

## **Predicting the Radiation Sensitivity of Male and Female Rhesus Macaques Using Gene Expression**

Authors: Ostheim, P., Majewski, M., Gluzman-Poltorak, Z., Vainstein, V., Basile, L. A., et al.

Source: Radiation Research, 195(1) : 25-37

Published By: Radiation Research Society

URL: <https://doi.org/10.1667/RADE-20-00161.1>

---

BioOne Complete ([complete.BioOne.org](https://complete.BioOne.org)) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at [www.bioone.org/terms-of-use](https://www.bioone.org/terms-of-use).

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

---

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

# Predicting the Radiation Sensitivity of Male and Female Rhesus Macaques Using Gene Expression

P. Ostheim,<sup>a</sup> M. Majewski,<sup>a</sup> Z. Gluzman-Poltorak,<sup>b,c</sup> V. Vainstein,<sup>b,d</sup> L. A. Basile,<sup>b</sup> A. Lamkowski,<sup>a</sup> S. Schüle,<sup>a</sup> H. L. Kaatsch,<sup>a</sup> M. Haimerl,<sup>e</sup> C. Stroszczyński,<sup>e</sup> M. Port<sup>a</sup> and M. Abend<sup>a,1</sup>

<sup>a</sup> Bundeswehr Institute of Radiobiology, Munich, Germany; <sup>b</sup> Neumedicines Inc., Pasadena, California; <sup>c</sup> Applied Stem Cell Therapeutics, Milpitas, California; <sup>d</sup> Hadassah Medical Center, Jerusalem, Israel; and <sup>e</sup> Department of Radiology, University Hospital Regensburg, Regensburg, Germany

Ostheim, P., Majewski, M., Gluzman-Poltorak, Z., Vainstein, V., Basile, L. A., Lamkowski, A., Schüle, S., Kaatsch, H. L., Haimerl, M., Stroszczyński, C., Port, M. and Abend, M. Predicting the Radiation Sensitivity of Male and Female Rhesus Macaques Using Gene Expression. *Radiat. Res.* 195, 25–37 (2021).

Radiosensitivity differs in humans and likely among closely-related primates. Reasons for variation in radiosensitivity are not well known. We examined preirradiation gene expression in peripheral blood among male and female rhesus macaques which did or did not survive (up to 60 days) after whole-body irradiation with 700 cGy (LD66/60). RNA samples originated from a blinded randomized Good Laboratory Practice study in 142 irradiated rhesus macaques. Animals were untreated (placebo), or treated using recombinant human IL-12, G-CSF or combination of the two. We evaluated gene expression in a two-phase study design where phase I was a whole genome screen [next generation sequencing (NGS)] for mRNAs (RNA-seq) using five RNA samples from untreated male and female animals per group of survivor and non-survivor (total  $n = 20$ ). Differential gene expression (DGE) was defined as a statistically significant and  $\geq 2$ -fold up- or downregulation of mRNA species and was calculated between groups of survivors and non-survivors (reference) and by gender. Altogether 659 genes were identified, but the overlapping number of differentially expressed genes (DGE) observed in both genders was small ( $n = 36$ ). Fifty-eight candidate mRNAs were chosen for independent validation in phase II using the remaining samples ( $n = 122$ ) evaluated with qRT-PCR. Among the 58 candidates, 16 were of significance or borderline significance ( $t$  test) by DGE. Univariate and multivariate logistic regression analysis and receiver operating characteristic (ROC) curve analysis further refined and identified the most outstanding validated genes and gene combinations. For untreated male macaques, we identified *EPX* ( $P = 0.005$ , ROC=1.0), *IGF2BP1* ( $P = 0.05$ , ROC=0.74) and the combination of *EPX* with *SLC22A4* ( $P = 0.03$ , ROC=0.85) which

appeared most predictive for the clinical outcome for treated and combined (untreated and treated) male macaque groups, respectively. For untreated, treated and both combined female macaque groups the same gene (*MBOAT4*,  $P = 0.0004$ , ROC = 0.81) was most predictive. Based on the probability function of the ROC curves, up to 74% of preirradiation RNA measurements predicted survival with a positive and negative predictive value ranging between 85–100% and associated odds ratios reflecting a 2–3-fold elevated risk for surviving per unit change (cycle threshold value) in gene expression. In conclusion, we identified gender-dependent genes and gene combinations in preirradiation blood samples for survival prediction after irradiation in rhesus macaques. © 2021 by Radiation Research Society

## INTRODUCTION

Ionizing radiation has been described as a two-edged sword (*1*) for its potential to cause deterministic [acute radiation syndrome (ARS)] as well as stochastic (e.g., tumors) health effects. Short- and long-term human health effects are some of the main issues in radiobiology (*2*). On the other hand, controlling tumor cell growth and minimizing radiation effects on nearby normal tissues are two major goals of radiotherapy. In this context, the term, radiation sensitivity, or “radiosensitivity”, plays an important role and describes the relative susceptibility of cells and tissues to the beneficial or harmful effect of ionizing radiation.

Understanding radiosensitivity and the individual human responses to radiation exposure with respect to tissue damage or developing radiation-related sequelae would be of benefit in several instances. For example, predictive methods to determine the radiosensitivity of patient tumors and normal tissue *a priori* would be important for personalized tumor therapy, thus widening the therapeutic window (*3–5*); astronauts traveling to Mars in 2030 would benefit from knowing their individualized radiation-related risk profile for working long-term in a space environment (*6, 7*); in radiological or nuclear emergency situations,

*Editor's note.* The online version of this article (DOI: <https://doi.org/10.1667/RADE-20-00161.1>) contains supplementary information that is available to all authorized users.

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

individuals exposed to the same magnitude of radiation might develop different degrees of the life-threatening ARS and knowing about the individual radiosensitivity could guide the treating physicians in choosing the appropriate therapy (8). Questions have remained about whether there are gender-dependent differences in radiosensitivity (9–12).

The nature of the underlying variation in radiosensitivity is poorly understood. Many published studies concerning radiosensitivity have used DNA repair and tumorigenesis as primary end points (13, 14). Cellular radiosensitivity depends on the cell cycle phase. It is well known that cells are least sensitive when in the S phase, then the G<sub>1</sub> phase, then the G<sub>2</sub> phase, and most sensitive in the M phase of the cell cycle (15). This was first described by the “law of Bergonié and Tribondeau” formulated in 1906 (16). Clonogenic survival of an irradiated cell correlates with its current cell cycle phase or the corresponding actual transcriptional status. Thus, in our opinion, the current transcriptional, e.g., functional status of a cell should also be understood in its sensitivity to ionizing radiation. This is why the end point of our current study is probably the most robust parameter in nature: cell death or, with reference to an organism, survival and non-survival. This variable as an end point becomes even more impressive when considering the current Corona pandemic: in the full range of Covid-infections from asymptomatic cases to severely ill patients, death appears to be the most robust phenotypical end point.

Early in the history of radiotherapy it was evident that individuals respond differently to radiation, reflected in the increased occurrence of acute side effects and/or of late effects such as secondary malignant diseases (13). The cause of these overreactions is most likely the insufficient repair of radiation-damaged genetic material of the cells. This can be shown by a mutated AT gene (ATM), leading to the rare clinical syndrome called “ataxia telangiectasia” and characterized by radiosensitivity of the individual (17). Because these rare individuals display a clinical phenotype, it is possible to adapt radiation treatment as appropriate. However, the great majority of radiosensitive individuals are inapparent and it is likely that radiosensitivity varies among people.

This underlines the importance of developing biomarkers/bioassays to identify highly radiosensitive individuals, e.g., with regard to radiotherapy, a long-term space flight or a large-scale radiation accident. We hypothesize that the individual transcriptional makeup of a single person at the time of the radiation exposure determines his or her radiation sensitivity (or resistance).

In collaboration with Neumedicines Inc. (Pasadena, CA), we measured gene expression in preirradiation peripheral blood samples of male and female rhesus macaques that survived or died after a whole-body irradiation to 700 cGy corresponding to a LD66/60 (37/56). We hypothesized the preirradiation transcription status of the cells would effect survival, thus, reflecting differences in radiosensitivity (identifying non-survivor) or radioresistance (identifying

survivor). For the current study, we performed a whole genome screening and identified protein coding mRNA genes by survival status and by gender. These mRNAs were then validated using qRT-PCR methodology for gene expression analysis. The rhesus macaques were either untreated (placebo) or treated using IL-12, G-CSF or a combination of both (18), which was evaluated analytically.

The main aim was to find genes predicting and identifying survivor and non-survivor and to examine the effect of gender and treatment using pre-exposure gene expression changes.

## MATERIALS AND METHODS

### *Animals and Irradiation*

Rhesus macaques were obtained from the Yongfu County Xingui Wild Animals Raising, China. The animals ranged in age from 2 years, 5 months to 3 years, 10 months and all animals weighed between 2.8 and 5.7 kg. Total body irradiation (TBI) comprised 700 cGy [60 cGy/min from a Theratron-1000 Co60 source (Best® Theratronics Ltd., Ottawa, Canada)]. Details regarding animal characteristics as well as the TBI procedure and other design details are described elsewhere (18). Non-human primates (NHP) were randomized (separated by sex; stratified by weight) and either untreated (vehicle) or treated by subcutaneously administering either recombinant human IL12 (175 ng/kg on day 1), G-CSF (10 mg/kg on days 1–18) or a combination [for more details see (18)]. Blood samples were taken during the week prior to irradiation (pre-treatment samples) as well as prior to death (days 8–35 postirradiation) or at 60 days for survivors. Those samples are referred to as post-treatment samples. Clinical signs were recorded twice a day and animals were euthanized at the occurrence of predefined criteria. On day 60, all surviving animals were euthanized. Details of clinical assessments, supportive care and euthanasia criteria are specified elsewhere (18). Regarding symptomatic palliative care, buprenorphine was given for pain; bupivacaine (0.25%) was administered topically for the management of mouth ulcers; Pepto-Bismol® was given for the management of diarrhea; topical hydrotherapy and/or iodine 1% was administered to wounds. Snacks or supplements (Rhesus Liquid Diet; Bio-Serv®, Flemington, NJ), EnsureVR (Abbott Laboratories, Abbott Park, IL), vegetables, juices or crushed cookies with banana) were given for anorexia. All procedures involving animals were conducted in compliance with the Good Laboratory Practice (GLP; 21 CFR Part 58) and performed at Citoxlab North America (Montreal, Canada) after approval by the Institutional Animal Care and Use Committee (IACUC) of Citoxlab North America, Inc. (CiToxLAB North America Study No. 1013-1033, Neumedicine Reference No. LAB-18B).

