Mutation Research 783 (2015) 23–35

Contents lists available at ScienceDirect

Mutation Research/Genetic Toxicology and Environmental Mutagenesis

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

The *in vivo Pig-a* assay: A report of the International Workshop On Genotoxicity Testing (IWGT) Workgroup☆

B. Bhaskar Gollapudi[a,1], Anthony M. Lynch[b,1], Robert H. Heflich[c,*,2], Stephen D. Dertinger[d], Vasily N. Dobrovolsky[c], Roland Froetschl[e], Katsuyoshi Horibata[f], Michelle O. Kenyon[g], Takafumi Kimoto[h], David P. Lovell[i], Leon F. Stankowski Jr.[j], Paul A. White[k], Kristine L. Witt[l], Jennifer Y. Tanir[m]

[a] *Exponent, Midland, MI, USA*
[b] *GlaxoSmithKline, Hertfordshire, UK*
[c] *US Food and Drug Administration, National Center for Toxicological Research, Jefferson, AR, USA*
[d] *Litron Laboratories, Rochester, NY, USA*
[e] *Federal Institute for Drugs and Medical Devices, Bonn, Germany*
[f] *National Institute of Health Sciences, Tokyo, Japan*
[g] *Pfizer Global Research and Development, Groton, CT, USA*
[h] *Teijin Pharma, Tokyo, Japan*
[i] *St. George's, University of London, UK*
[j] *BioReliance, Rockville, MD, USA*
[k] *Environmental Health Science and Research Bureau, Health Canada, Ottawa, ON, Canada*
[l] *National Institutes of Health, National Institute of Environmental Health Sciences, Division of the National Toxicology Program, Research Triangle Park, NC, USA*
[m] *Health and Environmental Sciences Institute, Washington, DC, USA*

ARTICLE INFO

Article history:
Received 13 September 2014
Accepted 15 September 2014
Available online 26 September 2014

Keywords:
Red blood cells
Reticulocytes
Glycosylphosphatidylinositol
Flow cytometry
CD59
Mutation

ABSTRACT

The *in vivo Pig-a* assay uses flow cytometry to measure phenotypic variants for antibody binding to cell surface glycosylphosphatidylinositol (GPI)-anchored proteins. There is good evidence suggesting that the absence of antibody binding is the result of a mutation in the endogenous X-linked *Pig-a* gene, which forms the rationale for the assay. Although the assay has been performed with several types of hematopoietic cells and in a variety of mammalian species, including humans, currently it is optimized only for measuring CD59-deficient (presumed *Pig-a* mutant) erythrocytes in the peripheral blood of rats. An expert workgroup formed by the International Workshop on Genotoxicity Testing considered the state of assay development and the potential of the assay for regulatory use. Consensus was reached on what is known about the *Pig-a* assay and how it should be conducted, and recommendations were made on additional data and refinements that would help to further enhance the assay for use in hazard identification and risk assessment.

Published by Elsevier B.V.

☆ The views expressed in this report are not necessarily shared by the institutions at which the workgroup members are employed. The information in these materials is not a formal dissemination of information by the U.S. Food and Drug Administration and does not represent agency position or policy. Data presented in this communication were partially funded by a grant from the National Institute of Health/National Institute of Environmental Health Sciences (NIEHS; grant no. R44ES021973). The contents of this report are solely the responsibility of the authors, and do not necessarily represent the official views of the NIEHS.

* Corresponding author at: U.S. Food and Drug Administration, NCTR, 3900 NCTR Road, Jefferson, AR 72079, USA. Tel.: +1 870 543 7493.
E-mail address: robert.heflich@fda.hhs.gov (R.H. Heflich).

[1] Co-Chairs.
[2] Rapporteur.

http://dx.doi.org/10.1016/j.mrgentox.2014.09.007
1383-5718/Published by Elsevier B.V.

1. Introduction

1.1. Background to this report

The *in vivo Pig-a* assay was first described for rats and mice in 2008 [1–4], and received almost immediate interest as a potential assay for evaluating the *in vivo* mutagenicity of new and existing substances. Subsequently, formal and informal workshops and presentations on the assay were held at meetings of the US Environmental Mutagenesis and Genomics Society and the Health and Environmental Sciences Institute (HESI) of the International Life Sciences Institute [5]. A Workgroup, made up of experts from academic, regulatory, and industrial laboratories, was formed in 2012 under the auspices of the International Workshop on Genotoxicity Testing (IWGT) to review the development of the *Pig-a* assay in the context of safety assessment strategies. The remit of the Workgroup was to consider the underlying science of the *Pig-a* assay and technical considerations for the assay protocol with a view to assay acceptance in a regulatory context and to recommend where and how further progress on developing the assay can be made. Subsequently, two meetings of the Workgroup were held in 2013, one in conjunction with the HESI Genetic Toxicology Technical Committee annual meeting in Washington, DC, on April 15 and the second at the IWGT meeting in Foz do Iguaçu, Brazil, on October 31–November 1. This report describes consensus statements on the *Pig-a* assay developed by the IWGT Workgroup.

1.2. Principle of the assay

The *Pig-a* assay is based on the identification of mutant cells that have an altered repertoire of cell surface markers (Fig. 1). The assay was developed from an understanding of the molecular nature of a rare human acquired genetic disorder, paroxysmal nocturnal hemoglobinuria (PNH) (reviewed [6–8]). The *Pig-a* gene (phosphatidylinositol glycan, class A gene) codes for the catalytic subunit of an *N*-acetyl glucosamine transferase that is involved in an early step of glycosylphosphatidylinositol (GPI) biosynthesis. GPI anchors an assortment of protein markers (*e.g.*, CD59, CD24, CD55) to the exterior surface of the cytoplasmic membranes of higher eukaryotes. In mammals, *Pig-a* (nomenclature: *PIG-A*

