RADIATION RESEARCH **190**, 000–000 (2018)
0033-7587/18 $15.00
©2018 by Radiation Research Society.
All rights of reproduction in any form reserved.
DOI: 10.1667/RR15013.1

Impact of Inter-Individual Variance in the Expression of a Radiation-Responsive Gene Panel Used for Triage

S. Agbenyegah,[a] M. Abend,[b,1] M. J. Atkinson,[c] S. E. Combs,[a,d] K. R. Trott,[a] M. Port[b] and M. Majewski[b]

[a] *Department of Radiation Oncology, Technical University of Munich, Munich, Germany;* [b] *Bundeswehr Institute of Radiobiology, Munich, Germany;* [c] *Institute of Radiation Biology, Helmholtz Zentrum München, Neuherberg, Germany; and* [d] *Institute of Innovative Radiotherapy, Helmholtz Zentrum München, Oberschleissheim, Germany*

Agbenyegah, S., Abend, M., Atkinson, M. J., Combs, S. E., Trott, K. R., Port, M. and Majewski, M. Impact of Inter-Individual Variance in the Expression of a Radiation-Responsive Gene Panel Used for Triage. *Radiat. Res.* 190, 000–000 (2018).

In previous studies we determined a gene expression signature in baboons for predicting the severity of hematological acute radiation syndrome. We subsequently validated a set of eight of these genes in leukemia patients undergoing total-body irradiation. In the current study, we addressed the effect of intra-individual variability on the basal level of expression of those eight radiation-responsive genes identified previously, by examining baseline levels in 200 unexposed healthy human donors (122 males and 88 females with an average age of 46 years) using real-time PCR. In addition to the eight candidate genes (*DAGLA*, *WNT3*, *CD177*, *PLA2G16*, *WLS*, *POU2AF1*, *STAT4* and *PRF1*), we examined two more genes (*FDXR* and *DDB2*) widely used in *ex vivo* whole blood experiments. Although significant sex- (seven genes) and age-dependent (two genes) differences in expression were found, the fold changes ranged only between 1.1–1.6. These were well within the twofold differences in gene expression generally considered to represent control values. Age and sex contributed less than 20–30% to the complete inter-individual variance, which is calculated as the fold change between the lowest (reference) and the highest Ct value minimum–maximum fold change (min–max FC). Min–max FCs ranging between 10–17 were observed for most genes; however, for three genes, min–max FCs of complete inter-individual variance were found to be 37.1 (*WNT3*), 51.4 (*WLS*) and 1,627.8 (*CD177*). In addition, to determine whether discrimination between healthy and diseased baboons might be altered by replacing the published gene expression data of the 18 healthy baboons with that of the 200 healthy humans, we employed logistic regression analysis and calculated the area under the receiver operating characteristic (ROC) curve. The additional inter-individual variance of the human data set had either no impact or marginal impact on the ROC area, since up to 32-fold change gene expression differences between healthy and diseased baboons were observed. **© 2018 by Radiation Research Society**

INTRODUCTION

High-dose radiation exposure in the hematopoietic tissues of bone marrow leads to hematological acute radiation syndrome (HARS) (*1–3*). The severity of the damage is not immediately apparent, complicating triage and treatment planning. With the medical treatment protocols for radiation accident victims (METREPOL), four HARS severity degrees can be discriminated using the METREPOL criteria of blood cell count changes during the first weeks after irradiation; low (H1), medium (H2), severe (H3) and fatal (H4) (*2*). Blood cell count changes lying within the normal range are classified as H0 and represent healthy individuals. In the management of large-scale radiation emergencies resources are limited, creating the need for a more rapid predictor of outcome for triage purposes. Based on the analysis of gene expression in the peripheral blood, several different panels of early response genes have been identified, although their predictive power remains uncertain (*3–8*).

In previous studies we identified a gene expression signature for the early prediction of later developing hematological acute radiation syndrome using peripheral blood from 18 irradiated baboons. These baboons were treated with partial- and total-body irradiation resulting in different degrees of the HARS (*9, 10*). Twenty-nine candidate genes were identified in the baboon study (*8*). Eight of these genes were subsequently confirmed in a panel of leukemia patients undergoing total-body irradiation (*11*). For the next validation step, in the current study we addressed the effect of inter-individual variance in gene expression on the discrimination of diseased and healthy baboons. Here we extended the sample size by approximately tenfold, from 18 baboons (in our previous study) to 200 unexposed healthy human donors examined within this

Editor's note. The online version of this article (DOI: 10.1667/RR15013.1) contains supplementary information that is available to all authorized users.

[1] Address for correspondence: Bundeswehr Institute of Radiobiology affiliated to the University of Ulm, Neuherbergstr. 11, 80937 Munich, Germany; email: michaelabend@bundeswehr.org.

study. In addition to the eight candidate genes (*DAGLA*, *WNT3*, *CD177*, *PLA2G16*, *WLS*, *POU2AF1*, *STAT* and *PRF1*) we analyzed two other genes, namely *FDXR* and *DDB2*. These genes are already known and widely used as radiation-responsive genes in the *ex vivo* whole blood experiments of several laboratories (*5, 6, 12–14*). Additional information on the inter-individual variance of these two widely used genes was of interest since, to date, this had not been available. In this study, we also sought to close some gaps in knowledge regarding gene expression changes caused by age and sex. However, the primary goal was to add another validation step to elucidate how inter-individual gene expression variance of our candidate genes might alter the discrimination of healthy (unexposed) from diseased (exposed) baboons. To do this, we replaced the published gene expression data from 18 healthy baboons with the newly generated gene expression data from 200 healthy humans, employing logistic regression analysis and calculating the area under the receiver operating characteristic (ROC) curve.