### *RNA Extraction and Quality Control*

Whole blood samples (2.5 ml) were processed following the PAXgene™ Blood RNA system (BD Diagnostics, PreAnalytiX GmbH, Hombrechtikon, Switzerland). In brief, blood was drawn into PAXgene blood RNA tubes at Citoxlab. The tubes were gently inverted (10 times), stored at room temperature overnight, then at –20°C. After all samples were collected, the PaxGene tubes were sent to Germany for further processing. After thawing, washing and centrifugation, cells in the supernatant were lysed (Proteinase K) followed by addition of lysis/binding solution taken from the mirVana™ kit (Life Technologies, Darmstadt, Germany). With the mirVana kit, total RNA, including small RNA species, was isolated by combining a phenol-chloroform RNA precipitation with further processing using a silica membrane. After several washing procedures, DNA residuals were digested on the membrane (RNAse-free DNase

Set; Qiagen, Hilden, Germany). RNA was eluted in a collection tube and frozen at  $-80^{\circ}\text{C}$ . Quality and quantity of isolated total RNA were measured spectrophotometrically (NanoDrop; PeqLab Biotechnology, Erlangen, Germany). RNA integrity was assessed using the 2100 Agilent Bioanalyzer (Life Science Group, Penzberg, Germany) and DNA contamination was controlled by conventional PCR using an actin primer. We used only RNA specimens with a ratio of  $A_{260}/A_{280} \geq 2.0$  (NanoDrop) and RNA integrity number (RIN)  $\geq 7.5$  for RNA-seq next generation sequencing (NGS) (IMGM Laboratories, Martinsried, Germany) or RIN  $\geq 7.3$  for qRT-PCR analysis.

### Phase I Screening

Whole genome screening for differentially expressed genes (protein coding mRNAs) was performed on 20 RNA samples (four groups comprised of five male and female and surviving or not rhesus macaques,  $n = 4 \times 5$ ) among the untreated arm (Fig. 1).

RNA-seq library preparation was based on the TruSeq® Stranded mRNA HT Library Prep Kit (Illumina®, San Diego, CA) following the manufacturer's recommendations (TruSeq® Stranded mRNA Sample Preparation Guide; cat. no. RS-122-9004DOC, part no. 15031047 Rev. E. 2013). In short, RNA was fragmented followed by cDNA synthesis, 3' adenylation and adaptor ligation-comprising indices for multiplex sequencing. After library quantification using the Qubit® dsDNA HS Assay Kit (Thermo Scientific, Waltham, MA) and assessment of length distribution by means of an Agilent 2100 Bioanalyzer system (Santa Clara, CA), equimolar pooling into a final sequencing library was performed. Prior to sequencing, adapter dimers were removed using AMPure® XP Beads (Beckman Coulter®, Brea, CA) followed by additional quality control by means of library quantification and electrophoresis, as described above. Subsequently, single-end sequencing ( $1 \times 75$  bp) was performed on the Illumina NextSeq 500 sequencing platform.

For in-depth analysis of differential gene expression, we used the CLC Genomics Workbench (version 9.0.1; CLC Bio-Qiagen, Aarhus, Denmark). After import of read data into the analysis tool, sequence reads were mapped against the *Macacca mulatta* genome (Mm1\_8.0.1; National Center for Biotechnology Information (NCBI), Bethesda, MD). The incorporated edgeR Bioconductor package (19) was used for identification of DGEs. In this study, only those gene transcripts that had a call "present" in at least 60% of RNA specimens were included in the analysis of gene expression and only genes with false discovery rate (FDR)-corrected  $P$  values  $\leq 0.05$  and with fold changes  $>2$ / $<-2$  among compared groups were considered to represent a candidate gene for validation in phase II. Due to the explorative nature of this study, the non-parametric test and the low sample size, we did not correct for multiple comparisons on the screening phase I of the study. We considered multiple comparisons in the bioinformatic approach in the validation part (phase II) of our study where the number of hypotheses tested in parallel was reduced from about 32,141 (phase I) to 58 mRNAs (see below).

### Phase II: Validation of Phase I Candidate Genes Using qRT-PCR

We validated the mRNA candidate genes from phase I (screening) using the remaining RNA samples ( $n = 122$ ; Fig. 1). We used a custom low-density array (LDA; high throughput qRT-PCR platform) for quantification of the 58 candidate genes using TaqMan® chemistry (Supplementary Table S1; <https://doi.org/10.1667/RADE-20-00161.1.S1>). A 1- $\mu\text{g}$  RNA aliquot of each RNA sample was reverse transcribed using a two-step PCR protocol (High Capacity Kit). cDNA (400  $\mu\text{l}$ ; 1  $\mu\text{g}$  RNA equivalent) was mixed with 400  $\mu\text{l}$  2  $\times$  RT-PCR master mix and pipetted into the fill ports of the LDA. Cards were centrifuged twice (1,200 rpm, 1 min; Multifuge3S-R; Heraeus, Wehrheim, Germany), sealed, and transferred into the 7900 qRT-PCR instrument. The qRT-PCR was processed following the qRT-PCR protocol for 384-well LDA format.

All technical procedures for qRT-PCR were performed in accordance with standard operating procedures implemented in our laboratory in 2008 when the Bundeswehr Institute of Radiobiology became certified according to DIN EN ISO 9001/2008. All chemicals for qRT-PCR using TaqMan chemistry were provided by Life Technologies (Darmstadt, Germany). For the LDA cycle threshold (Ct) values we used the median mRNA expression on each LDA for normalization purposes.

### Bioinformatics

For differentially expressed genes from phase I we performed gene set enrichment analysis using PANTHER pathway software version 11.0 [<http://www.pantherdb.org> (20–22)]. PANTHER groups genes with similar biological function based on their annotation [reference list was the current homo sapiens gene ontology (GO) database].  $P$  values were corrected for multiple testing according to the FDR algorithm (Table 1).

Furthermore, we identified potential target genes, respectively interacting proteins, of *MBOAT4* (membrane bound O-acyltransferase domain containing 4) and *SLC22A4* (solute carrier family 22 member 4), and summarized the predicted functional associations with the Search Tool for the Retrieval of Interacting Genes/Proteins [STRING; <https://string-db.org/>, version 11.0 (23)]. The interactions include direct (physical) and indirect (functional) associations and stem from computational predictions, from knowledge transfer between organisms, and from interactions aggregated from other (primary) databases. Gene enrichment analysis for gene ontologies and pathways are shown for these potential targets. Multiple test correction displays only the best results with FDR  $< 0.05$ .

Total and specific exon reads from NGS analysis were visualized using the Integrative Genomics Viewer [<http://software.broadinstitute.org/software/igv/home>, IGV 2.8.3 (24, 25)]. Reads could be annotated to specific exon regions, covered by the TaqMan probes, and quantified separately.

### Statistical Analysis

Quantitative gene expression data of the 58 candidate genes from phase II were examined using descriptive statistics [ $n$ , mean, standard deviation (SD)] and groups were evaluated using  $t$  tests or non-parametrical Kruskal-Wallis tests, where applicable. We assessed the assumptions of normality and equal variance and, if required, utilized either the pooled (equal variance) or the Satterthwaite variant of the  $t$  test (unequal variance). Unconditional logistic regression analysis was performed for each of the independent variables (genes) of interest separately (univariate) or multivariate on the binary outcome variable where the probability modeled was survival (yes/no), with non-survival as the reference. Using logistic regression, we calculated odds ratios (OR) and 95% confidence intervals. The OR describes fold changes in risk estimation per unit change (one Ct unit) in gene expression. For instance, a twofold increase in gene copy numbers of a certain gene (one Ct unit difference) associated with an OR of 4 indicates a fourfold increased probability of survival. We also determined the area under a receiver operating characteristic (ROC) curve providing a reasonable indication of overall diagnostic accuracy. ROC areas of 1.0 indicate complete agreement between the predictive model and the binary group comparison and thus a clear distinction between the survivor and the non-survivor group. Based on the probability function of the ROC curves reflecting true positives (TP; survivor), true negatives (TN; non-survivor), false positives (FP; non-survivor identified as survivor) and false negatives (FN; survivor identified as non-survivor) for each measurement, we calculated positive predictive values [(PPV) =  $\text{TP} \times 100 / (\text{TP} + \text{FP})$ ] and negative predictive values [(NPV) =  $\text{TN} \times 100 / (\text{TN} + \text{FN})$ ]. PPV and NPV are best thought of as the clinical relevance of a test (26). PPV is the accuracy of the test among cases that test positive and indicates how seriously to take a positive result: In this study, a PPV of 100% is a