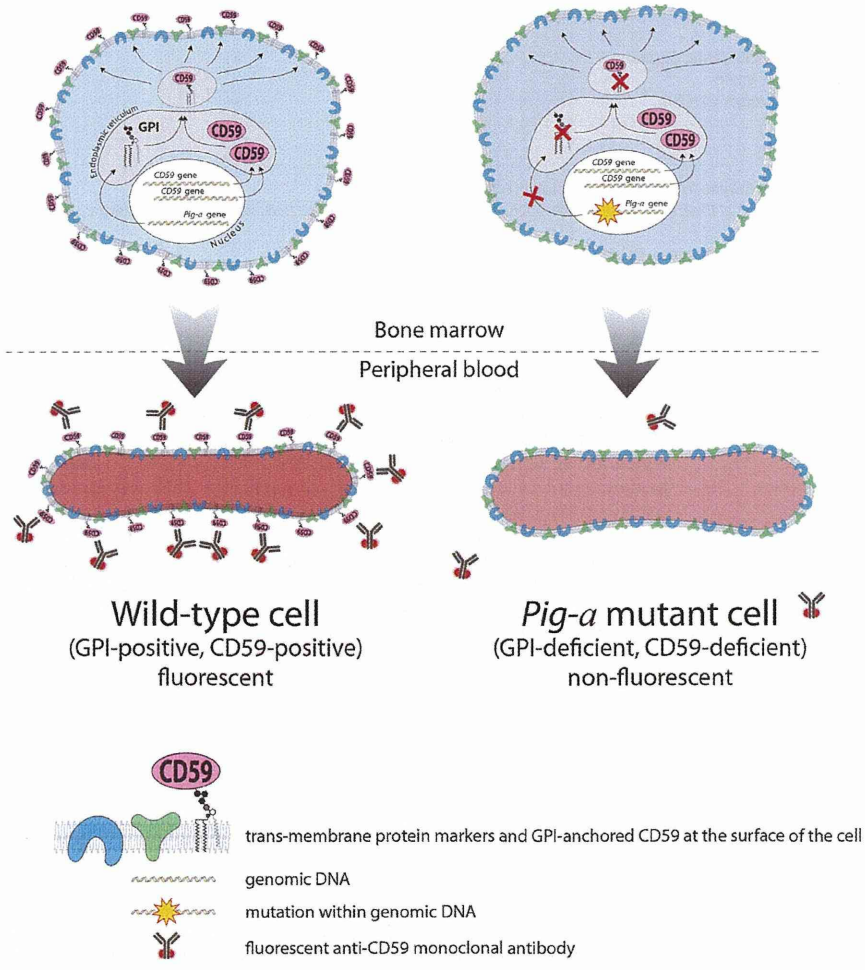

Fig. 1. Principle of the rat erythrocyte *Pig-a* assay. In wild-type nucleated erythroid precursor cells (upper left), glycosylphosphatidylinositol (GPI) anchors are synthesized in a series of steps at the endoplasmic reticulum (ER). The first step is catalyzed by the product of the X-linked *Pig-a* gene; all other steps of the GPI synthesis pathway are catalyzed by enzymes encoded by autosomal genes. The GPI core structure consists of 3 mannoses, 1 glucosamine, and 1 phosphatidylinositol, with 2 or 3 hydrophobic tails imbedded in the ER lipid bilayer membrane (a more detailed description of GPI structure and synthesis can be found in [7]). In the final stage of biosynthesis, a surface marker (*e.g.*, CD59 protein, which is synthesized independently) is attached to the GPI anchor and the assembly is transported to the surface of the cell. *Pig-a* mutant cells (upper right) are deficient in GPI anchor synthesis, and thus GPI-anchored markers are not transported to the cytoplasmic membrane. The *Pig-a* assay is conducted with peripheral blood erythrocytes having no nuclei (lower, left and right). At the surface of the cytoplasmic membrane, GPI anchored CD59 reacts with fluorescent anti-CD59 antibodies which make the whole wild-type cell fluoresce (lower left). *Pig-a* mutant cells do not express CD59 on their surface and thus are not fluorescently labeled (lower right). Flow cytometry is used to quantify the number of wild-type and mutant cells and estimate a *Pig-a* mutant frequency.

in humans and other primates; *Pig-a* in mice and rats) is an X-linked gene present as a single functional copy in cells from both males and females. Other genes involved in GPI biosynthesis are autosomal and have two functional copies. A single inactivating mutation in the *Pig-a* gene is sufficient to make a cell deficient in GPI anchors and, as a consequence, deficient in surface-bound GPI-anchored markers. Since it is exceedingly unlikely that anchor deficiency would occur due to inactivating mutations in both copies of the autosomal genes involved in GPI synthesis, measuring GPI deficiency is considered 'virtually equivalent' to measuring *Pig-a* mutation.

The assay distinguishes the *Pig-a* mutant phenotype from the wild-type phenotype by labeling cells with fluorescent antibodies against a GPI-anchored marker or markers [8] (Fig. 1). Wild-type cells bind marker-specific antibodies and fluoresce, while mutant cells do not bind the antibodies and do not fluoresce. Cells are analyzed using high throughput flow cytometry to quantify wild-type (*i.e.* labeled) and rare mutant (*i.e.* unlabeled) cells. There are multiple commercially available antibodies, stains, and reagents suitable for designing *Pig-a* assays in different cell types and in different mammalian species.

The use of flow cytometry for detecting cells deficient in surface markers imposes certain constraints on the tissues amenable for the *Pig-a* assay. For instance, the samples must be prepared as single-cell suspensions, and the cells should not be subjected to treatments that may alter the cell surface membrane (*e.g.*, proteolytic digestion). Currently, these two requirements have limited the *in vivo* application of the *Pig-a* assay to hematopoietic tissue (peripheral blood erythrocytes and white blood cells, and, to a lesser degree, bone marrow).

A combination of theory and experimental evidence indicates that *Pig-a* mutations occurring in bone marrow cells result in the emergence of the mutant cells that are measured in peripheral blood [9,10]. The appearance of *Pig-a* mutant cells in peripheral blood exhibits a time delay that is dependent upon the cell and GPI-anchor turnover rates, as well as the trafficking time from the bone marrow to the periphery for the specific cell type. Red blood cells (RBCs) are the most practical cell type for performing the assay since microliter volumes of blood contain sufficient quantities of RBCs for the enumeration of the rare mutants required for a successful *Pig-a* assay. The abundance of RBCs in peripheral blood allows serial blood collections from even small laboratory animals and thus permits longitudinal studies to be conducted on the same set of animals. In addition, the small fraction of newly formed erythrocytes, *i.e.*, reticulocytes (RETs), can be distinguished by appropriate staining. RETs have a rapid turnover rate (few days), and express the mutant phenotype originating in the bone marrow faster than the population of total RBCs in peripheral blood. Finally, at least for the rat assay, *Pig-a* mutant erythrocytes act as though they have a neutral phenotype [9], meaning that mutant cells are not subjected to negative selection and can accumulate with repeated dosing, potentially increasing the sensitivity of the assay. Observations concerning *Pig-a* mutant manifestation in the peripheral blood and mutant accumulation with repeat doses are consistent with what occurs in other *in vivo* gene mutation assays [11,12], and lend support to the mutational origin of the phenotype measured in the *Pig-a* assay.

1.3. Potential for translation of the endpoint from experimental models to humans

The pathway for GPI biosynthesis is conserved in most mammalian species, including common laboratory animals, such as mice, rats and monkeys, and in humans. In fact, approaches for scoring *Pig-a/PIG-A* mutant cells have been described for all these species [1–3,13–16]. Although credit for the original flow cytometric methodology for measuring mutant cells goes to David Araten and Lucio Luzzatto, who described *PIG-A* assays for human granulocytes and RBCs [17], human and monkey assays are now much less well developed than the rat assay. Also, while comparable methodology to the rat assay is available for the mouse, at present there is considerably less experience conducting the *Pig-a* assay in mice.