MATERIALS AND METHODS

Donors and Samples

Through the Bavarian Red Cross we received peripheral whole blood samples from 112 healthy males (56%) and 88 healthy females (44%) with an average age of 48.2 years and 42.3 years, respectively. To reflect a typical triage chain aliquots of 2.5 ml peripheral whole blood were collected into PAXgene™ blood RNA tubes (PreAnalytiX; QIAGEN, Hilden, Germany) from each individual. Tubes were inverted 10 times, kept at room temperature overnight, stored at –20°C during the collection phase, delivered to the Bundeswehr Institute of Radiobiology (Munich, Germany) on ice and stored there at –20°C. Before use, the samples were thawed at room temperature.

Candidate Genes

The ten genes used for analysis are coding for biological processes such as DNA damage repair (*DAGLA*, *DDB2*, *FDXR*, *PLA2G16*), cell cycle regulation (*PRF1*, *CD177*, *POU2AF1*) and immune response/developmental processes (*STAT4*, *WLS*, *WNT3*) (*15, 16*).

RNA Extraction and Quality Control

Whole blood samples from donors (2.5 ml) were processed according to the manufacturer's PAXGene QIAsymphony Kit protocol including a DNAse I treatment to eliminate possible genomic DNA contamination using automatic total RNA isolation robotics (QIAsymphony; QIAGEN). RNA was quantified spectrophotometrically (NanoDrop; PeqLab Biotechnology, Erlangen, Germany) and RNA integrity was assessed using the 2100 Agilent Bioanalyzer (Life Science Group, Penzberg, Germany). DNA contamination was controlled by conventional polymerase chain reaction (PCR) using β-actin primers. We used only RNA specimens with a ratio of $A_{260}/A_{280} \geq 2.0$ and RNA integrity number (RIN) ≥ 7.5 for qRT-PCR analyses.

Gene Expression Quantification Using Real-Time Polymerase Chain Reaction (qRT-PCR)

One μg RNA was converted into cDNA using a recommended protocol (High-Capacity cDNA Reverse Transcription kit). Based on this protocol equal volumes of template cDNA [10 ng RNA equivalent/10 μl (increased fivefold for detection of *WNT3*)] and TaqMan® Universal PCR Master Mix were mixed and a PCR reaction containing inventoried assays (TaqMan minor groove binder probes) for specific gene expression quantification were run in duplicate for amplification using the Applied Biosystems® 7900HT Fast Real-Time PCR System (Weiterstadt, Germany). The inventoried assays for each gene are given in parentheses as follows: *FDXR* (Hs01031617_m1), *DDB2* (Hs00172068_m1), *WNT3* (Hs00902257_m1), *POU2AF1* (Hs01573371_m1), *CD177* (Hs00360669_m1), *WLS* (Hs01553062_m1), *DAGLA* (Hs00391374_m1), *PLA2G16* (Hs00912734_m1), *STAT4* (Hs01028017_m1) and *PRF1* (Hs00169473_m1).

Eukaryotic 18S rRNA was used for normalization of expression data using the ΔΔCT approach. The calculated ratio between a reference and an observed gene of interest was referred to as ''differential gene expression.'' A ratio or fold change (FC) of one corresponds to a gene expression similar to a reference gene. A twofold change in gene expression over control was introduced to avoid false positive results and values falling within that range were considered to represent control values.

All materials used for qRT-PCR were obtained from Applied Biosystems. All technical procedures were performed in compliance with standard operating procedures implemented in the certification of the Bundeswehr Institute of Radiobiology according to DIN EN ISO 9001/2008.

Statistical Analysis

We calculated descriptive statistics of FCs in gene expression among groups and employed either the *t* test or the Mann-Whitney rank sum test (where applicable) to examine statistically significant differences between FC values of groups. For examinations of normal distribution we employed the Kolmogorov-Smirnov test. All gene expression data are normal distributed (data not shown). First-order-linear-regression analysis was used to examine an age dependency of gene expression over all age categories for each sex separately (Wald *P* value or *P* trend analysis). For age-dependent sex differences we also employed a paired *t* test. The ability of gene expression values to discriminate healthy from diseased binary groups (H0 vs. H1–3 or H2–3 HARS) was performed in univariate analysis using unconditional logistic regression and by calculating the area under the ROC curve.

An area of one under the ROC curve implies complete separation of both groups. Binary groups representing healthy (unexposed) and diseased (exposed) baboons as well as models comprising healthy (unexposed) human and diseased (exposed) baboons were examined using the corresponding gene expression data sets. We also calculated the coefficient of variation (CV), but based on duplicate gene expression measurements only. The CV can be converted into a relative measurement error [RME = abs(CT1–CT2)/mean corresponding to sqrt(2)*CV or CV*1.41] to adjust for the low number of replicate measurements. All calculations were performed using SAS software version 9.2 (Cary, NC).

RESULTS

Variability of RNA Isolation and qRT-PCR

On average, 5.8 μg (±1.2) total RNA/2.5 ml whole blood with RIN of 8.2 (±0.4) was isolated from the 200 samples. The methodological variance was addressed by calculating the coefficient of variation (CV) of duplicate qRT-PCR measurements relative to the mean cycle threshold (Ct) value. The CV of duplicate qRT-PCR measurements for all genes ranged between 0.3–0.5%, (RME = 0.4–0.7%) apart

from *WNT3* with a CV of 0.75% (RME = 1.06%, data not shown).

Gene Expression Variance Related to Age and Sex

Different statistics were employed for examinations on age- and sex-dependent gene expression changes, shown in Fig. 1 and Table 1. The left side of Fig. 1 reflects male (n = 112) and female (n = 88) gene expression values examined in all samples without considering age. Corresponding *P* values are provided on the left side of Table 1. We also calculated the gene expression for males and females per age category (right side of Fig. 1). By employing a paired *t* test we examined for sex-dependent gene expression changes per age category and over all age categories (*P* values are shown on the left side of Table 1). A *P* trend (first-order-linear regression) over all age categories was calculated for each sex to examine for an age dependency separately for each sex (*P* values are shown on the right side of Table 1). Our results fell into four categories, as shown below.