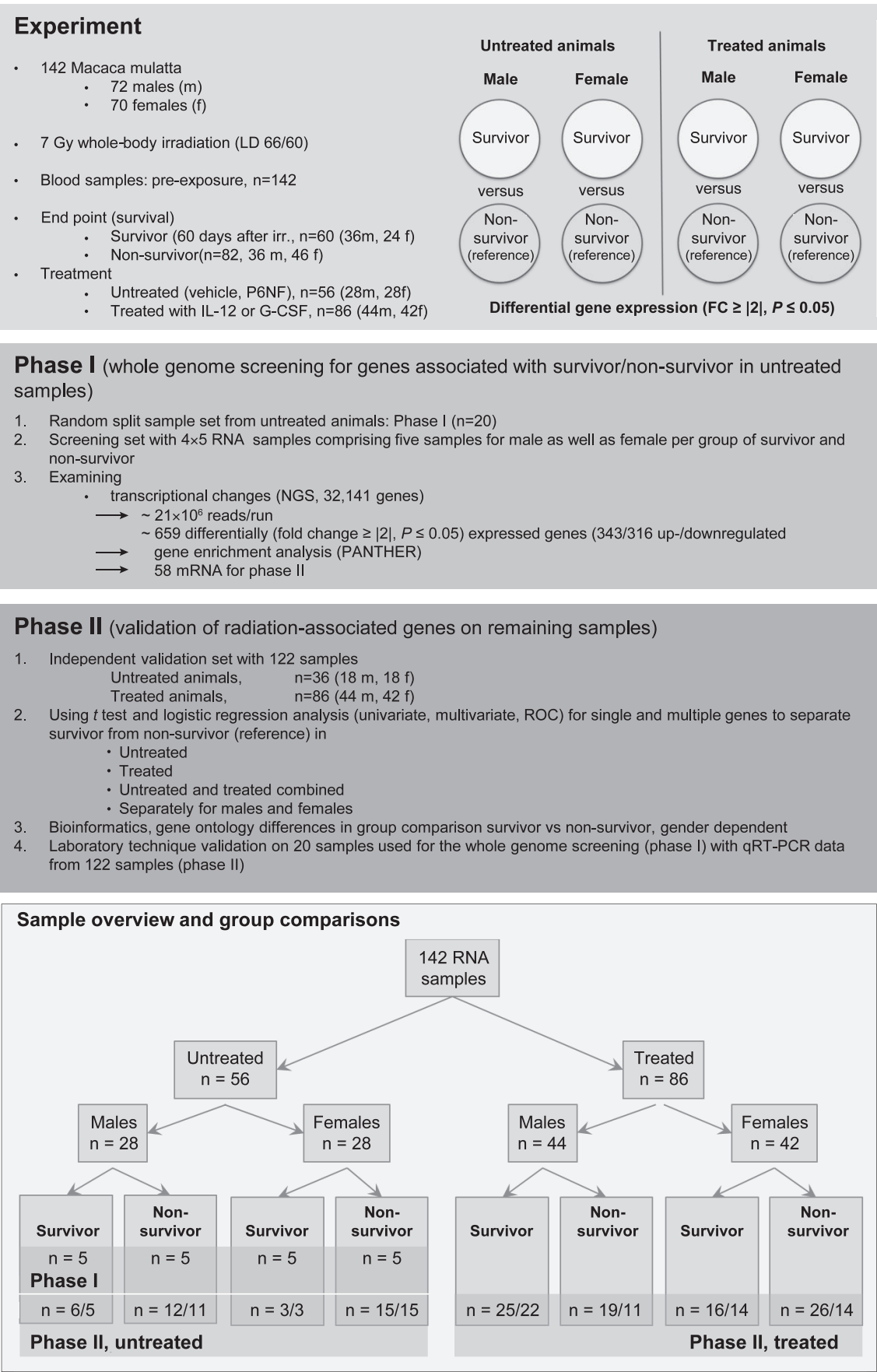


TABLE 1

Using The Differentially Expressed Genes (DEG) Found in Common (by Survival Status) from Irradiated Male and female Rhesus Macaque, a Classification of Over- and Underrepresented Gene Coding, e.g., Biological Processes or Protein Classes, was Conducted Using the PANTHER Pathway Software version 13.1 (<http://www.pantherdb.org>) which Comprises Gene Ontology (GO) Annotations Directly Imported from the GO Database

| Classification                                                 | Male            |                 | Female          |                 | Classification                                  | Male            |                 | Female          |                 |
|----------------------------------------------------------------|-----------------|-----------------|-----------------|-----------------|-------------------------------------------------|-----------------|-----------------|-----------------|-----------------|
|                                                                | Fold enrichment | <i>P</i> value* | Fold enrichment | <i>P</i> value* |                                                 | Fold enrichment | <i>P</i> value* | Fold enrichment | <i>P</i> value* |
| PANTHER protein class                                          |                 |                 |                 |                 | GO cellular component complete                  |                 |                 |                 |                 |
| <i>G-protein coupled receptor</i>                              | 5.4             | 4.4E-04         |                 |                 | <i>extracellular space</i>                      | 2.46            | 3.5E-02         |                 |                 |
| <b>receptor</b>                                                | <b>3.6</b>      | <b>1.7E-04</b>  | <b>2.15</b>     | <b>3.1E-02</b>  | <i>integral component of membrane</i>           | 1.59            | 4.3E-02         |                 |                 |
| <i>myelin protein</i>                                          |                 |                 | 21.04           | 9.8E-03         | <i>intrinsic component of membrane</i>          | 1.57            | 2.3E-02         |                 |                 |
| <i>structural protein</i>                                      |                 |                 | 5.83            | 2.6E-02         | <i>membrane part</i>                            | 1.48            | 2.4E-02         |                 |                 |
| <i>immunoglobulin superfamily cell adhesion molecule</i>       |                 |                 | 18.04           | 6.4E-03         | <i>nucleoplasm</i>                              | 0.24            | 3.7E-02         |                 |                 |
| <i>cell adhesion molecule</i>                                  |                 |                 | 3.88            | 5.7E-03         | <i>nuclear lumen</i>                            | 0.32            | 2.6E-02         |                 |                 |
| <i>membrane-bound signaling molecule</i>                       |                 |                 | 5.55            | 5.8E-03         | <i>intracellular organelle lumen</i>            | 0.33            | 2.4E-02         |                 |                 |
| <i>signaling molecule</i>                                      |                 |                 | 2.4             | 6.7E-03         | <i>intracellular part</i>                       | 0.72            | 3.2E-02         |                 |                 |
| <i>unclassified</i>                                            |                 |                 | 0.82            | 1.4E-02         | <i>organelle lumen</i>                          | 0.33            | 2.6E-02         |                 |                 |
|                                                                |                 |                 |                 |                 | <i>membrane-enclosed lumen</i>                  | 0.33            | 2.2E-02         |                 |                 |
| PANTHER GO-Slim molecular function                             |                 |                 |                 |                 | <i>organelle</i>                                | 0.67            | 2.3E-02         |                 |                 |
| <i>receptor activity</i>                                       | 2.7             | 3.3E-03         |                 |                 | <i>intracellular organelle</i>                  | 0.66            | 2.4E-02         |                 |                 |
| <i>structural constituent of myelin sheath</i>                 |                 |                 | 15.78           | 4.43E-02        | <i>nuclear part</i>                             | 0.32            | 2.3E-02         |                 |                 |
|                                                                |                 |                 |                 |                 | <i>nucleus</i>                                  | 0.48            | 2.2E-02         |                 |                 |
| GO molecular function complete                                 |                 |                 |                 |                 | <i>intracellular membrane-bounded organelle</i> | 0.59            | 2.9E-02         |                 |                 |
| <i>bradykinin receptor activity</i>                            | >100            | 9.0E-03         |                 |                 | <i>membrane-bounded organelle</i>               | 0.63            | 2.3E-02         |                 |                 |
| <i>G-protein coupled peptide receptor activity</i>             | 8.1             | 2.8E-03         |                 |                 | <b>extracellular region</b>                     | <b>2.19</b>     | <b>2.6E-02</b>  | <b>1.9</b>      | <b>3.0E-02</b>  |
| <i>peptide receptor activity</i>                               | 7.9             | 2.7E-03         |                 |                 | <i>hemoglobin complex</i>                       |                 |                 | 39.46           | 1.7E-03         |
| <i>signaling receptor activity</i>                             | 3.2             | 1.3E-04         |                 |                 | <i>extracellular matrix</i>                     |                 |                 | 3.62            | 2.4E-02         |
| <i>molecular transducer activity</i>                           | 3.1             | 9.6E-05         |                 |                 | <i>extracellular region part</i>                |                 |                 | 2.02            | 3.8E-02         |
| <i>transmembrane signaling receptor activity</i>               | 3.3             | 8.9E-05         |                 |                 | <i>plasma membrane</i>                          |                 |                 | 1.57            | 3.4E-02         |
| <i>hemoglobin binding</i>                                      |                 |                 | 47.35           | 3.1E-02         | <i>membrane</i>                                 |                 |                 | 1.31            | 3.7E-02         |
| <i>protein binding</i>                                         |                 |                 | 1.42            | 3.5E-02         | <i>cell periphery</i>                           |                 |                 | 1.55            | 3.6E-02         |
| <i>binding</i>                                                 |                 |                 | 1.25            | 9.7E-03         | <i>unclassified</i>                             |                 |                 | 0.66            | 3.7E-02         |
| <i>arachidonic acid binding</i>                                |                 |                 | 47.35           | 2.9E-02         | <i>catalytic complex</i>                        |                 |                 | 0.23            | 3.8E-02         |
| <i>icosatetraenoic acid binding</i>                            |                 |                 | 47.35           | 2.8E-02         |                                                 |                 |                 |                 |                 |
| <i>ion binding</i>                                             |                 |                 | 1.50            | 5.0E-03         | PANTHER GO-Slim biological process              |                 |                 |                 |                 |
| <i>icosanoid binding</i>                                       |                 |                 | 37.88           | 3.9E-02         | <i>regulation of catalytic activity</i>         | 3.5             | 5.6E-02         |                 |                 |
| <i>arylesterase activity</i>                                   |                 |                 | 37.88           | 3.7E-02         | <i>regulation of molecular function</i>         | 3.1             | 4.5E-02         |                 |                 |
| <i>oxygen binding</i>                                          |                 |                 | 27.06           | 1.7E-03         | <i>biological regulation</i>                    | 1.8             | 4.6E-02         |                 |                 |
| <i>oxygen carrier activity</i>                                 |                 |                 | 26.30           | 4.5E-03         | <i>regulation of biological process</i>         | 1.8             | 5.9E-02         |                 |                 |
| <i>molecular carrier activity</i>                              |                 |                 | 13.15           | 2.4E-02         |                                                 |                 |                 |                 |                 |
| <i>platelet-derived growth factor binding</i>                  |                 |                 | 22.96           | 2.6E-02         | PANTHER GO-Slim cellular component              |                 |                 |                 |                 |
| <i>carbonate dehydratase activity</i>                          |                 |                 | 19.42           | 3.2E-02         | <i>cell wall</i>                                |                 |                 | 17.22           | 3.9E-02         |
| <i>peroxidase activity</i>                                     |                 |                 | 10.18           | 4.1E-02         | <i>external encapsulating structure</i>         |                 |                 | 17.22           | 7.7E-02         |
| <i>oxidoreductase activity, acting on peroxide as acceptor</i> |                 |                 | 10.24           | 2.2E-02         |                                                 |                 |                 |                 |                 |
| <i>iron ion binding</i>                                        |                 |                 | 4.79            | <b>8.2E-03</b>  | GO biological process complete                  |                 |                 |                 |                 |
| <i>transition metal ion binding</i>                            |                 |                 | 2.35            | 5.2E-03         | <i>oxygen transport</i>                         |                 |                 | 26.3            | 1.3E-02         |
| <i>cation binding</i>                                          |                 |                 | 1.61            | 2.6E-02         | <i>gas transport</i>                            |                 |                 | 33.67           | 1.7E-05         |
| <i>unclassified</i>                                            |                 |                 | 0.64            | 3.2E-03         | <i>drug transport</i>                           |                 |                 | 7.01            | 1.2E-02         |
| <i>RNA binding</i>                                             |                 |                 | 0.23            | 3.9E-02         | <i>unclassified</i>                             |                 |                 | 0.59            | 7.4E-04         |
| <i>nucleic acid binding</i>                                    |                 |                 | 0.53            | 3.3E-02         |                                                 |                 |                 |                 |                 |

