THE GENE-EXPRESSION ANALYSIS OF TRANSGENERATIONAL INSTABILITY IN MICE

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

André Manuel Gonçalves França Gomes

Department of Genetics

University of Leicester

2013

Abstract

André Gomes: The gene-expression analysis of transgenerational instability in mice

Transgenerational instability is characterized by genome-wide increases in mutation rates observed in the somatic and germline tissues of non-exposed offspring of irradiated parents. An increasing body of experimental evidence from animal studies suggests that the phenomenon of transgenerational instability is attributed to the yet unknown epigenetic changes induced in the germline of irradiated animals and manifested in their offspring.

To establish whether paternal irradiation could result in the transgenerational changes in gene expression, CBA/Ca and BALB/c male mice were given 1 Gy of acute X-rays and mated 10 weeks after exposure to control females from the same strain. RNA samples were extracted from kidney, liver, spleen and brain of the first-generation offspring of irradiated and control parents. They were hybridised to the NimbleGen 12x135K multiplex expression array, containing 135,054 features on the expression profiles and covering 42,575 known mouse transcripts. Highly significant changes in gene expression at the FDR < 0.05 were detected in 39 transcripts, showing the same pattern in all four tissues. The GO analysis revealed three overlapping pathways showing highly significant enrichment ($P < 10^{-6}$) – rhythmic process, circadian rhythm and DNA-dependent regulation of transcription.

Differentially expressed genes of interest were further validated by qPCR. The pattern of DNA methylation of a number of genes was established by single-molecule and conventional bisulfite sequencing. Paternal irradiation caused no differences in DNA methylation of the genes analysed. Methylation can be transmissible through many cell divisions, however, the transcripts and size of expression change varied between tissue-types suggesting a more permissive/modifiable epigenetic alteration including, perhaps, the histone code.

Acknowledgements

Statistical analysis of microarray data was performed by Prof. Yuri Dubrova using Statistica and Systat software.

Dr. Ruth Barber performed qPCR analysis of differentially expressed genes of interest.

My lab colleague, James Wilson, performed 10 primary and 10 secondary PCR reactions for the conventional bisulfite sequencing of 1 *Arntl*, 1 *Npas2*, 3 *Per1* and 3 *Per2* regions, which together, include 3 BALB/c mice.

A big special thanks to my supervisor Prof. Yuri Dubrova for all the help and advice throughout the years. Thank you for being contagiously fun and kind to everyone in and outside your lab group and allowing me to work in such a cheerful place. You have given me a clearer view of life in science. Thank you for all the support and beers over the years.

Thanks to everyone that was in lab G7 for the continued support, advice and friendship. It was my pleasure to be both your colleague and friend. You've made my student life more enjoyable and memorable.

I would also like to thank Dr. Timothy Gant, Dr. Kate Phillips and Dr. Emma Taylor from the MRC Toxicology Unit for all the teaching involving DNA microarray work.

A general thanks to everyone in the Department of Genetics, particularly Rita Neumann who provided me with advice for work involving bisulfite sequencing.

Thanks to Centro de Ciência e Tecnologia da Madeira (CITMA) and to Agência Regional para o Desenvolvimento da Investigação Tecnológica e Inovação (ARDITI) for funding my studies.

The biggest thanks go to my family who supported and motivated me during my studies. Pai, Joana, Beatriz, Mahnaz e restante família, eu adoro-vos.

List of contents

List of Tables

List of Figures

List of Abbreviations

&	and
~	approximately
≈	approximately equal to
γH2AX	phosphorylation of histone 2AX
^{3}H-dCTP	[^{3}H]-labelled deoxycytidine triphosphate
5-me C	5-methyldeoxycytidine
ATM	ataxia telangiectasia mutated
ATR	ataxia telangiectasia and Rad3-related protein
Axin^{Fu}	Axin-fused
Bmal1	Brain and muscle ARNT-like protein
BORIS	Brother of the Regulator of Imprinted Sites
BP	blood pressure
CBP	CREB-binding protein
CCG	clock-controlled genes
CCRI	Clare Chemical Research Inc.
CK1ε	Casein kinase 1ε
Clock	circadian locomotor output cycles kaput
CNS	central nervous system
c-PTIO	2-(4-carboxyphenyl)-4,4,5,5-tetramethylimidazoline-1-oxyl-3-oxide
Cry1/2	Cryptochrome ½
C_t	cycle threshold value
CTCF	CCCTC-binding factor
CXCR4	CXC chemokine receptor 4
Dbp	D-box binding protein
DEPC	diethylpyrocarbonate
DIE	death inducing effect
DMSO	dimethyl sulfoxide
DNMT1/3a/3b/3L	DNA methyltransferase 1/3a/3b/3L
dNTP	deoxyribonucleotide triphosphate
dpc	days post *coitum*
DSB	double strand breaks
DTT	dithiothreitol
E'-box	Non-canonical E-box (5' – CANNTG – 3')
E4BP4	Nuclear factor, interleukin 3 regulated (Nfil3)
ESTR	expanded simple tandem repeats
F_0	parental generation
F_1	first filial generation (F_0 x F_0)
F_2	second filial generation (F_1 x F_1)
FEO	food-entrainable oscillator
FLS	Fermentas Life Sciences
FPG	Formamidopyrimidine glycosylase
GI	genomic instability
GM10115 cells	Chinese hamster-human hybrid parental cells
Gy	Gray
H3K4me3	Histone 3 lysine 4 tri-methylation
H3K9ac	Histone 3 lysine 9 acetylation
H3K9me3	Histone 3 lysine 9 tri-methylation

H3K27ac	Histone 3 lysine 27 acetylation
H3K27me3	Histone 3 lysine 27 tri-methylation
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
Hlf	Hepatic leukemia factor
HR	homologous recombination
Hm-2	human microsatellite-2
hMLH1	human mulL homolog 1
IAP	intra-cisternal A particle
ICR	Imprinting Control Region
IL-8	interleukin-8
IR	Ionizing radiation
KO	knockout
L-Bmal1$^{-/-}$	liver Bmal1 knockout mice
LET	linear energy transfer
LINE1	Long Interspersed Element-1
LS12 cells	chromosomally unstable cells
LTR	Long Terminal Repeat
M-W test	Mann-Whitney *U* test
MAPK	Mitogen activated protein kinase
MBD	methyl-CpG binding domain protein
MeCP2	Methyl CpG binding protein 2
mm^3	cubic millimetres
MM	master mix
Ms6-hm	Minisatellite 6 hypermutable
NAD+	nicotinamide adenine dinucleotide
NEB	New England Biolabs
NEC	no enzyme control
NF-H_2O	nuclease-free water
NO	nitric oxide
NTC	no template control
NUS	NanoDrop ND1000 UV-Vis Spectrophotometer
PC	positive control
PCNA	proliferating cell nuclear antigen
Per1/2/3	Period 1/2/3
PTM	post-translational modification
p^{un}	pink-eyed unstable allele
RBE	relative biological effectiveness
Rfx1	regulatory factor X
RIBE	radiation-induced bystander effects
RIGI	radiation-induced genomic instability
RITI	radiation-induced transgenerational instability
RNApII	RNA polymerase II
RNS	reactive nitrogen species
RORα/β/γ	retinoid acid receptor-related orphan receptor α/β/γ
ROS	reactive oxygen species
RPE	retinal pigment epithelium
RRE	Rev-erb/ROR-binding element
SCN	suprachiasmatic nuclei
SINE B2	Short Interspersed Nuclear Repeat B2

SSB	single strand breaks
STC	Sample Tracking Control
Tef	Thyrotroph embryonic factor
TI	transgenerational instability
TNF-α	tumour necrosis factor alpha
tRF	time-of-day restricted feeding
U	enzyme unit or number of enzymes (not concentration)
UCSC	University of California Santa Cruz
w/v	weight per volume
WSCL	Walker Safety Cabinets Ltd.

Chapter 1:

Introduction

1. INTRODUCTION

1.1. Ionizing radiation and DNA damage

Ionizing radiation (IR) is a ubiquitous environmental agent present everywhere in nature including in the soil and atmosphere (UNSCEAR, 2000 [review]). IR is a potent genotoxic agent capable of causing DNA damage such as cross linking, nucleotide base damage, single-strand breaks (SSB) and double-strand breaks (DSB) (Goodhead, 1994). IR can affect cellular processes by causing changes in gene expression, disruption of mitochondrial processes, cell cycle arrest and cell death by apoptosis (Valerie *et al.*, 2007). As IR passes through a tissue, it deposits energy in the biological material within a cell (**Figure 1.1**). This energy is manifested as excitations and ionizations to cellular molecules that are non-randomly distributed along localized tracks (Smith *et al.*, 2003 [review]). The quantity of ionization/excitation events varies depending on the quality of the radiation. The term "linear energy transfer" (LET) is used to classify radiation quality according to the average energy emitted per length of the transversal made by the radiation particle. X- and gamma rays are classified as low-LET radiation, protons and neutrons are intermediate-LET radiation and α-particles and heavy ions are high-LET radiation (UNSCEAR, 2006 [review]). The same dose of radiation with different qualities does not produce the same biological effect: neutrons

can produce seven-times more chromosomal and chromatid breaks than an equivalent dose of X-rays on one-cell mouse embryos (Weissenborn and Streffer *et al.*, 1988).

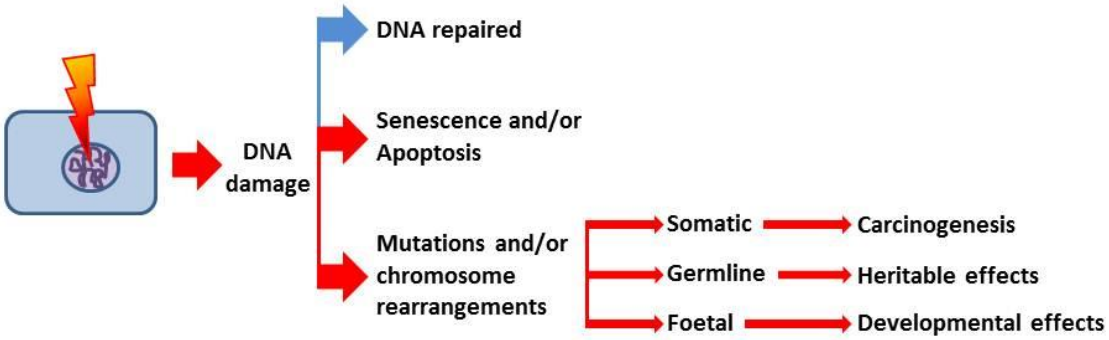

Figure 1.1: Direct effects of IR. IR passes through a cell depositing energy in its nucleus. This energy excites and ionises biological molecules causing damage. The cellular response to the damage determines whether the cell lives or dies. If the damage to DNA is small, the DNA repair machinery can recognize and fix the damage allowing survival. If the damage is overwhelming or if an error-prone repair system is activated one or more mutations may occur with or without production of clastogenic factors (defined in **section 1.6**). These in turn can trigger chromosomal rearrangements which may have different consequences in different cell types. The cellular response to IR can lead to malignant transformation in somatic cells, inheritance of genetic defects through germline cells or developmental defects in foetal cells. DNA damage can be incompatible with replication and cause cell cycle arrest at the G_1/S checkpoint. Accumulation of damage in a cell leads to aging or senescence. If the damage is extensive, cell death can occur by apoptosis, necrosis or delayed reproductive death (redrawn from UNSCEAR, 2006 [review]).

1.2. Non-targeted effects of exposure – challenging the paradigm

The main concern of radiation protection has always been the direct mutagenic effect of IR to exposed individuals. This was because the paradigm of radiation biology stated that genetic alterations were restricted to the cell that had been damaged by direct exposure to IR and that this damage, if left unrepaired, would be transmitted to all its progeny (UNSCEAR, 2001 [review]). However, a paradigm shift was caused by reports of radiation-induced effects in cells that were not directly irradiated (Lehnert and Goodwin, 1997; Zhou *et al.*, 2000; Shao *et al.*, 2004). Indeed, radiation affects non-targeted cells. Non-targeted effects of radiation lack a well-defined dose-response profile and include: (1) genomic instability occurring in the progeny of an irradiated cell – radiation-induced genomic instability (**Figure 1.2**) (reviewed by Lorimore *et al.*, 2003); (2) factors secreted from irradiated cells into the medium that, in turn, can cause the death of most clone cells – death inducing effect (Nagar *et al.*, 2003b); (3) effects manifested in cells that were not irradiated but are neighbours to cells that were irradiated or have received factors secreted from irradiated cells – bystander effects (Shao *et al.*, 2004); (4) clastogenic factors produced and released by irradiated cells into the blood plasma that can induce chromosomal damage to non-irradiated cells of animals or humans (Parsons *et al.*, 1954); and, (5) heritable effects of parental irradiation that can manifest across generations – transgenerational instability (Barber *et al.*, 2002; Shiraishi *et al.*, 2002).

It is unlikely that these non-targeted effects are directly induced by radiation. It is more likely that a currently unknown factor secreted by irradiated cells is capable of initiating these effects in non-irradiated cells or perpetuating genomic instability in the clonally expanded progeny of an irradiated cell (Morgan, 2003 [review]).

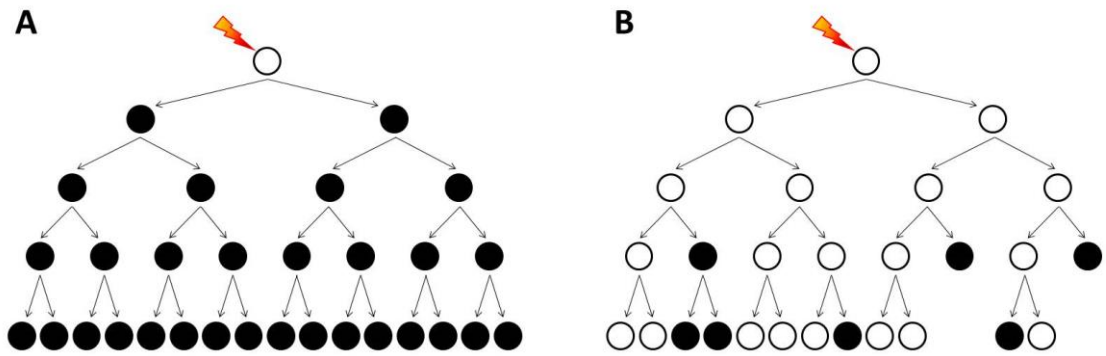

Figure 1.2: Models of response to IR by a cell and its progeny. Cells presenting genomic damage are shown as filled circles. Normal cells are shown as open circles. **(a)** Conventional paradigm of radiation exposure; Energy is deposited in the nucleus of a cell a mutation occurs and is transmitted to its clonal descendants. **(b)** Radiation-induced genomic instability; energy is deposited in the nucleus or cytoplasm of a cell and changes take place immediately following exposure. If damage is not repaired, the progeny is expected to show any transmissible radiation-induced genetic changes which can manifest themselves as multiple end-points (end-points listed on **section 1.3**). These can occur at varied clonal generations and may result in cellular death (redrawn from Lorimore *et al.*, 2003 [review]).

1.3. Radiation-induced genomic instability

Radiation-induced genomic instability (RIGI) is observed in cells at delayed times after irradiation and manifests in the progeny of exposed cells multiple generations after the initial insult (UNSCEAR, 2006 [review]). Genomic instability (GI) is measured by chromosomal rearrangements, changes in ploidy (chromosome number), gene mutations and amplifications, mini and microsatellite (tandem repeats) instability, micronucleus formation and reduced plating efficiency (late cell death) (Smith *et al.*, 2003 [review]). These are endpoints that may lead to development of diseases such as cancer (Ullrich and Ponnaiya, 1998). GI can be regarded as a marker for increased

cancer risk and can be used to predict the onset of cancer. When triggered by IR or chemicals, GI can be used as a marker for the onset of therapy-related second malignancy (Goldberg, 2003 [review]).

Unlike the direct DNA damage caused by radiation, the end-points of RIGI may not present linear dose responses. For example, there is conflicting evidence on whether the number of chromosomal aberrations, one of the end-points of RIGI, is dose-dependent (reviewed by Smith *et al.*, 2003).

RIGI can cause and be a consequence of bystander effects in the haemopoietic system where chromosomal instability was shown to produce bystander effects and vice versa (Lorimore and Wright, 2003 [review]). Radiation-induced apoptosis leads to an increased and long lasting monocyte/macrophage activity and non-specific inflammation (apoptosis is not an inflammatory process) that acts to remove dead or dying cells (Lorimore *et al.*, 2003 [review]). This is accompanied by the production of both cytokines and nitric oxide with either pro- or anti-inflammatory properties (Brune *et al.*, 1996; Giles *et al.*, 2000; Gregory, 2000 [review]). The long lasting non-specific inflammatory response could be the trigger for the onset of bystander effects and for RIGI (Neriishi *et al.*, 2001).

Chromosomal rearrangements (e.g. translocations, amplifications, numerical chromosome aberrations) could potentially cause RIGI as they involve amplification and recombination of large chromosomal regions (UNSCEAR, 2006 [review]). However, there is a study by Limoli *et al.* (1997) where treatment with four endonucleases did not cause delayed instability. It is apparent that whatever the endpoint process causing instability it must be controlled by an epigenetic mechanism. This observation is supported by Lorimore *et al.* (2003) who has concluded that the induced mutation frequency of 10^{-5} gene mutations per mammalian cell after exposure

to 1 Gy of X-rays does not correlate with studies where almost half of the clones produced by surviving cells present chromosomal instability.

1.4. The death inducing effect

Nagar *et al.* reported that growth medium from clones of GM10115 cells exhibiting radiation-induced chromosomal instability was cytotoxic to normal parental GM10115 cells. This effect was named the death inducing effect (DIE). It was observed that incubation in medium from chromosomally-unstable clones caused all parental GM10115 cells to die (Nagar *et al.*, 2003a).

In a later study it was observed that since there was micronuclei formation in surviving but chromosomally-unstable GM10115 cells there was a chance that this was a process resulting from stress-induced mitochondrial aggregation, apoptosis and release of lytic factors to the culture medium. However, GM10115 cells cultured with medium from irradiated LS12 cells also manifested DIE. LS12 cells are defective in the completion of apoptosis which means that after irradiation these cells must have secreted some factor into the culture medium (non-lytic to LS12 cells) which is responsible for the observed effect (Nagar *et al.*, 2003b).

Nagar *et al*. suggested that DIE may result from cytotoxic factors, probably cytokines, secreted by chromosomally unstable cells with the ability to cause DNA DSBs in recipient cells. These DSBs can then result in mitosis-linked cell death of the recipient cell as a result of chromosome breakage, micronucleus formation and/or apoptosis (Nagar *et al.*, 2003b). Alternatively, DIE can cause a heritable change in some of the survivors that can manifest itself, in close or distant clonal descendants, through one of the end-points of GI (Nagar and Morgan, 2005).

Curiously, DIE was observed only in normal parental cells after medium transfer from radiation-induced chromosomally unstable cells. Incubation with medium from irradiated but chromosomally stable GM10115 cells did not cause a reduction in the plating efficiency of the parental GM10115 cells. Therefore, it was concluded that irradiated GM10115 cells do not produce a bystander signal and that parental GM10115 cells are not responsive to the bystander-inducing factor. Radiation-induced DIE and radiation-induced bystander effects share a common end-point (i.e. cytotoxicity), however DIE is suggested to be triggered only by unknown factors released after the cleavage of nuclear DNA and the completion of apoptosis (Nagar *et al.*, 2003b).

There is a chance that DIE is restricted to GM10115 hybrid cells as there have been no reports of this effect in other cell lines. It is also possible that there is no such thing as a DIE and that GM10115 cells are unresponsive to factors in a bystander effect unless it can cause DNA breakage.

1.5. Radiation induced bystander effects

1.5.1. Cytoplasmic irradiation may require free radicals for bystander effect

Bystander effects have been shown to occur both *in vivo* and *in vitro* following exposure to low doses or low fluencies of IR (Azzam *et al.*, 2003 [review]). Absorbed dose represents the amount of radiation energy absorbed per unit mass of tissue (expressed in gray - Gy) and is being used to describe *in vivo* irradiation. Energy fluence represents the amount of radiation energy delivered to a unit area (expressed in millijoules per square centimetre – mJ/cm^2) and is being used to describe *in vitro*

irradiation (Sprawls, 1995 [book]; freely available at: http://www.sprawls.org/ppmi2/RADQU/).

Shao *et al.* (2004) carried out cytoplasmic helium ion microbeam irradiation of individual cells in a population composed of radioresistant glioma cells alone or in co-culture with fibroblasts. Micronuclei formation was observed in glioma cells alone and for both glioma and fibroblasts in co-culture. This proved that direct DNA damage was not essential for the onset of RIGI in the exposed cell or for occurrence of radiation-induced bystander effects (RIBE). Since cytoplasmic irradiation accomplishes mutagenicity with little killing of irradiated cells, it is considered to be more dangerous in the long-term than nuclear traversal as it allows cells to accumulate more damage (Wu *et al.*, 1999). It was also observed that nitric oxide (NO) was involved in the bystander-induced cellular damage (Shao *et al.*, 2004). Irradiation of 20% of cells resulted in a yield of micronuclei exceeding that expected by prediction when all cells are irradiated and no bystander effect is generated. Treatment of cell populations with c-PTIO, a NO scavenger, during and after microbeam irradiation reduced the formation of micronuclei in both populations of glioma and fibroblasts down to expected levels for 20% irradiated cell population. This proved that bystander damage arising from cytoplasmic irradiation was dependent on the generation of NO and that bystander effects were generated by extracellular release of a reactive nitrogen species (RNS) from the irradiated cells into the medium (Shao *et al.*, 2004). Medium transfer experiments have also observed that the first event following transfer of medium-containing cytotoxic factors to the unexposed recipient cells is a rapid (1-2 min) calcium pulse followed 30 min-2 h later by alterations in mitochondrial membrane permeability and reactive oxygen species (ROS) formation (reviewed by Lorimore *et al.*, 2003).

1.5.2. Cytokines as possible mediators of bystander effects

A study by Lehnert and Goodwin (1997) supports the involvement of ROS in mediating the bystander effects. They found that whole cell α-irradiation of human lung fibroblasts resulted in the generation of extracellular factors. Filtering the medium from irradiated cells and using it to incubate unexposed cells resulted in excessive sister chromatid exchanges (SCE) to a level comparable to those generated in directly irradiated cells, using the same radiation dose (Lehnert and Goodwin, 1997). Irradiation of the culture medium in the absence of cells caused the formation of short-lived (<30 min) SCE-inducing factors. Longer lasting (>24 h) SCE-inducing factors are generated from irradiated cells and were shown to survive freeze thawing but their action could be inhibited by exposing the medium to heat. Treatment of medium containing short or long lasting factors with superoxide dismutase resulted in an almost complete inhibition of the bystander effect. This study gives support to ROS as the soluble molecules involved in initiating bystander effects. ROS can also be accompanied by protein ligands that can increase the duration of the bystander effect (Lehnert and Goodwin, 1997). Cellular exposure to α-particles can cause an extracellular and intracellular increase in ROS that leads to a rise in interleukin-8 (IL-8) concentration in certain cell types (DeForge *et al.*, 1993). Alpha-irradiation of human lung fibroblasts in the presence of antioxidants considerably reduced extracellular IL-8 levels. This production of IL-8 occurs in parallel with elevated production of ROS and can initiate an inflammatory response (Narayanan *et al.*, 1999). This supports the observation that irradiation results in increased levels of ROS in the target cell that can induce the production of cytokines with SCE capability (e.g. TNF-α, IL-8) (Narayanan *et al.*, 1999; Azzam *et al.*, 2003 [review]). Some of these cytokines could then be released into the

medium and initiate bystander effects upon binding to receptors on non-irradiated cells (**Figure 1.3b**) (Lehnert and Goodwin, 1997; Narayanan *et al.*, 1999).