Age- and Sex-Independent Gene Expression Changes

Neither the expression of *WNT3* nor *WLS* showed a significant association with sex or age of the donors (Fig. 1; Table 1). Thus, the fold differences in gene expression between sexes (females served as the reference) were 1.03 (*WNT3*) and 1.23 (*WLS* (Table 1), respectively. Age categories did not influence these values for sex differences (Table 1). Also, insignificant gene expression changes of both genes (linear regression, Wald $P = 0.2$–0.5, Table 1) were observed, resulting in FC of 1.1–1.4 over the age groups (Table 1).

Sex-Dependent Gene Expression Changes

Six of the genes (*DAGLA*, *DDB2*, *FDXR*, *POU2AF1*, *PRF1* and *PLA2G16*) showed a significant association with the sex of the donor (Student's *t* test, $P \leq 0.001$, left side of Fig. 1, Table 1). When considering age groups, four genes (*DAGLA*, *DDB2*, *PRF1* and *PLA2G16*) showed a significant sex dependency (paired *t* test, $P = 0.02$–0.04). The fold difference in gene expression between the sexes ranged between 1.2–1.6 (Table 1, left side). All six genes showed insignificant age-dependent gene expression changes for both sexes (Table 1, right side) and the FC did not exceed 1.4 over all age categories for most of the genes, except *POU2AF1* (FC of 1.7, Table 1).

Age-Dependent Gene Expression Changes

CD177 showed no significant association with sex (Fig. 1, left side). The fold difference in gene expression was 0.80 (Table 1). This was neither significant without considering age groups (Student's *t* test, $P = 0.24$) nor when considering age groups (paired *t* test, $P = 0.54$, Table 1). Significant gene expression changes in *CD177* with age (linear regression, Wald $P = 0.03$) was found only in males with a negative slope of –0.28, resulting in a FC of 1.7 over all age categories (Table 1). Females showed no age-dependent gene expression changes ($P = 0.74$, Table 1).

Sex- and Age-Dependent Gene Expression Changes

A significant sex- and age-dependent association of gene expression was found for *STAT4* (left side, Fig. 1, Table 1). The fold difference in gene expression among both sexes was 1.41 (Table 1). This difference in gene expression was significant without considering age groups ($P \leq 0.001$) and remained statistically significant when age groups were considered ($P = 0.01$, Table 1). Also, a significant gene expression change with age was found in males (linear regression, Wald $P = 0.02$) with a FC of 1.1. Females showed no age-dependent gene expression changes ($P = 0.34$, Table 1).

Complete Gene Expression Variance

We examined the complete inter-individual variance in gene expression reflected by the difference between the (normalized) lowest (used as the reference) and the highest Ct values (min–max Ct values) and also the variance in gene expression of contributing factors such as age and sex (see above). Seven genes (*POU2AF1*, *PRF1*, *PLA2G16*, *STAT4*, *DAGLA*, *FDXR* and *DDB2*) showed min–max Ct differences ranging between 3–4 (8–16-fold difference in gene expression on a linear scale, Table 2). Two genes (*WLS* and *WNT3*) showed higher difference in Ct values ranging between 5–6 (32–64-fold difference on the linear scale). *CD177* reflected the highest inter-individual variance over all examined measurements with almost 11 min–max Ct values, converting into a 1,628-fold difference in gene expression observed between the min–max Ct measurements (Table 2).

Age- and sex-dependent gene expression differences contributed less than 20–30% to the complete inter-individual variance in all genes except *PLA2G16*; however, approximately twice as much higher age-dependent variance in gene expression was found for females compared to males (Table 2, middle part). For the gene *PLA2G16* the complete inter-individual variance was low (3.9 Ct values), and age (males, 28.3%; females 51.5%) and sex (18%) contributed approximately 70% of the observed inter-individual variance (Table 3). Thus, the residual inter-individual variance (excluding age- and sex-dependent variance) of approximately 70–80% of the complete observed inter-individual variance in gene expression originates from factors other than age or sex for all genes except *PLA2G16*.

Comparing the Human Data Set with Previous Data from Irradiated Baboons

In our previously reported analysis we used a different qRT-PCR platform to identify 29 candidate genes that were