Notes. Based on the comparison of observed vs. expected numbers of up- or downregulated genes (reference database was *Macaca mullata*), a fold enrichment in the number of genes annotated to this process and a corresponding *P* value [false discovery rate (FDR)-corrected] was calculated. Bold face indicates processes that were common between male and female rhesus macaques.

prediction that 100% of individuals survive when the test indicates that they will survive. An 85–100% range for PPV and NPV was considered a “good” test in this context. Samples identified in this range were superimposed in the ROC curves for better visualization of correct predictions of the logistic regression model in predicting a certain fraction of measurements. After logistic regression analysis we identified those genes and gene combinations that were the most predictive in males and females and separately for treated, untreated and both groups combined. To decrease the number of validated genes, we included further criteria such as the number of available measurements, *P* values and the area under the ROC curve. All

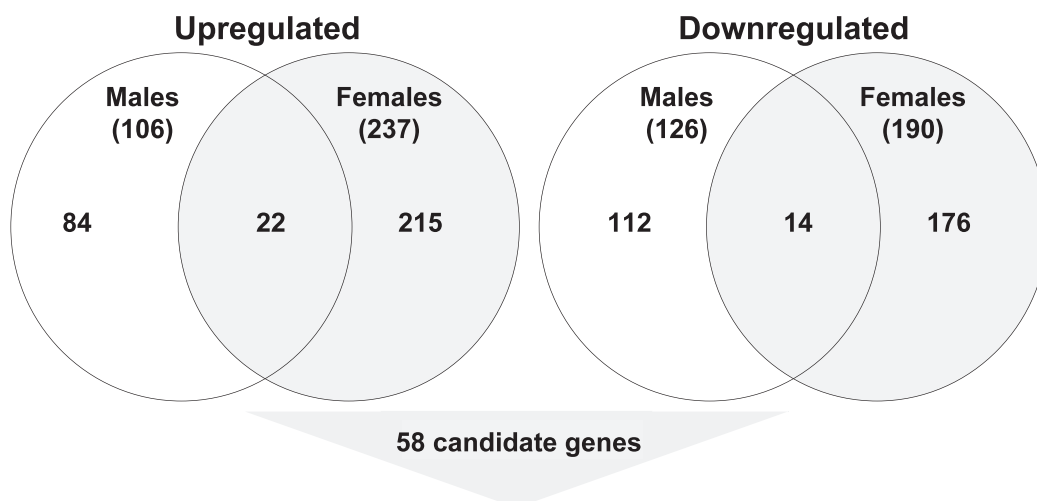
calculations were performed using SAS version 9.4 (Cary, NC) or Excel® (Microsoft® Corp., Redmond, WA).

## RESULTS

### Material Available for the Two-Phase Study Design

For this study, Neumedicines Inc. provided 142 preirradiation peripheral blood samples from 142 rhesus macaque comprising 72 samples from males and 70 samples from females (Fig. 1, upper panel).

**FIG. 1.** Flow diagram of included samples, split two-phase study design, gene expression results and bioinformatics. The overview shows the samples used in treated and untreated rhesus macaques and number of samples by survival status (survivor/non-survivor). Comparisons are shown separately for males and females in the lower panel for phase I and phase II. Numbers after the forward slash (/) (overview samples) indicate those samples ultimately included in the analysis after exclusion of samples due to methodological reasons (e.g., degraded RNA).



**FIG. 2.** Venn diagram of up- and downregulated genes that were differentially expressed by survival status (survivor/non-survivor) in male and female rhesus macaques. Overlapping number of genes that were similarly up- and downregulated in both genders are presented by the numbers inside the overlapping areas of the circles. The number of genes differentially expressed in males or females only is shown outside the overlapping area. Numbers in parenthesis represent the total number of differentially expressed genes in males and females. Based on the fold change,  $P$  value and the preference to genes differentially expressed in both genders, 58 candidate genes were selected and included in phase II for validation purposes.

During the screening approach (phase I), 20 blood samples were assessed using whole genome mRNA sequencing (RNA-seq). Those 20 samples were collected preirradiation and randomly selected from the untreated group (Fig. 1, overview upper panel and sample overview lower panel). The number of RNA samples was five for the group of survivor and non-survivor in males and females ( $n = 4 \times 5$ ).

For the validation (phase II) of mRNAs, we used the remaining 122 blood samples comprising another 36 and 86 RNA samples from the untreated and treated groups, respectively (Fig. 1, sample overview, phase II panel and lower panel). The number of 122 RNA samples available for validation was reduced because of RNA samples with low RNA amounts ( $n = 9$ ), RNA degradation ( $n = 6$ ) or undetermined signals during qRT-PCR even after repeated measurements ( $n = 12$ ), leaving 95 RNA samples eligible for phase II analysis [Fig. 1: Sample overview, numbers after the forward slash (/)]. In phase II we performed the validation in three ways in which blood samples from untreated and treated groups were examined separately and (if showing comparable results) in combination.

#### Phase I: RNA Isolation Screening Results

From 2.5 ml whole blood we isolated 6.6  $\mu\text{g}$  (SD 3.0  $\mu\text{g}$ ) total RNA on average before irradiation. RIN with a mean value of 9.1 (SD 0.4) suggested high-quality RNA sufficient for running both methods.

The quality parameter for a successful sequencing run expressed as the percentage of Q30 bases (translates into a 99.9% accuracy) is supposed to be  $>80\%$  and was on average 90.6% in our study. The average number of passed

filter (PF) reads was  $21.6 \times 10^6$  (SD  $3.8 \times 10^6$ ) and ranged between  $17.3\text{--}32.5 \times 10^6$  reads/run. In total, 87–91% of the reads mapped to known gene regions (exons and introns).

Based on our filter [fold change  $\geq |2|$ ,  $P < 0.05$ , calculated between groups of survivors and non-survivors (reference) and by gender] we identified 659 DEG (343 upregulated, 316 downregulated) with 22 and 14 of them being similarly upregulated (fold change  $\geq 2$ ,  $P < 0.05$ ) and downregulated (fold change  $\leq 0.5$ ,  $P < 0.05$ ) in both genders (Fig. 2). Bioinformatic analysis regarding the significant enrichment of genes to certain GO classifications revealed a small overlap between both genders, namely related to the protein class “receptor” and the GO cellular component called “extracellular region” (Table 1). The number of deregulated genes was comparable among males and females and approximately twice as many genes appeared downregulated relative to the upregulated genes in both genders (Fig. 2). Using the fold difference and the  $P$  value, we selected 58 most interesting candidate mRNAs for validation in phase II using qRT-PCR.