The ability to evaluate *Pig-a* mutants in both humans and laboratory animals means that hypotheses about the responsiveness of humans to potential mutagen exposures can be tested in animal models. Also, the human *PIG-A* assay may have value in clinical settings, where *Pig-a* assays conducted in laboratory animals could provide a translational biomarker for endogenous mutation *in vivo*. For example, the *Pig-a* and *PIG-A* assays could be used for monitoring the long-term effects of genotoxic chemotherapy [15] and aid in estimating the likelihood of developing secondary malignancies. Additionally, the *PIG-A* assay could be used in epidemiology studies for monitoring the health status (mutation accumulation) in populations exposed to potentially adverse environments, including occupational or accidental exposures to hazardous chemicals.

2. Topics discussed by the Workgroup

2.1. Placement of the assay within genotoxic testing strategies

The *Pig-a* assay should be considered an appropriate *in vivo* follow-up to positive results in bacterial and *in vitro* mammalian cell gene mutation assays. However, unless bone marrow exposure to the parent compound or to its metabolite(s) can be demonstrated directly or indirectly *via* plasma or by a reduction in the percentage of RETs, caution must be exercised in interpreting negative results. Moreover, based on analysis of the testing performed to date (see Section 2.2.3), the assay is not limited by any requirement of the test agent for metabolic activation. Although *Pig-a* is an endogenous gene, it is located on the X chromosome, and there is only one functional X chromosome. Thus, the Workgroup recognizes that the assay may not be suitable as a follow-up to either *in vitro* clastogenicity or aneugenicity findings since such events could potentially lead to cell death rather than mutation [18]. However, there currently is limited experience on the ability of the assay to detect clastogens or aneugens and the Workgroup considered the available information insufficient to justify a data-driven recommendation on this issue. Results from the *Pig-a* assay also could be used to build weight of evidence for the *in vivo* mutagenicity of a test agent, as might occur in cancer mode-of-action evaluations. Finally, the Workgroup noted the potential of using the *Pig-a* assay, as a measure of gene mutation, to complement the micronucleus (MN) assay, which measures clastogenicity/aneugenicity. Both can be readily included in routine *in vivo* safety evaluations, especially when the assays are integrated into subchronic (28 or 90 day) treatment protocols.

2.2. State of validation

Multi-laboratory trials initiated by Litron Laboratories [19], the Japanese Research Group [20], and the HESI initiative [5] have contributed to establishing protocols for the assay, testing the inter-laboratory reproducibility of the assay, and in expanding the number of agents tested.

2.2.1. Species, strain, and cell type

The assay has been investigated extensively only in rats. While the strain of rat may influence the absorption, distribution, metabolism, and elimination (ADME) of a test material and thus

affect the assay outcome, this is not anticipated to influence the biology of the endpoint; however, strain variation, if any, deserves further investigation.

Mammalian species used in other toxicological/investigational studies are amenable for use, theoretically, and there are a number of publications on *Pig-a* (or *PIG-A*) assays using other species (*e.g.*, human [15,17], mouse [2,13]). However, a recommendation for their routine use in safety assessment studies cannot be made at this time because standard protocols have not yet been evaluated extensively or published.

The assay has been extensively investigated using RBCs and RETs. Measuring the *Pig-a* mutant phenotype in RETs has the advantage of observing an induced response more quickly than in RBCs and has the theoretical advantage of mitigating any effects of immune lysis (which does not seem to be a major factor in rats [9]). For routine screening, however, it is recommended that RBCs also should be investigated for GPI-anchored protein deficiency because of the ability to score a much larger sample size of cells. Additional studies have used other cell types in the *Pig-a* assay (*e.g.*, lymphocytes [3,4], bone marrow erythroid cells [10]); but while these cell types are potentially useful, a recommendation for their routine use in safety assessment studies cannot be made at this time.

2.2.2. Intra/inter-laboratory reproducibility

The Workgroup could not identify any studies specifically intended to test the intra-laboratory reproducibility of the assay beyond assaying technical replicates (*e.g.*, [21]; also see results in Fig. 2 and Table 3). There is, however, good evidence for a high degree of inter-laboratory transferability and reproducibility based on the results with several potent and weak mutagens (and low doses of potent mutagens) that were tested in a systematic manner as part of the Litron trial [22–27]. This trial employed common protocols (refined for each stage of the trial), and common reagents, although the rat strains and flow cytometers used were at the discretion of the participating laboratory. Stage I consisted of information gathering, and Stage II tested the transferability of an anti-CD59 antibody-based method using data from Litron Laboratories as the comparator. Stage II was based on assays employing acutely administered doses of *N*-ethyl-*N*-nitrosourea (ENU), and involved 14 laboratories [22]. The results demonstrated a high level of agreement, and provided confidence in extending the studies to investigate the performance characteristics of the assay further. Stage III utilized a 28-day repeat-dose design, and experiments were performed with ENU [23], dimethylbenz[*a*]anthracene (DMBA; [24]), *N*-methyl-*N*-nitrosourea [25], benzo[*a*]pyrene (BaP; [26]), and 4-nitroquinoline-1-oxide (4-NQO; [27]) in at least 2 laboratories per agent. Results from these studies also showed a high degree of concordance among laboratories. In addition, these data provided the impetus for establishing a method to dramatically increase the number of cells evaluated per sample through the use of immunomagnetic separation technology [21]. Stage IV studies have been conducted with these updated methods and demonstrate improved statistical power [21,28,29], but only a fraction of the completed work has been published or available to the Workgroup in a form that could be used to assess intra- and inter-laboratory reproducibility.

Results from the Japanese trial, using a different staining protocol, have begun to appear, and demonstrate a similar degree of inter-laboratory reproducibility [20]. Therefore, the Workgroup concluded that the available data show that the assay is both robust and demonstrably reproducible within and across laboratories for test agents with a range of mutagenic potency. The primary variable influencing the assay outcome across laboratories appears to be sample labeling and processing for flow cytometry.

Regarding assessments of sample labeling and processing, the Workgroup acknowledged that performing independent studies with the same test agent is not always the most effective way of evaluating the impact of these factors on mutant frequency measurements. Rather, results from blood samples processed at one site on multiple occasions, and blood samples split among multiple sites, can often be more informative. For instance, Fig. 2 shows the results from 15 Sprague Dawley rats treated with either the vehicle (*i.e.*, negative control) or 1 of 2 dose levels of thiotepa. Four weeks after cessation of treatment for 28-days, blood samples were collected and processed by Litron personnel (at Rochester, NY) for same-day determination of mutant RET and RBC frequencies. A second aliquot of each whole blood sample was maintained in a refrigerator for next-day labeling and analysis at Litron, while a third set of coded aliquots was shipped overnight to Groton, CT to collaborators at Pfizer for labeling and flow cytometric analysis. All "next day" blood samples were maintained as whole blood in EDTA-coated vials until processing occurred. For the overnight shipment, the samples were kept cold, but not frozen. Both next-day analyses produced *Pig-a* mutant frequencies that were similar to those of the assays conducted on fresh samples (Fig. 2). Another study, using a different labeling protocol, demonstrated that the shipment of refrigerated blood samples had little effect on the measurement of *Pig-a* mutant frequencies [20]. These results strongly support the Workgroup's conclusion that, where there are sufficiently trained personnel, the sample labeling and flow cytometric analysis procedures are highly reproducible within and across laboratories.