Further studies have demonstrated that exposure of cell cultures to low fluency α-particle irradiation results in a metabolic process that increases the level of superoxide anion and hydrogen peroxide in bystander cells. These ROS are involved in the activation of stress-inducible pathways such as the p53 and MAPK pathways in bystander cells (Azzam *et al.*, 2002). ROS-activated kinases (including those in the MAPK superfamily) have been shown to have a role in the activation of gap-junctions (Lampe and Lau, 2000). Gap-junctions are channels that span the membranes of two cells and thus allow intercellular communication. Its basic constituent is a protein called connexin. These proteins have different expressions in different tissues and, because there is more than one connexin, their differential assembly also gives differential selectivity to the gap-junction channel (Veenstra, 1996). Irradiation of human fibroblast cultures presenting functional or deficient gap-junctions showed that these channels play a role in transmitting damage signals from irradiated to non-irradiated cells (**Figure 1.3a**). In fibroblasts with functional gap-junctions, irradiation resulted in the activation of the p53/p21^{Waf1} stress-induced signalling pathway and micronucleus formation in bystander cells. Pre-treatment of irradiated fibroblasts with gap-junction inhibitors (e.g. lindane) resulted in the inhibition of the bystander effect (Azzam *et al.*, 2001).

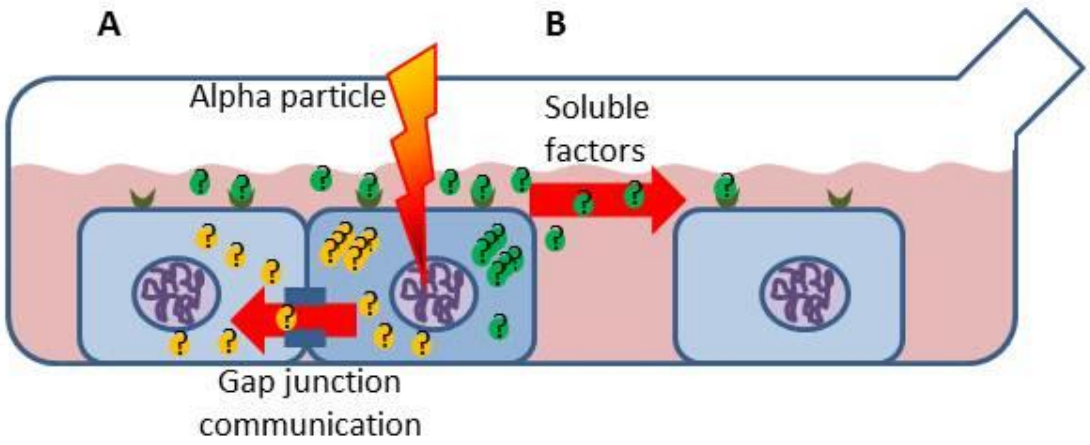

Figure 1.3: Radiation-induced factors control the bystander effect. After irradiation cells can generate cytotoxic signals that can be transmitted to other cells. Bystander effects can occur by: **(a)** passive transportation of small molecules/chemicals through gap junction channels; or, **(b)** by secretion of a soluble factor into the cell medium that can be transported to distant recipient cells (adapted from UNSCEAR, 2006 [review]).

1.5.3. Cytokines are not the only mediators of bystander effects

A study by Zhou *et al.* (2000) also supports the involvement of gap-junctions as mediators of bystander effects. Microbeam irradiation with 20 α-particles of 20% of randomly selected Chinese hamster ovary-human hybrid AL cells resulted in a mutation frequency three times higher than expected (Zhou *et al.*, 2000). Since most irradiated cells died from this lethal dose the majority of the mutations appeared in bystander cells (99.8% of progeny were from unexposed cells). Also since the class of mutations was different upon cytoplasmic and nuclear irradiation this suggested that different mechanisms of mutagenesis must be involved in both processes. Free radicals, particularly hydroxyl radicals, mediate the mutagenic process following cytoplasmic irradiation (Wu *et al.*, 1999). Treatment of cell cultures with the free radical scavenger DMSO, prior and after irradiation, did not inhibit the bystander effect (Zhou *et al.*,

2000). DMSO can scavenge hydroxyl radicals in the medium and, because DMSO is able to cross the plasma membrane, it can also scavenge inside irradiated and unexposed cells (De Ménorval *et al.,* 2012). This suggests that in AL cells the bystander effect transmitted to the unexposed cells is not mediated by hydroxyl radicals. However, there are other unknown long-lived free radicals, not scavenged by DMSO, which could be mediating the bystander effect (Koyama *et al.,* 1998). Lindane promotes the intracellular localization of Connexin 43, preventing the assembly of gap junction channels and impairing gap junction intercellular communication (Azzam *et al.,* 2001; Mograbi *et al.,* 2003). Treatment with lindane was able to reduce the bystander effect but was not able to eliminate it (Zhou *et al.,* 2000). This suggested that most of the bystander effect was mediated by gap junction communication and, to a lesser extent, by secretion of soluble factors without involving hydroxyl radicals.

1.5.4. Properties of a bystander pathway

These studies suggest that there are probably multiple bystander pathways and that the signalling pathway used depends on the genotype of the exposed organism. Antioxidants/free radical scavengers, drugs that prevent the collapse of mitochondrial membrane potential and gap-junction cell communication inhibitors can prevent cytotoxic effects on non-irradiated bystander cells (Azzam *et al.,* 2003 [review]). Another important observation is that a bystander signal can kill non-irradiated cells but cell to cell contact is not required to induce cell killing. However, not all irradiated cells can produce the toxic bystander factor and not all non-irradiated cells can receive and respond to the damage signal (UNSCEAR, 2006 [review]).

1.6. Clastogenic factors

Irradiation of animals or humans results in some irradiated cells releasing factors into the blood plasma that can cause chromosomal damage to unexposed cells, hence the name clastogenic factors. These factors were first described in 1954 when splenic irradiation of children with chronic granulocytic leukaemia resulted in sternum bone marrow damage (Parsons *et al.*, 1954). More recent studies reported the onset of chromosomal rearrangements in normal human lymphocytes that had been cultured in a medium with blood plasma (containing clastogenic factors) from an irradiated individual (**Figure 1.4**) (Morgan, 2003 [review]). Clastogenic factors have been detected in the blood plasma of victims from the Chernobyl disaster and in A-bomb survivors in Japan, 31 years after exposure (UNSCEAR, 2006 [review]). In all studies above, clastogenic factors were added, while diluted in the patient's blood plasma, into cell culture with cells from normal individuals. Then the clastogenic factors induced chromosomal damage in the normal cells which was observed by cytogenetic analysis. Clastogenic factors can also form, spontaneously, in patients with cancer-prone chromosome breakage disorders such as Ataxia telangiectasia, Bloom's syndrome and Fanconi anemia (Huang *et al.*, 2003 [review]). This raises questions as to the role of these factors on creating conditions for the onset of cancer.

It is still unknown what class of molecules clastogenic factors are but they are thought to be mediated by free radicals since superoxide dismutase, penicillamine, cysteine and other antioxidants reduce or eliminate clastogenic activity (UNSCEAR, 2006 [review]). Clastogenic activity is related to formation of lipid peroxidation products, inosine nucleotides and cytokines, all of which are suspected of mediating RIBE. Thus, clastogenic factors can be considered inducers of bystander effects (Lorimore *et al.*, 2003 [review]).

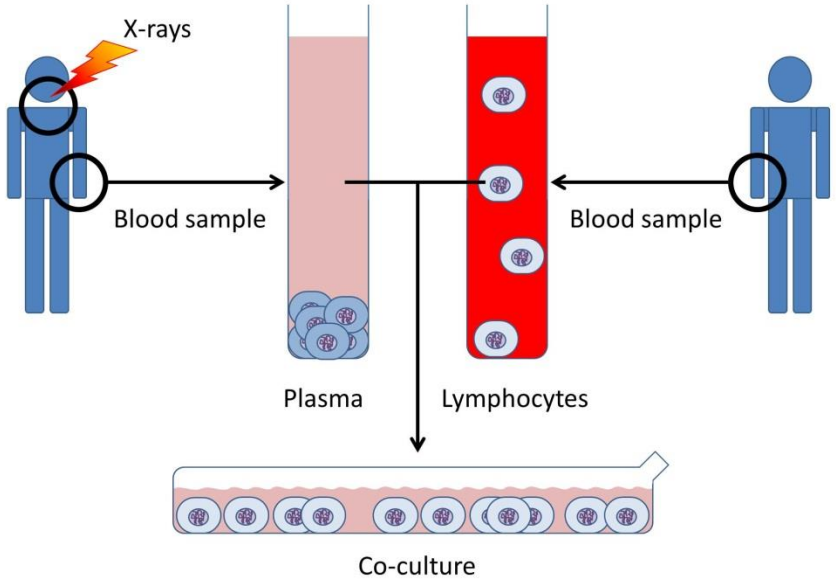

Figure 1.4: Scheme of the studies on clastogenic factors in cultures of human lymphocytes. Blood plasma from an irradiated patient can induce chromosomal rearrangements in the lymphocytes isolated from a normal individual after co-culture (adapted from Morgan, 2003 [review]).

1.7. Non-targeted and delayed effects of radiation on subsequent generations – transgenerational instability

Many *in vitro* studies provide evidence of the delayed genetic risks of radiation manifested as genomic instability in the progeny of irradiated cells (Barber and Dubrova, 2006 [review]). This instability is also present in animal models after direct or non-targeted irradiation of germ cells and may be transmitted to the next generation as determined by specific locus, dominant lethality and heritable translocation assays (Generoso *et al.*, 1980; Green *et al.*, 1987; Russell *et al.*, 1998). Indeed, unexposed offspring have shown elevated mutation rates in germline cells maintained at levels similar to those of the irradiated parent rather than those of control mice (Barber *et al.*,

2002). This suggests a multi-generational propagation of the radiation-induced damage. Somatic mutation rates are also elevated in the unexposed offspring of irradiated parents leaving them susceptible to carcinogenesis and other genetic disorders (Barber *et al.*, 2002; Barber and Dubrova, 2006 [review]).

1.7.1. Genomic instability in non-exposed offspring

The first case of transgenerational instability (TI) was demonstrated by Luning *et al.* (1976) which, upon injection of plutonium salts into male mice, observed elevated rates of dominant lethal mutations in the germline of non-exposed first filial generation (F_1) offspring. Since then, other studies have confirmed that parental irradiation can result in TI. Multiple techniques in association with a number of model systems have been used to study these effects with the most commonly used nowadays being the mouse (Barber and Dubrova, 2006 [review]). One traditional mutation scoring system in mice is the specific loci test for seven loci, also known as the Russell 7-locus test (Dubrova *et al.*, 1998).

1.7.1.1. Detection of germline mutation rate by specific locus test

The specific locus test, unlike others at the time, could directly measure germline mutation rate in the offspring of irradiated parents as the mice could survive past the embryonic stage into adulthood (Russell, 1951). This method has been designed to detect mutations at a particular set of loci in the mouse which were chosen because their alleles and corresponding mutations present visible and well known phenotypes. It consists in crossing a mouse that is homozygous for the recessive gene at the chosen loci with a wild-type mouse that is homozygous for the dominant gene. This means that the phenotype at the offspring will be imposed by the dominant allele supplied by the

wild-type mouse. Any mutation occurring at the dominant allele should cause the phenotype at the F_1 progeny to differ from the wild-type phenotype. Consequently, only mutations at the germline of wild-type mice can be detected (Russell, 1951). This method has been frequently used in radiation studies to determine spontaneous and radiation-induced mutation rates by crossing the homozygous recessive mouse with non-treated and irradiated wild-type mice, respectively (Searle, 1974). Because of the low mutation rate at these loci, this method requires enormous numbers of samples to detect any changes in germline mutation rate. The low sensitivity of this method led to the development of a new technique by Yauk *et al.* (2002) to detect germline mutation induced by IR or chemical mutagens (see **section 1.7.1.2**).

1.7.1.2. Detection of mutation rate using ESTR loci mutation induction

Highly unstable expanded simple tandem repeats (ESTR) can be used as a reporter of mutation induction in mice. Initially, ESTR mutation induced in the germline of exposed mice were analysed through the somatic DNA of their offspring (pedigree analysis) (Dubrova *et al.*, 1998). The development of single-molecule PCR allowed the detection of ESTR mutations in DNA samples extracted from sperm of the directly exposed mice. This reduced experimental time by bypassing the need to wait for mating and birth. Using this technique, C57BL/6×CBA/H F_1 hybrid males directly exposed to 1Gy of X-rays were shown to have increased ESTR mutation rates in sperm, brain and spleen tissues. This study was the first to use ESTR mutation to measure somatic mutation rates in directly irradiated mice (Yauk *et al.*, 2002). The ESTR mutation rates were consistent with those obtained under similar conditions using pedigree analysis (Barber *et al.*, 2000).

A single molecule-PCR approach accurately detects mutation in an ESTR, such as *Ms6-hm* or *Hm-2,* following amplification from a single template molecule. Mutation is detected by an increase/decrease in ESTR length (Dubrova *et al.*, 1998). This PCR is repeated many times for an irradiated mouse, or an offspring of the irradiated mouse, and repeated in many individuals to derive a mutation rate for the loci from a selected tissue (Yauk *et al.*, 2002). The ESTR loci was initially classified as a minisatellite due to its large array size and high mutation rate but its low repeat size (4-6 bp) and the homogeneous arrangement of its repeats were similar to microsatellites causing it to be re-named. ESTR loci have high mutation rates in somatic and germline tissues allowing detection of mutation induction using small samples and at doses far lower than those used with previous methods (Dubrova, 2003 [review]).

1.7.2. Properties of transgenerational instability

1.7.2.1. TI is not strain-specific or sex-specific and occurs in multiple generations

Using ESTR loci to measure mutation rate, Dubrova *et al.* (2000) showed that exposure of male CBA/H mice to fission neutrons resulted in a transgenerational increase in germline mutation rate. This effect was observed in both male and female offspring of paternally irradiated mice and provided the first evidence for the transmission of GI through the germline to non-exposed offspring. The high mutation rate in the F_2 germline was shown to be partly due to increased mutational mosaicism in the germline of F_1 mice suggesting that transgenerational destabilization must have occurred after fertilization or in the early stages of the developing F_1 germline (Dubrova *et al.*, 2000). Another study demonstrated increased ESTR mutation rates in the

offspring of irradiated CBA/H (2 Gy acute X-rays), BALB/c (1 Gy acute X-rays) and C57BL/6 (0.4 Gy chronic fission neutrons) mice. Acutely irradiated mice received the complete radiation dose at one point in time while chronically irradiated mice received the complete radiation dose distributed over a period of time. The irradiated paternal or maternal mice were mated with unexposed mice from the same strain and opposite sex. Both the F_1 and F_2 generations of all strains presented increased germline mutation rates that did not return to the spontaneous levels in unexposed controls and were similar to levels in directly exposed mice. This result demonstrated that TI was not strain-specific or sex-specific and could occur across multiple generations. However, there was an inter-strain difference in the susceptibility to TI as F_1 and F_2 offspring of BALB/c and CBA/H mice presented higher levels of ESTR mutations than those of the C57BL/6 strain (Barber *et al.*, 2002).

1.7.2.2. TI is not tissue-specific and affects all genome

Since GI could be transmitted to the germline of the F_1 offspring of irradiated males, there was also the possibility of the offspring presenting GI in somatic tissues. In a study by Barber *et al.* (2006) male CBA/Ca and BALB/c mice were irradiated with 2 and 1 Gy of acute X-rays, respectively. After breeding the irradiated mice with females of the same strain, it was noticed that ESTR mutation rates in the offspring were equally elevated in both somatic tissues (i.e. bone marrow, spleen) and germline tissues (i.e. sperm). This result suggested that TI was not tissue-specific and could occur in multiple tissues. It was also noticed that transgenerational increases in ESTR mutation rates were correlated with elevations in somatic mutation rate at protein-coding genes. An elevation of thioguanine-resistant mutations at the hypoxanthine guanine phosphoribosil (*hprt*) locus has been observed in cultured spleen cells of offspring from irradiated male

CBA/Ca and BALB/c mice. A highly significant 3.7-fold and 3.3-fold increase in the frequency of *hprt* mutations was found in these strains as a consequence of TI. This gene codes for an enzyme that plays a role in salvaging purines from fragmented DNA and is present on the X chromosome inherited from the non-exposed mothers. Thus, its mutation in the offspring of irradiated males demonstrates that both the maternal and the paternal alleles are affected by TI. This result was observed in multiple strains and demonstrated that TI affected not only inter-genic regions (e.g. ESTR) but also genic regions (**Figure 1.5**) (Barber *et al.*, 2006). Mosaicism occurs when one group of cells differs, genetically, from a neighboring group within a tissue and within an organism. Mosaics can form due to DNA mutations, mutation reversions and epigenetic changes occurring spontaneously in a cell or induced by mutagens such as radiation. Somatic mosaics are not transmissible to the offspring, but germline mosaics are. If the event leading to mosaicism occurs during development both somatic and germline cells will become mosaic (Chial, 2008 [review]). Mating irradiated CBA/H males 3 and 6 weeks after exposure (pre- and post-meiosis of spermatogenesis, respectively) and determination of paternal mutation rate allowed the identification of when the radiation-induced mutations were most likely to occur. There were no differences in mutation rate within F_1 and F_2 generation offspring formed from pre- and post-meiotically exposed sperm. Considering the increased mutational mosaicism observed in F_1 and F_2 generations, the TI is thought to occur at mitosis following fertilization or at the early stages of embryonic development (**Figure 1.5**) (Barber *et al.*, 2002).

Several techniques, in addition to the specific locus and ESTR loci mutation induction, have been used to study transmission of RIGI across generations. One example was the analysis of somatic reversions of the pink-eyed unstable allele (p^{un}) (Shiraishi *et al.*, 2002). The p^{un} mutation results from a tandem duplication of a 71 kb

region that disrupts the wild-type function of the gene. This autosomal recessive mutation causes a dilute, light colour to the coat and the retinal pigment epithelium (RPE). The p^J allele expresses the same abnormal phenotype as p^{un} but results from a simple deletion mutation. While the phenotype at the p^J allele cannot be reversed, the p^{un} gene is frequently subject to homologous recombination (HR) that can delete the duplicated region and restore the normal *p* gene producing a black pigment in the coat and RPE (Niwa, 2003 [review]). This event was explored to demonstrate that the frequency of somatic p^{un} reversions was significantly elevated in the offspring of irradiated males (Shiraishi *et al.*, 2002). Crossing one mouse homologous for the p^{un} allele with another mouse homologous for the p^J allele resulted in an offspring with the abnormal phenotype. Shiraishi *et al.*, observed that irrespective of the parent providing the p^{un} allele, irradiation of the male parent increased the frequency of p^{un} reversion in the offspring. This report gives further support to a genome-wide destabilization event occurring after fertilization.

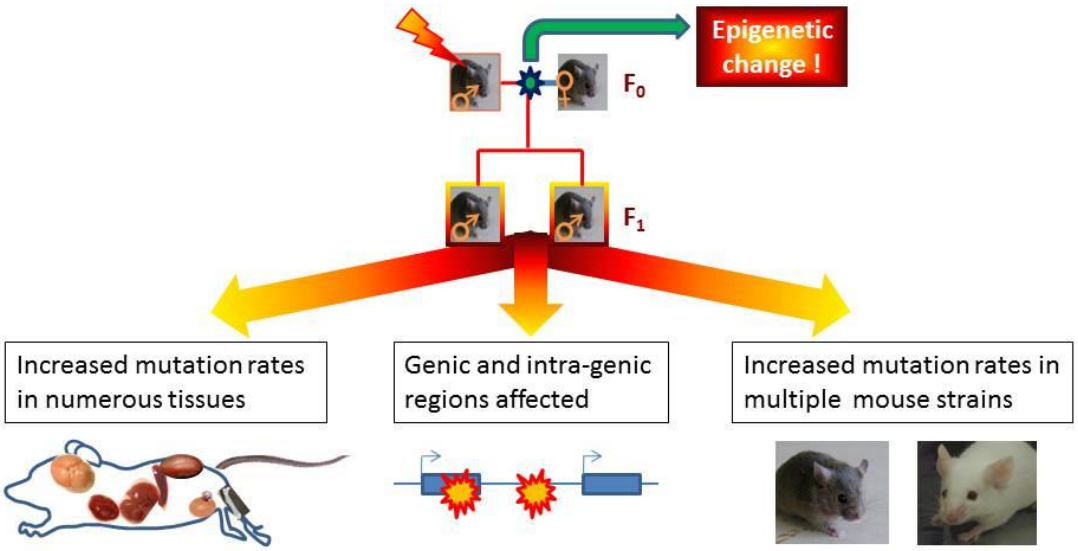

Figure 1.5: The epigenetic mechanism of transgenerational instability. See text for details (figure constructed based on Barber *et al.*, 2006; Barber *et al.*, 2009).

1.7.2.3. TI may be sex-specific when parental irradiation is *in utero*

The impact of *in utero* irradiation on ESTR mutation levels has been determined for exposed mice and offspring comparing GI formation and inheritance in male and female mice. It was observed that the frequency of ESTR mutations in the somatic and germline tissues of male and female mice irradiated at 12 days post *coitum* (dpc) was elevated during adulthood (Barber *et al.*, 2009). GI was successfully transmitted to all tissues of the F_1 offspring of *in utero* irradiated paternal mice. It was not, however, transmitted to the F_1 offspring of *in utero* irradiated maternal mice. This result suggests that RIGI is transmitted to the F_1 generation by an epigenetic signal that survives developmental reprogramming in male parents but is erased in female parents (Barber *et al.*, 2009).

1.7.2.4. TI can be caused by radiation-induced factors

Paternal irradiation of mice 6-7 weeks prior to mating results in changes to embryonic cell proliferation in the F_1 and F_2 generations. To study this effect, a mouse chimera assay was used in which an irradiated embryo was fused with a normal embryo prior to implantation. The cell proliferation changes were concluded to result from competitive cell proliferation disadvantage caused by uneven growth inhibition across the aggregation chimera (Wiley *et al.*, 1997). In a later study, the aggregation chimera was found to be salvaged from competitive cell proliferation disadvantage by inhibiting gap junction intercellular communication. Signalling between the cells of the irradiated and normal embryos was responsible for the abnormal cell proliferation. The chimera was salvaged by gap junction inhibition which, curiously, did not affect blastocyst formation and only worked when applied after the first 15 h of cell culture (after gap junction development) (Vance and Wiley, 1999).

Cranial paternal 20 Gy X-ray irradiation of rats resulted in the accumulation of DNA strand breaks and global loss of DNA methylation in the shielded mature sperm cells (Tamminga *et al.*, 2008a). It was suggested that if the sperm cell possesses genetic and epigenetic damage this would interfere with the epigenetic reprogramming after fertilization and alter the DNA methylation profile of the offspring. In the F_1 progeny, there was no increase in DNA damage but there was global loss of DNA methylation in the bone marrow, thymus and spleen tissues (but not in the liver). Hypomethylation in the bone marrow could be explained by the observed decrease in the expression of DNMT1, DNMT3a, DNMT3b and MeCP2. Since PCNA was down-regulated in the bone marrow, the changes in protein expression and hypomethylation, mentioned above, were not due to increased cellular proliferation. The expression of DNMT3b was slightly decreased in the spleen and thymus of the unexposed offspring. No significant changes in protein expression were registered in the liver of the offspring. The DNA damage and global hypomethylation seen in the distant and protected germline of the irradiated rats were caused by bystander effects. This means that bystander effects possess the ability to induce genetic and epigenetic changes in the offspring and initiate TI. The signal for the bystander effect had to travel from the irradiated brain cells and through the blood. Thus, cytokines, lytic factors or small regulatory RNAs released from the apoptotic, necrotic or surviving irradiated cells could play a role in the signal (Tamminga *et al.*, 2008a).