#### Phase II: Validation Using qRT-PCR Measurements

From 58 candidates, 16 appeared expressed and showed significant or borderline significant DGE in survivors relative to non-survivors in separate analysis of gender and treatment groups (Supplementary Table S2; <https://doi.org/10.1667/RADE-20-00161.1.S2>). Univariate and multivariate logistic regression analysis, ROC curve analysis as well as considering the number of measurements per group ( $\geq 60\%$ , in general) allowed for a further selection and identification of the genes and gene combinations with the most robust validation results (Table 2). The number of

**TABLE 2**  
**Overview of the Genes Selected from Supplementary Table S2 (<https://doi.org/10.1667/RADE-20-00161.1.S2>) that were Most Predictive for Discriminating Survival Status (Survivor/Non-survivor) Based on Gene Expression Differences among Both Groups**

| mRNA species               | Data source | n (survivor) | n (non-survivor) | Mean Ct value | Mean Ct value  | Fold change | P value  | Logistic regression |            |         |       |
|----------------------------|-------------|--------------|------------------|---------------|----------------|-------------|----------|---------------------|------------|---------|-------|
|                            |             |              |                  | (survivor)    | (non-survivor) |             |          | Odds ratio          | 95% CI     | P value | ROC   |
| Males                      |             |              |                  |               |                |             |          |                     |            |         |       |
| Untreated                  |             | (5/11)       |                  |               |                |             |          |                     |            |         |       |
| <i>EPX</i>                 | NGS         |              |                  |               |                | 0.1         | 0.003    |                     |            |         |       |
|                            | LDA-IA      | 5            | 4                | 12.3          | 15.7           | 10.7        | 0.005    | Complete separation |            | 1.0     |       |
| <i>SLC22A4</i>             | NGS         |              |                  |               |                | 0.14        | 0.0003   |                     |            |         |       |
|                            | LDA-IA      | 5            | 5                | 16.0          | 14.1           | 0.27        | 0.02     | nd                  | nd         | 0.34    | 0.96  |
| <i>LCN2</i>                | NGS         |              |                  |               |                | 0.21        | 0.016    |                     |            |         |       |
|                            | LDA-IA      | 5            | 5                | 15.5          | 11.3           | 0.05        | 0.03     | 2.810               | 0.51–15.6  | 0.240   | 0.92  |
| <i>MBOAT4</i>              | NGS         |              |                  |               |                | 0.86        | 0.95     |                     |            |         |       |
|                            | LDA-IA      | 5            | 4                | 14.6          | 15.7           | 2.1         | 0.3      | 0.7                 | 0.35–1.57  | 0.4     | 0.7   |
| Treated                    |             | (22/11)      |                  |               |                |             |          |                     |            |         |       |
| <i>IGF2BP1</i>             | LDA-IB      | 15           | 7                | 15.9          | 14.9           | 0.5         | 0.05     | 2.91                | 0.89–9.52  | 0.08    | 0.74  |
| <i>CHI3L1</i>              | LDA-IB      | 17           | 8                | 13.9          | 15.0           | 2.1         | 0.08     | 0.54                | 0.26–1.1   | 0.09    | 0.74  |
| <i>MBOAT4</i>              | LDA-IB      | 17           | 5                | 13.9          | 14.3           | 1.2         | 0.7      | 0.9                 | 0.51–1.61  | 0.7     | 0.6   |
| Untreated/treated combined |             | (27/22)      |                  |               |                |             |          |                     |            |         |       |
| <i>CHD5</i>                | LDA-IA+IB   | 16           | 10               | 16.6          | 15.0           | 0.3         | 0.006    | 2.610               | 1.12–6.06  | 0.026   | 0.790 |
| <i>LCN2</i>                | LDA-IA+IB   | 21           | 10               | 15.0          | 12.8           | 0.22        | 0.007    | 1.720               | 1.06–2.81  | 0.029   | 0.71  |
| <i>EPX</i>                 | LDA-IA+IB   | 23           | 11               | 12.3          | 14.2           | 3.8         | 0.01     | 0.61                | 0.40–0.93  | 0.02    | 0.77  |
| <i>SLC22A4</i>             | LDA-IA+IB   | 22           | 11               | 15.4          | 14.5           | 0.53        | 0.03     | 2.29                | 1.05–5.01  | 0.038   | 0.74  |
| <i>DYSF</i>                | LDA-IA+IB   | 22           | 12               | 13.1          | 11.5           | 0.3         | 0.04     | 1.49                | 0.99–2.23  | 0.05    | 0.65  |
| <i>CHI3L1</i>              | LDA-IA+IB   | 22           | 14               | 14.1          | 15.0           | 1.9         | 0.046    | 0.570               | 0.32–1.01  | 0.055   | 0.714 |
| <i>MBOAT4</i>              | LDA-IA+IB   | 22           | 9                | 14.1          | 14.9           | 1.7         | 0.3      | 0.8                 | 0.51–1.23  | 0.3     | 0.6   |
| <i>EPX + SLC22A4</i>       |             |              |                  |               |                |             |          | 0.64                | 0.42–0.96  | 0.03    | 0.85  |
|                            |             |              |                  |               |                |             |          | 2.4                 | 0.97–5.82  | 0.058   | 0.029 |
| Females                    |             |              |                  |               |                |             |          |                     |            |         |       |
| Untreated                  |             | (3/15)       |                  |               |                |             |          |                     |            |         |       |
| <i>EPX</i>                 | NGS         |              |                  |               |                | 0.14        | 5.9E–05  |                     |            |         |       |
|                            | LDA-IA      | 3            | 15               | 13.0          | 13.8           | 1.8         | 0.5      | 0.8                 | 0.37–1.58  | 0.5     | 0.6   |
| <i>SLC22A4</i>             | NGS         |              |                  |               |                | 1.58        | 2.75E–07 |                     |            |         |       |
|                            | LDA-IA      | 3            | 15               | 16.0          | 16.1           | 1.1         | 0.9      | 0.89                | 0.19–4.12  | 0.89    | 0.51  |
| <i>LCN2</i>                | NGS         |              |                  |               |                | 0.15        | 0.0004   |                     |            |         |       |
|                            | LDA-IA      | 3            | 15               | 15.8          | 15.2           | 0.7         | 0.6      | 1.442               | 0.44–4.75  | 0.55    | 0.578 |
| <i>MBOAT4</i>              | NGS         |              |                  |               |                | 5.4         | 0.00002  |                     |            |         |       |
|                            | LDA-IA      | 3            | 15               | 13.7          | 15.8           | 4.2         | 0.03     | 0.32                | 0.09–1.19  | 0.089   | 0.822 |
| Treated                    |             | (14/14)      |                  |               |                |             |          |                     |            |         |       |
| <i>IGF2BP1</i>             | LDA-IB      | 8            | 8                | 15.3          | 15.2           | 1.0         | 0.9      | 1.03                | 0.54–1.98  | 0.92    | 0.5   |
| <i>CHI3L1</i>              | LDA-IB      | 8            | 11               | 14.1          | 14.3           | 1.1         | 0.8      | 0.9                 | 0.46–1.77  | 0.76    | 0.59  |
| <i>MBOAT4</i>              | LDA-IB      | 9            | 12               | 13.4          | 14.9           | 2.8         | 0.02     | 0.45                | 0.21–0.977 | 0.041   | 0.75  |
| Untreated/treated combined |             | (17/30)      |                  |               |                |             |          |                     |            |         |       |
| <i>CHD5</i>                | LDA-IA+IB   | 5            | 11               | 15.6          | 15.6           | 1.0         | 0.9      | 1.05                | 0.44–2.53  | 0.91    | 0.58  |
| <i>LCN2</i>                | LDA-IA+IB   | 9            | 24               | 14.9          | 14.1           | 0.6         | 0.4      | 1.2                 | 0.81–1.77  | 0.36    | 0.59  |
| <i>EPX</i>                 | LDA-IA+IB   | 14           | 26               | 13.0          | 13.2           | 1.2         | 0.7      | 0.98                | 0.70–1.37  | 0.8979  | 0.53  |
| <i>SLC22A4</i>             | LDA-IA+IB   | 11           | 22               | 15.9          | 15.8           | 1.0         | 0.9      | 1.08                | 0.48–2.44  | 0.85    | 0.55  |
| <i>DYSF</i>                | LDA-IA+IB   | 12           | 27               | 13.3          | 12.4           | 0.5         | 0.2      | 1.27                | 0.89–1.81  | 0.196   | 0.599 |
| <i>CHI3L1</i>              | LDA-IA+IB   | 11           | 26               | 14.0          | 14.2           | 1.1         | 0.7      | 0.91                | 0.54–1.51  | 0.71    | 0.54  |
| <i>MBOAT4</i>              | LDA-IA+IB   | 12           | 27               | 13.5          | 15.4           | 3.7         | 0.0004   | 0.39                | 0.21–0.74  | 0.0041  | 0.81  |
| <i>EPX + SLC22A4</i>       |             |              |                  |               |                |             |          |                     |            |         |       |

*Notes.* Data are presented separately for male (upper) and female (lower) rhesus macaques and in order by untreated, treated, and untreated/treated-combined groups. The number of measurements and the mean gene expression [cycle threshold (Ct)] values are provided by survival status as well as the fold-change difference (delog of inverse log2 transformed Ct values). The *P* value of this group comparison originates from a parametric *t* test or a non-parametric Kruskal-Wallis test, where applicable. Logistic regression analysis was performed. Data are presented in standard text for the gender showing significant differences and are underlined using a rippled line for the gender for which significance was not achieved. Genes are ordered per treatment group starting with genes that were significantly associated with both genders followed by genes with increasing *P* value. The most significant genes per treatment group are shown in bold.



these genes was three (males) and one (females) for the untreated group, two (males) and one (females) for the treated group and six (males) and one (females) when merging untreated and treated groups. For untreated male macaques, we identified *EPX* ( $P = 0.005$ ,  $\text{ROC} = 1.0$ , eosinophil peroxidase) as the most predictive gene and *MBOAT4* ( $P = 0.03$ ,  $\text{ROC} = 0.82$ ) was most predictive for the females in this untreated group. In the treatment group and for males, *IGF2BP1* ( $P = 0.05$ ,  $\text{ROC} = 0.74$ , insulin like growth factor 2 mRNA binding protein 1) showed strongest results concerning ROC analysis and *MBOAT4* ( $P = 0.02$ ,  $\text{ROC} = 0.75$ ) was again most predictive for females. When merging measurements of both, untreated and treated groups, it was the combination of the genes *EPX* with *SLC22A4* ( $P = 0.03$ ,  $\text{ROC} = 0.85$ ) which appeared most predictive for survival for the male group and, again, *MBOAT4* ( $P = 0.0004$ ,  $\text{ROC} = 0.81$ ) for the female group. Graphical representation of the most predictive genes shows the potential of normalized gene expression (Ct values) for discriminating survival in untreated, treated and both groups combined (Fig. 3). However, some samples could not discriminate survival based on differences in gene expression measurements. For example, we observed similar gene expression values in survivor and non-survivor groups (e.g., *SLC22A4*, treated and untreated, Fig. 3). This was also true for *MBOAT4*, but during five comparisons: 1. screening (NGS) validation; 2. Untreated; 3. treated and 4. untreated-treated groups combined; and 5. all samples combined, gene expression values constantly and significantly discriminated a large fraction of survivors from non-survivors and the other way around (Fig. 4).