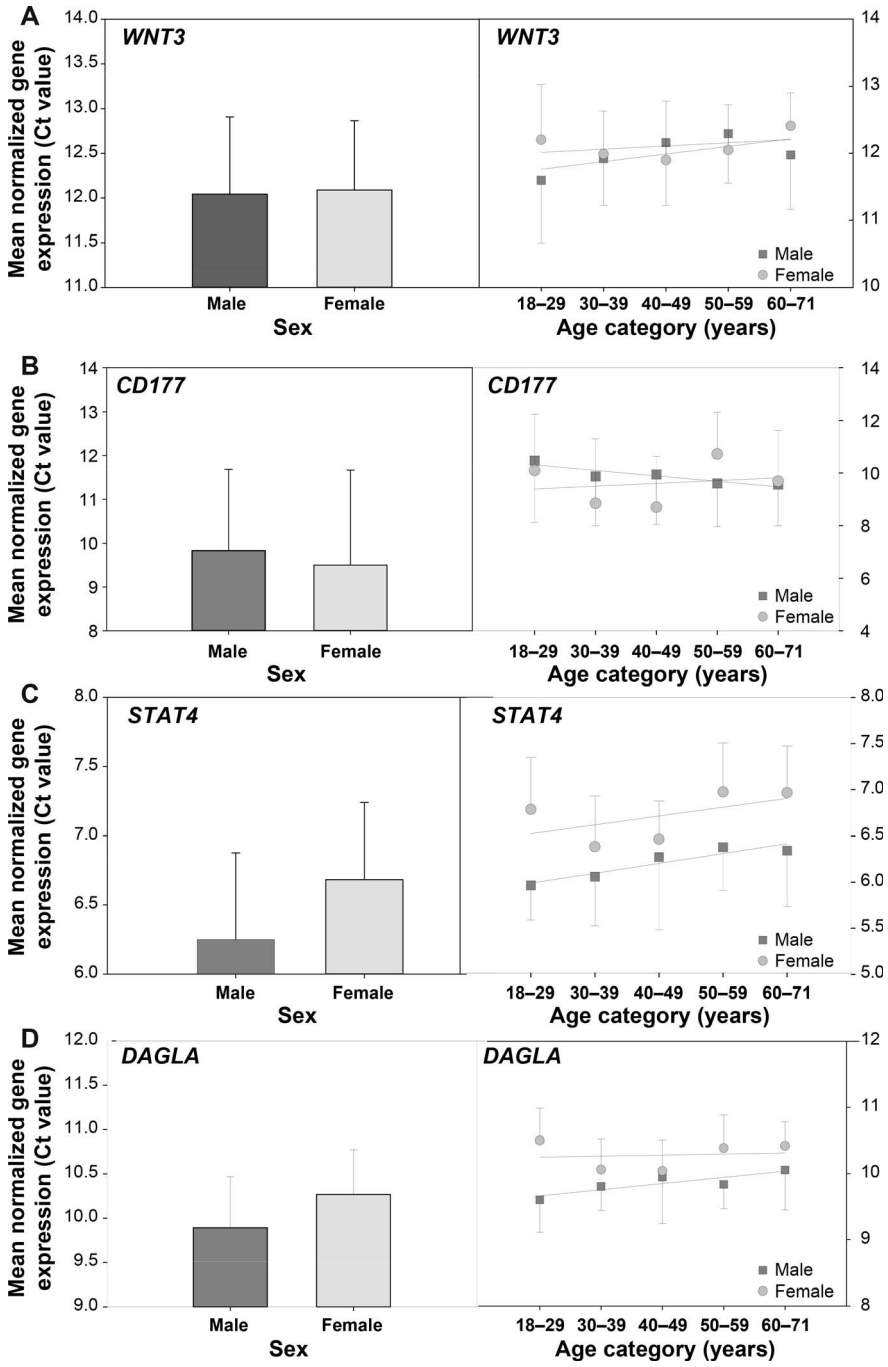

FIG. 1. Examination of age- and sex-dependent gene expression changes of eight genes. Based on age- and sex-dependent gene expression changes we identified four categories. One example of each is provided here. Panel A: *WNT3* for age- and sex-independent gene expression changes. Panel B: *CD177* for age-dependent gene expression changes only. Panel C: *STAT4* for both sex- and age-dependent gene expression changes. Panel D: *DAGLA* for sex-dependent gene expression changes only. The left side panels show comparisons of the mean normalized gene expression [normalized Ct values (ΔCt)] between males and females, without considering age. The right side panels show comparisons of gene expression between males and females over five age categories employing the paired *t* test. To examine the age dependency over all age categories, we used a first order linear regression model to fit gene expression versus age categories for each sex separately. Symbols represent mean values and error bars the standard deviation.

able to discriminate between the HARS outcome categories H1–3 and H2–3 (*8*). In the current study we examined the gene expression distribution of eight of these genes in our 200 human samples from healthy donors, representing H0 HARS (no HARS). To compare the human and baboon data sets we shifted the observed gene expression distribution to reach the same mean Ct values for each data set without altering the gene expression distribution of both H0 groups.

TABLE 1
Overview of Statistical Values to Examine the Effect of Age and Gender on Baseline Gene Expression Measurements of Eight Genes

	Gender dependency			Age dependency		
	Not considering age		Considering age groups			
Gene ID	*P* value	Fold difference	*P* value	Gender	*P* value	Slope
Age- and sex-independent gene expression changes						
WNT3	0.695	1.03	0.532	Males	0.21	0.11
				Females	0.53	0.04
WLS	0.099	1.23	0.35	Males	0.15	0.10
				Females	0.47	0.11
Sex-dependent gene expression changes						
DAGLA	1.0E–03	1.37	0.04	Males	0.05	0.09
				Females	0.85	0.02
DDB2	1.0E–03	1.39	0.02	Males	0.09	0.08
				Females	0.81	0.02
FDXR	1.0E–03	1.29	0.10	Males	0.06	0.05
				Females	0.79	0.02
POU2AF1	9.23E–03	1.19	0.10	Males	0.06	0.16
				Females	0.40	0.08
PRF1	1.0E–03	1.37	0.02	Males	0.22	0.06
				Females	0.39	0.08
PLA2G16	1.3E–10	1.60	2.00E–03	Males	0.80	0.01
				Females	0.59	0.03
Age-dependent gene expression changes						
CD177	0.25	0.80	0.54	Males	0.03	–0.28
				Females	0.75	0.10
Sex- and age-dependent gene expression changes						
STAT4	1.0E–03	1.41	0.01	Males	0.02	0.10
				Females	0.34	0.09

Notes. On the left side of the table, sex-dependent gene expression changes are shown, with or without consideration of age. Genes are categorized into four subgroups depending on their age and sex-related gene expression dependency (described in the subheaders). Sex-dependent gene expression changes without considering age were examined using either the *t* test or Mann-Whitney rank sum test where applicable, but the paired *t* test was applied when age categories were considered. Fold difference in gene expression was calculated using females as reference. The right side shows age-dependent gene expression changes separately for each sex, which were investigated using linear regression to determine Wald *P* value and the slope.

As a result, the data behaved as if all samples were generated using a low-density array (data not shown).

For a better comparison of an overlap in gene expression between H0 and clinical outcome groups (H1–3 or H2–3 HARS), we plotted the converted qRT-PCR and microarray baboon data as a scatter plot on the left side of a graph (baboon model) and on the right side of the graph we replaced the original H0 baboon data with the human H0 data and compared it with the same baboon clinical outcome groups for each gene (human-baboon model, Fig. 2). We also employed logistic regression analysis and calculated the area under the ROC curve (Table 3).