2.2.3. Test agent coverage

A list of 41 agents tested in the rat *Pig-a* assay was compiled by the Workgroup (Table 1). Most of the agents identified by the Workgroup were genotoxic in one or more tests, including 26 that were Ames' test positive. The testing that has been conducted identified most of the agents expected to be positive as positive in the assay. Besides direct-acting simple alkylating agents, positive responses were detected from a number of chemicals that require metabolic activation in order to manifest their genotoxicity, including 2-acetylaminofluorene, aflatoxin B1, aristolochic acids, BaP, cyclophosphamide, diethylnitrosamine (DEN), dibenzo[*a*,*l*]pyrene, DMBA, and urethane. It is apparent from these observations that sufficient reactive metabolites from these agents damage the DNA of erythroid precursor cells to produce a response in the assay, even when other tissues may be responsible for the metabolism of these compounds. None of the unanticipated negative responses in the assay were clearly associated with a requirement for metabolic activation. Of note, the magnitude of positive responses indicates that the dynamic range of the *Pig-a* assay is at least 2–3 orders of magnitude: while background frequencies were $\leq 5 \times 10^{-6}$ (see Section 2.5.1), induced frequencies for the most potent mutagen tested to date, ENU, were (depending on the dose) in excess of 1000×10^{-6} (*e.g.*, [9]).

The Workgroup found a few examples of responses produced by genotoxic agents that illustrated both the strengths and the limitations of the assay. Urethane, a genotoxicant requiring metabolic activation not present in many *in vitro* test systems [30], was positive in the assay. DEN, which is considered to be a liver-specific genotoxicant that is often negative for assays of bone marrow genotoxicity (like the erythrocyte MN assay), initially tested negative in the assay using the 'basic' protocol [24]. Subsequent testing using an immunomagnetic separation protocol produced a positive response [31]. Etoposide, azidothymidine (AZT), and 5-fluorouracil tested negative in the assay but all have mutagenic mechanisms which would be anticipated to be difficult to detect with an X-linked reporter gene, and only the basic protocol and/or acute dosing have been used for testing to date. Finally, aflatoxin B1 was positive in the assay, but only following subchronic, and not acute dosing,

Table 1
Agents tested in the rat *Pig-a* assay.

Compound	CAS no.	Chemical class	Expected outcome[a]	Rat *Pig-a* result		References	Comments
				Acute exposure[b]	Subchronic exposure[c]		
2-Acetylaminofluorene	53-96-3	Aromatic amine	+	+	+	[33]; Novartis (unpublished)	Requires metabolic activation; 2 labs
Acetaminophen	103-90-2	Hydroxyaniline	–	–HT	–HT	Jannsen (unpublished)	Potent liver toxicant
Acrylamide	79-06-1	α,β-unsaturated amide	–/Weak+		–/? HT	FDA-NCTR (unpublished)	TGR, *Hprt* lymphocyte positive (2 mo treatment)
Aflatoxin B1	1162-65-8	Mycotoxin	+	–HT	Weak+	Jannsen (unpublished)	Requires metabolic activation by CYP3A4 (not highly expressed in rats compared to humans)
o-Anthranilic acid	118-92-3	Aromatic amine	–	–HT	–HT	[33]	Non-alerting structure
Aristolochic acids	313-67-7	Aromatic nitro	+	+	+	[32]; Novartis (unpublished)	2 labs; weak *in vivo* MN
Azathioprine	446-86-6	Aromatic nitro	–/Weak +	Weak+	Weak+	[33]	Immunotoxicant
Azidothymidine	30516-87-1	Nucleoside analog	–/Weak+	–Basic (7 days)		[60]	*In vivo* MN positive; *Tk* mutation positive (mouse)
Benzo[*a*]pyrene	50-32-8	PAH	+	+	+	[26,29,37,52]	Requires metabolic activation; 3 labs
2-Butoxyethanol	111-76-2	Ethylene glycol derivative	–	–HT	–HT	Pfizer (unpublished)	Hemolytic agent
Chlorambucil	305-03-3	Alkylator; nitrogen mustard	+	+	+	[33,53]; Novartis (unpublished)	2 labs; twice at one lab
4-Chloro-1,2-diaminobenzene	95-83-0	Aromatic amine	–(Male) +(Female)	–HT (male)	–HT (male)	GSK (unpublished)	*In vivo* MN positive acute exposure only; Females not tested
Cisplatin	15663-27-1	Antineoplastic; crosslinker	+	+	+	[54,58]	2 labs
Cyclophosphamide	6055-19-2	Alkylator, nitrogen mustard	Weak+	+	+	[33,56,58]	Requires metabolic activation; strong *in vivo* MN positive; 3 labs
Dibenzo[*a,l*]pyrene	191-30-0	PAH	+	+		Roche (unpublished)	Requires metabolic activation
Diethylnitrosamine	55-18-5	Nitrosamine	–(Blood)		–Basic; weak + HT	[24,31]	Requires metabolic activation; weak *in vivo* MN
[1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride, (EDAC)	25952-53-8	Carbodiimide	–		–HT	[61]	Ames and *in vitro* MN positive; Tested up to MTD, rapid degradation into non-mutagenic EDAU in acid conditions and in aqueous solutions (SD rats)
7,12-Dimethyl-benz[*a*]anthracene	57-97-6	PAH	+	+	+	[1,20,24,37,52]; Janssen (unpublished)	Requires metabolic activation; 5 labs; twice at one lab (SD & Wistar rats)
1,2-Dimethylhydrazine hydrochloride	306-37-6	Hydrazine	–(Blood)	–Basic		Pfizer (unpublished)	
Ethylmethane sulfonate	62-50-0	Alkylator	+	+	+	[56,62]; FDA-NCTR (unpublished)	3 labs
N-Ethyl-*N*-nitrosourea	759-73-9	Alkylator	+	+	+	[1,3,9,20,22,23,37,51,56,57,62]; PBR Laboratories (unpublished)	Multiple labs (13+ for 3 day; 5 for 28-day)

Table 1 (*Continued*)