ROC curves of the genes/gene combinations with the most robust validation results for males (*EPX* and *SLC22A4*) and females (*MBOAT4*) reflect their ability to discriminate survivors and non-survivors (Fig. 5). Sensitivity, specificity as well as PPV (percentage of correctly predicted survivor) and NPV (percentage of correctly predicted non-survivor) were calculated and results are shown in Supplementary Table S3 (<https://doi.org/10.1667/RADE-20-00161.1.S3>). PPV and NPV between 85–100% were superimposed in the ROC curves for improved visualization of correct logistic regression model predictions performed in a certain fraction of measurements. Correct prediction in the number of RNA samples [TP (survivor) and TN (non-survivor)] increased from 58.8% and 57.6% for *EPX* and *SLC22A4* in separate analysis to 71% for both genes combined (Fig. 5). This model was particularly robust in predicting male survival as indicated by the large overlapping PPV region of single measurements in the corresponding graph of Fig. 5. For *MBOAT4*, 74.4% of all measurements correctly predicted survival (Fig. 5, last graph). Other than the male model, *MBOAT4* was robust for predicting female survival after radiation exposure, as indicated by the large overlapping NPV region of single measurements in the last graph of Fig. 5. Associated OR of, e.g., 0.4 for *MBOAT4*, indicate an approximately 2.6-fold

increased risk for dying per unit change (Ct value), which converts into a twofold reduction of *MBOAT4* copy numbers relative to the survivor (Table 2).

#### *Bioinformatic Analysis of Potential Target Genes of MBOAT4 and SLC22A4*

We examined the potential target genes of *MBOAT4* and *SLC22A4* for functional associations using STRING (23). With this tool, we further identified Gene enrichment analysis for Gene Ontologies and pathways from the next neighbors (Supplementary Fig. S1; <https://doi.org/10.1667/RADE-20-00161.1.S4>).

For *MBOAT4*, it appeared that most of the associated genes coded for growth hormones, appetite-regulating hormones, etc. *SLC22A4*, being a solute carrier family 22 member itself, is mainly involved in the solute transporter network. For both genes, no associations with apoptosis, cell cycle mechanisms, DNA damage repair or radiosensitivity were detected.

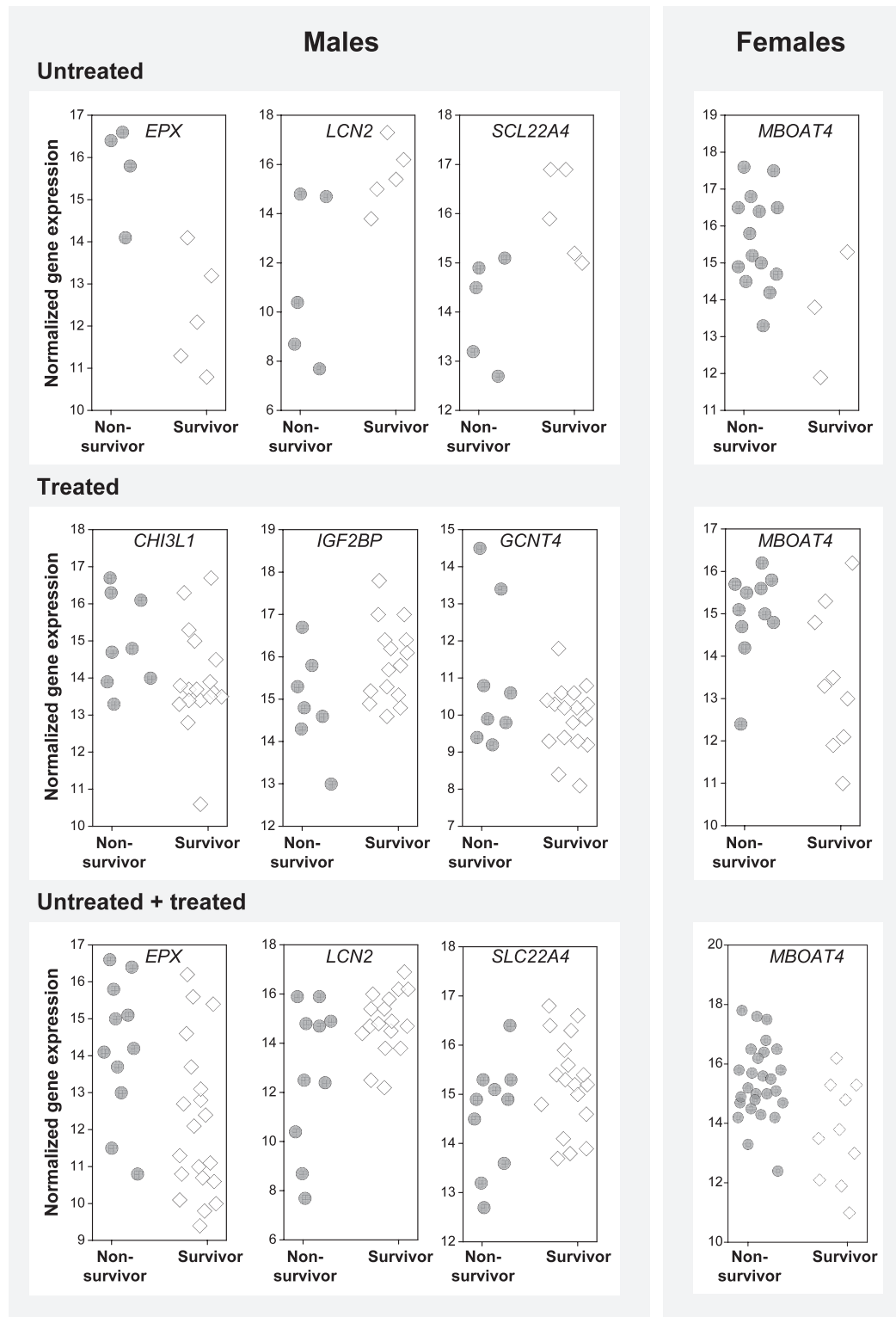
#### *Comparison of Gene Expression Measurements Using NGS and qRT-PCR*

A successful validation of NGS data using qRT-PCR was only reached when considering two aspects. 1. NGS methodology provides the total number of reads per gene, but qRT-PCR using TaqMan Chemistry can verify only reads at a certain exon region due to the intron-spanning primer probe design inherent to this technology. 2. This necessitates the identification of the gene's exon regions where radiation-induced differences in copy numbers appear, and validating it using a corresponding TaqMan assay. These prerequisites could be fulfilled for, e.g., *MBOAT4*, *LCN2*, *DYSF* and *SLC22A4*, resulting in differential gene expression values of comparable magnitude (Supplementary Fig. S2; <https://doi.org/10.1667/RADE-20-00161.1.S4>).

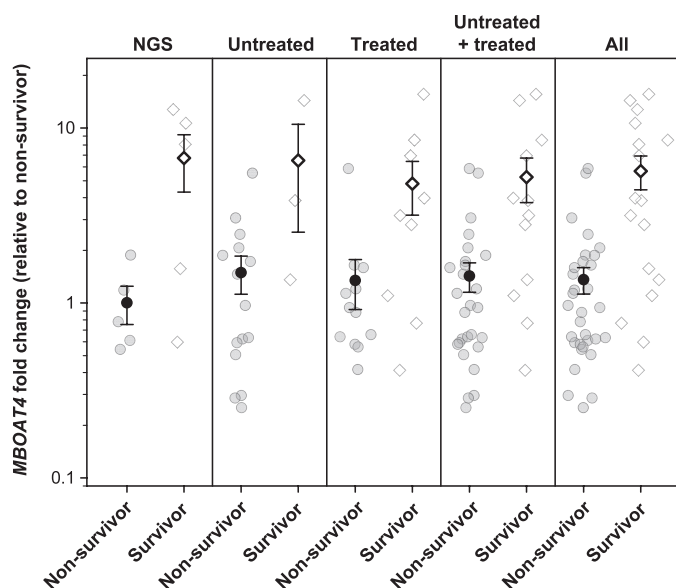
## DISCUSSION

Reasons for range of variation in radiosensitivity are not well known, but of high significance, e.g., concerning second tumors or first tumors after radiotherapy (13). Knowledge of a patient's individual radiosensitivity would be very helpful, especially in the case of exposure to moderate and high doses (e.g., radiotherapy, radio-nuclear accident). Biomarkers and bioassays for estimating individual radiation sensitivity could help to identify individuals with increased risk of short- or long-term health effects (3). For instance, predictive assays for identification of radiosensitive radiotherapy patients, astronauts or occupational radiation workers would be of great utility. Gender differences in radiosensitivity have been observed and are not entirely understood (9).