Substituting the much larger human data set for the original baboon baseline data allowed us to apply a much larger range for the distribution of values in unexposed samples (Fig. 2). Comparisons between the baboon model and the human-baboon model revealed marginal alterations (second–third digit after the decimal point) in the area under the ROC curve, resulting in complete or almost complete separation of both HARS outcome groups with ROC values ranging between 0.96–1.0 (Table 3). Overlapping gene expression values of a few unexposed human samples with the baboon HARS outcome groups still resulted in high overall accuracy (ROC $\geq$ 0.96) given the high number (n = 200) of unexposed human samples considered by the ROC calculation. When focusing on a similar discrimination ability of the diseased (exposed) baboons in comparison with either the baboon H0 or the human H0 data set (Fig. 2), we identified four genes behaving in that manner, namely *WNT3*, *DAGLA*, *CD177* and *WLS* (Fig. 2, upper part). For the remaining four genes the larger variance of the H0 human data set no longer agreed with the discrimination based on irradiated and unirradiated animals. This difference was caused by interference from approximately 5% of the H0 human gene expression measurements (Table 3). In other words, 95% (Fig. 2, horizontal bar) of the H0 human data set overlapped with the gene expression distribution of the H0 baboon data set; thus, interference by approximately 5% of the human H0 data set caused a slightly reduced discrimination ability of unexposed (H0) versus diseased (exposed) cases (Table 3 and Fig. 2).

DISCUSSION

In previously published studies we identified 29 candidate genes for the early prediction of later developing HARS

TABLE 2
Contribution of Age and Gender on the Completely Observed Inter-Individual Variance of Eight Genes

	CD177	WLS	WNT3	POU2AF1	PRF1	PLA2G16	STAT4	DAGLA	FDXR	DDB2
Complete inter-individual variance										
Min	5.2	5.2	9.5	6.9	3.1	6.9	4.9	8.5	9.4	7.5
Max	15.9	10.9	14.8	11.0	7.1	10.8	8.4	11.9	12.7	10.7
Max–min	10.7	5.7	5.2	4.1	4.1	3.9	3.5	3.4	3.3	3.2
Fold change	1,627.8	51.4	37.1	17.0	16.7	14.7	11.5	10.7	9.9	9.4
Percentage	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Age dependent variance, males										
Min	9.2	6.7	11.8	7.9	4.5	9.1	6.0	9.7	10.5	8.4
Max	10.0	6.8	12.2	8.4	4.5	10.2	6.2	9.7	10.5	8.5
Max–min	0.8	0.1	0.4	0.5	0.0	1.1	0.2	0.0	0.0	0.1
Fold change	1.7	1.1	1.3	1.4	1.0	2.1	1.1	1.0	1.0	1.1
Percentage	7.5	1.8	7.7	12.2	0.0	28.3	5.7	0.0	0.0	3.1
Age dependent variance, females										
Min	8.1	7.0	11.7	7.9	4.8	8.1	6.4	10.1	10.6	8.8
Max	9.8	7.5	12.1	8.7	5.3	10.1	6.9	10.4	11.0	9.1
Max–min	1.7	0.5	0.4	0.8	0.5	2	0.5	0.3	0.4	0.3
Fold change	3.2	1.4	1.3	1.7	1.4	4.0	1.4	1.2	1.3	1.2
Percentage	15.9	8.8	7.7	19.6	12.3	51.5	14.2	8.8	12.1	9.3
Gender-dependent variance (reference, females)										
Min	9.0	6.8	11.9	8.2	4.5	9.1	6.1	9.7	10.5	8.4
Max	9.7	7.1	12.03	8.3	5.0	9.8	6.6	10.2	10.8	9.0
Max–min	0.7	0.3	0.1	0.1	0.5	0.7	0.5	0.5	0.3	0.6
Fold change	1.6	1.2	1.1	1.1	1.4	1.6	1.4	1.4	1.2	1.5
Percentage	6.7	5.3	2.5	2.4	12.3	18.0	14.2	14.6	9.1	18.6
Residual inter-individual variance (excluding age and gender variance)										
Residual Ct value	8.3	4.9	4.7	3.2	3.1	1.2	2.5	2.6	2.6	2.3
Percentage	77.4	85.9	89.8	78.0	75.4	30.4	71.6	76.6	78.9	72.1

Notes. This table reflects the completely observed inter-individual gene expression variance of 10 genes, calculated as the difference between the lowest (reference) and the highest Ct value (min–max Ct value) for each gene. The min–max Ct value differences are delogged and converted into a fold change (FC) on a linear scale and are set to 100%. Corresponding values related to age and sex are given below. The contribution of both factors to the inter-individual variance is calculated in percentage of the corresponding min–max Ct values to the complete inter-individual variance. The residual inter-individual variance was calculated after subtraction of age (female values) and sex min–max Ct values and the residual inter-individual variance is finally given in percentage to the complete inter-individual variance Ct values.

using 18 irradiated baboons. The most promising eight candidate genes were further validated in appropriate clinical models (*11*). For the next validation phase, in the current study we sought to address the effect of inter-individual gene expression variance on the discrimination of diseased and healthy baboons. We also questioned the applicability of our baboon candidate genes for radiation-exposed humans. Exposures to humans are likely in the context of possible radiological or nuclear events (e.g., nuclear power plant disasters, dirty bomb). We established the gene expression baseline for ten selected ionizing radiation responsive genes in 200 healthy (unexposed) human donor whole blood samples using qRT-PCR with a particular focus on age- and sex-dependent changes in gene expression. In previous studies, it has been discussed that factors such as sex, age, body mass index and smoking might play a role in the inter-individual variation in the gene expression of ionizing radiation-responsive genes (*17–19*). These studies usually depended on microarray data using low sample numbers. To our surprise, and although sometimes significant, age and sex showed $\leq$1.7-fold changes in gene expression over the examined age categories and between the sexes and only in some of the genes studied here. These FCs fall within the twofold change in gene expression over control and are, therefore, considered to represent control values.