1.7.3. Cancer predisposition

Germline and somatic mutations are elevated in the offspring of irradiated mice. Because carcinogenesis is a process dependent on the accumulation of mutations in somatic cells, it is only logical to assume that TI may increase cancer susceptibility.

Indeed, many of the phenotypic studies analyzing radiation-induced TI (RITI) have concentrated on evaluating cancer susceptibility in the offspring of irradiated parents (reviewed by Dubrova, 2007, pp.139-154 [book]). The first study of this kind was carried by Nomura who reported that X-irradiation of ICR strain parental mice resulted in a significant increase in lung tumours (mostly benign) in the F_1 offspring (Nomura, 1982). This result was confirmed by some but not all studies (UNSCEAR, 2008 [review]). In other studies Nomura was unsuccessful in finding an increase in leukaemia in offspring after acute irradiation of spermatogonia from an ICR mouse parent (Nomura, 1990). The same study was successful when using the N5 mouse strain in which a 10-fold increase in the incidence of acute lymphocytic leukaemia was detected in the offspring of irradiated parents (Nomura, 1990). These results demonstrate the inter-strain variability in radiation response, which is commonly observed in these experiments. They also support the view that irradiation of mice generally results in increased predisposition to cancer in subsequent generations. However, how the TI works in the cells of the non-exposed offspring to produce these effects is currently unknown.

In addition to somatic mutations, genome-wide hypomethylation can also be caused by exposure to IR. Genome-wide hypomethylation is often viewed as a sign of genome destabilization and is observed in pre-cancerous stages and established cancers (Kovalchuk *et al.*, 2004). It is very often combined with localized *de novo* methylation, often at promoter regions, and transcriptional repression (Ehrlich, 2002 [review]). Altered DNA methylation, if transmitted to the F_1 generation, could increase susceptibility to carcinogenesis.

1.8. TI is an epigenetic phenomenon

"An epigenetic trait is a stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence" (Berger *et al.*, 2009 [review]). Although the mechanism of TI remains to be identified considerable evidence suggests that it is not controlled by DNA sequence but by epigenetic traits.

The TI studies mentioned in the previous sections suggest that inheritability of radiation-induced effects cannot be explained by Mendelian segregation of a mutant allele. The strongest experimental evidence to support this hypothesis include the following observations: (1) almost all offspring (~100%) of irradiated parents manifest genomic instability; (2) unexposed F_1 offspring present a similar amount of damage on both ESTR alleles from the irradiated father and unexposed mother; and, (3) transgenerational genomic instability equally manifests in the F_1 and F_2 offspring of irradiated parents (Barber *et al.*, 2002; Barber *et al.*, 2006).

1.8.1. Non-coding RNA as a possible mechanism of TI

One hypothesis is that irradiation of germ cells could cause a radiation-induced accumulation of RNA molecules that upon fertilization would be transmitted to the zygote and change gene expression. This would then initiate the TI observed in the offspring of irradiated parents. Indeed, there is evidence that sperm can package RNA and, though its role is not known, it is suspected of carrying epigenetic information for the zygote (Miller *et al.*, 2009). Crosses between wild-type mice homozygous for *Kit* ($Kit^{+/+}$) and mice heterozygous for the null mutant Kit^{tm1Alf} ($Kit^{tm1Alf/+}$) support this hypothesis (Rassoulzadegan *et al.*, 2006). $Kit^{tm1Alf/tm1Alf}$ homozygotes die at birth but $Kit^{tm1Alf/+}$ heterozygotes survive and manifest a heritable white tail and paws phenotype. This phenotype is manifested in the $Kit^{tm1Alf/+}$ progeny resulting from $Kit^{tm1Alf/+} \times Kit^{+/+}$

or $Kit^{tm1Alf/+} \times Kit^{tm1Alf/+}$ crosses, independently of the gender combination. Surprisingly, the $Kit^{+/+}$ progeny (*Kit**) from these crosses also presented the white patch phenotype. This indicated a heritable epigenetic modification to the Kit^{+} allele by cross-talk with the Kit^{tm1Alf} allele (i.e. paramutation) taking place in heterozygous parents. The Kit^{tm1Alf} mutant was thought to be transcribed pre-meiotically in germ cells of the $Kit^{tm1Alf/+}$ heterozygous parent and transmitted into the zygote following fertilization. This view was supported by the observation that either total RNA from $Kit^{tm1Alf/+}$ or Kit^{+}-specific microRNAs, if microinjected into wild-type fertilized eggs, could induce the white patch phenotype (Rassoulzadegan *et al.*, 2006).

The prevalence of mutant Kit^{tm1Alf} transcript in the offspring demonstrates that RNA can be stably inherited from the mutant paternal/maternal mouse. Similarly, an inheritance of radiation-induced transcripts could take place between an irradiated parent and its non-exposed offspring (Dubrova, 2007, pp.139-154 [book]). However, the mechanism of TI cannot be fully dependent on the transmission of RNA as it is dependent on transcription and condensed spermatids and spermatozoa are transcriptionally inactive (Leduc *et al.*, 2008). This statement is supported by reports observing TI in non-targeted offspring of male mice irradiated during the late post-meiotic stages of spermatogenesis (Vorobtsova, 2000; Shiraishi *et al.*, 2002).

Non-coding RNA may modulate TI when transcription is possible. Indeed, it has been shown that piwi-interacting RNAs in foetal male germ cells, undergoing *de novo* methylation, can direct chromatin modifiers in the nucleus to methylate transposable elements and eventually silence them (Aravin and Bourc'his, 2008).

1.8.2. Oxidative stress and cytokine signalling as a mechanism of TI

Another idea is that the DNA damage in the offspring of irradiated parents could be generated by free radicals or cytokine signalling. It has been reported that irradiation *in vitro* results in an enhanced non-specific inflammatory response characterized by the release of high levels of cytokines and NO due to a prolonged phagocytic activity (Brune *et al.*, 1996; Giles *et al.*, 2000; Gregory, 2000 [review]). Cytokines and free radicals were suspected of contributing to an epigenetic and/or bystander effect causing increased mutation rates in the progeny and neighbours of irradiated cells (Lorimore *et al.*, 2003 [review]). Indeed, ROS are a major source of endogenous damage capable of causing SSBs, DSBs, abasic sites and various nucleotide modifications (Jackson and Loeb, 2001). The release of free radicals and cytokines could explain the amount and variety of the DNA damage in the cells of offspring of irradiated parents. However, there are observations that contradict a mechanism of TI based on an oxidative stress response: (1) Formamidopyrimidine glycosylase (FPG) comet assays have not detected any oxidatively damaged nucleotides in bone marrow cells of offspring of irradiated males (Barber *et al.*, 2006); and, (2) the very small cytoplasmic component of spermatozoa would make it very difficult for sufficient amounts of long-lived ROS, or other radiation-induced species, to be transported to the zygote to initiate TI (Rakyan *et al.*, 2001; Dubrova, 2003 [review]).

1.8.3. Histone code as a mechanism of TI

Approximately 1% of histones can be inherited through germ cells (van der Heijden *et al.*, 2006). Because these proteins can be modified by radiation and stably inherited, there is a chance that they may modulate TI (Barton *et al.*, 2005; Barber and Dubrova, 2006 [review]).

The active or repressed condition of chromatin does not depend solely on DNA methylation. In fact, DNA methylation just supports the assembly of proteins to form the chromatin (Clark, 2007 [review]). The collective action of histones in nucleosomes plus histone post-translational modifications (PTM) and proteins involved in gene function and DNA replication will determine whether the chromatin is transcriptionally active (euchromatin) or not (heterochromatin) (Miranda and Jones, 2007 [review]).

Erroneous epigenetic marks responsible for diseases such as cancer can be acquired during life. Reprogramming is essential in this aspect as it erases these epigenetic marks during gametogenesis and embryogenesis (Morgan *et al.*, 2005). At the spermiogenesis stage of spermatogenesis in males, histones are replaced by protamines and other nuclear basic proteins (Wouters-Tyrou *et al.*, 1998). However, histones are not completely replaced by protamines in mature sperm. It has been estimated that about 1% of the DNA in the mouse sperm remains bound to histones in nucleosomes (van der Heijden *et al.*, 2006). These regions are thought to contain heterochromatin and are located in the nuclear periphery (includes retrotransposons and telomeres) and in the centre of the nucleus (includes centromeres) (Palmer *et al.*, 1990; Haaf and Ward, 1995; Pittoggi *et al.*, 1999; Zalenskaya *et al.*, 2000). It is also thought that these regions, once the male genome is decondensed, help direct the formation of chromatin in other regions of the genome (van der Heijden *et al.*, 2006). Their importance is such that disruption of this heterochromatin is suggested to result in genome instability (Peters *et al.*, 2001; de Lange, 2002 [review]). Histone H4 was reported to be acetylated at K8 or K12 prior to full decondensation of the male pronucleus in the fertilized egg. This observation suggests that these PTMs were established prior to fertilization and that nucleosomes can survive the packaging of the

genome during spermiogenesis and be transmitted to the zygote (van der Heijden *et al.*, 2006).

The effects of radiation on chromatin were analysed by irradiation of F_0 ATM heterozygous ($ATM^{+/-}$) male mice with 0.1 Gy acute ^{137}Cs gamma-rays and mating with non-irradiated ATM homozygous ($ATM^{+/+}$) females (Baulch *et al.*, 2007). Irradiated A/B spermatogonia gave rise to non-exposed male $ATM^{+/-}$ offspring which were mated under similar conditions for the next 3 filial generations. The irradiated and non-irradiated $ATM^{+/-}$ males had less ATM in the germline to stop the cell cycle at G_1/S and S/G_2 checkpoints meaning more damage is caused by radiation and GI, respectively, than in $ATM^{+/+}$ males of the corresponding filial generation. Radiation-induced chromatin alteration was visualized in the mature sperm of $ATM^{+/-}$ F_0 males, after culling, 45 days after irradiation. This chromatin alteration was detected as a reduction in the electrophoretic mobility of sperm DNA using the neutral pH sperm comet assay and was observed until the F_3 generation (Baulch *et al.*, 2007). This report provided support to the view that chromatin states can be inherited and that the severity of radiation/GI/TI may be dependent on the potential for DNA repair in the irradiated mouse and its offspring (more detail on DNA repair in irradiated mice and offspring on **section 1.8.4**).

Histones could also be involved in transmitting RIGI across generations. This would depend on their ability to survive spermiogenesis which in turn would depend on how strongly they bind to the DNA. It has been reported that in the chromocenter of spermatids, which is a mass of heterochromatic chromatin at the centre of the nucleus, the histones are hypoacetylated, bind AT-rich DNA, survive the maturation process and are transmitted to the zygote (Widlund *et al.*, 1997; van der Heijden *et al.*, 2006). The histones at the posterior region of the spermatids are hypoacetylated but are removed

during spermiogenesis which suggests that hypoacetylation is not enough to secure histones at this stage (van der Heijden *et al.*, 2006). The combination of weak DNA binding ability (histone hypoacetylation) and intrinsic nucleosome stability (provided by the AT-rich sequences) can protect nucleosomes (and its histones) from being removed (van der Heijden *et al.*, 2006).

1.8.4. DNA methylation as a mechanism of TI

IR causes different effects in directly irradiated tissue and non-irradiated bystander tissue, as shown in a study involving unilateral 1 Gy X-ray irradiation of mice (Koturbash *et al.*, 2006). A significant increase in γH2AX foci indicated the presence of DSBs in directly irradiated tissues of mice whole-body-exposed or unilaterally-exposed and in the non-exposed bystander tissues of unilaterally-exposed mice. Unilateral irradiation caused a significant decrease in DNA methylation of directly-exposed cutaneous tissue as demonstrated by increased ^{3}H-dCTP incorporation. This effect was smaller and non-significant in bystander tissues. The levels of DNMT3a and 3b were slightly down-regulated in both directly-irradiated and bystander tissues. DNMT1 was up-regulated in directly-exposed tissues of the whole-body and unilaterally-exposed mice with a higher DNMT1 up-regulation being detected in bystander tissues. There was a small up-regulation of MeCP2 in the directly exposed tissues of both irradiated models and a large up-regulation of MeCP2 and MBD in bystander tissues of unilaterally-exposed mice (Koturbash *et al.*, 2006). In another study, a correlation between DNA hypomethylation and down-regulation of DNMT3a and 3b was shown in mouse liver and spleen following whole-body X-ray irradiation (Raiche *et al.*, 2004). Rats cranially or whole-body irradiated with 20 Gy of X-rays were shown to possess

genome-wide hypomethylation and significant decrease in protein levels of DNMT1, DNMT3a, DNMT3b and MeCP2 in spleen (Koturbash *et al.*, 2007).

A final hypothesis for the epigenetic mechanism of TI is that DNA methylation in genomic region(s), involved in the maintenance of genome stability, could be altered by radiation in germ cells and be transmitted to the next generation. Indeed, DNA methylation in germ cells can be modified by radiation exposure (Tamminga *et al.*, 2008a; Tamminga *et al.*, 2008b; Filkowski *et al.*, 2010). Altered DNA methylation patterns are capable of surviving reprogramming occurring during gametogenesis and embryogenesis for both females and males (Rakyan *et al.*, 2001; Rakyan *et al.*, 2003; Cropley *et al.*, 2010) and can be transmitted through many cell divisions (Roemer *et al.*, 1997; Rakyan *et al.*, 2003). There is evidence that DNA repair is unaffected by radiation exposure, and subsequent alterations in DNA methylation, as it remains active in irradiated mice and their unexposed offspring (Barber *et al.*, 2006; Koturbash *et al.*, 2006). Radiation-induced alterations in the DNA methylation pattern of gene(s) or genomic region(s), with GI potential, have not been consistently shown across studies in germline tissues.

DNA methylation is known to maintain genomic stability, allowing a normal development and cell proliferation (Pogribny *et al.*, 2004). Genotoxic agents such as IR can affect DNA methylation patterns in directly exposed tissues. IR exposure can cause a dose-dependent, sex- and tissue-specific global DNA hypomethylation and a sex- and tissue-specific promoter hypermethylation (Kovalchuk *et al.*, 2004; Pogribny *et al.*, 2004). Both extremes of DNA methylation are thought to contribute towards carcinogenesis (Jones and Baylin, 2002 [review]). Radiation-induced promoter hypermethylation is thought to involve *de novo* DNMTs but the signalling leading to this event is currently unknown (Jones and Baylin, 2002 [review]; Kovalchuk *et al.*,

2004; Raiche *et al.*, 2004). Global hypomethylation is thought to be a consequence of DNA repair under low DNMT1 expression (Pogribny *et al.*, 2004). DSBs are a very common type of DNA damage induced by IR that is effectively repaired by HR or non-homologous end joining (Baulch *et al.*, 2007). In some DNA repair processes, like HR, DNA synthesis may take place and the polymerases add non-methylated cytosine to the elongating strand leaving the DNMT1 to return the chemical modification. If DNA synthesis under radiation-induced down-regulation of DNMT1 is followed by replication and cell division it could cause the loss of the methylation mark on the synthesised DNA (Pogribny *et al.*, 2004). Following whole body X-ray irradiation, an inverse correlation between DNA methylation and SSB levels was demonstrated in spleen, liver and lung tissues (Pogribny *et al.*, 2004). There has also been a report of DNMT1 down-regulation in non-exposed spleen following cranial X-ray irradiation (Koturbash *et al.*, 2007). These observations suggest that radiation can induce global hypomethylation patterns in directly-irradiated and distant non-exposed tissues of the irradiated mouse and that these could explain the increased mutation rates (i.e. GI) and cancer.

Changes to the DNA methylation pattern could mediate the formation of TI if detected in the germline of the irradiated mice and if the alteration could be produced following direct irradiation or following a non-targeted effect of radiation (i.e. RIBE). The Brother of the Regulator of Imprinted Sites (BORIS) is an epigenetic regulator expressed in primary spermatocytes and round spermatids and has been found to be down-regulated following irradiation and loss of methylation in mice testis (Loukinov *et al.*, 2002; Tamminga *et al.*, 2008b). BORIS regulates DNA methylation levels, genomic imprinting and testis-specific expression during spermatogenesis (Tamminga *et al.*, 2008b; Suzuki *et al.*, 2010). *BORIS* expression is repressed by p53, CTCF and CpG

methylation (Pugacheva *et al.*, 2010). X-ray irradiation of mouse testis has been shown to cause a change in the expression of many microRNAs including miR-709 (Tamminga *et al.*, 2008b). miR-709 increase was found to be DNA damage-induced, ATR/Rfx1-mediated and was found to target and silence *BORIS*. Because radiation-induced DNA damage causes DNA hypomethylation that can up-regulate BORIS, it was concluded that the increased miR-709 expression was a protective mechanism by the irradiated testis to prevent further erasure of DNA methylation (Tamminga *et al.*, 2008b). IR has been found to lead to a significant accumulation of DNA damage in the exposed and bystander testis tissues leading to deregulation of *BORIS* expression and causing an altered DNA methylation in the testis (Tamminga *et al.*, 2011).

Alteration of DNA methylation patterns could be the end-point of the mechanism of RITI and BORIS one of the regulators of this event. However, its dependency on DNA hypomethylation for up-regulation suggests that BORIS requires the action of other cellular components and processes in order to participate in RITI (e.g. DNMT1, DNMT3b and/or DNA repair) (Pugacheva *et al.*, 2010). Condensed spermatids and spermatozoa are unable to express *BORIS* or repair DNA (Leduc *et al.*, 2008; Pugacheva *et al.*, 2010). Further studies need to investigate whether, after mature sperm irradiation and fertilization, *BORIS* expression is possible in the zygote. The danger of radiation-induced BORIS down-regulation also needs to be studied before BORIS can be considered a player in the mechanism of RITI.

1.9. Differential expression of circadian genes

The analysis of gene expression in offspring of irradiated males belonging to the CBA/Ca and BALB/c strains of mice have shown significantly differential expression in genes which usually have 24 hour non-constant patterns of expression. The setting of

mouse culling and tissue extraction to a fixed day time allowed these circadian genes to be accurately and reliability measured for transcript level. Thus, background literature on circadian genes was included in the introduction.

1.10. The mammalian circadian clock

Many organisms on Earth (i.e. animals, plants, fungi and cyanobacteria) possess biological processes in which genes obey a rhythm of expression. This biological rhythm obeys a 24-hour period and is controlled by a circadian (*circa*-about and *dies*-day) clock that is autonomous and endogenous to a particular type of cell/tissue (Loudon, 2012 [review]). The circadian clock can also be reset by an external stimulus, called zeitgeber, to match a local time and ensure that physiological processes are performed at an optimum time for the organism. The resulting biological rhythm is known as a circadian rhythm (Lamia *et al.*, 2008) but can only truly be called circadian if the oscillation is maintained in the absence of zeitgebers (Stow and Gumz, 2011 [review]).

1.10.1. Environmental versus systemic cues

The mammalian circadian timing system obeys a hierarchical network of molecular oscillators driven by a central pacemaker known as the suprachiasmatic nuclei (SCN). The SCN is located in the anterior hypothalamus in the brain and is composed of multiple circadian oscillator cells. These cells are synchronized, and the SCN is entrained, each day by exposure to environmental stimuli, mainly light (Zuber *et al.,* 2009). The light/dark cycle is captured by the retina and transmitted through the retinohypothalamic tract to the SCN in which a rhythmic signal is formed and distributed to all tissues of the body via hormones and the autonomous nervous system

(Buijs *et al.*, 1998). This signal sets the rest/activity cycles which, in turn, impose the fast/feeding cycle responsible for setting the circadian rhythms in the peripheral tissues in association with the peripheral circadian clock (Schibler *et al.*, 2003; Vollmers *et al.*, 2009). The fast/feeding cycle is set by the light-entrained SCN, if food is always present. However, when food is scarce or absent, the feeding rhythms are not synchronised to the light/dark cycle or set by the SCN but are instead set by a food-entrainable oscillator (FEO) composed by many regions of the CNS, excluding the SCN (Atwood *et al.*, 2011).

Usually the combination of zeitgebers like temperature and light can lead to a stable phase and maximum amplitudes when synchronised naturally (i.e. warm/light and cold/dark) to entrain the circadian rhythm. If two zeitgebers oppose each other, the phase of the circadian rhythm is determined by one or both zeitgebers depending on the relative strength of both stimuli (on the size and direction of amplitude of rhythm) and on the susceptibility of the cellular make-up to the stimuli. Opposed stimulation by both zeitgebers usually causes a phase jump/shift at a point of conflict between the two zeitgeber cycles. In ectothermic vertebrates, temperature variations can affect and entrain melatonin production cycles in the pineal gland, which in turn, causes entrainment by phase shifting (Rensing and Ruoff, 2002).

1.10.2. The SCN regulates peripheral clocks

Peripheral tissues such as the retina, adipose tissue, heart, liver, kidney and spleen also contain circadian clocks and can be photo-entrained by the SCN (Reppert and Weaver, 2002 [review]; Young, 2006; Lamia *et al.*, 2008; Keller *et al.*, 2009; Zuber *et al.*, 2009; Geyfman *et al.*, 2012). The SCN and peripheral tissues, except for the retina, cannot be entrained directly by the light stimulus alone. This statement is

supported by the observation that circadian biochemical and physiological processes persist in the retina in the event of SCN lesion or transected optic nerves (Terman *et al.*, 1993; Tosini and Fukuhara, 2002 [review]). Also, mice lacking functional rods and cones and lacking the photopigment melanopsin, from their retinas, are incapable of light-induced phase shifts (Hattar *et al.*, 2003; Panda *et al.*, 2003). Finally, bilateral eye removal can abolish photo-entrainment altogether in rodents (Nelson and Zucker, 1981).

The fast/feeding cycles allows peripheral tissues, such as liver, heart, kidney, and pancreas, to anticipate the daily periods of food availability and optimize processes involved in the food ingestion and metabolism and in utilization and storage of energy (Schibler *et al.*, 2003; Lamia *et al.*, 2008). The core oscillators behind physiological processes provide circadian rhythms of expression to output genes. Between 5-20% of the transcriptome of peripheral tissues is expressed in a circadian fashion and the vast majority of these genes is tissue-specific (Froy, 2011 [review]). There is very little overlap between circadian genes across tissues as most of these genes are involved in rhythmic/cyclical processes that are specific to the local physiology (Atwood *et al.*, 2011).

The SCN can, through systemic signals, regulate a number of behavioural and physiological circadian rhythms throughout the body. The circadian rhythms in the peripheral tissues are, however, generated independently of the SCN. In fact, about 90% of peripheral circadian genes are regulated only by local autonomous oscillators and in the absence of signals transmitted by blood or neurons (Zuber *et al.*, 2009). Peripheral oscillators can receive environmental phase-setting information directly or through the SCN. The information from the SCN can come as a strong entrainment signal or as a weak signal for fine-tuning of the circadian rhythm. The latter seems to be achieved

through crosstalk between peripheral and SCN oscillators (Barnard and Nolan, 2008 [review]).

1.10.3. Molecular components of the circadian clock

Central and peripheral oscillators share a similar molecular core clock based on a set of transcriptional-translational feedback loops (**Figure 1.6**). Eight core circadian genes form and maintain the basic function of the mammalian circadian clock: Casein kinase 1ε (*CK1ε*), Cryptochrome 1 (*Cry1*), Cryptochrome 2 (*Cry2*), Period 1 (*Per1*), Period 2 (*Per2*), Period 3 (*Per3*), circadian locomotor output cycles kaput (*Clock*) and Brain and muscle ARNT-like protein (*Bmal1*, also known as *Arntl* and *MOP3*) (Taniguchi *et al.*, 2009).