Compound	CAS no.	Chemical class	Expected outcome[a]	Rat *Pig-a* result		References	Comments
				Acute exposure[b]	Subchronic exposure[c]		
Etoposide	33419-42-0	Topoisomerase inhibitor	–	–Basic	–Basic	Pfizer (unpublished)	Bacterial specific mutagen
5-Fluorouracil	51-21-8	Pyrimidine analog	–/?	–HT		NCDSER (unpublished)	Toxicity to red blood cells observed; positive in concurrent *in vivo* MN test
Furan	110-00-9	Heterocycle	–		–Basic (56 days)	[63]	Requires metabolic activation by CYP2E1; concurrent spleen lymphocyte *Hprt* and *Pig-a* and *cII* liver mutation assays were negative
Hydroxyurea	127-07-1	Inhibits DNA replication	–	–HT	–HT	[33]	
Ionizing radiation	N/A		+	+(1 and 4 days)		[58]; Litron Laboratories (unpublished)	2 labs
Isopropylmethane sulfonate	926-06-7	Alkylator	+	+(1 and 3 days)	+	Pfizer (in preparation)	
Isopropyltoluene sulfonate	2307-69-9	Alkylator	+	+(1 day)		Pfizer (unpublished)	
Melamine	108-78-1	Triaminotriazine	–	–HT		NCDSER (in preparation)	
Melphalan	148-82-3	Alkylator; nitrogen mustard	+	+	+	[33,53]; Novartis (unpublished)	2 labs; twice at one lab
Methylmethane sulfonate	66-27-3	Alkylator	Weak+	+	+	[33]; Roche (unpublished)	2 labs
Methylphenidate	113-45-1	Phenylethylamine derivative (psycho-stimulant)	–		–Basic (21 days)	[55]	
4-Nitroquinoline-1-oxide	56-57-5	Aromatic *N*-oxide	+	+	+	[20,27,37,52,56,64]; AbbVie (unpublished)	4 labs; twice at one lab
N-Nitroso-N-methylurea	684-93-5	Alkylator	+	+	+	[25,37,52,65]	3 labs; twice at one lab (basic and HT)
1,3-Propane sultone	1120-71-4	Alkylator	+	+	+	[28,53]; Novartis (unpublished)	2 labs; twice at one lab
Pyrene	129-00-0	PAH	–	–HT	–HT	[29]	
Sodium chloride	7647-14-5	Salt	–	–HT	–HT	[33]	
Sulfisoxazole	127-69-5	Aryl sulfonamide	–	– HT	–HT	[33]	
Temozolomide	85662-93-1	Imidazotetrazine; (Alkylating agent)	+	+(5 days)		Roche (unpublished)	
Thiotepa	52-24-4	Aziridine	+	+	+	[33,53]; Novartis (unpublished); Litron/Pfizer (summarized within)	2 labs; twice at one lab and a shipment study with a third lab
Urethane	51-79-6	Carbamate	+		+	[66]	Requires metabolism that is not available *in vitro*; TGR positive; weak effects

[a] Based on expert judgment of Ames and *in vitro* mammalian gene mutation data, mechanism of genotoxicity, and tissue distribution *in vivo*.

[b] Acute is 3 consecutive days dosing unless otherwise noted.

[c] Subchronic is 28 consecutive days dosing unless otherwise noted.

HT, immunomagnetic separation; only noted when a compound is negative; basic, no immunomagnetic separation; only noted when a compound is negative; TGR, transgenic rodent; –, negative; +, positive;?, equivocal; N/A, not applicable; ND, no data; PAH, polyaromatic hydrocarbon; NCDSER, National Shanghai Center for Drug Safety Evaluation and Research.

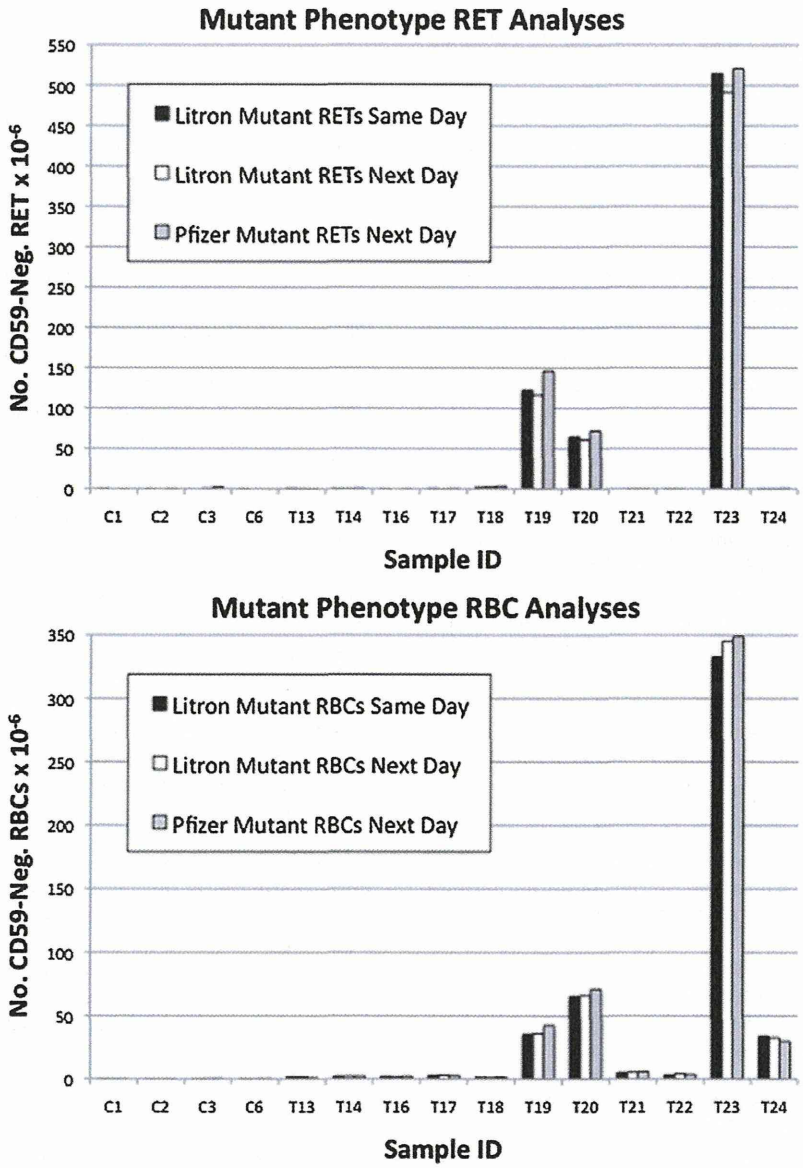

Fig. 2. *Pig-a* data from 15 individual male Sprague Dawley rats treated with vehicle (C) or 1 of 2 dose levels of thiotepa (T). All animals were treated daily for 28 consecutive days. 'C' animals were treated with the vehicle, T13-T18 animals were given 0.492 mg thiotepa/kg/day, while animals T19–T24 were given 10 mg thiotepa/kg/day for 17 days and 5 mg/kg/day for the remaining 11 days. Blood samples were collected 28 days after the last treatment and processed by Litron Laboratories' personnel for same-day determination of *Pig-a* mutant RET and RBC frequencies. A second aliquot of each whole blood sample was maintained in a refrigerator overnight prior to labeling and analysis at Litron (Rochester, NY), while a third set of aliquots was shipped to collaborators at Pfizer (Groton, CT) for labeling and flow cytometric analysis. All "next day" blood samples were maintained as whole blood in EDTA-coated vials until processing occurred. In the case of the overnight shipment, vials were kept cold, not frozen. Each bar represents the *Pig-a* mutant frequency for each individual rat's blood sample at one of the 3 conditions tested. Note that some bars for the control and low-dose-treated rats are hardly visible on the figure. The results, which show extreme rat-to-rat variability among the high-dose animals at this late blood collection time point, indicate that across a wide dynamic range, *Pig-a* mutant frequencies are highly reproducible within and between laboratories.

while acrylamide was at best inconclusive after subchronic testing; acrylamide was not tested using acute treatment.

