (Greenwood et al., 2004). Using calculated population doublings (see Glossary) to estimate the 50% growth reduction level, it was demonstrated that the frequency of positive results with compounds that are not mutagenic or carcinogenic is reduced, while agents that act via direct interaction with DNA are reliably positive.
11. In certain cases it can be useful to examine chromosome aberrations at metaphase in lymphocytes cultured from test animals after one or more administrations of test compound, just as bone marrow metaphase cells can be used. Since circulating lymphocytes are not replicating, agents that require replication for their genotoxic effect (e.g., some nucleoside analogs) are not expected to be detected in this cell type. Because some lymphocytes are relatively long-lived, in principle there is the potential for accumulation of un-repaired DNA damage *in vivo* that would give rise to aberrations when the cells are stimulated to divide *in vitro*. The *in vivo* lymphocyte

assay can be useful in following up indications of clastogenicity, but in general another tissue such as liver is a more informative supplement to the micronucleus assay in hematopoietic cells because exposure to drug and metabolite(s) is often higher in liver.

12. The inclusion of a second *in vivo* assay in the battery is to provide assurance of lack of genotoxicity by use of a tissue that is well exposed to a drug and/or its metabolites; a small number of carcinogens that are considered genotoxic gave positive results in a test in liver but were negative in a cytogenetics assay *in vivo* in bone marrow. These examples likely reflect a lack of appropriate metabolic activity or lack of reactive intermediates delivered to the hematopoietic cells of the bone marrow.

Assays for DNA strand breaks, DNA adducts, and mutations in transgenes have the advantage that they can be applied in many tissues. Internationally agreed protocols are not yet in place for all the *in vivo* assays, although considerable experience and published data and protocol recommendations exist for DNA strand break assays (Comet and alkaline elution assays), DNA adduct (covalent binding) measurements, and transgenic rodent mutation assays, in addition to the UDS assay. For a compound that is positive *in vitro* in the MLA and induces predominantly large colonies, and is also shown not to induce chromosome breakage in an *in vitro* metaphase assay, an *in vivo* assay for mutation, such as a transgenic mouse mutation assay, should be considered in preference to a DNA strand break assay. The UDS assay is considered useful mainly for compounds that induce bulky DNA adducts or are positive in the Ames test. Because cytotoxicity induces DNA strand breakage, careful cytotoxicity assessment is needed to avoid confounding the results of DNA strand break assays. This has been well characterized for the *in vitro* alkaline elution test (Storer et al., 1996) but not yet fully validated for the Comet assay. In principle the DNA strand break assays can be used in repeat-dose toxicology assays with appropriate dose levels and sampling times.

Since liver of mature animals is not a highly mitotic tissue, often a non-cytogenetic endpoint is used for the second assay, but when dividing hepatocytes are present, such as after partial hepatectomy, or in young rats (Hayashi et al., 2007), micronucleus analysis in liver is possible, and detects known genotoxic compounds.

13. For common vehicles like aqueous methyl cellulose this would usually be appropriate, but for vehicles such as Tween 80, the volume that can be administered could be as much as 30 fold lower than that given acutely.
14. Caution is appropriate if the toxicological study design includes additional blood sampling, e.g., for measurement of exposure. Such bleeding could perturb the results of micronucleus analysis since erythropoiesis stimulated by bleeding can lead to increases in micronucleated erythrocytes.
15. Increases in micronuclei can occur without administration of any genotoxic agent, due to disturbance in erythropoiesis (such as regenerative anemia; extramedullary hematopoiesis), stress, and hypo- and hyperthermia (reviewed by Tweats et al., 2007, I). In blood, changes in spleen function that affect clearance of micronucleated cells from the blood could lead to small increases in circulating micronucleated red blood cells.
16. Positive controls for either short-term or repeat dose genotoxicity studies:
For micronucleus (and other cytogenetic) assays, the purpose of the positive control is to verify that the individuals scoring the slides can reliably detect increases in micronuclei. This can be accomplished by use of samples from periodic studies (every few months) of small groups of animals (one sex) given acute treatment with a

positive control. For manual scoring such slides can be included in coded slides scored from each study. Positive control slides should not be obvious to readers based on their staining properties or micronucleus frequency. For automated scoring, appropriate quality control samples should be used with each assay.

For other *in vivo* genotoxicity assays, the purpose of positive controls is to demonstrate reliable detection of an increase in DNA damage/mutagenicity using the assay in the chosen species, tissue and protocol. After a laboratory has demonstrated that it can consistently detect appropriate positive control compounds in multiple independent experiments, carrying out positive control experiments periodically is generally sufficient provided experimental conditions are not changed. However, currently it is considered that for the Comet assay concurrent positive controls are advisable.