The primary loop that drives circadian transcription includes a heterodimer formed by basic-helix-loop-helix-Per-Arnt-Sim (bHLH-PAS) domain-containing transcription factors. Through the interaction of Per-Arnt-Sim (PAS) domains, Bmal1 can associate with either neuronal PAS domain protein 2 (Npas2) or Clock to regulate gene expression and maintain molecular rhythms (Ishida *et al.*, 2008). Npas2 can replace Clock in the SCN and is the common binding partner for Bmal1 in forebrain clocks (Reick *et al.*, 2001; DeBruyne *et al.*, 2007). Although both Npas2 and Clock perform similar roles they possess different properties that allow them to execute special functions. For example, Npas2-Bmal1 heterodimers repress Bmal1 gene expression (Reick *et al.*, 2001) while Clock has intrinsic histone acetyltransferase activity that allows it further control of transcription (Doi *et al.*, 2006). Clock/Npas2-Bmal1 heterodimer bind E-box elements (or E-box-like, E'-box) and drive the transcription of clock-controlled output genes as well as that of the Period (*Per1, Per2* and *Per3*) and Cryptochrome genes (*Cry1* and *Cry2*). The increased abundance of Per

and Cry proteins induces a negative feedback in which Per and Cry proteins, dimerise, translocate to the nucleus inhibit the Clock-Bmal1 activity and, consequently, inhibit their own transcription (Geyfman *et al.*, 2012).

Ck1ε/γ and the F-box protein Fbxl3 control the stability of Per1/2 and Cry1/2, respectively (Froy, 2011 [review]). The proteosomal degradation of Per and Cry proteins allows Clock and Bmal1 levels to rise again completing the cycle of circadian rhythm (Mullenders *et al.*, 2009). An auxiliary loop is formed by the activator ROR (RORα, RORβ and RORγ) and the inhibitor Rev-erb (Rev-erbα and Rev-erbβ) which modulate the expression of *Bmal1* and other target genes that contain a Rev-erb/ROR-binding element (RRE) in their promoter. Another auxiliary loop is composed by the activator Dbp and the repressor E4bp4 which regulate transcription through D-box elements (Atwood *et al.*, 2011).

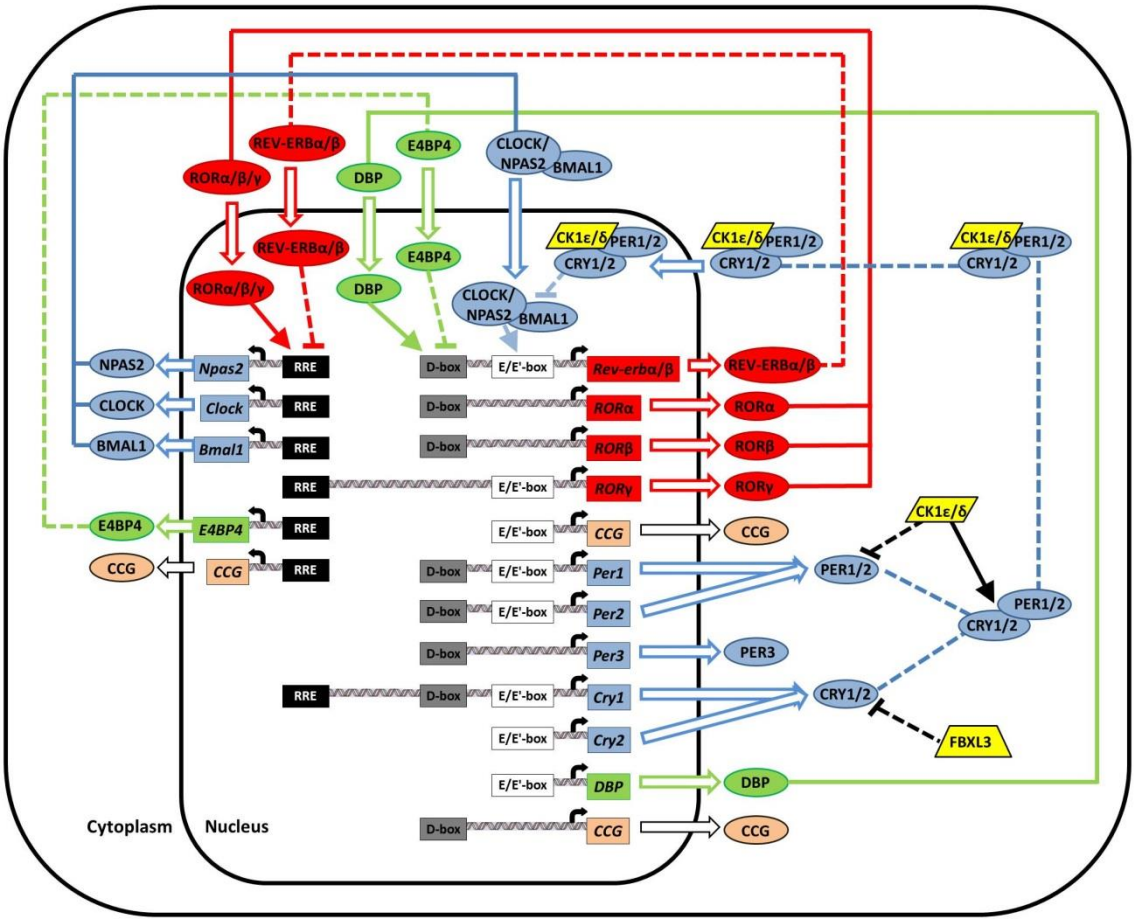

Figure 1.6: Core molecular circadian clock in central and peripheral oscillators in mice. In blue, the primary and negative feedback loops of circadian transcription. In red and green the auxiliary loops. Proteins in charge of directing phosphorylation of Per and initiating ubiquitination of Cry, leading to degradation, are displayed as yellow quadrilaterals. Morning (E/E'-box), day-time (D-box) and evening (RRE) clock-controlled DNA elements drive expression of clock-controlled genes (CCG; in beige) accounting for all of the circadian transcriptome in a tissue involved in processes like metabolism, locomotor activity, behaviour, etc. Activator and repressor loops of transcription are displayed as straight-line arrows and end-capped dashed-lines, respectively. See text for details (figure constructed based on Preitner *et al.*, 2002; Ueda *et al.*, 2002; Barnard and Nolan, 2008 [review]; Liu *et al.*, 2008; Ukai-Tadenuma *et al.*,

2008; Yamajuku *et al.*, 2010; Ukai-Tadenuma *et al.*, 2011; Takeda *et al.*, 2011; Ye *et al.*, 2011; Korenčič *et al.*, 2012; Shi *et al.*, 2013).

1.10.4. Transcription factors increase the output of the clock

Post-translational modifications control the stability and nuclear entry of clock components and represent a stage of the entrainment process by which zeitgebers and endogenous stimuli exert their control over the clock (Stow and Gumz, 2011 [review]). Components of the clock are the same in peripheral and nuclei oscillators but the target genes with rhythmic expression are usually tissue-specific. There is support for the idea that the Clock-Bmal1 heterodimer amplifies the transcriptional cascade by the rhythmical expression of transcription factors that are then responsible for the majority of the rhythmic, tissue-specific expression (Miller *et al.*, 2007). This view is supported by studies where *Clock* knockout mice and mice devoid of *Dbp*/*Hlf*/*Tef*, a triple knockout of transcription factors expressed by Clock-Bmal1, exhibit a decrease in the amplitude/loss of oscillation on the same output genes (i.e. *Gilz*, *Usp2*, *Mapre2*) (Zuber *et al.*, 2009).

1.10.5. The importance of cell-autonomous clocks in peripheral tissues

The circadian clock allows rhythmic gene expression in the liver and confers rhythmicity to hepatic processes including glucose and lipid homeostasis, xenobiotic metabolism and liver regeneration. Microarray analysis of gene expression detected 1,130 transcripts in Met murine hepatocytes-D3 cells maintaining circadian rhythms *in vitro* and in the absence of systemic signals including glucocorticoids, insulin, growth hormone and food intake. Amongst the circadian genes driven by the cell-autonomous

hepatic clock, there were two essential enzymes which were involved in the polyamine synthesis. Polyamines are themselves important modulators of protein-protein and protein-DNA interactions for the initiation of transcription and may, therefore, influence circadian output. Disrupting the cell-autonomous clock has been shown to cause delays in regeneration and desynchronisation of cell proliferation in liver (Atwood *et al.*, 2011).

The role and importance of the cell-autonomous clock on the circadian control of genes and processes in liver has also been shown *in vivo*. Lamia *et al.* (2008) observed that mice with liver-specific disruption of Bmal1 (*L-Bmal1*$^{-/-}$) present hypoglycaemia during the fasting phase of the daily cycle. Fasting time in wild-type mice corresponds to the circadian time 0-12 h which also corresponds to a peak rise in glucose transporter 2 (*Glut2*) transcript and protein levels. Glut2 is involved in transporting glucose from hepatocytes into the blood to counterbalance the daily cycle of glucose ingestion. In *L-Bmal1*$^{-/-}$ mice, circadian expression of *Glut2* is lost and is replaced by a flat-lined expression. This causes a deficiency of Glut2 during the fasting stage and onsets hypoglycaemia. *L-Bmal1*$^{-/-}$ mice also presented faster glucose clearance from the blood after a meal, when compared to control mice, but had a normal or possibly reduced total body fat content. It was suggested that due to *Bmal1* being rhythmically expressed in the pancreas and brain, the *L-Bmal1*$^{-/-}$ mice had, respectively, a normal insulin production and sensitivity, and a normal feeding behaviour and locomotion. This was concluded after considering other experiments where mice lacking *Bmal1* in all tissues (*Bmal1*$^{-/-}$) had no daily rhythms of locomotor and feeding activities, had reduced production of insulin, insulin hypersensitivity, high blood glucose concentration and slow glucose clearance after a meal. Signals for hyperglycaemia overwhelmed those for hypoglycaemia causing *Bmal1*$^{-/-}$ mice to have a

high total fat content (Lamia *et al.*, 2008). These observations show the importance of *Bmal1* in the regulation of multiple circadian parameters (i.e. glucose homeostasis, insulin production) and inform on the role of *Bmal1* on the circadian clock mechanism of the SCN tissue and at least one peripheral tissue (i.e. liver). This study shows that clock disruption *in vivo* of *L-Bmal1*$^{-/-}$ and *Bmal1*$^{-/-}$ mice disrupts circadian processes in the liver of mice but that intact distant clocks in *L-Bmal1*$^{-/-}$ mice, through their normal oscillatory activity, prevented further physiological degradation. Thus, local cell-autonomous clocks are essential in generating the circadian rhythm of tissue-specific genes that cannot be replicated solely by systemic signals from distant clocks. The SCN and other distant clocks can however regulate liver clocks in response to changing external conditions.

Obesity and insulin resistance are examples of metabolic disorders that can be caused by disrupting circadian rhythms. A study combining gas and liquid chromatography with mass spectrometry and computational analysis has successfully modelled how the metabolome, transcriptome and proteome stages might work together to achieve metabolic homeostasis. Node validation analysis detected that, unlike wild-type mice, *Clock*$^{-/-}$ mice lost circadian oscillations in many metabolites (e.g. uracil, uridine) and enzymes (e.g. uridine phosphorylase 2) but of particular interest was the loss of oscillation of the NAD^+ metabolite (Eckel-Mahan *et al.*, 2012). Clock-Bmal1 expresses Nampt which synthesises NAD^+ and NAD^+ activates Sirt1 which deacetylates Per1/2 and Bmal1 in an important step of the removal of Per-Cry repression (Froy, 2011 [review]).The NAD^+ metabolite, produced during oxidative phosphorylation, is known for increasing in liver during fasting and decreasing after feeding. During the late phase of the resting period (ZT15-ZT21; includes fasting period), NAD^+ increased in the liver of wild-type but not *Clock*$^{-/-}$ mice: for *Clock*$^{-/-}$/wild-type, NAD^+ fold-change

concentration was 1.29, 0.71, 0.72 and 0.79 for ZT3, ZT9, ZT15 and ZT21, respectively. This study establishes a link between the circadian transcriptome and specific metabolic pathways in liver (Eckel-Mahan *et al.*, 2012).

In normal mice kept in constant darkness, 15% of liver transcripts still possess circadian rhythms of expression. Rhythmic transcription in mouse liver seems to be generated by both the hepatic autonomous oscillators and the food/fasting cycle. A study has demonstrated the link between these two factors in the generation of hepatic circadian rhythms by comparing the effects of *ad libitum* or time-of-day restricted feeding (tRF) on transcript and protein levels in wild-type and oscillator-deficient *Cry1*$^{-/-}$; *Cry2*$^{-/-}$ mice (Vollmers *et al.*, 2009). Daily rhythm in feeding can drive the rhythmic activity of food-induced and fasting-induced transcripts even in the absence of a clock. However, during prolonged fasting, there is the down-regulation of circadian transcriptional outputs with the exception of core clock oscillations. In double-*Cry*-deficient mice, rhythmic feeding salvages the rhythmic expression of 617 out of 4,960 transcripts tRF-driven in normal mice. The absence of a functional clock causes the anticipation of feeding by prior expression of tRF-driven transcripts to be lost in double-*Cry*-deficient mice. Rhythmic feeding increases the amplitude of rhythmic expressions of core clock components and output target genes. This was observed in wild-type and double-*Cry*-deficient mice by comparing *ad libitum* and tRF conditions. In agreement with these findings, other studies have concluded that both the phase and amplitude of circadian transcripts are influenced by the feeding activity in peripheral tissues (Vollmers *et al.*, 2009).

Spleen, lymph nodes and isolated macrophages also contain an autonomous intrinsic clock that can operate without systemic signals. *Ex vivo* and *in vitro* assays, were used to administer bacterial endotoxin to spleen cells at different circadian times

and allowed the detection of circadian rhythms in the secretion of TNF-α and IL-6 (Keller *et al.*, 2009). This rhythmic cytokine secretion was not driven by rhythmic glucocorticoid levels and was inconsequential of fluctuations in cellular numbers in spleen. Thus, these events were concluded to be driven by a cell-autonomous clock in spleen. Global transcription analysis revealed that >8% of genes expressed in peritoneal macrophages were circadian and included key elements in the Lipopolysaccharide-triggered toll-like-receptor 4/TNF-α pathway which are in charge of mediating immune responses (Keller *et al.*, 2009).

Circadian rhythms have been observed in kidney during the urine formation/excretion of water and electrolytes. These rhythms are generated by an intrinsic renal clock and can survive in abnormal feeding/sleeping behaviours for a long time. *Clock*$^{-/-}$ mice present a disruption of renal excretory rhythms caused by severe alterations to the circadian oscillations in secretion and reabsorption functions carried by the kidney nephron. The observed reduction in the reabsorption of sodium and water was associated with changes in the expression of key regulatory genes mediating secretion/reabsorption of sodium and water. *Clock*$^{-/-}$ mice also presented a significantly lower arterial and systolic blood pressure (BP) but a normal daily pattern of change in BP. An increased water intake and diluted urine in *Clock*$^{-/-}$ mice was suggested to be caused either by primary polydipsia or diabetes insipidus (Zuber *et al.*, 2009). The kidney is crucial in the maintenance of a healthy BP by controlling processes such as aldosterone production and sodium secretion/reabsorption. BP increases in the early morning, plateaus during the day and decreases during sleep (Stow and Gumz, 2011 [review]). Patients showing an absence of the 10-20% dipping in BP during night are significantly susceptible to various cardiovascular disorders and death (Routledge *et al.*, 2007). Salt-sensitive hypertension and non-dipping BP are just some of the conditions

that usually accompany the deterioration of renal function (Stow and Gumz, 2011 [review]).

1.10.6. Sub-circadian cycles – ultradian rhythms

Circadian oscillations are produced in peripheral tissues by autonomous clocks and regulated/maintained via a number of physiological signals including core body temperature and hormones produced/transmitted by other peripheral tissues and the SCN. Another class of cycling rhythms named ultradian rhythms have been observed in liver tissue and include transcripts with 8h or 12h rhythms of expression. Some transcripts with 12h cycles of expression have shown to have 24h rhythms in other tissues implying a role by tissue-specific factors in the generation of ultradian rhythms. Similar to circadian genes, there is a complete loss of ultradian rhythms under *ex vivo* conditions caused by reduced amplitude of the rhythm. Tissue culture was regarded unfit for the detection/measurement of oscillations of circadian output genes. Only core clock and first order circadian output genes were found to be able to maintain 24h rhythms during several hours under *ex vivo* conditions. Restrictive feeding is another cue that can affect both circadian and ultradian rhythms but while in circadian genes the period of fasting needs to be long to affect their oscillation, in ultradian transcripts the response is more prompt and their rhythmicity is lost by phase shifting of the cycle (Hughes *et al.*, 2009). Thus, ultradian rhythms are generated by both cell-autonomous clocks and systemic signals.

1.11. Project aims

The aim of this thesis is to analyse and compare the pattern of gene expression in the non-exposed offspring of control and irradiated male mice. Using the NimbleGen

12x135K multiplex expression array, the pattern of gene expression was established in the first-generation offspring of CBA/Ca and BALB/c male mice. To ensure that the transgenerational effects affecting the offspring of irradiated males are related to the observed changes in gene expression, RNA samples extracted from brain, kidney, liver and spleen were profiled. A number of statistical approaches were used to assess the quality of the microarray data and to compare the data obtained by profiling RNA samples taken from the offspring of control and irradiated males (**Chapter 4**).

Genes with significantly altered expression were further validated by qPCR quantification of transcript level (**Chapter 5**). This technique uses the geometric mean of the concentration of reference genes to accurately measure the relative expression of the differentially expressed genes in test and control samples. Microarray measurements were validated and demonstrated that the phenomenon of TI affects specific genes in the non-exposed offspring of paternally irradiated mice.

The epigenetic changes behind the phenomenon of TI were likely to occur in transcription regulatory elements along the DNA sequence of these differentially expressed genes. In order to test this hypothesis the promoter, exonic and intronic regions of differentially expressed genes were measured for changes in DNA methylation pattern using single-molecule and conventional bisulfite sequencing (**Chapter 6**).

Chapter 2:

Materials

2.1. Chemicals/reagents

Chemicals were obtained from Fisher Scientific (Division of Thermo Fisher Scientific, Loughborough, UK) and Sigma-Aldrich Company Ltd. (Poole, UK).

Molecular biology reagents were obtained from Ambion Inc. (Division of Life Technologies, Paisley, UK), Fisher Scientific (Division of Thermo Fisher Scientific, Loughborough, UK), Invitrogen UK (Division of Life Technologies, Paisley, UK), Kapa Biosystems Inc. (supplied by Anachem Ltd., Luton, UK), Lonza (Basel, Switzerland), New England Biolabs (NEB) (Hitchin, UK), Qiagen Ltd. (Crawley, UK) and Sigma-Aldrich Company Ltd. (Poole, UK).

Specialized equipment was obtained from Agilent Technologies UK Ltd. (Wokingham, UK), Ambion Inc. (Division of Life Technologies, Paisley, UK), Axon Instruments Ltd. (Inverurie, UK), C. Gerhardt UK Ltd. (Brackley, UK), Clare Chemical Research Inc. (CCRI) (Dolores, USA), Corning Inc. (Division of Sigma-Aldrich Company Ltd., Poole, UK), DuPont (supplied by Anachem Ltd., Luton, UK), EdgeBio (Gaithersburg, USA), Eppendorf UK Limited (Stevenage, UK), Fermentas Life Sciences (FLS) (St. Leon-Rot, Germany), Fisher Scientific (Division of Thermo Fisher Scientific, Loughborough, UK), Gilson Scientific Ltd. (Luton, UK), Karlan (Phoenix, USA), Labtech International Ltd. (Uckfield, UK), Millipore UK Ltd. (Livingston, UK),

MJ Research (Division of Bio-Rad Laboratories Ltd., Hemel Hempstead, UK), NanoDrop (Division of Thermo Fisher Scientific, Loughborough, UK), Hemel Hempstead, UK), Nikon UK Ltd. (Surrey, UK), PRO Scientific Inc. (Oxford, USA), Progen Scientific (London, UK), Purite Ltd. (Thame, UK), Qiagen Ltd. (Crawley, UK), Roche Diagnostics Ltd. (Burgess Hill, UK), Savant (Division of Thermo Fisher Scientific, Loughborough, UK), Scientific Laboratory Supplies Ltd. (Nottingham, UK), Sigma-Aldrich Company Ltd. (Poole, UK), Sterilin Ltd. (Division of Thermo Fisher Scientific, Loughborough, UK), Swann-Morton Ltd. (Sheffield, UK), Syngene UK (Cambridge, UK), Techne (supplied by Thermo Fisher Scientific, Loughborough, UK), Thermo Scientific (Division of Thermo Fisher Scientific, Loughborough, UK), TreffLab (Distributed by Anachem Ltd., Luton, UK), VWR International Ltd. (Lutterworth, UK), Walker Safety Cabinets Ltd. (WSCL) (Glossop, UK) and Weber Scientific International Ltd. (Teddington, UK).

2.2. Kits

The following kits were used: Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Division of Life Technologies, Paisley, UK); EpiTect Bisulfite kit (Qiagen Ltd., Crawley, UK); Expand High Fidelity PCR System dNTPack (Roche Diagnostics Ltd., Burgess Hill, UK); High Capacity RNA-to-cDNA Master Mix (Invitrogen UK, Division of Life Technologies, Paisley, UK); Maxima SYBR Green qPCR Master Mix (2x), ROX solution provided (FLS, St. Leon-Rot, Germany); NimbleGen One-Colour DNA Labeling kit, NimbleGen Sample Tracking Control kit, NimbleGen Wash Buffer kit (Roche Diagnostics Ltd., Burgess Hill, UK); RNA 6000 Nano kit (Agilent Technologies UK Ltd., Wokingham, UK); SuperScript Double-Stranded cDNA Synthesis kit (Invitrogen UK, Division of Life Technologies, Paisley,

UK); TURBO DNA-free kit (Ambion Inc., Division of Life Technologies, Paisley, UK); and, Zymoclean Gel DNA Recovery kit (Zymo Research, Division of Cambridge Bioscience Ltd., Cambridge, UK).

2.3. Solutions

Ethanol and Industrial Methylated Spirits (IMS) were supplied by the Chemical Store (Chemistry Department, University of Leicester). 10% w/v Sodium Dodecyl Sulfate (SDS), 20x Sodium Chloride Sodium-Citrate (SSC) buffer, 10x Tris-Acetate-EDTA (TAE) electrophoresis buffer and 10x Tris-Borate-EDTA (TBE) electrophoresis buffer were supplied by the Media Kitchen (Department of Genetics, University of Leicester).

Table 2.1 shows all non-commercial solutions used and their preparation method. Unless indicated, solutions were stored at room temperature.

Table 2.1: List of solutions used in this thesis.

Solution	Preparation
5x DNA Loading Dye	10 ml 50x TAE, 12.48 g Ficoll 400, 0.1 g bromophenol blue. Final volume 100 ml with nuclease-free water (NF-H_2O).
10x TAE	0.4 M Tris-Ac, 0.05 M NaAc, 0.001 M EDTA. Adjusted to pH 7.8 with acetic acid.
10x TBE	44.5 mM Tris-borate [pH 8.3], 1 mM EDTA.
10% SDS	10 g of SDS in 80 ml NF-H_2O in fumehood. Heat at 68°C, if required. Final volume 100 ml with NF-H_2O.
20x SSC	3 M NaCl, 0.3 M Tri-Sodium Citrate.
Diethylpyrocarbonate (DEPC) water	1 ml Diethyl Pyrocarbonate in 1 l NF-H_2O. Autoclaved.
Ethidium bromide (EtBr)	10 mg/ml (dissolved in dH_2O)
Lysis Solution A	0.1 M NaCl, 0.025 M EDTA, 0.02 M Tris-HCl [pH 8.0].
Lysis Solution B	1 g SDS, 0.0125 M EDTA, 0.01 M Tris-HCl [pH 8.0]. Final volume 100 ml with NF-H_2O.
PBS	1 tablet of Phosphate buffered saline dissolved in 200 ml of dH_2O.
Proteinase K	25 mg/ml (dissolved in dH_2O). **Stored at -20^oC**.