We also explored the contribution of age and sex gene expression variance to the complete inter-individual variance for each gene, which did not exceed 20–30% for all genes except *PLA2G16*. Also, the age dependency in females appeared approximately two times higher compared to males. These values were calculated based on mean gene expression values per age category and even decreased to less than 12% contribution on the complete inter-individual variance in linear regression models with age employed on a continuous scale (data not shown). While we can only speculate about the reason for these sex- and age-dependent differences in gene expression, it is well known that both sexes differ in their metabolic and hormonal status, which results in a sex bias concerning immunological and inflammatory diseases (*20*).

It is of particular interest that at least 70–80% of the complete inter-individual variance is caused by parameters other than age and sex. We received blood samples from healthy donors via the Bavarian Red Cross, who examines the blood for certain diseases and accepts only blood samples from healthy donors. Nevertheless, donors could suffer from chronic diseases and lifestyle differences

TABLE 3
Impact of the Human H0 Data Set for Discrimination from Exposed (Diseased) Baboons

				Logistic regression analysis				Ct values	
Gene name	Model	Clinical outcome	n [H1(2)–3/H0]	OR	95% CI	ChiSq	ROC	Baboons	Human, 95% percentile
WNT3	Baboon	H2–3 vs. H0	20/17		Complete separation		1.0	15.5	15.5
	Baboon-human	H2–3 vs. H0	20/200		Complete separation		1.0		
POU2AF1	Baboon	H2–3 vs. H0	17/15		Complete separation		1.0	16.6	16.4
	Baboon-human	H2–3 vs. H0	17/200	106.1	6.0–>999.9	0.0015	0.999		
DAGLA	Baboon	H2–3 vs. H0	7/16	0.001	<0.001–2.567	0.09	0.991	16.2	16.6
	Baboon-human	H2–3 vs. H0	7/200	0.004	<0.001–0.173	0.002	0.991		
STAT4	Baboon	H1–3 vs. H0	14/15		Inconclusive	nd	0.995	16.8	17.0
	Baboon-human	H1–3 vs. H0	14/200	46.4	7.6–282.5	<0.0001	0.98		
PLA2G16	Baboon	H1–3 vs. H0	15/20		Complete separation		1.0	15.6	16.1
	Baboon-human	H1–3 vs. H0	15/200	30.8	7.4–128.5	<0.0001	0.977		
PRF1	Baboon	H1–3 vs. H0	17/21	17.9	2.8–113.4	0.002	0.955	17.5	17.2
	Baboon-human	H1–3 vs. H0	17/200	22.7	6.6–78.5	<0.0001	0.959		
CD177	Baboon	H1–3 vs. H0	9/16	0.1	0.01–1.01	0.05	0.986	13.6	12.4
	Baboon-human	H1–3 vs. H0	9/200	0.057	0.007–0.47	0.008	0.991		
WLS	Baboon	H1–3 vs. H0	8/15		Inconclusive	nd	0.992	12.3	12.0
	Baboon-human	H1–3 vs. H0	8/200		Inconclusive	0.01	0.989		

Notes. To examine the impact of the human H0 data set on the discrimination of our previously examined diseased baboon data we at first reexamined the discrimination of the healthy (unexposed) baboons (H0) from the diseased (H1–3 or H2–3 category) baboons employing logistic regression analysis and calculating the odds ratio (OR), 95% confidence interval (95% CI, where applicable), the chi-square *P* value as well as the area under the ROC curve (left and middle side) for each gene. The upper row depicts the reexamined baboon model and the lower row shows the human-baboon model where we replaced the healthy (unexposed) H0 baboon data set by the healthy (unexposed) human H0 data set employing the same statistics for each gene. The right part of the table provides a comparison of min–max H0 baboon Ct values with the corresponding 95% percentile Ct values of the human H0 data set.

(smoking or drinking habits), and genetic differences might be contributing factors, but data regarding these were not available for our study.

When considering similar genetics, a much lower variability was found among genetically identical twins than among unrelated individuals (*21*). Also, the variation in gene expression patterns observed in healthy individuals was smaller than the variation observed among samples from individuals with either cancers or a bacterial infection (*22*). We were unable to find further evidence in the literature of which factors might be involved. In a final approach we compared the published gene expression distribution of the H0 baboon data set with the newly generated H0 human data set. Given the ten-times larger sample size of the human (n = 200) relative to the H0 baboon data set (n = 16–17), we were uncertain about the discrimination of the exposed from the diseased animals (Fig. 2) when using the H0 human data set. Surprisingly, although a larger variance of human H0 data was detected, this had only a marginal effect on the discrimination of the healthy human (unexposed, H0) versus the diseased (exposed) animals when calculating the area under a ROC curve (Table 3). Also, 95% of the human H0 data set behaved like the H0 baboon data set regarding the discrimination ability of the exposed (diseased animals) and only 5% of the H0 human data set data contributed to a marginally reduced discrimination ability in four (*POU2AF1*, *STAT4*, *PLA2G16* and *PRF1*) of the eight genes examined while no differences could be detected for the other four genes (Fig. 2). This was also caused by the several-fold gene expression differences of our candidate genes observed between both groups so that additional gene expression variance of the human samples could not alter the complete or almost complete discrimination results examined in the baboons. These results indicate that a misclassification [e.g., healthy, unexposed H0 individuals are considered to represent diseased individuals (H1–3) or the other way around] can be expected in only four genes and based on 5% of all gene expression measurements. Adjusting the discrimination baseline accordingly will prevent this kind of misclassification. Another phenomenon was a very high variance in gene expression of unexposed donors as observed for *CD177* (1,628-fold). We wondered whether *CD177* would still be predictive for exposed and diseased individuals. For *CD177* unexposed samples the Ct values increased while decreased Ct values were predictive for the exposed and diseased baboons (Fig. 2). Thus, the variance primarily caused by higher Ct values in unexposed baboons resulted in a limited effect on the discrimination of the exposed and diseased baboons as reflected by a human-baboon ROC = 0.991 (Table 3). Irrespective of that it can be assumed that lower gene expression variance results in an improved discrimination of unexposed over exposed and diseased categories.