Based on these observations, the Workgroup concluded that definitive negative calls should be based on data from the most sensitive protocols (*i.e.*, using immunomagnetic separation), and tests conducted to the maximum tolerated dose (MTD) or the limit dose (see Section 2.4.2). This recommendation extends to agents anticipated to be negative, including Ames' negative agents. Of the 8 agents tested that generally are considered non-genotoxicants, 5 have been tested to the above standards (*o*-anthranilic acid, 2-butoxyethanol, pyrene, sodium chloride, sulfisoxazole), while 3 have been tested incompletely (furan, melamine, methylphenidate). All 8 have tested negative in the assay. Of these 8 chemicals, 2-butoxyethanol is of particular significance because it caused marked intravascular lysis of RBCs and induced a strong compensatory erythropoiesis, yet it did not affect *Pig-a* mutant frequencies.

Within the group of agents that have been tested, the Workgroup noted two reports of assays (using 3- and 28-day treatments with aristolochic acids and 28-day treatments with 4-NQO) that were negative for erythrocyte MN induction, but positive for *Pig-a* mutation [27,32]. There also were 3 agents identified by the Workgroup (AZT, hydroxyurea and 5-fluorouracil) that induced micronucleated RETs but did not increase *Pig-a* mutant

erythrocyte frequencies. The experiments with hydroxyurea employed both 3- and 28-day study designs with immunomagnetic separation (Table 1) [33]. Similarly, etoposide was negative in limited *Pig-a* testing, and previous work with this non-DNA-reactive compound indicates that it would have tested positive for MN induction (*e.g.* [34]). It was noted that these observations support the routine use of the *Pig-a* assay in conjunction with the *in vivo* MN assay, especially when subchronic treatments are conducted, as the information from these 2 assays could be complementary and/or confirmatory.

The Workgroup recommended expanding the number of tested agents in general and Ames' negative agents in particular (see Section 3.1.3). The Workgroup further recommended testing more Ames' negative agents that might interfere with an erythrocyte-based assay (*e.g.*, those that might induce erythropoiesis and/or lyse RBCs, like 2-butoxyethanol [Table 1]).

2.3. Evaluating mutations responsible for Pig-a mutants

Guidance provided by the Organization for Economic Cooperation and Development (OECD; [35]) indicates that assay validation should include a demonstration that the assay measures what it is designed to measure. In this case, it is important that a *Pig-a* gene mutation assay measures mutation induction in the *Pig-a* gene rather than, for instance, gene silencing or the clonal expansion of pre-existing *Pig-a* mutations. Establishing the mutational basis of the response is a challenge, however, because the assay measures the mutant phenotype in enucleated erythrocytes. To date, there is no direct evidence that *Pig-a* mutation is responsible for producing *Pig-a* mutant erythrocytes. However, there is good evidence from nucleated bone marrow and spleen cells from rodents, and extensive experience from human PNH patients (see below), that supports the conclusion that GPI-anchored protein-deficient cells, in fact, have mutations in the *Pig-a* gene.

Studies indicate that proaerolysin (a bacterial toxin that potentiates its toxicity through binding to GPI) selects T-lymphocyte clones from ENU-treated rats that contain mutations in the *Pig-a* gene [4,36]. In addition, flow-sorted erythroid mouse bone marrow cells that are deficient in GPI-anchored markers also contain mutations in the *Pig-a* gene [10]. In addition, it has been demonstrated that rat CD59-negative erythrocytes are indeed GPI anchor-deficient (*i.e.*, resistant to aerolysin) [37]. Furthermore, Lemieux et al. [38] reported similar dose–response relationships for BaP-induced *Pig-a* mutant erythrocytes compared with bone marrow *lacZ* mutations using a transgenic mouse model. The Workgroup noted, however, that these data fall short of demonstrating that the individual mutants identified by flow cytometry (from either lymphocytes or erythrocytes) contain (or contained) mutations in the *Pig-a* gene. There is also an extensive literature indicating that PNH patients contain large clonal expansions of lymphocytes with *Pig-a* mutations and large numbers of phenotypically mutant erythrocytes [39–41].

The Workgroup concluded that the weight of evidence was consistent with a direct association between *Pig-a* mutation (rather than gene silencing or enzymatic inactivation) and the phenotype measured in the assay, and indicated that lack of absolute proof should not preclude use of the *Pig-a* assay for regulatory purposes. The Workgroup, however, encourages further work in this area to better characterize the mutations presumably measured by the assay (see Section 3.1.1). Until this proof is provided, the Workgroup advised that the endpoint measured by the assay should be referred to as 'GPI-anchored protein-deficient erythrocytes' or, more simply, '*Pig-a* mutant erythrocytes'.

2.4. Assay protocol

2.4.1. Test animals/sex

The assay has been investigated extensively and standardized only in the rat. Methods exist, however, for measuring *Pig-a* mutation in other species, in particular, the mouse. However, a recommendation for their routine use in safety assessment studies cannot be made at this time as standard protocols have not yet been published.

So far, nearly all the rat studies have been conducted in males, with only limited data in females. Consequently, no definitive statement can be made as to whether X-chromosome silencing in females (*viz.*, "Lyonization") has any influence on assay sensitivity. Nevertheless, evidence from the X-linked hypoxanthine-guanine phosphoribosyltransferase (*Hprt*) locus indicates the frequency of *Hprt* mutants in circulating T-lymphocytes is similar in male and female humans [42] and mice (*e.g.* [43,44]). Female human carriers of *HPRT* germ line mutations are healthy but their somatic cells are mosaics in terms of enzyme activity because of the random inactivation of the X-chromosome; in contrast *HPRT*-mutant males develop one of two clinical syndromes: Lesch-Nyhan syndrome or the Kelley–Seegmiller syndrome (reviewed by [45]). The incidence of Lesch–Nyhan syndrome (*HPRT* nullizygous) is very rare in females, with only a handful of women described [46]; the mechanisms responsible are complex and non-random X-inactivation has been observed in a single female patient [47]. There are no notable publications indicating whether Lyonization has any influence on the *HPRT* cell mutation assay. Nevertheless, the Workgroup suggested establishing the relative sensitivity of male and female animals to agents with known genotoxic potential (see Section 3.1.2). In addition, it was suggested that future research should explore the response of the assay to chemicals that affect the methylation status of the genome, as such changes may be more problematic for an assay using a reporter on the X-chromosome. Until these questions are resolved, it is advisable to use results from females with a certain degree of caution. However, where human exposure to chemicals may be sex-specific, as with some pharmaceutical agents, the Workgroup also recommends that the test should be performed with animals of the appropriate sex.