RNase A	10 mg/ml (dissolved in dH_2O). **Stored at -20°C**.
Salmon Sperm (SS) gDNA	3 mg/ml (dissolved in 5 mM Tris [pH 7.0]). **Stored at -20°C**.
SS dilution buffer	SS gDNA solution diluted with NF-H_2O to 16 µg/ml. **Stored at -20°C**.

2.4. Enzymes

KAPA *Taq* DNA polymerase was produced by Kapa Biosystems Inc. (supplied by Anachem Ltd., Luton, UK). Exonuclease I (Exo I) restriction enzyme and Shrimp Alkaline Phosphatase (SAP) were supplied by FLS (St. Leon-Rot, Germany). Proteinase K and RNase A were supplied by Sigma-Aldrich Company Ltd. (Poole, UK).

2.5. Oligonucleotides

Oligo dT18 primer (5'-TTTTTTTTTTTTTTTTTT-3') used in cDNA synthesis and primers for PCR, SM-PCR and qPCR (described in **Appendix I**; **Tables S1-S6**) were obtained from Sigma-Aldrich Company Ltd. (Poole, UK).

The molecular weight markers φX174 DNA digested with *Hae*III and λ DNA digested with *Hind*III were supplied by NEB (Hitchin, UK).

2.6. Mouse Strains

BALB/c and CBA/Ca mice were obtained from Harlan UK Ltd. (Loughborough, UK) and housed at the Division of Biological Services, University of Leicester. They were kept at a standard daily 12:12 light:dark cycle, with the light period starting at 8 am and ending at 8 pm. Mice were fed *ad libitum* with standard laboratory chow diet and water. All animal experimentation was carried out under guidelines issued by the Medical Research Council in "Responsibility in the use of animals for medical research" (July 1993) and under the UK Home Office project license PPL No 80/2267.

2.7. Arrays

NimbleGen 12x135K Mouse Expression Arrays (MM8 genome build) were supplied by Roche Diagnostics Ltd. (Burgess Hill, UK). The slide consisted of 12 sub-arrays with 135,054 60mer probes targeting 42,575 different mouse transcripts (3-4 probes per transcript). Probes were printed onto 13x13 μm features, organized in 8.9x6.5 mm arrays, on a 25x76 mm glass slide.

2.8. Specialized software

The following software was used: GenePix Pro 6.0 (Molecular Devices, Sunnyvale, USA); NimbleScan v2.5 (Roche Diagnostics Ltd., Burgess Hill, UK); Statistica v7.1 (StatSoft Ltd., Bedford, UK); Systat 13 (Systat Software Inc., London, UK); Opticon Monitor v3.1 (Bio-Rad Laboratories Ltd., Hemel Hempstead, UK); and, Chromas v2.1 (Technelysium Pty Ltd., South Brisbane, Australia).

Chapter 3:

Methods

3.1. Mouse irradiation and mating

Seven week-old BALB/c and CBA/Ca male mice were exposed to whole body, dorsal, acute irradiation with 1 Gy of X-rays. X-rays were delivered at 0.5 Gy/min (250 kV constant potential, Half Value Layer 1.2 mm Cu, Pantak industrial X-ray machine, Connecticut, USA). Male mice were mated at 15 weeks of age with two unexposed female mice of the same strain and age. Tissue samples were taken from the eight-week old male offspring of irradiated and non-exposed males.

3.2. Mouse tissue extraction

Mice were killed by cervical dislocation in a room under sterile conditions. Organs were removed, with those destined for RNA extraction being snap frozen in liquid nitrogen and stored at -80°C. Organs destined for DNA extraction were put on ice before storage at -80°C. All mice culling and tissue extraction were performed between 9 am and 11:30 am.

3.3. Preparation of total RNA for gene expression analysis

3.3.1. RNA extraction and quantitation

Mouse tissues were homogenised in 7 ml of Tri reagent (Sigma-Aldrich) using a PRO Multi-Gen 7 blender (PRO Scientific). Kidney and brain were homogenised

whole. Liver and spleen were cut into small pieces prior to homogenisation using a Petri dish (Sterilin) and a scalpel blade (Swann-Morton). When cutting, dry ice was used to keep both tissue and tools cold, avoiding tissue defrosting and RNA degradation.

The lysate was divided into 1.5 ml microcentrifuge tubes (TreffLab) and left for 5 min at room temperature to guarantee complete denaturation of cellular and nucleoprotein complexes. A $1/10^{th}$ volume of 1-bromo-3-chloro-propane (Sigma-Aldrich) was added and the sample, emulsified by inverting the tube for 15 seconds and mixed for 10 sec using a Vortex (Fisher Scientific). The sample was left another 2-3 min at room temperature to facilitate the separation of nucleic acids from proteins. Centrifugation at 12,000 rcf, at 4°C for 14 min, separated the sample into three phases: a top aqueous phase containing RNA; an intermediate gel-like phase containing gDNA; and, an organic phase containing proteins. The top aqueous phase was transferred to a new 1.5 ml microcentrifuge tube and 1 volume of isopropanol (Sigma-Aldrich) was added and mixed by inversion. Tubes were left for 5 min at room temperature and RNA was pelleted by centrifugation at 12,000 rcf, 4°C for 10 min. The supernatant was discarded and the RNA pellet was washed twice with 75% ethanol and 25% of DEPC. Following removal of the ethanol, the pellet was left to air-dry for 5 min before being dissolved in 60-200 µl of NF-H_2O for kidney, liver and brain RNA. Spleen RNA was dissolved in 10-40 µl of NF-H_2O. The concentration of RNA was measured on the NanoDrop ND1000 UV-Vis Spectrophotometer (NUS) (NanoDrop). Ideally, RNA samples had a concentration of 1000-2100 ng/µl, an A_{260}/A_{280} ratio ≥1.8 and an A_{260}/A_{230} ratio ≥1.8. RNA samples were aliquoted for storage at -80°C.

3.3.2. Quality control of the RNA on the Bioanalyzer

All RNA samples were submitted to On-chip gel electrophoresis using the Agilent 2100 Bioanalyzer (Agilent Technologies) to confirm their integrity and purity (for details see Agilent RNA 6000 Nano Assay Protocol – Edition April 2007). All RNA samples had an RNA integrity number ≥8.1 (the threshold quality of RNA for microarray work is ≥7.0) (Raman *et al.*, 2009).

3.4. Sample preparation and quantitation

3.4.1. First strand cDNA synthesis

cDNA synthesis was performed using a SuperScript Double-Stranded cDNA Synthesis Kit (Invitrogen). 100 pmol oligo dT_{18} primer (Sigma-Aldrich), 10 µg of total RNA and DEPC water were mixed to a final volume of 11 µl in a 0.2 ml microcentrifuge tube. The mixture was denatured at 70°C for 10 min using a DNA Engine Tetrad 2 (DET2) PCR machine (MJ Research). Tubes were briefly spun and placed on ice for 5 min. First-strand synthesis was carried in this mixture by 400 U (unit/number of enzymes) of Superscript II at 42°C for 60 min in the presence of 1x First Strand Buffer, 0.01 M Dithiothreitol (DTT) and 0.5 mM deoxyribonucleotide triphosphate (dNTP) mix.

3.4.2. Second strand cDNA synthesis

The first reaction was submitted to second strand synthesis by incubating at 16°C for 2 h in the presence of 1x Second Strand Buffer, 0.2 mM dNTP mix, 10 U DNA ligase, 40 U DNA polymerase I, 2 U RNase H and DEPC water to a final volume of 150 µl. A final incubation at 16°C for 5 min with 10 U of T4 DNA polymerase

allowed extension and 5' overhang complementarity forming a blunt-ended cDNA duplex. Addition of 1/16th volume of 0.5 M EDTA inactivated polymerases and completed the synthesis reaction.

3.4.3. RNase A cleanup

Samples were gently mixed and incubated with 2.1 U of RNAse A (Sigma-Aldrich) at 37°C for 1 h 30 min. 2 ml Phase-lock tubes (Qiagen) were centrifuged at 12,000 rcf for 2 min. The samples were transferred into the phase-lock tubes and mixed, by inversion, with 1 volume of 1:1 phenol:chloroform (Sigma-Aldrich) before being centrifuged at 12,000 rcf for 5 min. The upper aqueous layer, containing the purified cDNA, was pipetted into a 1.5 ml microcentrifuge tube.

3.4.4. cDNA precipitation

A 1/10th volume of 7.5 M ammonium acetate was added to the sample, mixed by inversion and centrifuged briefly in a microcentrifuge. Co-precipitation was achieved by adding 35 µg of mussel glycogen (Sigma-Aldrich) to the samples, mixed and centrifuged briefly. 2 volumes of 100% ethanol were added and mixing by repeated inversion and the cDNA was precipitated by centrifuging at 12,000 rcf for 20 min. The supernatant was carefully removed and the pellet washed by adding 80% ethanol and 20% water and mixing by repeated inversion. The tubes were centrifuged at 12,000 rcf for 5 min. The supernatant was removed and the washing was repeated. The pellet was dried in a SpeedVac (Savant) for 5 min, rehydrated with 20 µl of NF-H_2O and stored at -20°C.

3.4.5. Quantitation and quality control of cDNA

cDNA samples were quantified on the NUS and met the requirements of concentration ≥100 ng/µl, A_{260}/A_{280} ratio ≥ 1.8 and A_{260}/A_{230} ratio ≥ 1.8. The quality of the cDNA fragments was also analysed by running the samples in a 0.8% agarose gel (see **section 3.10.7**). The spectrum of cDNA fragments in each sample was required to have a median size ≥400 bp.

3.5. cDNA sample labelling

Labelling of cDNA samples was done using the NimbleGen One-Colour DNA Labelling Kit (Roche). 1.75 µl of β-mercaptoethanol (Sigma-Aldrich) and 998.25 µl Random primer buffer were mixed and 924 µl of this mixture was used to re-suspend the Cy3-random nonamer primers. The primers were aliquoted into 40 µl volumes in 0.2 ml PCR tubes and stored at -20°C, for use at a later date, or rested on ice for immediate use.

In a 0.2 ml PCR tube, 1 µg of cDNA, 40 µl of Cy3-random nonamers and NF-H_2O were added to a final volume of 80 µl. Samples were heat-denatured at 98°C for 10 min and cooled on ice for 10 min. A mixture of 20 µl containing 5 mM dNTPs, NF-H_2O and 100 U of Klenow fragment was added to each denatured sample. The samples were mixed, centrifuged and incubated at 37°C for 2 h. The labelling reaction was stopped by adding 10 µl of 0.5 M EDTA. To each sample was added 11.5 µl of 5 M NaCl followed by vortexing, centrifugation and transfer into a 1.5 ml tube containing 110 µl of isopropanol. Sample vortexing and incubation for 10 min at room temperature was followed by precipitation by centrifuging at 12,000 rcf for 10 min. The supernatant was removed and the pink pellet was washed with ice-cold 80% ethanol and 20% water while being dislodged from the bottom of the tube by pipetting. After centrifuging, at

12,000 rcf for 2 min, the supernatant was removed and the pellets were dried in a SpeedVac for 5 min. The pellets were dissolved by adding 25 µl of NF-H_2O and vortexing. Each sample was quantified using the NUS and 4 µg of each Cy3-labelled cDNA was transferred into a new 1.5 ml tube. The contents were completely dried in a SpeedVac on low heat for 15-60 min, depending on the volume, and submitted for hybridization.

Light-sensitive primers and amplification product were protected from light while resting on ice and in the SpeedVac. Light in the laboratory was reduced during work.

3.6. Hybridization, washing & scanning

3.6.1. Sample tracking and denaturation

The NimbleGen Hybridization System 4 (Roche) was set to 42°C, 3 h before use, for the temperature to stabilize. Each dried sample pellet was re-suspended in 3.3 µl of a unique NimbleGen Sample Tracking Control (STC) solution (Roche). The STC is a Cy3-labelled cDNA that hybridizes to a known location on the array. Each of the 12 different STCs was used to spike a different sample.

After adding the STC, the samples were vortexed and centrifuged. Then, a hybridization solution master mix consisting of 88.5 µl of 2x Hybridization Buffer, 35.4 µl of Hybridization Component A and 3.6 µl of Alignment Oligo was prepared. Each sample received 8.7 µl of this hybridization solution and was mixed by vortexing, centrifuged and denatured by incubating at 95°C for 5 min. The samples were then transferred to a heat block (Techne) at 42°C for a minimum of 5 min before being

vortexed, centrifuged and loaded onto the corresponding sub-array on the 12x135K NimbleGen slide.

3.6.2. Mixer-slide assembly

The HX12 mixer (Roche) is an adhesive solid cover which, when in contact with the slide, forms 12 individual chambers allowing the samples to hybridize onto a single sub-array on the 12x135K slides. A Precision Mixer Alignment Tool (PMAT) (Roche) and forceps (Millipore) were used to prepare a mixer-slide assembly following detailed instructions on the NimbleGen Arrays User's Guide v3.2 (Roche NimbleGen, 2009). Once assembled, the mixer-slide assembly was removed from the PMAT and placed in the slide bay of the Hybridization System. A Mixer Brayer (Roche) was used to rub the mixer and facilitate adhesion of the gasket to the slide. A CP10 Microman pipette (Gilson Scientific) was used to carefully dispense 6 µl of sample into the fill port of a sub-array chamber, pushing the chamber air out through the vent port. Before loading the samples, the pipette tip was inspected for bubbles. During loading, the pipette tip was kept perpendicular to the slide in order to avoid leakage at the fill port and gentle pressure was applied between the tip and the port to ensure a tight seal while loading the sample. After loading, a gummy port seal (Roche) was used to seal both the fill and the vent ports of each chamber. Lastly, the bay clamp was closed, the mixing panel on the hybridization system was switched on and set to mode B, and the samples were left to hybridize at 42°C for 19 h.

3.6.3. Washing of hybridized arrays

NimbleGen Wash Buffer Kit and NimbleGen Array Processing Accessories (Roche) were used to wash the 12x135K slides from the hybridization solution and unbound labelled cDNA.

Washes I, II and III were prepared, separately, by adding 243 ml of NF-H_2O, 27 ml of 10x Wash Buffer I, II or III and 27 µl of 1 M DTT to the corresponding wash tank and kept at room temperature. An additional washing solution I was prepared on a shallow dish and heated in a microwave. Once the temperature dropped to 42.5°C, the mixer-slide assembly was removed from the Hybridization System, inserted into the Mixer Disassembly Tool and submerged into the warm Wash I. While submerged, the mixer was carefully peeled from the slide with the slide being removed from the Disassembly Tool and washed in the warm Wash I by gently agitating the slide for 10-15 sec. The slide was placed into a slide rack and tank containing room temperature Wash I and washed with vigorous and constant agitation by resting on a platform shaker (C. Gerhardt) at 110 rpm for 2 min. Wash I was concluded by vertically displacing the rack from the tank 10 times followed immediately by resting on paper towel, to remove excessive buffer carryover, and transfer into a tank with Wash II. The tank and slide were shaken at 110 rpm for 1 minute. The rack was vertically displaced 10 times, rested on a paper towel, transferred to the Wash III tank and washed by shaking at 110 rpm for 15 sec. The slide was removed from wash III, placed in an open basket slide carrier and dried by centrifugation (Thermo Scientific) at 1000 rpm for 4 min. Finally, the slide was inserted into a slide box and covered with the lid, protected from light, until ready to be scanned.

3.6.4. One-Colour Array Scanning

The slides were cleaned of any dust using an inert gas spray (VWR International) and scanned using an Axon 4200 AL scanner (Axon Instruments) and the GenePix Pro 6.0 program (Molecular Devices). Using GenePix Pro 6.0, the following settings for scanning were used: wavelength at 532 nm, for Cy3; PMT Gain ranging 318-378 units; laser power at 100%; pixel size at 5 μm; single line to average; and, focus position at 0 μm.

The quality of the scan was determined by the histogram generated from the array image. A good scan produced a histogram with the baseline of the curve in the range $8.5x10^{-4}$ – $2.5x10^{-5}$ of normalized counts at the 65,000 intensity level (near saturation). Values above or below the range would require re-scanning the slide at a lower or higher PMT, respectively.

3.7. NimbleScan Data Analysis

3.7.1. Feature identification plus sample purity and consistency checks

The scanned image was analysed using the NimbleScan v2.5 program (Roche NimbleGen, 2008). This program was used to divide the image into 12 sections each containing one sub-array. Design and gene description files were then loaded into the program and each sub-array image was fitted and aligned with a grid, allowing probes to be identified. Sub-arrays were considered uncontaminated with samples from other sub-arrays if each sub-array had a single STC present.

Other tools for assessing the data included the Experimental Metrics Report which provided a set of summary statistics that could help identify potential problems during the hybridization. The experimental metrics varied between arrays even if the

sample was the same but that was acceptable, to some extent, as these metrics were designed to inform on consistency by comparing arrays within and between slides. The experimental metrics set for this study and their respective values included: Signal Range <0.5; Uniformity Mean at 1000-8200 intensity; Uniformity Coefficient of Variation <0.13; Mean empty at 400 intensity; Mean experimental at 1000-8200 intensity; and, Mean Random at 450 intensity. Experimental metrics are described on the Guide to Interpreting the Experimental Metrics Report v3.0 (Roche NimbleGen, 2011).

3.7.2. Normalization and extraction of expression data

The 12 grid-aligned image files from a single slide and the corresponding design file (*.ndf), were used to create Pair Reports. The Pair Report contains the raw values for the probe intensities of each sub-array.

The data in the Pair Reports was then submitted to the Robust Multichip Average (RMA) algorithm (Irizarry *et al.*, 2003). In this analysis, the NimbleScan software normalized expression data across slides for a single tissue (e.g. all sub-arrays in all slides for kidney) using quantile normalization (Bolstad *et al.*, 2003). This analysis gave rise to two files for each sub-array. The first, named Pair File, contained information about each probe, their location within the gene, and the normalized signal intensity. In the second file, named Calls File, the probes were averaged and named with their respective genes. The Calls File contained the normalized gene expression value and accession number for each gene interrogated by the array.

Non-normalized and normalised microarray data were deposited on the Gene Expression Omnibus (GEO) database, submission number GSE42621, using a web-based interface (http://www.ncbi.nlm.nih.gov/geo/).

3.8. Validation of microarray data by qPCR

Microarray data was validated by quantitative real-time PCR (qRT-PCR/qPCR) which, in turn, was set up taking into account The Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines (Bustin *et al.*, 2009).

3.8.1. DNase treatment of total RNA

Approximately 10 µg of total RNA was submitted for DNase treatment using the Ambion TURBO DNA-free kit (Ambion). DNase treatment was carried in a 1.5 ml microcentrifuge tube by mixing the RNA with 5 µl of TURBO DNase buffer, 2 U of TURBO DNase and water to a final volume of 50 µl and incubating the mixture at 37°C for 30 min. To inactivate and precipitate the DNases, 5 µl of DNase Inactivation Reagent was added, mixed and incubated at room temperature for 5 min. The contents were spun at 10,000 rcf for 1.5 min and the top layer containing the DNase-treated RNA was transferred into a fresh tube. The RNA sample was quantified using the NUS and stored at -80°C.

3.8.2. First-strand cDNA synthesis

The High Capacity RNA-to-cDNA Master Mix (MM) (Invitrogen) was used to convert 1 µg of DNase-treated RNA into single-stranded cDNA. Each positive reverse transcriptase (+RT) reaction involved mixing the RNA sample with 10 µl of 2x RT buffer, 1 µl of 20x Enzyme mix and water to a final volume of 20 µl in 0.2 ml PCR tubes (Thermo Scientific). Each negative RT (-RT) reaction was prepared similarly to the +RT reaction but the Enzyme mix was replaced with water. The reactions were incubated in a DET2 PCR machine at 37°C for 60 min. The reaction was stopped by

heating at 95°C for 5 min, left at 4°C until collection and stored at -20°C until use. To ensure consistency during qPCR, +RT reactions were triplicated for each RNA sample and pooled after cDNA synthesis.

The 20x Enzyme mix contains a MultiScribe MuLV RT enzyme and an RNase inhibitor protein. The 2x RT Buffer mix includes dNTPs, random octamers and oligo dT-16 primers.

3.8.3. Optimization of qPCR

3.8.3.1. Primer testing using KAPA *Taq* polymerase

The primers for qPCR were designed by Dr Ruth Barber following standard recommendations for qPCR primers (i.e. amplicons of approximately 100 bp, annealing temperature of 60°C) to produce amplicons with minimal secondary structure as determined by the mfold software (Zuker, 2003). Primers were designed for the *Aprt* and *Actb* reference genes and for the *Per2* and *Arntl* target genes (**Appendix I**; **Table S1**).

Regular PCR was performed by adding 1 µl of cDNA to a 10 µl reaction containing 1x Kappa buffer, 0.2 mM dNTPs, 0.2 µM of forward primer, 0.2 µM of reverse primer, 0.5 U of KAPA *Taq* polymerase (Kapa Biosystems) and water. A no template control (NTC) and no enzyme control (NEC) were also prepared. The reactions were then incubated at 95°C for 5 min, followed by 30 cycles of 95°C for 30 sec, 60°C for 30 sec and 72°C for 1 min and a final extension at 72°C for 10 min using the DET2 PCR machine. Half of the sample was loaded on a 2% agarose gel (**section 3.10.7**). If a single clear amplicon of pre-determined size was detected only for the

positive reaction and no primer-dimers were detected, the optimization would be considered successful and repeated using the 2x Maxima SYBR Green qPCR MM.

3.8.3.2. Primer testing using the 2x Maxima SYBR Green qPCR Master Mix

qPCR was performed, using the 2x Maxima SYBR Green qPCR MM (FLS), for 1 µl of cDNA (50 ng) in a 15 µl reaction containing 7.5 µl of Maxima SYBR green qPCR MM (2x), 0.2 µM of forward primer, 0.2 µM of reverse primer and water. Control reactions were also prepared and included a NTC, a NEC and a positive control (PC) reaction using a gDNA template. Amplifications for each cDNA sample were performed in triplicate and incubated in a DNA Engine Opticon 2 PCR machine (MJ Research) with cycling conditions comprising 95°C for 10 min and 40 cycles of 95°C for 15 sec, 60°C for 30 sec and 72°C for 30 sec.

Each sample was analysed for the presence of a single peak using the melting curve option on the Opticon Monitor v3.1 software (MJ Research). The single peak displayed by the software infers the amplification of a unique cDNA sequence.

The 2x MM contains a Hot Start *Taq* polymerase, SYBR Green qPCR Buffer, the fluorescent dye SYBR green I and a combination of dUTP and dTTP. To troubleshoot PCR contamination, dUTP was incorporated in the PCR product. In the event of contamination, treatment of the reaction with Uracil-DNA glycosylase prior to PCR cleaves amplicons carried over from previous reactions.

3.8.3.3. qPCR efficiency test

A qPCR efficiency test was performed to confirm a constant PCR amplification efficiency of gene expression in all biological samples analysed and to ensure a reliable

comparison between samples and genes (Filion, 2012, p.58 [book]). For accurate gene expression quantification, it is essential for the control and test samples (controls and paternally irradiated offspring) to have similar RT and PCR efficiencies (Kennedy and Oswald, 2011, p.75 [book]). There are multiple factors able to reduce the efficiency of RT and PCR reactions which can be avoided mainly by using high quality reagents and materials and, by careful design, standardisation and execution of RNA extraction and qPCR procedures. Other factors are sporadic and unavoidable meaning that the reaction efficiency will never exactly be 100% (Filion, 2012, pp.56, 59 [book]). For this reason, a qPCR efficiency correction is included in the qPCR analysis software used (Pfaffl, 2001).

The qPCR efficiency test was carried for each target and reference gene using cDNA generated as described on **section 3.8.2** but from a pool of RNA samples including 6 control and 6 test group individuals belonging to the same strain and tissue. The resulting qPCR efficiency data was submitted as the reaction efficiency for each gene. Reactions with 0.5, 1, 5, 10 and 50 ng of cDNA template were prepared using the Maxima SYBR Green qPCR MM. Control reactions were also prepared and included a NTC, NEC and PC. All reactions were prepared (as described on **section 3.8.3.2**) and run in triplicate. To generate an accurate standard curve, all reactions belonging to the same strain, tissue and amplifying the same sample mixture, regardless of the cDNA input/dilution, were run on the same PCR plate. This made the fluorescence threshold the same for all reactions (see below) and avoided introducing more variation into the assay.