Our study included two additional genes (*FDXR*, *DDB2*) widely used in *ex vivo* whole blood experiments. The demonstration of their inter-individual gene expression variance is of importance for this field of research and has been discussed elsewhere (*19, 21*). For the other candidate

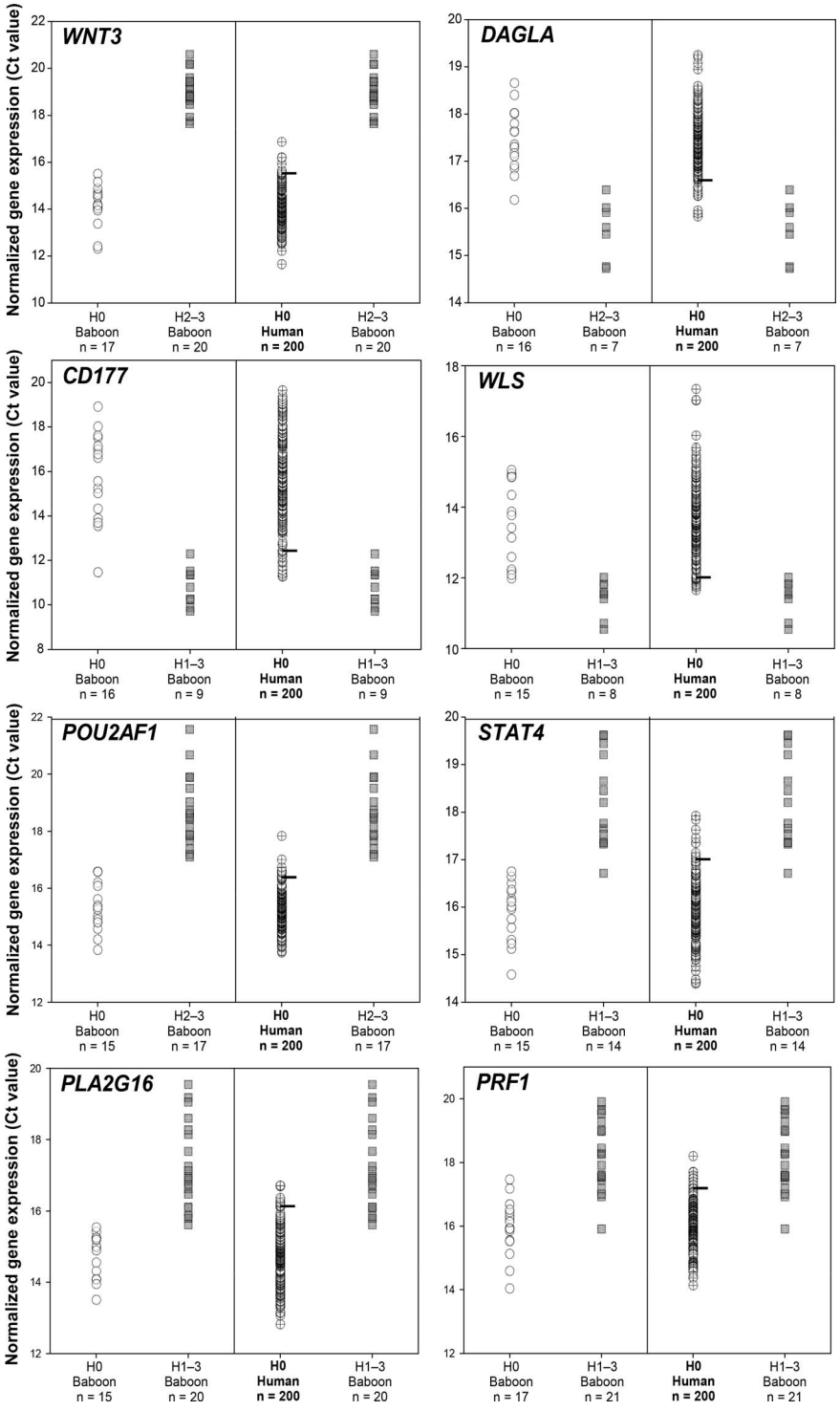

FIG. 2. Effect of increased inter-individual gene expression variance (human data set) for discrimination of clinical outcome groups. We compared the human data sets with the baboon clinical outcome groups separately for each gene. The smaller H0 baboon data set is followed by the larger human H0 data set. The corresponding distribution of gene expression values examined in diseased irradiated baboons allows a comparison on the discrimination ability using either the H0 baboon or H0 human data set. A quantitative comparison is provided in Table 3. The crosshair reflects the 95% percentile of the human H0 data set. H0 = no HARS; H1–3/H2–3 = HARS degree one to three or two to three.

genes, however, we did not find specific information in the literature.

The current study has some advantages and limitations. The CV below 1% of our qRT-PCR duplicate measurements underlines that our gene expression measurements are not heavily altered by methodological variance, but truly reflective of inter-individual variance. With the examination of 200 human samples we extended our previously reported baboon analysis about tenfold. To our knowledge, this is the largest study to systematically examine the gene expression baseline of selected genes in healthy humans. However, to compare the unexposed baboon data set (developed with a TLDA qRT-PCR platform) with the newly generated human data set of healthy donors (developed in a 96-well qRT-PCR platform) we analyzed both normally distributed data sets assuming the same mean value, but differences in the distribution of gene expression values remained. This prerequisite allowed a comparison spanning both species and different qRT-PCR technologies in the absence of completely comparable data sets.