2.4.2. Number of doses, maximum doses, and age and number of animals

The Workgroup recommends following criteria similar to those given in OECD Test Guidelines (TGs) for conducting acute and/or repeat dose *in vivo* toxicology and genetic toxicology studies (*e.g.*, TG407, TG408, and TG488 [48–50]). In general, young adult animals should be used (note that evidence indicates that treating animals *in utero* produces unstable responses [51]). A minimum of 3 dose levels of the test article should be employed, with the high dose being the MTD or, if toxicity does not limit the dose, 1000 mg/kg/day for subchronic (14 days or more) and 2000 mg/kg/day for acute treatments. Range-finding studies to select dose levels should be conducted in the same species, strain, and sex as the main test. The MTD is defined as the highest dose that will be tolerated without evidence of excessive toxicity such as death, pain, suffering, excessive weight loss, and/or distress. At this time, there is an inadequate understanding on how perturbations in hematopoiesis affect the performance of the *Pig-a* assay. It is important, therefore, that the results of the assay are interpreted in the context of potential assay confounders.

Group sizes of 6 are recommended but 5 analyzable animals per group may be acceptable if justified (*e.g.*, based on power calculations, see Table 2). The use of an animal group for a positive control is not considered mandatory provided an appropriate standard that "mimics" mutants [52] is used each time flow cytometric analysis is performed. The standard serves to establish the flow cytometer

Table 2
Power analyses for 2 laboratories' vehicle control data (Litron, $n = 114$ Sprague Dawley rats; and Pfizer, $n = 35$ Wistar Han rats).

Parameter	Increase over baseline	Chance of detecting ($n = 6$/group)*		Smallest increase detectable with 80% power ($n = 10$/group)	
		Litron	Pfizer	Litron	Pfizer
CD59-negative RETs	2×	32%	45%	3.1-fold	2.4-fold
	3×	60%	65%		
	4×	77%	92%		
CD59-negative RBCs	2×	50%	60%	2.3-fold	2-fold
	3×	84%	92%		
	4×	95%	98%		

Based on log(10)-transformed mutant phenotype RET and RBC frequencies; these power calculations were from the one-way analysis option within PROCpower (SAS v9.2), 4 group test (control *vs.* three other treatment groups), $\alpha = 0.05$, one-sided test. All data were generated using immunomagnetic separation technology.

* With corrected *t*-tests (*e.g.*, Dunnett), need $n = 10$/group to achieve a similar power for this design, based on current estimates of variability.

settings and provides a rational and consistent means of identifying the *Pig-a* mutant cell populations.

2.4.3. Prescreening animals

Measuring *Pig-a* mutant frequencies in animals one or a few days before the start of test agent treatment can be useful for eliminating outlier animals with unusually high or low mutant phenotype frequencies. The elimination of outlier animals may serve to increase the sensitivity of the assay, since rare animals with high frequencies can adversely affect statistical power. The Workgroup considers this practice to be potentially useful, but optional. In addition, if used, the criteria for excluding animals should be established prior to conducting the assay.

2.4.4. Treatment and sampling schedules

Since *Pig-a* mutant cells accumulate with repeated dosing, the Workgroup recommends using a subchronic treatment protocol, ideally a 28-day daily dosing schedule similar to that recommended in OECD TG488 for performing the transgenic rodent (TGR) gene mutation assay [50]. This schedule has been used to generate much of the current *Pig-a* data and a 28-day treatment schedule is consistent with integrating the assay into standard subchronic general toxicology studies ([48]; see Section 2.4.5). The Workgroup recommends exploring the feasibility of using a 90-day treatment schedule for the assay, as this treatment duration also is used for subchronic general toxicology studies [49]. In addition, acute and subacute dosing is considered useful in some instances, with justification (*e.g.*, adequate exposure relative to the cumulative exposure achieved by subchronic treatments, and/or the intended or expected human exposure is acute in nature). However, with short treatment schedules, it is especially important to allow sufficient time for a positive result to be expressed in peripheral blood cell population(s).

For both acute and 28-day dosing studies, the Workgroup recommends conducting mutant analyses at approximately 28–31 days after the initiation of treatment. Earlier, optional sampling times can be informative, especially in the case of acute studies and/or to bolster confidence in a weak effect. If optional recovery groups are included (or if the animals are not euthanized soon after the 28-day sampling), a sample at later time points (*e.g.*, 2–4 weeks after cessation of treatment for subchronic treatments and 4–6 weeks after cessation of treatment for acute treatments) may be included, as these later time points have been shown to more fully manifest responses in the total erythrocyte population [9,33,52,53]. Caution should be exercised in assaying animals at intervals >56 days after starting the treatment, especially after toxic treatments, as greater variation in mutagen-induced responses have been observed [9,54].

2.4.5. Assay integration

One of the most attractive features of the *Pig-a* assay is its potential for integration into repeat dose toxicology studies and with other genetic toxicology assays. Integration helps in the assessment of a potential mutagenic response in the context of other genetic and general toxicology endpoints (*e.g.*, target tissue exposure and toxicity). The combination of several assays in one set of test animals is also consistent with the 3Rs principles for animal welfare. The assay requires only microliter samples of peripheral blood, which can be obtained in a minimally invasive manner without disturbing the assessment of other endpoints. Unlike the TGR gene mutation test, the *Pig-a* assay is not limited to specific strains of animals, and the timing of sample collection is not as critical as it is for the *in vivo* Comet assay. This makes the *Pig-a* assay more attractive than these other *in vivo* assays for combination with established toxicology protocols. The fact that mutant frequencies can be determined before, during, and after the treatment schedule, makes it possible to collect longitudinal data, which may help the interpretation of mutant frequency responses in relation to the other endpoints that are collected.

2.4.6. Sample analysis

There are several published methods for analyzing blood samples for *Pig-a* mutants, but currently two major approaches are employed. The most common method used in North America and Europe is available as a commercial kit (*In Vivo* MutaFlow®, Litron Laboratories, Rochester, NY; [33]). The protocol employed by this method starts by leukodepleting the blood, labeling *Pig-a* wild-type cells with a fluorescent antibody against a GPI-linked marker (anti-CD59 is used for rats) and a fluorescent nucleic acid dye that distinguishes mature erythrocytes, RNA-positive RETs, and residual leukocytes. An immunomagnetic separation technique is used to deplete samples of wild-type cells prior to flow cytometric analysis. Immunomagnetic separation dramatically increases the number of cell equivalents analyzed for mutants while simultaneously reducing data acquisition times.