17. Determination of whether micronucleus induction is due primarily to chromosome loss or to chromosome breakage could include staining micronuclei *in vitro* or *in vivo* to determine whether centromeres are present, e.g., using Fluorescent *in situ* Hybridization (FISH) with probes for DNA sequences in the centromeric region, or a labeled antibody to kinetochore proteins. If the majority of induced micronuclei are centromere positive, this suggests chromosome loss. (Note that even potent tubule poisons like colchicine and vinblastine do not produce 100% kinetochore positive micronuclei, but more typically 70 to 80%, and are accepted as primarily aneugens for assessing risk). An alternative approach is to carry out an *in vitro* or *in vivo* assay for metaphase structural aberrations; if negative this would imply that micronucleus induction is related to chromosome loss.
18. Standard induced S9 mix has higher activation capacity than human S9, and lacks phase two detoxification capability unless specific cofactors are supplied. Also, non-specific activation can occur *in vitro* with high test substrate concentrations (see Kirkland et al., 2007). Genotoxicity testing with human S9 or other human-relevant activation systems can be helpful. Analysis of the metabolite profile in the genotoxicity test incubations for comparison with known metabolite profiles in preclinical species (in uninduced microsomes or hepatocytes, or *in vivo*) or in preparations from humans can also help determine the relevance of test results (Ku et al., 2007), and follow-up studies will usually focus on *in vivo* testing in liver. A compound that gives positive results *in vitro* with S9 might not induce genotoxicity *in vivo* because the metabolite is not formed, is formed in very small quantities, or is metabolically detoxified or rapidly excreted, indicating a lack of risk *in vivo*.

7. GLOSSARY

Alkaline elution assay:

See *DNA strand break assay*.

Aneuploidy:

Numerical deviation of the modal number of chromosomes in a cell or organism.

Base substitution:

The substitution of one or more base(s) for another in the nucleotide sequence. This can lead to an altered protein.

Cell proliferation:

The ability of cells to divide and to form daughter cells.

Centromere/kinetochore:

Structures in chromosomes essential for association of sister chromatids and for attachment of spindle fibers that move daughter chromosomes to the poles and ensure inclusion in daughter nuclei.

Clastogen:

An agent that produces structural breakage of chromosomes, usually detectable by light microscopy.

Cloning efficiency:

The efficiency of single cells to form clones. It is usually measured after seeding low numbers of cells in a suitable environment.

Comet assay:

See *DNA strand break assay*.

Culture confluency:

A quantification of the cell density in a culture by visual inspection.

Cytogenetic evaluation:

Chromosome structure analysis in mitosis or meiosis by light microscopy or micronucleus analysis.

DNA adduct:

Product of covalent binding of a chemical to DNA.

DNA repair:

Reconstitution of the original DNA sequence after DNA damage.

DNA strand breaks:

Single or double strand scissions in the DNA.

DNA strand break test:

Alkaline treatment that converts certain types of DNA lesions into strand breaks that can be detected by the alkaline elution technique, measuring migration rate through a filter, or by the single cell gel electrophoresis or Comet test (in which cells embedded in a thin layer of gel on a microscope slides are subjected to electric current, causing shorter pieces of DNA to migrate out of the nucleus into a “Comet tail”). The extent of DNA migration is measured visually under the microscope on stained cells.

Frameshift mutation:

A mutation (change in the genetic code) in which one base or two adjacent bases are

added to (inserted in) or deleted from the nucleotide sequence of a gene. This can lead to an altered or truncated protein.

Gene mutation:

A detectable permanent change within a single gene or its regulating sequences. The changes can be point mutations, insertions, or deletions.

Genetic endpoint:

The precise type or class of genetic change investigated (e.g., gene mutations, chromosomal aberrations, DNA strand breaks, DNA repair, DNA adduct formation, etc).

Genotoxicity:

A broad term that refers to any deleterious change in the genetic material regardless of the mechanism by which the change is induced.

Micronucleus:

Particle in a cell that contains nuclear DNA; it might contain a whole chromosome(s) or a broken centric or acentric part(s) of chromosome(s).

Mitotic index:

Percentage of cells in the different stages of mitosis amongst the cells not in mitosis (interphase) in a preparation (slide).

Numerical chromosome changes:

Chromosome numbers different from the original haploid or diploid set of chromosomes; for cell lines, chromosome numbers different from the modal chromosome set.

Plasmid:

Genetic element additional to the normal bacterial genome. A plasmid might be inserted into the host chromosome or form an extra-chromosomal element.

Point mutations:

Changes in the genetic codes, usually confined to a single DNA base pair.

Polychromatic erythrocyte:

An immature erythrocyte in an intermediate stage of development that still contains ribosomes and, as such, can be distinguished from mature normochromatic erythrocytes (lacking ribosomes) by stains selective for RNA.