The qPCR data was analysed using Opticon Monitor v3.1 which plotted fluorescence intensity during the amplification reaction against the cycle of PCR. A fluorescence threshold, for cycle threshold value (C_t) determination, was automatically

set up by the software as close as possible to the baseline of the exponential phase of amplification, while avoiding background noise. The C_t values were averaged for the replicates in the same assay.

To calculate qPCR efficiency, cDNA inputs were log10 transformed and plotted against the corresponding C_t value to generate a scatter plot called the standard curve. A line of best fit uses Pearson's correlation coefficient (R^2) to measure the linearity of the standard curve and should have an R^2 value >0.985 for the line to be a good fit to the standard curve (Kennedy and Oswald, 2011, pp.76-77 [book]). The efficiency of the qPCR reaction (E) was derived from the slope of the standard curve, as described by the equation $E = 10^{-1/slope}$.

qPCR reactions ideally had an efficiency of 100%, calculated as $(E - 1) \times 100$, which corresponds to a $slope = -3.32$ and $E = 2$, meaning that after each cycle the amount of initial cDNA doubled. The qPCR was considered optimised at 90-110% efficiency ($slope = [-3.58, -3.10]$) (Kennedy and Oswald, 2011, pp.75-77 [book]). qPCR efficiency data for each gene, generated from a pool of all samples from both groups but from the same strain and tissue, was submitted to the Relative Expression Software Tool (REST) 2009 software (Pfaffl *et al.*, 2002).

3.8.4. Relative quantification of gene expression using qPCR data

After obtaining the optimum amplification efficiency for each gene, the level of gene expression was determined.

Gene expression was analysed by qPCR for 6 individuals from each experimental group at a common cDNA input/dilution (5 ng). Reactions were prepared using the Maxima SYBR Green qPCR MM (as described on **section 3.8.3.2**) and included 3 cDNA template reactions (which also inform on intra-assay variation), 2

NEC and 1 NTC for each gene and animal sample combination. Each PCR plate contained 2 control samples and 2 test samples from the same strain and tissue. For each sample, 2 reference genes and 2 target genes were analysed. All reactions were repeated on a separate run (technical replicate) to measure inter-assay variation or reproducibility. A manual fluorescence threshold fixed at 0.05 fluorescence intensity for all reactions allowed reliable comparison between genes and samples and accurate gene expression quantification. The C_t values were extracted and averaged for same assay replicates and technical replicates. qPCR gene expression data for each experimental group, belonging to the same strain and tissue, was submitted to the REST 2009 software.

3.9. DNA extractions for SM-PCR

All materials, including solutions, used for DNA extraction and PCR and were clean to single molecule PCR (SM-PCR) standards. All work was conducted in a Class II Microbiological Safety Cabinet (WSCL).

3.9.1. SM-clean somatic DNA extractions

On day 1 of extraction, 15 ml High Density Phase Lock tubes (Qiagen) were spun in a swing out rotor at 1500 rcf for 5 min and then 1 ml of Lysis Solution B and 22.5 U of Proteinase K was added (Sigma-Aldrich). A whole somatic tissue was divided into 2-4 sections and chopped finely on a Petri dish using a scalpel blade. In the case of spleen, only $1/5^{th}$ (about 15.6 mm^3 volume) of the tissue was used for extraction and divided into 2 sections. Each section received 1 ml of Lysis Solution A on the Petri dish, was chopped further and transferred into a 15 ml Phase Lock tube using a Pasteur pipette (Fisher Scientific). The tubes were inverted and incubated at 55°C overnight or

for a minimum of 5 h. On the following morning the homogenate was significantly less viscous and more translucent otherwise, an additional 22.5 U of Proteinase K was added and the incubation prolonged for 1-2 h.

Purification, re-suspension and quantification of the somatic DNA were performed as described elsewhere (see **sections 3.9.3** and **3.9.4**).

3.9.2. SM-clean sperm DNA extractions

Extraction of DNA from sperm was carried for two individuals at the same time with both caudal epididymides being used for each. The excess fat was cut off on a Petri dish using a scalpel blade and the tissue finely chopped in all directions until it was mushy. To release the tissue from the dish, 1 ml of PBS (Sigma-Aldrich) was added and mixed. A metal mesh (Sigma-Aldrich) was folded and bent into a cone shape and placed on top of a screw top 1.5 ml tube (TreffLab). The tissue in PBS was transferred into the centre of the cone and gently pipetted up and down.

A 5 µl aliquot of the filtrate was placed onto the haemocytometer (Weber Scientific International) and visualised under a light microscope (Nikon) to establish sperm concentration, mobility and somatic tissue contamination. The remaining filtrate was spun at 15,700 rcf for 2 min, to form a yellow pellet. After the supernatant was completely removed, 1 ml of ice-cold 1x SSC was added and the pellet completely re-suspended by vortexing. Lysis of residual somatic cells was concluded by adding 10 µl of 10% SDS and mixing by inverting several times. The solution was centrifuged at 15,700 rcf for 2 min to form a pure white pellet, containing only sperm cells. The supernatant was completely removed and the pellet re-suspended in 960 µl of 0.2x SSC by vortexing. The confirmation of somatic cell lysis was done on a light microscope by loading a 5 µl aliquot onto the haemocytometer. Sperm cell lysis was performed by

adding 70 µl of β-mercaptoethanol (Sigma-Aldrich), 100 µl of 10% SDS and 15 U of Proteinase K to the remaining tissue solution, mixing by inversion and incubating at 37°C for 2-3 h, with occasional mixing. If the tissue digestion was incomplete, 15 U of Proteinase K was added and the mixture was incubated for an additional 1 h.

Purification, re-suspension and quantification of the sperm DNA were performed as described elsewhere (see **sections 3.9.3** and **3.9.4**).

3.9.3. Removal of RNA and proteins from DNA

The homogenate containing somatic DNA was purified of RNA by adding 32 U of RNase A and 1/10th volume of RNase buffer, mixing by inversion and incubating at 37°C for 1 h.

For either the somatic and sperm DNA, purification was achieved by adding 1 volume of 1:1 phenol and chloroform into each tube, mixing by inversion and centrifuging at 1500 rcf for 5 min. An additional 1 volume of chloroform was added, to remove traces of phenol from the aqueous phase, mixed by inversion and centrifuged at 1500 rcf for 5 min. The resulting aqueous solution was submitted to an ethanol wash to remove any unwanted salts (as described in **section 3.9.4**).

3.9.4. Ethanol washing of DNA

The aqueous phase containing DNA was transferred into clean 15 ml tube (Corning) containing 3 volumes of 100% ethanol and 10% 3M NaAc, pH 5.2. The DNA pellet was transferred to 1.5 ml tube and washed with 500 µl of 80% ethanol and 20% water and spun at 15,700 rcf for 30 sec. The supernatant was removed and DNA samples were vacuum-dried for 5 min in the SpeedVac, re-suspended with NF-H_2O and

stored at -20 °C. The DNA concentration was measured on the NUS and the quality of DNA samples was verified on a 0.8% agarose gel after running for 30 min at 120 V.

3.10. Bisulfite Sequencing of DNA

3.10.1. SM-Sodium Bisulfite Conversion of DNA

Sodium bisulfite conversion of DNA was performed as described in the EpiTect Bisulfite Handbook but under SM-PCR-clean conditions (Qiagen, 2009a). According to Rita Neumann (see **Acknowledgements**), the Qiagen protocol does not deliver the publicised conversion efficiency of >99% meaning that modifications were made to the original protocol in order to improve conversion efficiency.

In the original protocol, the DNA is denatured 3 times in the presence of sodium bisulfite at 95°C for 5 min before being incubated at 60°C for 25, 85 and 175 min (Qiagen, 2009a). The most robust temperature for bisulfite conversion is 55°C with complete conversion occurring between 4-18 h (Grunau *et al.*, 2001). Both deamination and DNA degradation proceed faster at higher temperatures, as stated by Grunau *et al.* (2001) who observed a time window of 14 h for 55°C and 1 h for 95°C. After the first DNA denaturation, during bisulfite conversion, there is no need to denature for long periods of time on the following cycles because at >25°C the rate of deamination is superior to the rate of DNA renaturation (Grunau *et al.*, 2001). Thus, the denaturation time (at ≥95°C) was approximately the same but was divided into more cycles while the total reaction time was increased from ~5 h to ~9 h. The DNA was incubated at 99°C for 3 min before being cycled 25 times at 96°C for 30 sec and 55°C for 20 min.

From Rita Neumann's (*personal communication*) suggested 1-2 µg range, only 1 µg of DNA was submitted for conversion in order to ensure complete deamination of

the DNA. Clean-up of the bisulfite converted DNA was performed as described on the EpiTect Bisulfite Handbook and yielded 40 µl of sample (Qiagen, 2009a). All samples were aliquoted to reduce damage from repeated freeze/thaw cycles and stored at -20°C.

3.10.2. Primer design for PCR and sequencing primers

DNA sequences were obtained from the University of California Santa Cruz (UCSC) Genome Browser (found at: http://genome.ucsc.edu/). These were used to design primer pairs and confirm sequencing results. Primers were designed following the Marmur formula, $T_m = 2\,°C \times (A+T) + 4\,°C \times (C+G)$ (Marmur and Doty, 1962). Primers were designed with an optimum length of 19-25 bp, melting temperature of 60°C, a GC content of 50-60%, with a maximum of 3 G or C nucleobases in the last 5 bases at the 3' end of the primer, absence of secondary structure in the primer and template, avoiding annealing to repeat or homopolymeric sequences and with a primer pair melting temperature mismatch of ≤ 2°C. Bisulfite converted SM-PCR primer pairs were validated by checking their unconverted sequences for individual specificity. Primers were confirmed for specificity and affinity using the UCSC *in silico*-PCR tool (freely available at: http://genome.ucsc.edu/cgi-bin/hgPcr?command=start).

3.10.3. Nested PCR circumvents problems associated with bisulfite SM-PCR

Bisulfite SM-PCR uses a chemically altered ssDNA sample that is degraded and sensitive and, as a result, leads to an overestimation of the amount of intact DNA template being submitted for PCR. Furthermore, due to low sequence complexity, primers can mispair and amplify random templates usually of smaller length than the target (Neumann, 2007, p.97 [book]).

Nested PCR overcomes this issue with an initial/primary PCR (1° PCR) that detects and increases the availability of the single DNA template of interest (Neumann, 2007, p.103 [book]). A small portion of this reaction is then fed into a secondary PCR (2°PCR) with primers that anneal onto the extremities of the copied target DNA sequence. The 2°PCR amplifies the target sequence further and generates enough PCR products to pass the visible threshold of gel electrophoresis.

3.10.4. Nested SM-PCR of multiple locus

SM-PCR at a pre-established SM dilution of the template often generates inconsistent numbers of positive reactions. To increase success of SM-PCR at a selected dilution, multiple sets of primers with non-intercalating target sequences were used to amplify these targets to levels detectable by 2°PCR. This method allows multiple dilutions to be profiled for each target sequence with reduced workload and cost. The 2°PCR used a single primer pair to extract one region of interest and amplify it to a level detectable by gel electrophoresis. No primary product from another sequence was detected in a secondary amplification at any time.

Primer pairs with non-intercalating targets were designed to amplify regions of similar length in order to be used in the same SM-PCR reactions. These primers targeted promoter, exonic and intronic regions of the *Per1*, *Per2*, *Per3*, *Arntl* and *Npas2* loci. Primer sets were tested together to ensure that there was no preferential amplification. The combinations started with 5 primer sets and were decreased as one or more sets were found not to work. Re-shuffling of primer sets between groups was also done. In conclusion, 3 primer sets was found to be the maximum for an effective 1°PCR (primer set groupings are listed in **Appendix I**; **Table S6**).

3.10.5. SM-PCR of diluted bisulfite converted DNA

3.10.5.1. Converted DNA dilution

DNA longer than 1000 nucleotides hardly survives the conversion reaction so SM is achieved at small sample dilutions: usually around 1 volume of converted DNA (at ~25 ng/µl) plus 7 volumes of SS dilution buffer (i.e. 1/8 dilution), and rarely went beyond 1/20.

The SS dilution buffer was mixed with the bisulfite converted DNA in a 1.5 ml DNA Lobind tube (Eppendorf), by pipetting up and down, and 1 µl of the mixture was used per PCR. The 16 µg/ml SS dilution buffer (see **Table 2.1**) spiked each PCR reaction with SS gDNA (Sigma-Aldrich) to reduce the adsorption of the converted DNA to the tubes. The amounts of SS gDNA in each PCR reaction varied depending on the dilution for the converted DNA (e.g. 14 ng of SS gDNA per PCR for a 1/8 dilution of converted DNA). The mixture was kept on ice before being added into PCR tubes (Thermo Scientific) containing the remaining components for the SM-PCR reaction (described on **section 3.10.5.2**). Initial testing of a sample dilution was executed in 4 PCR reactions but increased to 24 reactions, if 1 or 2 reaction(s) showed amplification.

3.10.5.2. Primary SM-PCR using multiple primer pairs

SM-PCR was performed using the Expand High Fidelity PCR System (Roche). The DNA sample was diluted (as described in **section 3.10.5.1**) and 1 µl was used for PCR. A MM was prepared and after addition each 1° PCR reaction contained 1/10th final volume of 10x buffer, 0.2 mM dNTPs, 0.02 mg/ml BSA, 1 M betaine, 1.05 U of enzyme mix (containing *Taq* DNA polymerase and *Tgo* DNA polymerase), maximum primer set concentration of 1.2 µM and water to a final volume of 10 µl. The primer

concentration depended on the number of primer sets, but all primers added were at equal molar concentration in the reaction: 3 primer sets were at a concentration of 1.2 µM per reaction in which each primer contributed 0.2 µM; 2 primer sets were at 0.8 µM per reaction with each primer at 0.2 µM; and 1 primer set was at 0.8 µM per reaction with each primer at 0.4 µM. The PCR enhancers/additives BSA (NEB) and betaine (Sigma-Aldrich) considerably improved the reliability of the 1° and 2° PCR reactions.

The 1° PCR reactions were incubated at 95°C for 5 sec, 58°C for 1 min, 65°C for 3 min, 95°C for 10 sec, 58°C for 30 sec, 65°C for 3 min followed by 40 cycles of 96°C for 20 sec, 58°C for 30 sec and 65°C for 2 min. The reactions were left at 10°C until use or stored at -20°C.

3.10.5.3. Secondary SM-PCR

On the 2° PCR, the MM was added first followed by 0.5 µl of 1° PCR product. After adding the MM, each 2° PCR reaction contained $1/10^{th}$ final volume of 10x buffer, 0.2 mM dNTPs, 0.02 mg/ml BSA, 1 M betaine, 1.05 U of enzyme mix, 0.8 µM of a single primer set, and water to a volume of 9.5 µl.

The 2° PCR was carried out at 95 °C for 1 min followed by 40 cycles of 96 °C for 20 sec, 56 °C for 1 min and 66 °C for 2 min. The reactions were left at 10 °C until use or stored at -20 °C. The 2° PCR products were analysed for amplicon size and the presence of primer dimers on an agarose gel (**section 3.10.7**).

3.10.6. PCR of undiluted bisulfite converted DNA

This stage did not require using materials/reagents from SM-PCR however, doing so saved stock DNA sample and reagents. Thus, all work was executed under SM-PCR-clean conditions.

Conventional 1° PCR was performed as described before for the SM-PCR (**section 3.10.5.2**) but differed in that 1 µl of undiluted DNA was added into a PCR tube, followed by a single pair of primers and the MM containing SS dilution buffer (made to provide 1.4 ng SS gDNA per PCR). The conventional 2° PCR was performed as described before for the SM-PCR (**section 3.10.5.3**) but differed in that a pair of primers were added followed by the remaining MM and the 1° PCR feed. The 2° PCR products were analysed for amplicon size and the presence of primer dimers, on an agarose gel (**section 3.10.7**).

Bisulfite conventional PCR involved the preparation of reactions containing a single pair of primers and different DNA samples, making the reaction assembly less dependent on the MM.

3.10.7. Detecting PCR amplicons on an agarose gel

Depending on the size, PCR products were resolved on 0.8-2% agarose gel in 0.5 x TBE buffer. 2-8 µl of a PCR product in $1/5^{th}$ final volume of 5x Loading Dye (total volume 10 µl) were loaded and run at 225 V for 45 min-2 h. PCR products were visualised on a UV-transilluminator (Syngene).

3.10.8. Successful SM amplification determined by Poisson distribution

Each bisulfite converted DNA, was submitted to 24 reactions of 1° and 2° PCR. A minimum of 9 reactions without PCR product (negative reaction) was indicative that SM level had been achieved. Following the Poisson distribution it is possible to determine the mean number of DNA templates for a reaction with PCR product (positive reaction). The 9/24 ratio for negative reactions represents an average of

approximately 0.98 DNA templates in each positive reaction (as described in **section 6.1.4**).

3.10.9. Preparation of PCR products for sequencing

3.10.9.1. Purification of PCR amplicons

Two methods were used to clean amplified DNA from low molecular weight components of PCR reaction. Before selecting a cleaning method, the presence of primer dimers was analysed by submitting 2 µl of a PCR product to gel electrophoresis (**section 3.10.7**). If the primers were present, PCR products were excised from the gel and DNA was extracted using the Zymoclean Gel DNA Recovery Kit (Zymo Research) with the purified PCR amplicon being eluted in 10 µl of NF-H_2O. If the primers were not present, then the DNA only needed to be purified from unincorporated primers and nucleotides. The clean-up reaction involved mixing 3 µl of PCR product with 6 U of Exo I (FLS) and 0.9 U of SAP (FLS). The contents were mixed briefly by pipetting or by using a plate mixer, and incubated at 37 °C for 15 min. The reaction was stopped by heating the mixture at 85 °C for 15 min.

3.10.9.2. Quantification of PCR amplicons

The quantification of amplified DNA was performed by running 1 µl of the purified DNA on a 1% agarose gel with mass DNA ladders λHindIII and/or φX. The gel was then taken to a UV-transilluminator and a printed photo and/or JPEG file was generated. The intensity of the DNA bands from the PCR product and ladder(s) were compared to calculate the concentration of the PCR product (Barbas *et al.*, 2007). The

equation below describes how this quantification was achieved, using the λHindIII DNA ladder as an example.

$$\frac{(A)bp\ of\ \lambda\text{HindIII fragment similar to amplicon}}{(B)\ bp\ of\ \ \lambda\text{ plasmid before digestion}} \times (C)\ ng\ of\ ladder\ in\ gel\ well = (D)\ ng\ of\ PCR\ product\ in\ the\ gel \quad \textbf{(3.1)}$$

Then,

$$\frac{(D)}{(E)\ \mu l\ of\ PCR\ product\ loaded} = (F)\ ng/\mu l\ PCR\ concentration \quad \textbf{(3.2)}$$

If there was a very large amount of amplicon and a thick band/smear was observed, either the sample was diluted and run again on a gel or a correction factor would have to be applied as described in the equation below.

$$(F) \times \frac{thickness\ of\ amplicon\ (mm)}{thickness\ of\ ladder\ fragment\ (mm)} = (G)\ ng/\mu l\ PCR\ concentration \quad \textbf{(3.3)}$$

In which the terms (A, B, C, D, E, F or G) represent numerical values of mass, volume or length depending on position and equation used.

3.10.10. Sequencing of amplified DNA

Sequencing was performed using the Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). In a 0.2 ml microcentrifuge tube, 10-30 ng/kb length of purified PCR amplicon was mixed with 1/20th volume of Big Dye Mix v3.1, 1/5th volume of 5x Big Dye Buffer, 3.2 pmol of a forward or reverse primer and NF-

H_2O to a final volume of 20 µl. The mixture was briefly vortex and centrifuged. Sequencing was carried by incubation of the mixture in a PCR machine at 96°C for 1 min followed by 36 cycles at 96°C for 10 sec, 50°C for 5 sec, and 60°C for 4 min. Samples were left at 4°C until use.

The procedure for removal of buffer, unincorporated fluorophore labelled ddNTPs and unincorporated primer was initiated by first adding a $1/10^{th}$ volume of 2.2% SDS, mixing thoroughly using a plate mixer and incubating at 98°C for 5 min and 25°C for 10 min. The Performa DTR Gel Filtration Cartridge and 1.5 ml tube (EdgeBio) were centrifuged at 1000 rcf for 3 min to concentrate the water into the gel matrix. The Cartridge was transferred into a new 1.5 ml tube, the sequencing product loaded into the Cartridge and centrifuged at 1000 rcf for 3 min. The purified eluate was then submitted to the PNACL laboratory (Centre for Core Biotechnology Services, University of Leicester) and the sequence was read from the labelled DNA by an ABI 3730 sequencer (Applied Biosystems).

Chapter 4:

Gene expression analysis in offspring of paternally irradiated mice

4.1. Data transformation and analysis

Probe intensity values from the expression summary report were $\log_2$ transformed since the distribution of raw data was substantially skewed to the right (Field *et al.*, 2012, p.192 [book]). Upon log2 transformation, each sample was checked for whether the data distribution fitted a normal distribution by: (1) visually inspecting a frequency histogram of the data; and, (2) by submitting the data to the Shapiro-Wilk normality test (described in Sokal and Rohlf, 2012, pp.422-425 [book]), which calculated a *p*-value for the hypothesis of the data being drawn from a normal distribution. The transformation also attributed equality of variance across all samples as demonstrated by the Levene's test for homogeneity of variance (described in Sokal and Rohlf, 2012, pp.414-417 [book]). These tests were carried using the Systat 13 software.

The logarithmic transformation successfully made the data distribution symmetrical and thus suitable to carry out parametric statistical tests. As already mentioned, to ensure the reproducibility of microarray data, each RNA sample was hybridised to two arrays and the mean values for each independent technical replicate was estimated from 135,054 probes. As each known mouse transcript was represented on the array by at least 3 probes, the mean values for 42,575 transcripts were then

estimated. To establish the differences between groups, the mean values and standard errors of means were estimated for each transcript on each microarray.

4.2. Microarray data validation

To ensure the reproducibility of the microarray data, each RNA sample was hybridised twice on different arrays (i.e. technical replicates). The same RNA sample has a highly similar gene expression profile across two independent replicate arrays while two different samples have less similarity, in comparison. If the microarray material is accurate and reliable and the technique is flawless, the same animal sample hybridised onto independent arrays (i.e. located on different slides) should share more similarity than even two different animal samples belonging to the same group, strain and tissue hybridised onto dependent arrays (i.e. located on the same slide). This similarity between samples can be measured, partly, through the covariance of the data points in their expression profiles.

4.2.1. Correlation between technical replicates

The reproducibility of microarray data was evaluated using Pearson's correlation coefficient, r_{xy}, represented by the equation below.

$$r_{xy} = \frac{\sigma_{xy}}{\sqrt{\sigma_x}\sqrt{\sigma_y}} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad \textbf{(4.1)}$$

where, x_i and y_i are the $\log_2$ transformed values for two samples, $\bar{x}$ and $\bar{y}$ are their respective means, n is the number of transcripts, and σ_x and σ_y are the standard deviations of the first and second sample, respectively. r_{xy} is always between [-1; 1]; value 1 represents a perfectly positive correlation, 0 represents no correlation and -1 is a

perfectly negative or inverse correlation (Drăghici, 2003, p.269 [book]; Field *et al.*, 2012, p.209 [book]).

The coefficients of correlation were estimated for each tissue within and between all technical replicates (**Table 4.1**, **Figure 4.1**). The values of correlation within technical replicates reflect the stochastic variation of the microarray technique, whereas those within biological replicates reflect, in addition to stochastic variation, individual variation between animals belonging to the same group, strain and tissue. Stochastic variation, group, strain and tissue all contribute towards variation in the pattern of gene expression.