In Japan, many studies have been conducted using a fluorescent antibody to detect erythroid cells (HIS49 for rats [55,56]), plus fluorescent antibodies to detect a GPI-anchored marker (CD59 in rats) and transferrin receptor (CD71, which is used as a marker for RETs). This method has also begun incorporating a magnetic separation step to increase the number of cells scored in the assay, in this case, CD71-positive RETs [57]. Both immunomagnetic separation approaches have been used in multi-laboratory studies. Also, a study that compared the two methods found that they produced similar *Pig-a* mutant frequencies [58].

The Workgroup concluded that there is no compelling reason to prefer one protocol over another, provided that a sufficient number of cells are assayed for *Pig-a* mutants. Whichever protocol is used, it is advisable to perform reconstruction-type experiments,

where samples having a known range of mutant erythrocytes are analyzed, before undertaking studies with uncharacterized chemical(s). Reconstruction experiments provide a benchmark for evaluating the degree to which a cell processing protocol and analytical scoring platform produce expected results over a relevant range of mutant frequencies.

2.4.7. Minimum cells scored

The sample size should be large enough to contain at least 1 GPI-anchor-deficient variant event per animal. From a practical standpoint, this typically means assaying a minimum of $1–5 \times 10^6$ RETs or RBCs per animal. However, in order to avoid the chance occurrence of zero variants in the sample and to increase the robustness of the test, assaying samples that are as large as practically feasible is recommended for both the control and treated animals. As a guide, if the mutant frequency (p) and the number of cells scored (n) is adjusted so that the value $n \times p = 5$, then <0.7% of the samples would have zero mutants. Based upon a Poisson distribution with a mean of 5, the proportion of samples with a zero count is less than 0.7%. This assumes no inter-animal variability. If there is, then although the group mean may be 5, some animals will have mean values lower than 5 so a sample from them will have more chance of being zero. This could lead to a higher proportion of samples with zero counts.

As shown below, background *Pig-a* mutant frequencies can be low, and collecting data from considerably more than the recommended minimum cells may be necessary to avoid having samples with no mutants detected (*i.e.*, zero values) and to be able to conduct assays with suitable statistical power (see Table 2). While it is possible to extend flow cytometer analysis to hours, this makes it impractical to study the required number of samples in a standard study, *e.g.*, 20–24 samples or more per sampling time (as suggested in Section 2.4.2). Thus, the Workgroup encourages methodologies (*e.g.*, immunomagnetic separation [21,57]) that increase the throughput of the assay when needed.

2.4.8. Preservation/shipping of samples

The preserving and possibly shipping of blood samples is useful when it is inconvenient to perform the mutant analysis on the day of sample collection, for performing follow up/confirmatory assays, or when the mutant analysis is to be performed at a site some distance from where the animal work was performed. Although methods for the fixation and long term storage of unprocessed blood samples have not yet been proposed, the Workgroup reviewed data indicating that unstained leukodepleted blood samples and whole blood maintained in EDTA-coated tubes can be stored for several days without affecting measurements made in the *Pig-a* assay – percent RETs, frequency of mutant RETs, and frequency of mutant RBCs (Table 3). Furthermore, as illustrated by Fig. 2, data also show the feasibility of shipping blood samples under cold but not frozen conditions. Similarly, it is expected that the preservation/shipping of blood samples can be applied for any antibody labeling protocol. Whether storage occurs for delayed on-site analyses, or whether shipment(s) occur to an analytical site, the Workgroup recommends each laboratory develop their own supporting data before using these options.

2.5. Data acceptance and analysis

2.5.1. Minimum experiment and background frequencies

The Workgroup recommends that a minimum study for measuring *Pig-a* mutant frequency should consist of 3 appropriately spaced doses of the test agent and a vehicle control, with each group containing 6 animals (although 5 analyzable animals may be acceptable if justified by power analysis), and with mutant frequencies analyzed at approximately 28–31 days after the beginning of the treatment. Additional sampling times are not required, but are encouraged. A positive control group is not necessary, but a demonstration of the laboratory's ability to detect mutants using mutant mimic standards is required as a part of instrument calibration.

The available data indicate that the background or vehicle control mutant frequencies in the assay differ slightly with species [13,14] and staining method [58], but are generally low in experienced laboratories ($\leq 5 \times 10^{-6}$). Each laboratory should establish its own background distribution for each strain and sex of animal used; the negative controls for assays should be within acceptable limits based on generally accepted quality control methods (see [59]).

2.5.2. Statistical analysis

Appropriate statistical methods should be used in evaluating test results. One such method is an analysis of variance (ANOVA) followed by pair-wise comparisons of mutant frequencies in treated groups to the vehicle control group. Other methods may be suitable but should be justified. When parametric methods such as ANOVA are used, it is important to ensure that the data meet the assumptions underlying the ANOVA (such as residuals being approximately normally distributed). A log(10) transformation of mutant frequencies often has been found satisfactory for data to meet these assumptions. If the objective of the analysis is to declare a comparison significant, the tests should be performed with a predefined alpha value, and should generally be conducted as one-sided tests, that is, with the goal of detecting a treatment-related increase in mutant cells. Since zeroes may be observed occasionally, prior to some transformations, *e.g.*, log-transformation, an offset of +0.1 (*i.e.*, addition of 0.1 to each mutant cell frequency expressed as mutants $\times 10^{-6}$) may be required. While there are statistical approaches that are capable of simultaneously considering mutant RET and mutant RBC data, and these approaches may be useful in certain instances, the Workgroup notes that these two endpoints most often have been considered separately in the published literature. There are statistical approaches that consider treatment-related effects as a repeated-measure (longitudinal). These methods may be useful and appropriate in some cases, but are more complex to carry out and interpret, and the majority of published reports have considered each time point separately. As described below, in addition to pair-wise tests, evidence of dose-related increases should be evaluated with appropriate statistical tests.

2.6. Criteria for classifying responses

The criteria for classifying responses in the *Pig-a* assay are similar to those currently recommended in OECD Test Guidelines for other genotoxicity assays (*e.g.*, TG488 [50]). While the biological relevance of the response should be a major consideration, appropriate statistical methods also should be used for evaluating the size of induced responses. A positive response is characterized by a dose-related increase in mutant frequency, a significant increase in at least one dose, and an induced mutant frequency greater than the historical negative control distribution. Biological relevance is supported by increases in both total RBC and RET mutant frequencies, especially increases that conform with the expected temporal manifestation of mutants in these two erythrocyte cohorts, and increases that occur at multiple sampling points.

Because of the large dynamic range of induced *Pig-a* mutant frequencies, the Workgroup agreed that a doubling over the concurrent negative control may not be sufficient evidence of a biologically relevant treatment-induced effect. Since the power to detect an X-fold increase in mutant RBCs *vs.* mutant RETs tends to