In collaboration with Neumedicine, Inc., we received 142 blood samples originating from a clinical study to examine



**FIG. 3.** The jitter plots represent the genes with the most robust validation results for discriminating survival status (survivor/non-survivor) in untreated, treated and both groups combined and separately for males (three genes on the left side) and females (one gene on the right side) using normalized gene expression values (Ct values). Each symbol represents gene expression measurements from one rhesus macaque.



**FIG. 4.** The jitter plots represent fold changes in *MBOAT4* gene expression (relative to non-survivors) by survival status (from left to right) during the screening phase (stage I) using NGS (next generation sequencing), and qRT-PCR-based independent validation (stage II) in untreated, treated, both groups combined and all study samples (stage I + II). Each symbol (without an error bar) represents gene expression measurements in one rhesus macaque [fold change =  $2^{\Delta(\text{Ct survivor} - \text{mean Ct non-survivors})}$ ]. Symbols comprising error bars (standard error of the mean) reflect the mean per group.

the effect of IL-12 and G-CSF on the survival of the rhesus macaque after whole-body irradiation. Male and female rhesus macaques were treated in the course of this study, providing the opportunity to examine gender-specific differences in radiosensitivity/-resistance regarding the clinical end point survival. We examined gene expression using blood samples collected in rhesus macaques before irradiation to predict survival. We hypothesized that preirradiation transcriptome activity would affect clinical outcome (survivor, non-survivor), because this individual genetic activity determines, e.g., differences between cell lines known to respond differently to the same radiation dose (27).

Our study took advantage of a clinical trial that allowed an assessment of gene expression for prediction purposes among various combinations of treatment groups. We detected similar results in untreated and treated groups.

During whole genome mRNA screening, we identified 659 genes differentially expressed by survivor status. From the 58 selected candidate genes, which were independently validated using qRT-PCR on the remaining 122 samples, a combination of two genes (*EPX* and *SLC22A4*) in males and one gene (*MBOAT4*) in females appeared predictive regarding identification of survivor (radioresistant) or non-survivor (radiosensitive; Fig. 5).

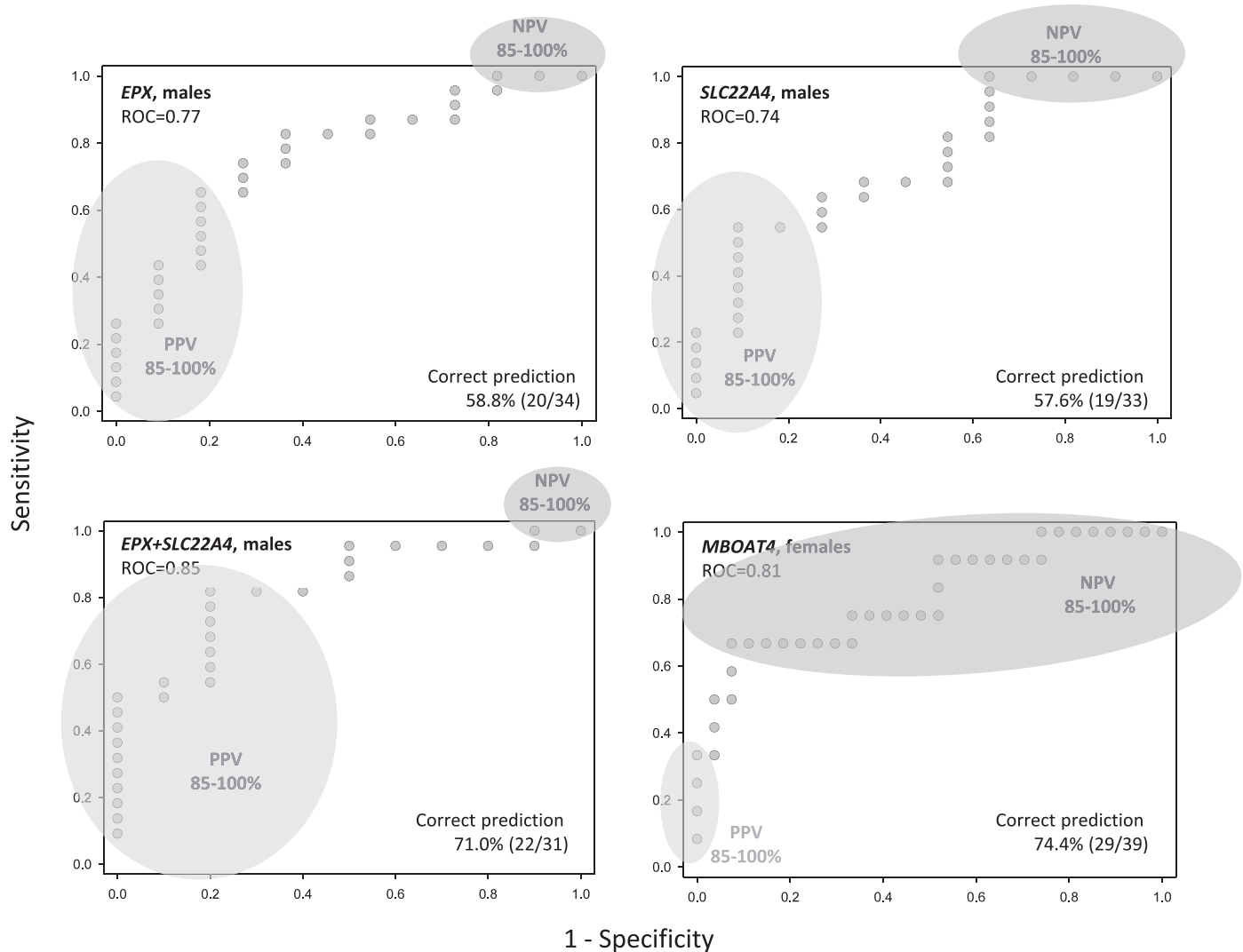
We found only a limited number of genes differentially expressed in both males and females (Fig. 2). Bioinformatic analysis on biological classifications showing enriched numbers of genes revealed a small number in common

between both genders (Table 2). Thus, different genes in both genders will be associated with radiosensitivity or radioresistance. None of these genes has been previously reported to show an association with radiation responsiveness.

Recently, we reported results of our examinations in preirradiation blood samples from irradiated baboons (27). Here, we evaluated radiosensitivity using another clinical end point, namely severity of acute radiation syndrome (ARS), because baboons were irradiated with non-lethal doses. Also, only male baboons and no females were examined in that study. *SLC22A4* appeared inversely regulated in a comparison of the two interspecies data sets (upregulated in baboons and downregulated in rhesus macaques) after irradiation (Table 2). Similarly inconsistent results were previously published (28) for a gene (*FDXR*) where it appeared to be inversely regulated after irradiation in a comparison between baboon and human data. This exemplifies the need for validation of each gene in different species. The other most outstanding and validated genes from the current study (e.g., *MBOAT4*, *EPX*, *IGF2BP1*) were not differentially expressed in the previous baboon study. This could be due to the fact that the baboon study used only males, whereas *MBOAT4* was predictive for the females only in this rhesus macaque study. With regard to the small baboon sample size, another reason might be that in the current study only a fraction of samples could discriminate survival status based on differences in gene expression measurements in nearly all candidate genes (Fig. 3).

Other than in the current rhesus macaque study, we did not find many validated candidate mRNAs in the male baboons. This may reflect the approximately 10× lower sample size (18 preirradiation peripheral blood samples) in baboons vs. 142 rhesus macaques in our current study. However, even our validated genes in rhesus macaques sometimes showed similar gene expression values by survival status (Figs. 3 and 4), rendering discrimination of clinical outcome difficult, suggesting that larger sample sizes may be needed. Due to sample size limitations, only a certain fraction of the samples was predictive with a 85–100% certainty (PPV, NPV) in our models (Fig. 5). For instance, based on *MBOAT4* gene expression changes in female survivor and non-survivor, 74.4% (29/39) of NHP could be correctly identified with an NPV and PPV lying between 85–100%. On the contrary, no prediction could be made for the remaining approximately 26% of irradiated female NHP.

Because a smaller number of rhesus macaques were untreated ( $n = 56$ ) vs. treated ( $n = 86$ ) (Fig. 1), a separate analysis was needed with a concomitant loss of power. The increased number of validated genes (from 3 to 9 in males; Table 2) and the lowered  $P$  values (e.g., from 0.03 to 0.0004 for *MBOAT4* in females; Table 2) as the sample size increased in groups of untreated, treated and both groups combined clearly demonstrates this. Among females,



**FIG. 5.** Receiver operating characteristic (ROC) curves were built using logistic regression analysis for the most predictive genes and gene combinations (*EPX* and *SLC22A4*) in male and (*MBOAT4*) in female rhesus macaques. Each dot represents gene measurements of one rhesus macaque reflecting sensitivity and specificity to discriminate survivor from non-survivor. The corresponding area in gray reflects survival status (survivor/non-survivor) given a negative (NPV) and positive predictive value (PPV) ranging between 85–100%.

*MBOAT4* appeared predictive over all treatment group comparisons (Fig. 4). This may suggest that *MBOAT4* could be a key gene from a clinical perspective because the relationship with survival was independent of treatment group. This is also reflected by OR lying in the same range (OR 0.32–0.45) for *MBOAT4* either in untreated, treated or both groups combined (Table 2). Translating that finding into situations with radiosensitivity known ahead of time, the exposure from, e.g., a mission to Mars becomes more significant, because treatment would not change the predicted risk. These findings suggest that host transcriptional status/activity before radiation exposure might affect clinical outcome.

To better understand pathomechanisms, we examined the gene ontology of these genes more carefully. We found *MBOAT4* is a protein-coding gene and GO annotations related to this gene include, e.g., ghrelin O-acyltransferase.