Within technical replicates correlations measure the similarity of the same animal sample hybridised onto independent arrays.

Between technical replicates correlations measure the similarity of different animal samples hybridised onto dependent or independent arrays and includes correlations between samples from the same/different group, same/different strain and the same tissue.

As stated previously, all correlations are pairwise.

Table 4.1: Correlation within and between technical replicates.

	Within replicates		Between replicates		M-W test[1]	
Tissue	N	Mean ± s.e.m.	N	Mean ± s.e.m.	χ^2, df = 1	Prob
Brain	24	0.9539±0.0033	528	0.9529±0.0008	0.01	0.9218
Kidney	24	0.9444±0.0053	528	0.9344±0.0010	6.97	0.0083
Liver	24	0.9680±0.0011	528	0.9645±0.0003	7.40	0.0065
Spleen	24	0.9687±0.0022	528	0.9614±0.0004	22.58	2.02×10^{-6}
All tissues	96	0.9588±0.0019	2208	0.9533±0.0004	15.25	0.0001

For each tissue the mean values and standard error were calculated for coefficients of correlation within and between technical replicates. When the data for all tissues was

[1] M-W test – Mann-Whitney test.

aggregated, the mean values within replicates significantly exceeded those between replicates (the Mann-Whitney *U* test, χ^2 = 15.25 (df = 1), *P* = 0.0001). The same result was observed when the correlations were carried in a single tissue, with brain being the only exception.

Technical microarray replicates can be used to measure the stochastic variation in the experiment so that statistical tests can be applied to evaluate differences. Stochastic variation may have the potential to produce false changes in gene expression profiles of samples which may, in turn, ruin an experiment. The confirmation of the highest level of similarity within technical replicates assists in the confirmation of low levels of experimental variation. Consequently it is important to test the equality of the two classes (i.e. within and between replicates). Parametric tests are more powerful than non-parametric tests if the four assumptions of the parametric test are met: normally distributed data, homogeneity of variance, interval or ratio data for independent variables, and independence of observations across samples (Field *et al.*, 2012, pp.667, 923 [book]; Sokal and Rohlf, 2012, pp.410-425 [book]). Because the coefficient of correlation is a discrete variable, the statistical analysis must be carried by a non-parametric test. For this set of data, the Mann-Whitney *U* (M-W) test is a robust non-parametric alternative to the *t*-test for independent samples (Sokal and Rohlf, 2012, p.446 [book]).

The M-W test was used to run a test of equality of location of data points for within and between technical replicates. The null hypothesis (H_0) is that within technical replicates do not differ in the location of their observations/data points to the between technical replicates while the alternative hypothesis (H_1) is that they do. The observations on both classes were listed from smallest to largest and submitted to the

M-W test, retrieving a *U*-statistic. With samples larger than 20, the sampling distribution of the *U*-statistic rapidly approaches the normal distribution (Sokal and Rohlf, 2012, p.451 [book]). Thus, the *U*-statistic will be accompanied by a *z*-value (normal distribution random variable), and the respective *P*-value (Field *et al.*, 2012, p.659 [book]). The z-value can be calculated by using the equation below.

$$Z = \frac{U - \frac{1}{2}n1 + n2}{\sqrt{\frac{n1n2(n1 + n2 + 1)}{12}}} \quad \textbf{(4.2)}$$

Both classes seem to differ in the location of their observations for each tissue according to the M-W test ($P < 0.05$). To facilitate interpretation of the M-W test result the chi-squared (χ^2), correspondent to the *U*-statistic, is presented. This is possible because a χ^2-distribution can be generated from a normal distribution and the *U*-statistic has a normal distribution (Sokal and Rohlf, 2012, p.154 [book]). Resorting to a contingency table, the test statistic and a *P*-value were computed (**Table 4.1**). The null hypothesis was rejected as the mean values within replicates significantly exceeded those between replicates when considering all tissues together ($P = 0.0001$). Within replicate correlations were stronger than between replicates in kidney ($P = 0.0083$), liver ($P = 0.0065$), spleen ($P = 2.02 \times 10^{-6}$), but not in brain ($P = 0.9218$).

The Kolmogorov-Smirnov test was used to compare the distributions of coefficients of correlation between and within technical replicates (**Figure 4.1**).

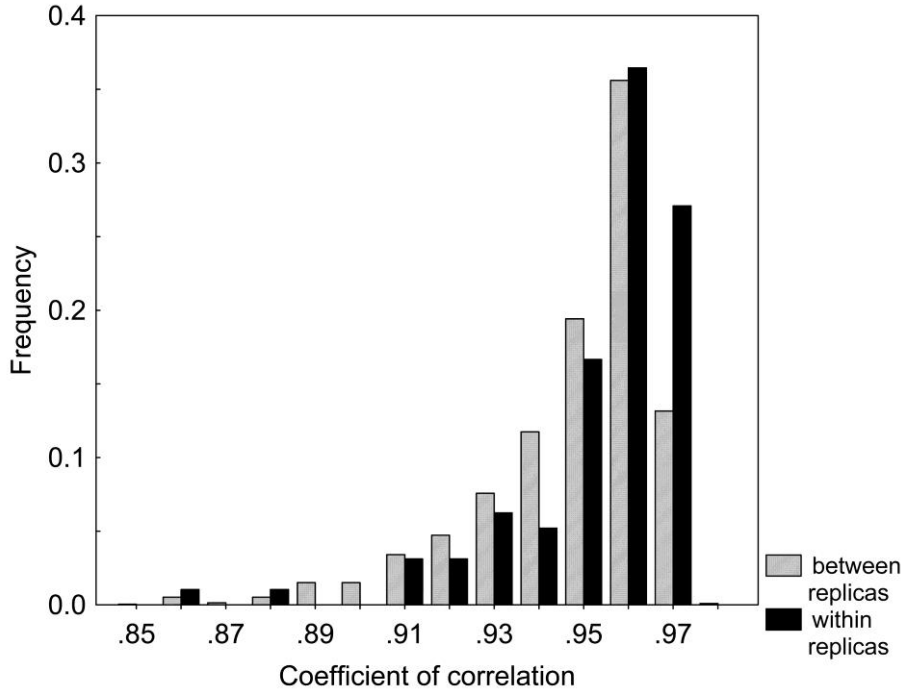

Figure 4.1: Distribution of coefficients of correlation between and within technical replicates for all tissues (the Kolmogorov-Smirnov test, $P = 0.0007$).

The Kolmogorov-Smirnov two-sample test is a non-parametric test that tests differences between two distributions. The correlation values were pooled for each class being compared and their relative cumulative frequencies were calculated. The absolute value of the difference between the relative cumulative frequencies in both samples was computed and the largest unsigned difference (D) was located. The critical value, at $\alpha = 0.05$ ($D_{.05}$) for the D-statistic, was computed (Sokal and Rohlf, 2012, pp.455-460 [book]). Because the D-statistic was larger than $D_{.05}$ the null hypothesis, that the coefficients of correlation from the within and between replicates classes were distributed identically, was rejected ($P = 0.0007$).

Given these results have validated the reproducibility of the microarray technique, the mean transcript values for the expression profile of each technical replicate was used for further analysis, as described below.

4.2.2. Correlation between mouse strains

The study of the transgenerational effects of radiation using microarray gene expression profiling required the use of two strains of mice and three biological replicates. Two strains of mice are used in the same experiment in order to ensure that any difference observed is an effect of paternal irradiation and is not strain-specific. Establishing that there are differences in the expression profiles of mice from different strains also assists in the validation of the technique. The coefficients of correlation within and between inbred strains were therefore estimated for each tissue (**Table 4.2**, **Figures 4.2-4.5**).

Within strain correlations measure the similarity of the same animal sample hybridised onto independent arrays or different animal samples from the same strain and tissue hybridised onto dependent or independent arrays. Within strain correlations include within and between technical replicates correlations of samples from the same strain and tissue.

Between strains correlations measure the similarity of animal samples from different strains, but from the same tissue, hybridised onto dependent or independent arrays. Thus, between strains correlations include, solely, between technical replicates correlations of samples from different strains but from the same tissue.

Table 4.2: Correlation within and between strains.

	Within strains		Between strains		M-W test	
Tissue	N	Mean ± s.e.m.	N	Mean ± s.e.m.	χ^2, df = 1	Prob
Brain	132	0.9753±0.0005	144	0.9681±0.0006	66.04	$< 10^{-16}$
Kidney	132	0.9606±0.0013	144	0.9521±0.0011	35.15	3.06×10^{-9}
Liver	132	0.9801±0.0003	144	0.9696±0.0003	187.84	$< 10^{-16}$
Spleen	132	0.9763±0.0005	144	0.9644±0.0005	145.43	$< 10^{-16}$
All tissues	528	0.9731±0.0005	576	0.9636±0.0004	15.25	$< 10^{-16}$

For each tissue the mean values and standard error were calculated for coefficients of correlation within and between strains. For each comparison between the two classes, the mean correlation values within strains significantly exceeded those between strains.

Similarly to the previous analysis, the M-W test was used for assessing whether the two classes of correlations had the same location of data points. The null hypothesis was rejected as the correlations in within strains were demonstrated to be stronger than those in between strains considering all tissues together ($P < 10^{-16}$) and separately for brain ($P < 10^{-16}$), kidney ($P = 3.06 \times 10^{-9}$), liver ($P < 10^{-16}$) and spleen ($P < 10^{-16}$).

The Kolmogorov-Smirnov test was used to establish whether the distributions of correlation coefficients differed between the two classes (**Figures 4.2-4.5**).

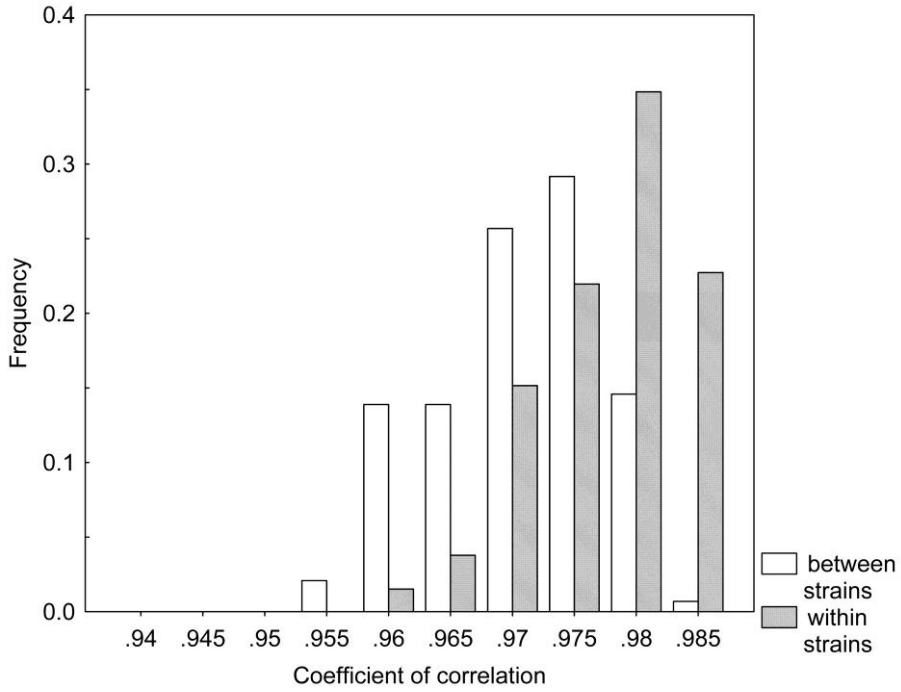

Figure 4.2: Distribution of coefficients of correlation between and within strains for RNA samples extracted from brain (the Kolmogorov-Smirnov test, $P = 1.12 \times 10^{-7}$).

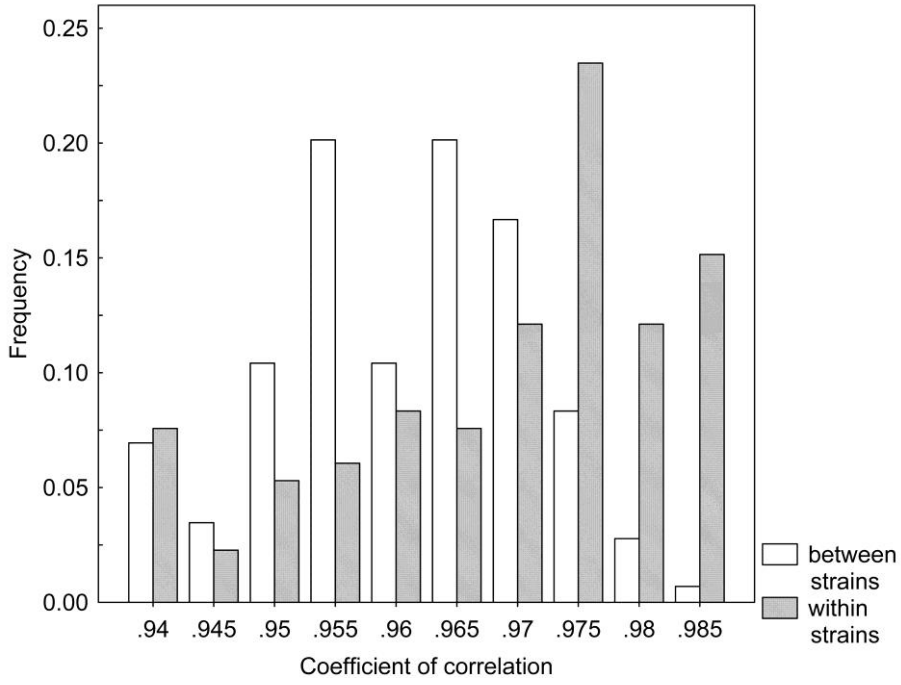

Figure 4.3: Distribution of coefficients of correlation between and within strains for RNA samples extracted from kidney (the Kolmogorov-Smirnov test, $P = 1.12 \times 10^{-7}$).

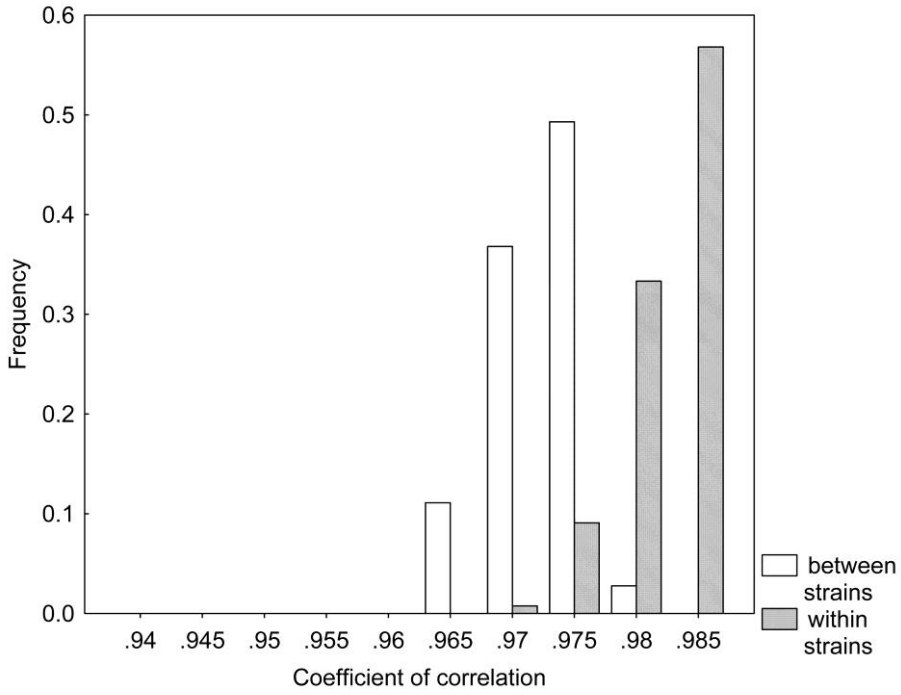

Figure 4.4: Distribution of coefficients of correlation between and within strains for RNA samples extracted from liver (the Kolmogorov-Smirnov test, $P = 1.12x10^{-7}$).

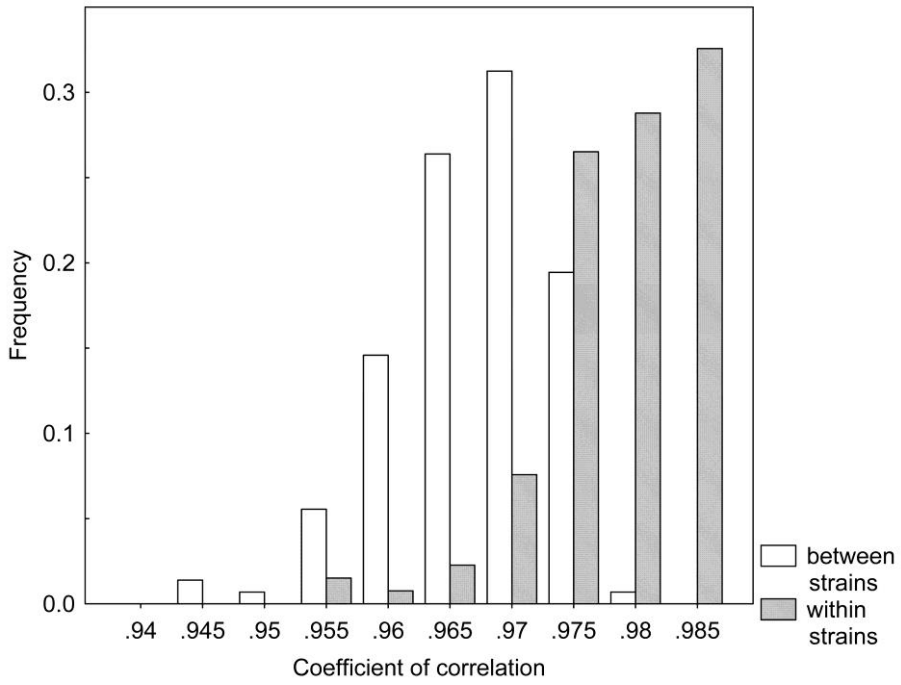

Figure 4.5: Distribution of coefficients of correlation between and within strains for RNA samples extracted from spleen (the Kolmogorov-Smirnov test, $P = 1.12x10^{-7}$).

Taken together, the results presented in this section clearly showed that the inter-strain (between strains) variation in the pattern of gene expression in all four tissues significantly exceeded the intra-strain (within strain) variation (i.e. individual variation).

The hierarchy of influence towards expression profile variability was established for strains and tissues, as described below.

4.2.3. Correlation between all RNA samples

The values of the coefficients of correlation are expected to be larger for groups of samples sharing the same biological conditions (same tissue, strain, group). To establish whether this was the case for the current set of microarray data, the coefficients of correlation between all RNA samples were estimated and compared (**Table 4.3**, **Figure 4.6**).

Within tissue correlations include within and between technical replicate correlations or, alternatively, include within and between strain correlations, from one tissue.

Between tissues correlations measure the similarity of biological samples that may or may not belong to the same animal but do belong to different tissues and have been hybridised onto independent arrays.

Table 4.3: Coefficients of correlation between all RNA samples.

	N	Range	Mean ± s.e.m	Median
Within strain&within tissue	528	0.9010 – 0.9873	0.9731 ± 0.0005	0.9766
Between strains&within tissue	576	0.8822 – 0.9800	0.9634 ± 0.0004	0.9659
Between tissues	3456	0.6122 – 0.8383	0.7190 ± 0.0008	0.7009
All samples	4560	0.6122 – 0.9873	07793 ± 0.0017	0.7454

Mean and standard error values were calculated for pairwise sample correlations including all RNA samples. The highest mean correlation was observed in within strains followed by between strains, all samples and between tissues. The median also follows

this decreasing size order for these 4 classes. The all samples class contains all sample correlations so it has both the weakest and strongest correlations creating a wide range of correlations. The between tissues class contains the weakest correlations.

Figure 4.6 presents the distribution of correlation coefficients for the three abovementioned groups. For this set of data, the null hypothesis that the classes do not differ in the location of their data points was tested using the Kruskal-Wallis test. The Kruskal-Wallis test is a non-parametric version of analysis of variance and therefore able to analyse >2 distributions. The H-statistic produced by this test is approximately χ^2 distributed since $n \geq 5$ for each class (Sokal and Rohlf, 2012, p.446 [book]). The Kruskal-Wallis test rejected the null hypothesis and confirmed that there are significant differences between the correlations from within strain, between strains and between tissues classes.

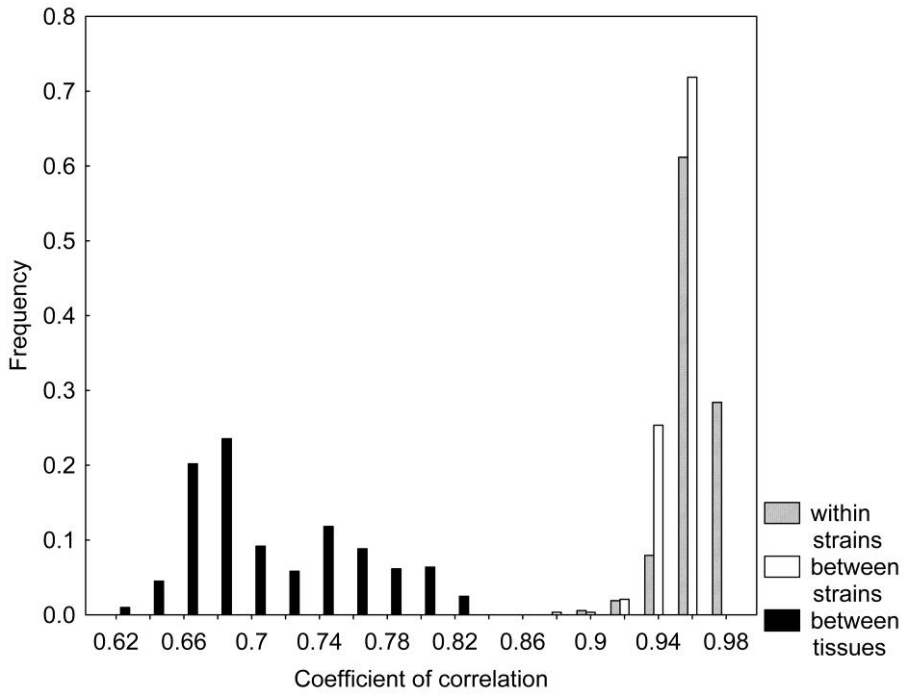

Figure 4.6: Distribution of coefficients of correlation between all RNA samples (the Kruskal-Wallis test, $\chi^2 = 2527.27$ (df = 2), $P < 10^{-16}$). The Kruskal-Wallis test estimated

that the distributions of correlation coefficients from within strains, between strains and between tissues had significantly different location of correlations.

4.2.4. Conclusion

The main objective of this part of the project was to validate the microarray technique used. As already mentioned, the differences in the expression profiles within technical replicates reflect the stochastic variation of microarray technique. The results described above clearly show that such variation was minimised as the highest values of correlation coefficients between RNA profiles were found within technical replicates. The expression profiles were shown to be closer in samples from the same strain or tissue than in samples from different strains or tissues. Thus, group, strain and tissue variation were also minimal and did not affect the results.

4.3. Global differences in the expression profiles

The key objective of this study was to establish the transgenerational effects of paternal irradiation on the pattern of gene expression in their first-generation offspring. Similarly to previous analyses, the mean values for 42,575 transcripts were used and technical replicates were not averaged. These data were used to re-measure the inter-strain and inter-tissue influence on variation of expression profiles using a different approach. Additionally, this new approach was used to measure the influence of paternal irradiation on gene expression profiles.