Polyploidy:

Numerical deviation of the modal number of chromosomes in a cell, with approximately whole multiples of the haploid number. Endoreduplication is a morphological form of polyploidy in which chromosome pairs are associated at metaphase as “diplochromosomes”.

Population doubling or culture growth:

This can be calculated in different ways; one example of an appropriate formula is: Population doublings (PDs) = the log of the ratio of the final count (N) to the starting (baseline) count (X₀), divided by the log of 2. That is: $PD = [\log(N \div X_0)] \div \log 2$.

Recombination:

Breakage and balanced or unbalanced rejoining of DNA.

RTG (relative total growth):

This measure of cytotoxicity takes the relative suspension growth (based on cell loss and cell growth from the beginning of treatment to the second day post-treatment) and multiplies it by the relative plating efficiency at the time of cloning for mutant

quantization.

Single Cell Gel Electrophoresis assay:

Comet assay. *See DNA strand break assay.*

Survival (in the context of mutagenicity testing):

Proportion of living cells among dead cells, usually determined by staining or colony counting methods after a certain treatment interval.

Transgene:

An exogenous or foreign gene inserted into the host genome, either into somatic cells or germ line cells.

Unscheduled DNA synthesis (UDS):

DNA synthesis that occurs at some stage in the cell cycle other than S-phase in response to DNA damage. It is usually associated with DNA excision repair.

8. REFERENCES

- Ashby J, Paton D. The influence of chemical structure on the extent and sites of carcinogenesis for 522 rodent carcinogens and 55 different human carcinogen exposures. *Mutat Res* 1994;286:3-74.
- Corvi R, Albertini S, Hartung T, Hoffmann S, Maurici D, Pfuhler S et al. ECVAM Retrospective Validation of the *in vitro* micronucleus test (MNT). *Mutagenesis* 2008;23:271-283.
- Gatehouse D, Haworth S, Cebula T, Gocke E, Kier L, Matsushima T et al. Report from the working group on bacterial mutation assays: international workshop on standardisation of genotoxicity test procedures. *Mutat Res* 1994;312:217-33.
- Goodman & Gilman. The pharmacological basis of therapeutics. Hardman JG, Limbird LE, Gilman AG, editors. 10th ed. New York: McGraw-Hill Professional 2001.
- Greenwood SK, Hill RB, Sun JT, Armstrong MJ, Johnson TE, Gara JP et al. Population Doubling: A simple and more accurate estimation of cell growth suppression in the *in vitro* assay for chromosomal aberrations that reduces irrelevant positive results. *Environ Mol Mutagen* 2004;43:36-44.
- Hamada S., Sutou S, Morita T, Wakata A, Asanami S, Hosoya S et al. Evaluation of the rodent micronucleus assay by a 28-day treatment protocol: summary of the 13th collaborative study by the collaborative study group for the micronucleus test (CSGMT)/Environmental Mutagen Society of Japan (JEMS)–Mammalian Mutagenicity Study Group (MMS). *Environ Mol Mutagen* 2001;37:93-110.
- Hartmann A, Agurell E, Beevers C, Brendler-Schwaab S, Burlinson B, Clay P et al. Recommendations for conducting the *in vivo* alkaline Comet assay. *Mutagenesis* 2003;18:45-51.
- Hayashi M, MacGregor JT, Gatehouse DG, Adler I, Blakey DH, Dertinger SD et al. *In vivo* rodent erythrocyte micronucleus assay. II. Some aspects of protocol design including repeated treatments, integration with toxicity testing, and automated scoring. *Environ Mol Mutagen* 2000;35:234-52.
- Hayashi M, MacGregor JT, Gatehouse DG, Blakey DH, Dertinger SD, Abramsson-Zetterberg L et al. *In vivo* erythrocyte micronucleus assay. III. Validation and regulatory acceptance of automated scoring and the use of rat peripheral blood reticulocytes, with discussion of non-hematopoietic target cells and a single dose-level limit test. *Mutat Res* 2007;627:10-30.
- Heddle JA, Dean S, Nohmi T, Boerrigter M, Casciano D, Douglas GR et al. *In vivo* transgenic mutation assays. *Environ Mol Mutagen* 2000;35:253-9.
- Hotchkiss CE, Bishop ME, Dertinger SD, Slikker W, Moore MM, MacGregor JT. Flow cytometric analysis of micronuclei in peripheral blood reticulocytes IV: an index of chromosomal damage in the rhesus monkey (*macaca mulatta*). *Toxicol Sci* 2008 ;102:352-8.
- Kasper P, Uno Y, Mauthe R, Asano N, Douglas G, Matthews E et al. Follow-up testing of

rodent carcinogens not positive in the standard genotoxicity testing battery: IWGT workgroup report. *Mutat Res* 2007;627:106-116.