In addition to its influence on the appetite-associated hormone ghrelin, no association between *MBOAT4* and radiosensitivity/resistance has been described before. The same applies for *SLC22A4*, which is mainly associated with transporter networks. Nevertheless, following an agnostic approach and applying a specific end point of survival status that has rarely been applied previously in radiosensitivity studies, novel findings are expected. For example, *EPX* is a protein coding gene associated with eosinophil peroxidase, its activities include the oxidation of halide ions to bacteriocidal reactive oxygen species (29). We observed an upregulation of *EPX* in male survivors compared to non-survivors. Knowing that radiation also induces reactive oxygen stress (30), there could be a causal association of *EPX* and radiosensitivity caused by its influence on reactive oxygen stress.



NGS results could only be successfully validated with qRT-PCR when identifying the exon-region contributing primarily to the radiation-induced fold differences in gene expression and using a TaqMan assay that specifically covers this region (Supplementary Fig. S2; <https://doi.org/10.1667/RADE-20-00161.1.S4>). This reduced the number of successfully validated candidate genes and represents a weakness of our study, since several more of the screened genes over the few successfully validated genes might exist.

In summary, our examinations in rhesus macaques identified gender-dependent genes and gene combinations in preirradiation blood samples for predicting radiosensitivity (non-survivor) and radioresistance (survivor) after irradiation. These genes were validated in an independent study using animals and technology different from those used in the screening phase. Predictive assays might be developed to guide decisions about situations in which individuals are exposed to radiation in work-related or accidental scenarios.

## SUPPLEMENTARY INFORMATION

Table S1. Overview of the 58 selected candidate genes from phase I regarding the fold change, gender and the TaqMan assay used for performing qRT-PCR.

Table S2. Overview of all genes that were predictive for discriminating survival status based on gene expression differences. Data are presented separately for male (left side) and female (right side) rhesus macaques and ordered starting with untreated, treated, and untreated and treated groups combined. The number of measurements and the mean gene expression values (Ct values) are provided by survival status as well as the fold-change difference (delog of inverse log2 transformed Ct values). The *P* value of this group comparison originates from a parametric *t* test or a non-parametric Kruskal-Wallis test, where applicable. Data are presented in black for the gender showing significant differences and in gray for the gender where significance was not achieved. Genes are ordered per treatment group starting with genes that were significantly associated with survival in both genders followed by genes with increasing *P* value.

Table S3. Based on the probability function of the receiver operating characteristic (ROC) curves reflecting true positives (surviving), true negatives (not-surviving), false positives (not-surviving identified as surviving) and false negatives (surviving identified as not-surviving) for each measurement, a quantification of the correct/incorrect allocation of gene expression measurements/comparisons was made. The following test characteristics are shown for most predictive genes and gene combinations (*EPX* and *SLC22A*) in males and (*MBOAT4*) in females: sensitivity, 1-specificity, positive predictive value (PPV), negative predictive value (NPV). Rows with bold numbers represent the best prediction per category.

Fig. S1. Network of predicted associations for *MBOAT4* and *SLC22A4* and potential target genes, respective proteins, with additional aggregation of their coding function using their gene ontology data in the table.

Fig. S2. NGS-based fold changes recalculated by only taking reads from the area, covered by the TaqMan assay and the qRT-PCR-based fold changes. Data are shown for four of the most predictive candidate genes (from Table 2) in which differential gene expression values are of comparable magnitude (*MBOAT4*, *LCN2*, *DYSF* and *SLC22A4*). Additionally, the coverage of the TaqMan assay is rated by providing the percentage of total reads in the assay area for the specific genes. Furthermore, the total exon reads per gene and the exon reads in the probe area are shown, as well as the number of exons per gene.

## ACKNOWLEDGMENTS

The carefully performed technical assistance of Eva Grumpelt, Thomas Müller and Sven Doucha-Senf are very much appreciated. Furthermore, we appreciate the sophisticated support when working on NGS-data by Reinhard Ullmann. We thank Jamie L. Tom and Alice Trinh who facilitated handling of blood samples. This work was supported by the Biomedical Advanced Research and Development Authority (BARDA), U.S. Department of Health and Human Services (contract grant nos. HHSO100200800060C and HHSO100201100037C).

Received: July 1, 2020; accepted: September 18, 2020; published online: November 12, 2020

## REFERENCES

- Hall EJ. Radiation, the two-edged sword: Cancer risks at high and low doses. *Cancer J* 2000; 6:343–50.
- Blakely EA. Biological effects of cosmic radiation: Deterministic and stochastic. *Health Phys* 2000; 79:495–506.
- Herskind C, Talbot CJ, Kerns SL, Veldwijk MR, Rosenstein BS, West CML. Radiogenomics: A systems biology approach to understanding genetic risk factors for radiotherapy toxicity? *Cancer Lett* 2016; 382:95–109.
- Wisdom AJ, Kirsch DG. Functional assay to guide precision radiotherapy by assessing individual patient radiosensitivity. *EBioMedicine* 2019; 41:26–7.
- Moding EJ, Kastan MB, Kirsch DG. Strategies for optimizing the response of cancer and normal tissues to radiation. *Nat Rev Drug Discov* 2013; 12:526–42.
- Durante M, Cucinotta FA. Heavy ion carcinogenesis and human space exploration. *Nat Rev Cancer* 2008; 8:465–72.
- Furukawa S, Nagamatsu A, Neno M, Fujimori A, Kakinuma S, Katsube T, et al. Space radiation biology for “Living in Space.” *Biomed Res Int* 2020; 2020:4703286.
- Cullings HM, Fujita S, Funamoto S, Grant EJ, Kerr GD, Preston DL. Dose estimation for atomic bomb survivor studies: Its evolution and present status. *Radiat Res* 2006; 166:219–54.
- Narendran N, Luzhna L, Kovalchuk O. Sex difference of radiation response in occupational and accidental exposure. *Front Genet* 2019; 10:260.
- Fendler W, Malachowska B, Meghani K, Konstantinopoulos PA, Guha C, Singh VK, et al. Evolutionarily conserved serum microRNAs predict radiation-induced fatality in nonhuman primates. *Sci Transl Med* 2017; 9:eal2408.
- Holland JM, Mitchell TJ. The relationship of strain, sex, and body weight to survival following sublethal whole body X irradiation. *Radiat Res* 1976; 66:363–72.

12. Carnes BA, Gavrilova N, Grahn D. Pathology effects at radiation doses below those causing increased mortality. *Radiat Res* 2002; 158:187–94.
13. Habash M, Bohorquez LC, Kyriakou E, Kron T, Martin OA, Blyth BJ. Clinical and functional assays of radiosensitivity and radiation-induced second cancer. *Cancers (Basel)* 2017; 9:147.
14. Mozdarani H. Biological complexities in radiation carcinogenesis and cancer radiotherapy: impact of new biological paradigms. *Genes (Basel)* 2012; 3:90–114.
15. Hall EJ, Giaccia AJ. *Radiobiology for the radiologist*: 7th ed. Philadelphia: Lippincott Williams & Wilkins; 2012.
16. Bergonie J, Tribondeau L. Interpretation of some results of radiotherapy and an attempt at determining a logical technique of treatment. *Radiat Res* 1959; 11:587–8.
17. Rothblum-Oviatt C, Wright J, Lefton-Greif MA, McGrath-Morrow SA, Crawford TO, Lederman HM. Ataxia telangiectasia: A review. *Orphanet J Rare Dis* 2016; 25:159.
18. Gluzman-Poltorak Z, Vainstein V, Basile LA. Recombinant interleukin-12, but not granulocyte-colony stimulating factor, improves survival in lethally irradiated nonhuman primates in the absence of supportive care: Evidence for the development of a frontline radiation medical countermeasure. *Am J Hematol* 2014; 89:868–73.
19. Robinson MD, McCarthy DJ, Smyth GK. edgeR: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2009; 26:139–40.
20. Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, et al. PANTHER: A library of protein families and subfamilies indexed by function. *Genome Res* 2003; 13:2129–41.
21. Mi H, Dong Q, Muruganujan A, Gaudet P, Lewis S, Thomas PD. PANTHER version 7: Improved phylogenetic trees, orthologs and collaboration with the Gene Ontology Consortium. *Nucleic Acids Res* 2009; 38:D204–10.
22. Mi H, Muruganujan A, Casagrande JT, Thomas PD. Large-scale gene function analysis with the panther classification system. *Nat Protoc* 2013; 8:1551–66.
23. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019; 47:D607–13.
24. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative Genome Viewer. *Nat Biotechnol* 2011. <https://doi.org/10.1038/nbt.1754>.
25. Robinson JT, Thorvaldsdottir H, Wenger AM, Zehir A, Mesirov JP. Variant review with the integrative genomics viewer. *Cancer Res* 2017; 77:e31–4.
26. Trevethan R. Sensitivity, specificity, and predictive values: Foundations, pliabilitys, and pitfalls in research and practice. *Front Public Health* 2017; 5:307.
27. Port M, Herodin F, Valente M, Drouet M, Ullmann R, Majewski M, et al. Pre-exposure gene expression in baboons with and without pancytopenia after radiation exposure. *Int J Mol Sci* 2017; 18:541.
28. 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.
29. Silveira JS, Antunes GL, Kaiber DB, da Costa MS, Marques EP, Ferreira FS, et al. Reactive oxygen species are involved in eosinophil extracellular traps release and in airway inflammation in asthma. *J Cell Physiol* 2019; 234:23633–46.
30. Azzam EI, Jay-Gerin JP, Pain D. Ionizing radiation-induced metabolic oxidative stress and prolonged cell injury. *Cancer Lett* 2012; 327:48–60.