As mentioned previously, we can only speculate about sources of inter-individual variance other than age and sex. As for age, we covered a range of 18–71 year old individuals. Thus, data on juveniles and children are missing. Also, individuals suffering from common diseases such as acute or chronic infections are missing or unknown. In particular, these two aspects must be addressed within the next validation round.

In summary, inter-individual gene expression variance is an important factor to consider when using gene expression changes to discriminate healthy (unexposed) from diseased (exposed) individuals. For the eight previously identified candidate genes neither age nor sex considerably interfered with the discrimination of either HARS group, but the residual inter-individual variance (excluding age and sex) in four genes required slight adjustments of thresholds to avoid misclassification of exposed (diseased) as unexposed (H0) and the other way around. Consequently, probabilities of our logistic regression models that do not allow for discrimination between the groups of healthy and diseased individuals will be considered as uninformative.

ACKNOWLEDGMENTS

We highly appreciate Sven Doucha-Senf, Eva Grumpelt, Thomas Mueller and Oliver Wittmann for their professional and sensitive introduction of our radiobiology masters student Sandra Agbenyegah to the technical procedures, including RNA isolation, RNA quality/quantity controls and qRT-PCR. For statistical advice we are very thankful to Herbert Braselmann from the Helmholtz Research Center for Environmental Health (GmbH), Research Unit Radiation Cytogenetics, Munich, Germany.

Received: December 18, 2017; accepted: May 24, 2018; published online: Month 00, 2018

REFERENCES

1. Fliedner TM, Graessle DH, Meineke V, Feinendegen LE. Hemopoietic response to low dose-rates of ionizing radiation shows stem cell tolerance and adaptation. Dose Response 2012; 10:644–63.
2. Gourmelon P, Benderitter M, Bertho JM, Huet C, Gorin NC, De Revel P. European consensus on the medical management of acute radiation syndrome and analysis of the radiation accidents in Belgium and Senegal. Health Phys 2010; 98:825–32.
3. Amundson SA, Grace MB, McLeland CB, Epperly MW, Yeager A, Zhan Q, et al. Human in vivo radiation-induced biomarkers. Cancer Res 2004; 64:6368–71.
4. Grace MB, McLeland CB, Blakely WF. Real-time quantitative RT-PCR assay of GADD45 gene expression changes as a biomarker for radiation biodosimetry. Int J Radiat Biol 2002; 78:1011–21.
5. Abend M, Badie C, Quintens R, Kriehuber R, Manning G, Macaeva E, et al. Examining radiation-induced in vivo and in vitro gene expression changes of the peripheral blood in different laboratories for biodosimetry purposes: First RENEB Gene Expression Study. Radiat Res 2016; 185:109–23.
6. Manning G, Macaeva E, Majewski M, Kriehuber R, Brzoska K, Abend M, et al. Comparable dose estimates of blinded whole blood samples are obtained independently of culture conditions and analytical approaches. Second RENEB gene expression study. Int J Radiat Biol 2017; 93:87–98.
7. Brzoska K, Kruszewski M. Toward the development of transcriptional biodosimetry for the identification of irradiated individuals and assessment of absorbed radiation dose. Radiat Environ Biophys 2015; 54:353–63.
8. Port M, Herodin F, Valente M, Drouet M, Lamkowski A, Majewski M, et al. First generation gene expression signature for early prediction of late occurring hematological acute radiation syndrome in baboons. Radiat Res 2016; 186:39–54.
9. Port M, Herodin F, Valente M, Drouet M, Lamkowski A, Majewski M, et al. Gene expression signature for early prediction of late occurring pancytopenia in irradiated baboons. Ann Hematol 2017;96:859–70.
10. Valente M, Denis J, Grenier N, Arvers P, Foucher B, Desangles F, et al. Revisiting biomarkers of total-body and partial-body exposure in a baboon model of irradiation. PLoS One 2015;10:e0132194.
11. Port M, Majewski M, Herodin F, Valente M, Drouet M, Forcheron F, et al. Validating baboon ex vivo and in vivo radiation-related gene expression with corresponding human data. Radiat Res 2018; 189:389–98.
12. Paul S, Amundson SA. Development of gene expression signatures for practical radiation biodosimetry. Int J Radiat Oncol 2008; 71:1236–44.e76.
13. Knops K, Boldt S, Wolkenhauer O, Kriehuber R, et al. Gene expression in low- and high-dose-irradiated human peripheral blood lymphocytes: possible applications for biodosimetry. Radiat Res 2012; 178:304–12.
14. Manning G, Kabacik S, Finnon P, Bouffler S, Badie C. High and low dose responses of transcriptional biomarkers in ex vivo X-irradiated human blood. Int J Radiat Biol 2013; 89:512–22.
15. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000; 25:25–9.
16. Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, et al. The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res 2004; 32:258D–61.
17. Abend M, Azizova T, Muller K, Dorr H, Senf S, Kreppel H, et al. Independent validation of candidate genes identified after a whole genome screening on Mayak workers exposed to prolonged occupational radiation. Radiat Res 2014; 182:299–309.
18. Eady JJ, Wortley GM, Wormstone YM, Hughes JC, Astley SB, Foxall RJ, et al. Variation in gene expression profiles of peripheral blood mononuclear cells from healthy volunteers. Physiol Genomics 2005; 22:402–11.

19. Paul S, Amundson SA. Gene expression signatures of radiation exposure in peripheral white blood cells of smokers and non-smokers. Int J Radiat Biol 2011; 87:791–801.

20. Verthelyi D. Sex hormones as immunomodulators in health and disease. Int Immunopharmacol 2001;1:983–93.

21. Correa CR, Cheung VG. Genetic variation in radiation-induced expression phenotypes. Am J Hum Genet 2004; 75:885–90.

22. Whitney AR, Diehn M, Popper SJ, Alizadeh AA, Boldrick JC, Relman DA, et al. Individuality and variation in gene expression patterns in human blood. Proc Natl Acad Sci U S A 2003; 100:1896–901.