Kenelly JC, Waters R, Ashby J, Lefevre PA, Burlinson B, Benford DJ et al. *In vivo* rat liver UDS assay. In: Supplementary Mutagenicity Tests, UKEMS Recommended Procedures. Kirkland DJ, Fox M, editors. Cambridge University Press 1993;52-77.

Kirkland DJ, Pfuhler S, Tweats D, Aardema M, Corvi R, Darroudi F et al. How to reduce false positive results when undertaking *in vitro* genotoxicity testing and thus avoid unnecessary follow-up animal tests: report of the ECVAM workshop. *Mutat Res* 2007;628:31-55.

Kirsch-Volders M, Sofuni T, Aardema M, Albertini S, Eastmond D, Fenech M et al. Report from the *in vitro* micronucleus assay working group. *Mutat Res* 2003;540:153-63.

Kissling GE, Dertinger SD, Hayashi M, MacGregor JT. Sensitivity of the erythrocyte micronucleus assay: dependence on number of cells scored and inter-animal variability. *Mutat Res* 2007;634:235-40.

Ku WW, Bigger A, Brambilla G, Glatt H, Gocke E, Guzzie PJ et al. Strategy for genotoxicity testing-metabolic considerations. *Mutat Res* 2007;627:59-77.

MacGregor JT, Bishop ME, McNamee JP, Hayashi M, Asano N, Wakata A et al. Flow Cytometric analysis of micronuclei in peripheral blood reticulocytes: II. An efficient method of monitoring chromosomal damage in the rat. *Toxicol Sci* 2006;94:92-107.

Moore MM, Honma M, Clements J, Harrington-Brock K, Awogi T, Bolcsfoldi G et al. Mouse lymphoma thymidine kinase locus gene mutation assay: follow-up international workshop on genotoxicity test procedures—New Orleans, Louisiana, April 2000. *Environ Mol Mutagen* 2002;40:292-9.

Moore MM, Honma M, Clements J, Bolcsfoldi G, Burlinson B, Cifone M et al. Mouse lymphoma thymidine kinase gene mutation assay: follow-up meeting of the international workshop on genotoxicity testing—Aberdeen, Scotland, 2003—Assay acceptance criteria, positive controls, and data evaluation. *Environ Mol Mutagen* 2006;47:1-5.

Müller L, Kasper P. Human biological relevance and the use of threshold arguments in regulatory genotoxicity assessment: experience with pharmaceuticals. *Mutat Res* 2000;464: 9-34.

OECD Guidelines for Genetic Toxicology 1997.

Scott D, Galloway SM, Marshall RR, Ishidate M Jr, Brusick D, Ashby J et al. Genotoxicity under extreme culture conditions. A report from ICPEMC task group 9. *Mutat Res* 1991;257:147-204.

Storer RD, McKelvey TW, Kraynak AR, Elia MC, Barnum JE, Harmon LS et al. Revalidation of the *in vitro* alkaline elution/rat hepatocyte assay for DNA damage: improved criteria for assessment of cytotoxicity and genotoxicity and results for 81 compounds. *Mutat Res* 1996;368:59-101.

Suzuki H, Ikeda N, Kobayashi K, Terashima Y, Shimada Y, Suzuki T et al. Evaluation of

liver and peripheral blood micronucleus assays with 9 chemicals using young rats. A study by the Collaborative Study Group for the Micronucleus Test (CSGMT)/Japanese Environmental Mutagen Society (JEMS) - Mammalian Mutagenicity Study Group (MMS). *Mutat Res* 2005;583:133-45.

Thybaud V, Aardema M, Clements J, Dearfield K, Galloway S, Hayashi M et al. Strategy for genotoxicity testing: hazard identification and risk assessment in relation to *in vitro* testing. *Mutat Res* 2007;627:41-58.

Tweats DJ, Blakey D, Heflich RH, Jacobs A, Jacobsen SD, Morita T et al. Report of the IWGT working group on strategies and interpretation of regulatory *in vivo* tests. I. Increases in micronucleated bone marrow cells in rodents that do not indicate genotoxic hazards. *Mutat Res* 2007;627:78-91.

Tweats DJ, Blakey D, Heflich RH, Jacobs A, Jacobsen SD, Morita T et al. Report of the IWGT working group on strategy/interpretation of regulatory *in vivo* tests. II. Identification of *in vivo*-only positive compounds in the bone marrow micronucleus test. *Mutat Res* 2007;627:92-105.

Wakata, A, Miyamae Y, Sato S, Suzuki T, Morita T, Asano N et al. Evaluation of the rat micronucleus test with bone marrow and peripheral blood: summary of the 9th collaborative study by CSGMT/JEMS. MMS. *Environ Mol Mutagen* 1998;32:84-100.