Microarrays measure genes in isolation but, subsequent analysis can identify relationships between genes and identify genes or samples that behave in a similar or coordinated manner. Analysis of microarrays can concentrate on the similarity between genes or on the similarity between samples. The choice of analysis depends on the

experimental objectives but the analytical methods are the same for both. For this study, we were interested in the similarity of genes meaning that each gene was measured by the samples (Stekel, 2003, pp.139-141 [book]).

Establishing the pattern of differences of the transcription profiles requires a distance metric for measuring the difference between samples. A distance metric takes two points in space and calculates a positive number that appraises the proximity of those two points. Distance metrics must consider three theoretical properties: symmetry, positivity and triangle inequality. That is, for a distance metric d with two points x and y and a third point z in an *n*-dimensional space of real numbers, the equations below must be true (Drăghici, 2003, pp.264-265 [book]).

$$d(x, y) = d(y, x) \quad \textbf{(4.3)}$$

$$d(x, y) \geq 0 \quad \textbf{(4.4)}$$

$$d(x, y) \leq d(x, z) + d(z, y) \quad \textbf{(4.5)}$$

These three properties are shared by all distance metrics (Drăghici, 2003, p.264 [book]). Some examples of distance metrics include: Euclidean distance, Manhattan, Squared Euclidean, Correlation distance and Chebychev distance. These distances are calculated between two *n*-dimensional vectors, $x = (x_1, x_2, \dots, x_n)$ and $y = (y_1, y_2, \dots, y_n)$ (Drăghici, 2003, pp.265-270 [book]).

The Manhattan distance (**equation 4.6**) is calculated along directions parallel to the *x* and *y* axis and, consequently, has the disadvantage of changing dramatically the properties of the genes. When compared to Euclidean distance, the Manhattan distance tends to retrieve a large numerical value for the same relative position of the data points, leading to a slight emphasis on the outliers of a data set (Drăghici, 2003, pp.266-268 [book]).

$$d_M(x,y) = \sum_{i=1}^{n} |x_i - y_i| \quad \textbf{(4.6)}$$

The Squared Euclidean distance (**equation 4.7**) lacks the square root of the Euclidean distance and tends to form less compact clusters, emphasizing the distances and, consequently, the outliers (Drăghici, 2003, pp.270, 275 [book]).

$$d_{E^2}(x,y) = \sum_{i=1}^{n} (x_i - y_i)^2 \quad \textbf{(4.7)}$$

The Chebychev distance (**equation 4.8**) is used when the objective is to detect any big difference between vectors. It is resistant to small amounts of noise and to outliers but it will be affected by a large change in the coordinates of one vector caused by noise or measurement error (Drăghici, 2003, p.268 [book]).

$$d_{max}(x,y) = \max_i |x_i - y_i| \quad \textbf{(4.8)}$$

The correlation distance can use the Pearson correlation coefficient (parametric; **equation 4.1**) or the Spearman correlation coefficient (non-parametric; **equation 4.1** but using ranked data) (Kimmel and Oliver, 2006, pp.195-197 [book]). The correlation coefficient (r_{xy}) is a measure of similarity but can be converted into a distance metric using the formula: $d_R(x,y) = 1 - r_{xy}$ (Drăghici, 2003, p.269 [book]; Stekel, 2003, p.143 [book]). The correlation distance will take values between 0 and 2 and positive correlations will take distances in the interval [0, 1[, no correlation takes a distance of 1 and inverse correlations take distances in the interval]1, 2]. The Pearson correlation distance looks for similar variation as opposed to similar numerical values (Drăghici, 2003, pp.269, 275 [book]). It is powerful, spots positive and negative correlations and is

scale-invariant on centred data (i.e. correlation between two samples containing raw or $\log_2$ values will produce the same result if the data on both samples is centred). It has the disadvantage of requiring linear relationship between two interval variables and being susceptible to outliers (Stekel, 2003, p.152 [book]; Field *et al.*, 2012, p.219 [book]). The Spearman correlation distance also spots positive and negative correlations and, because it is based on ranking data, it is robust to outliers and completely scale-invariant, even in the absence of centering. The Spearman correlation is less powerful than the Pearson correlation if the data obeys parametric assumptions and has low noise (Stekel, 2003, pp.147-148, 152 [book]).

The Euclidean distance (**equation 4.9**) measures the distance between two points, x and y, in a multi-dimensional space and is described by the equation below.

$$d_E(x,y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad \textbf{(4.9)}$$

The Euclidean distance can take values in the interval [0, +∞] where 0 represents no distance/difference between RNA samples and a positive value quantifies the distance (Drăghici, 2003, pp.264-265 [book]). Euclidean distance has geometric interpretation, can retain up-and down-regulation information with appropriate scaling and can detect magnitude of changes, if used without scaling. The disadvantages are that it is not scale-invariant, since the results vary with different scales, and it cannot detect negative correlations (Stekel, 2003, p.152 [book]).

In this experiment, an outlier is a raw data point equal to 65,536 intensity value and is obtained when a probe is saturated. The combined use of saturated control probes and histogram assessment provide control over scanning and avoids image files with saturated test probes. Additionally, the RMA analysis is capable of correcting these

outliers using a median polish prior to background correction and normalization (Irizarry *et al.*, 2003). Following the RMA analysis, outliers were not found in the Pair File that, in turn, contains the normalised probe values for a sample. Additionally, box plots showed that all samples were identically distributed and, again, were lacking outliers. In a frequency histogram of probes values from any sample, the raw intensity distribution is skewed to the right as lowly-expressed transcripts are more common than highly-expressed transcripts. Highly-expressed transcripts were observed to have a wider range of intensity values across samples, meaning that there is higher variation in the data points located at the high intensity tail of the distribution. The variation between all samples was assessed in Systat 13 using the Levene's test for homogeneity of variance (as described in Field *et al.*, 2012, pp.415-416 [book]) and it was concluded that there was heterogeneity of variance (heteroscedasticity) across some samples. This variation at the higher intensity data points would drive cluster formation, depending on the linkage rule and distance metric, during hierarchical clustering. This could be compensated by using a linkage rule based on the smaller distances such as Single linkage; however, the clusters from microarray data tend to be divided into more subclusters with single or average linkage (Kimmel and Oliver, 2006, p.205 [book]). The other option would be to use the Standardised Euclidean distance which is capable of reducing variation in either the x or y axis by dividing each Euclidean distance by the standard deviation of each dimension/data point (Drăghici, 2003, pp.270-272 [book]). However, because the variation is present along both the x and y axis (i.e. in all samples) this distance cannot be used. Log_2 transformation attributed homogeneity of variance (homoscedasticity) across all samples as demonstrated by Levene's test. This was possible because log_2 transformation can reduce the weight of the high intensity

values in all samples since larger values will be decreased further than smaller values upon transformation (Field *et al.*, 2012, pp.191-192 [book]).

In agglomerative hierarchical clustering the cluster formation is usually dependent on the larger distances between samples. Log_2 transformation allows these samples to be clustered according to the differences not only at the higher expressed transcripts but also at the lower expressed transcripts (i.e. the whole distribution contributes towards cluster formation between samples). Manhattan, Squared Euclidean and Chebychev distances emphasize the larger distances and would work against the log_2 transformation so, for that reason, cannot be used as the distance metric. Using the Pearson correlation for microarray validation confirmed that there are only positive correlations between samples; so, its use as a distance metric would not add new information. Because the current set of data was transformed, the Euclidean distance would not be able to detect the real magnitude of differences between vectors. However, it would be able to detect magnitude differences relative to the log_2 transformation. This is not possible for correlation distances as the denominator of **equation 4.1** erases the magnitude difference between vectors (Drăghici, 2003, p.269 [book]).

Thus, the Euclidean distance was used in this project to measure the distance between the mean expression values of each gene from two animal samples.

4.3.1. Measurement of differences in expression profiles between experimental groups using Euclidean distance

It is important to quantify the effects of paternal irradiation on the expression profiles of the offspring before focusing the analysis on single genes. To measure differences between expression profiles of every two RNA samples, the Euclidean distance was estimated and the resulting matrix of distances was subjected to

agglomerative hierarchical clustering. Agglomerative hierarchical clustering is a method for the exploratory analysis of gene expression microarray data that does not specifically test any particular hypothesis (Kimmel and Oliver, 2006, p.198 [book]). For this reason, the Euclidean distances between samples from the same or different groups were first subjected to a number of statistical tests.

Table 4.4 presents the mean values of Euclidean distances within and between the experimental groups. Within each inbred strain, the distances between the offspring of irradiated and control male mice were separately analysed. **Figure 4.7** shows the distribution of distances for all four tissues.

The Euclidean distance is a discrete variable and, therefore, can only be analysed by non-parametric tests. Both the M-W and Kruskal-Wallis tests are based on ranks and measure the differences in location of data points. The Kolmogorov-Smirnov test is a more comprehensive test that measures differences in both location and shape of two distributions but is weaker than the M-W test in detecting differences in location (Sokal and Rohlf, 2012, pp.457, 460 [book]). The Kolmogorov-Smirnov test is sensitive to differences in location, dispersion and skewness and is useful in detecting kurtosis and bimodal distributions (Sokal and Rohlf, 2012, p.455 [book]).

Within group distances measure the similarity of the same animal sample hybridised onto independent arrays or different animal samples from the same group, strain and tissue (i.e. biological replicates) hybridised onto dependent or independent arrays.

Between group distances measure the similarity of different animal samples from different groups, but from the same strain and tissue, hybridised onto dependent or independent arrays.

Table 4.4: Euclidean distances within and between experimental groups.

	Within groups		Between groups		M-W test	
Tissue	N	Mean ± s.e.m.	N	Mean ± s.e.m.	χ^2, df = 1	Prob
Brain	60	82.12 ± 1.44	72	82.83 ± 1.36	0.02	0.8837
Kidney	60	113.95 ± 3.85	72	119.58 ± 4.04	1.04	0.3081
Liver	60	81.01 ± 1.26	72	82.69 ± 1.00	1.66	0.1975
Spleen	60	116.18 ± 3.57	72	110.39 ± 3.18	1.66	0.1975
All tissues	240	98.31 ± 1.76	288	98.87 ± 1.66	0.06	0.8072

Mean and standard error of the Euclidean distances within and between experimental groups, within each inbred strain and tissue, were estimated and compared using the M-W test. No differences in the location of sample distances were detected between the 2 classes of distributions for each single/group tissue(s).

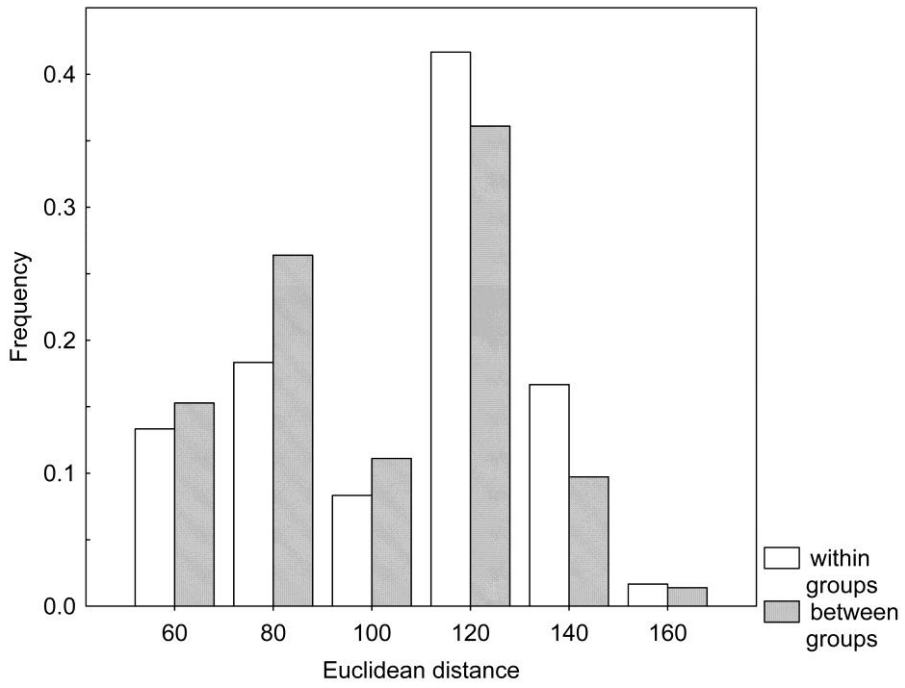

Figure 4.7: Distribution of Euclidean distances within and between experimental groups for all RNA samples (the Kolmogorov-Smirnov test, P = 0.8072). No significant differences were detected between the two distributions.

According to the M-W and Kolmogorov-Smirnov tests, the within groups and between groups distance distributions did not differ significantly when the data was

aggregated for all tissues. Additionally, the M-W test detected no differences in location between both distributions for each tissue. In conclusion, the two classes of Euclidean distances did not differ significantly because the global gene expression profiles in offspring of irradiated and offspring of control mice did not differ significantly.

4.3.2. Distance between mouse strains

Similarly to the Pearson's correlation coefficient, the Euclidean distance was used to measure the influence of strain type on the pattern of gene expression.

Table 4.5 presents the mean values of Euclidean distances within and between the two inbred strains. **Figures 4.8-4.11** compare the distributions of intra- and inter-strain distances in each tissue. The statistical analysis of these data confirmed the very significant level of inter-strain variation in all tissues.

Table 4.5: Euclidean distances within and between inbred strains.

	Within strains		Between strains		M-W test	
Tissue	N	Mean ± s.e.m.	N	Mean ± s.e.m.	χ^2, df = 1	Prob
Brain	132	82.51 ± 0.99	144	100.33 ± 0.86	118.64	$< 10^{-16}$
Kidney	132	117.02 ± 2.81	144	135.63 ± 1.94	39.74	2.89×10^{-10}
Liver	132	81.93 ± 0.79	144	112.23 ± 0.46	204.39	$< 10^{-16}$
Spleen	132	113.02 ± 2.38	144	138.38 ± 1.87	48.41	3.45×10^{-12}

Mean and standard error of the Euclidean distances within and between inbred strains were estimated and compared using the M-W test. The distances within strains were significantly smaller than those between strains for each tissue analysed.

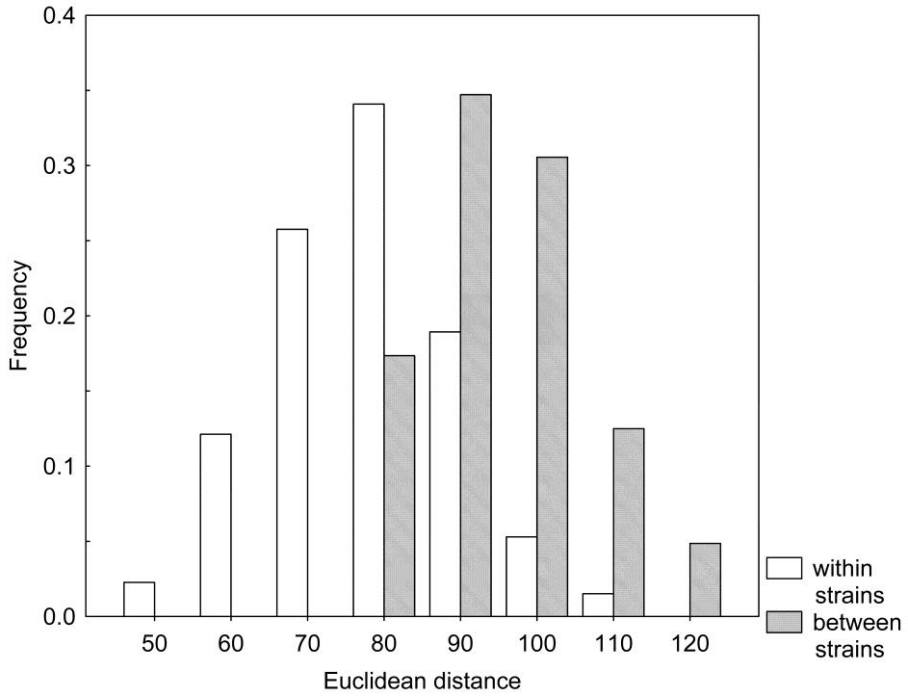

Figure 4.8: Distribution of Euclidean distances within and between strains for RNA samples extracted from brain (the Kolmogorov-Smirnov test, $P = 1.19 \times 10^{-7}$).

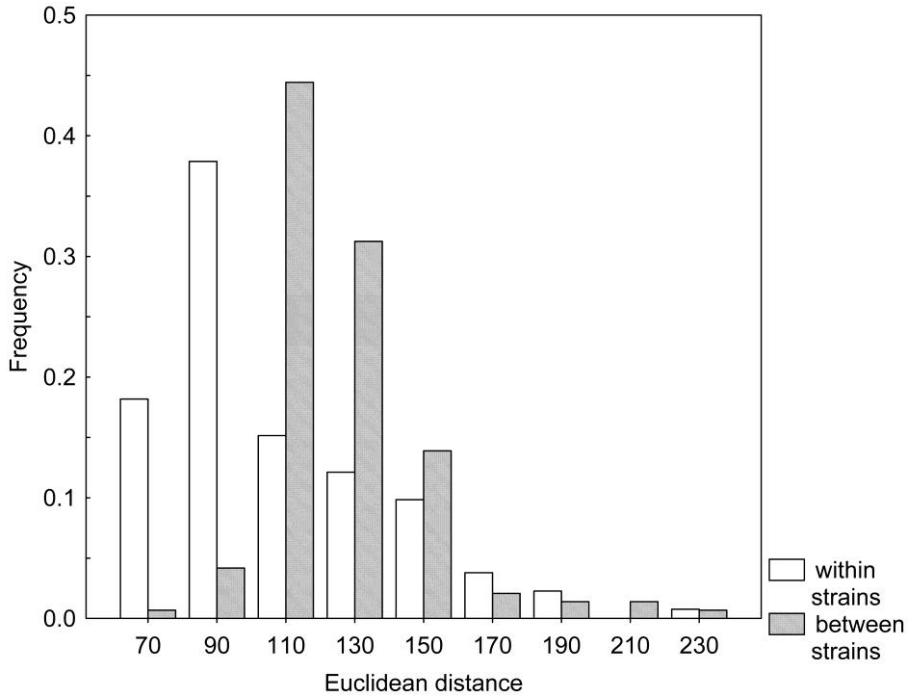

Figure 4.9: Distribution of Euclidean distances within and between strains for RNA samples extracted from kidney (the Kolmogorov-Smirnov test, $P = 1.19 \times 10^{-7}$).

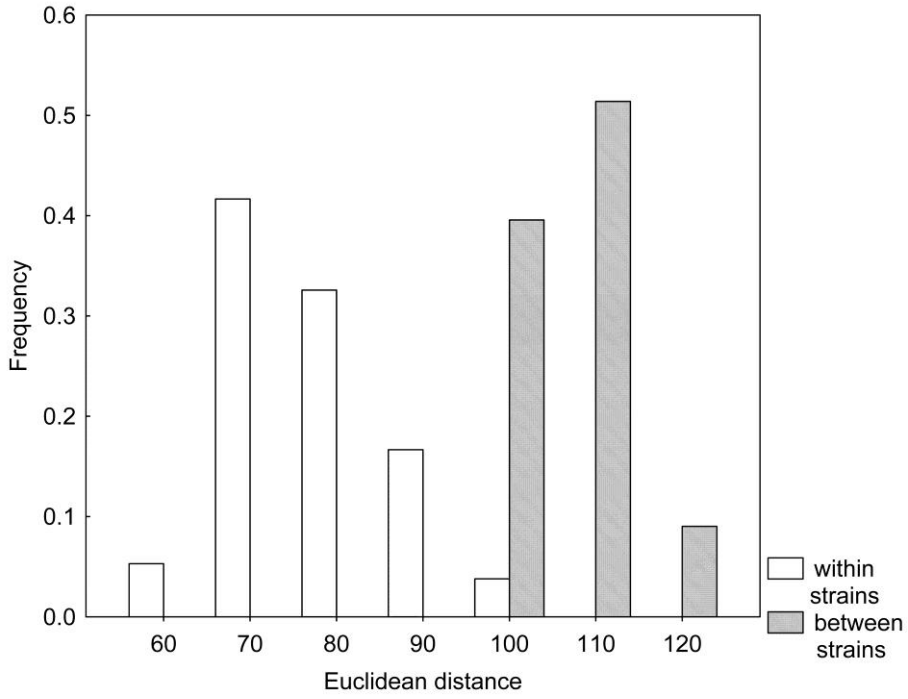

Figure 4.10: Distribution of Euclidean distances within and between strains for RNA samples extracted from liver (the Kolmogorov-Smirnov test, $P = 1.19 \times 10^{-7}$).

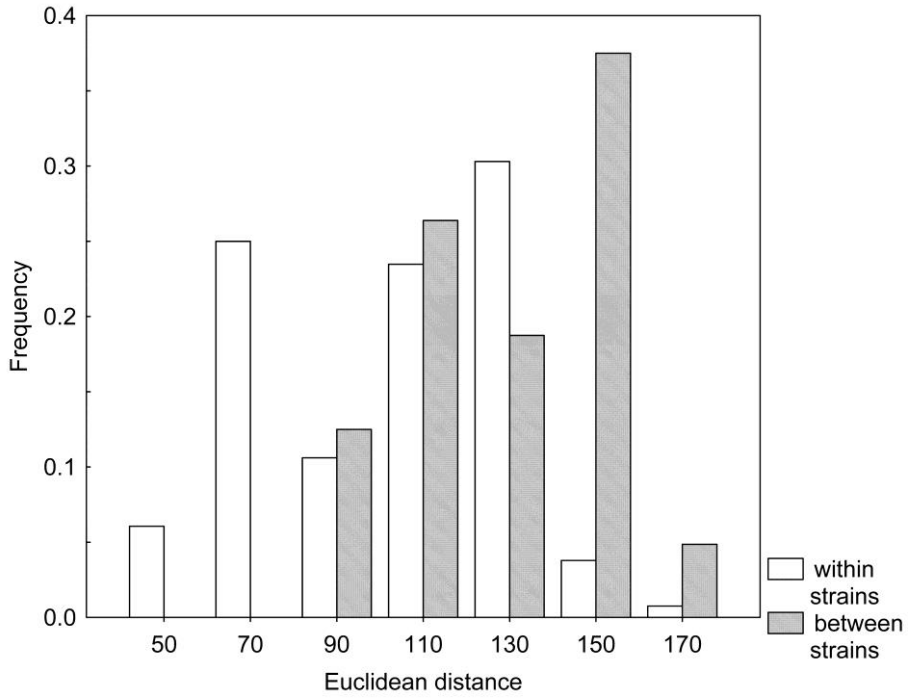

Figure 4.11: Distribution of Euclidean distances within and between strains for RNA samples extracted from spleen (the Kolmogorov-Smirnov test, $P = 1.19 \times 10^{-7}$).

According to the M-W and Kolmogorov-Smirnov tests, the within strains and between strains distance distributions differed significantly in each tissue analysed.

4.3.3. Distance between all RNA samples

The values of Euclidean distance are expected to be smaller for groups of samples sharing the same biological conditions.

Table 4.6 and **Figure 4.12** present a global comparison of the Euclidean distances for the whole set of experimental data. As expected, the highest level of dissimilarity was found between the four tissues, followed by the between strains and within strains classes. The Kruskal-Wallis test confirmed the differences between these 3 classes of distances (see **Figure 4.12**).

Table 4.6: Euclidean distances for all samples.

	N	Range	Mean ± s.e.m	Median
Within strains&within tissue	528	58.33 – 235.25	98.62 ± 1.21	89.27
Between strains&within tissue	576	81.22 – 247.44	121.64 ± 0.98	115.75
Between tissues	3456	336.47 – 502.53	429.94 ± 0.69	431.03
All samples	4560	58.33 – 502.53	352.63 ± 2.10	414.64

Mean and standard error values were calculated for sample distance between all RNA samples. The lowest mean Euclidean distance was observed in within strains followed by between strains, all samples and between tissues. The median also follows this increasing distance order for these 4 classes. The all samples class contains the distances between all samples so it has both the smallest and largest distances. The between tissues class contains the largest distances.

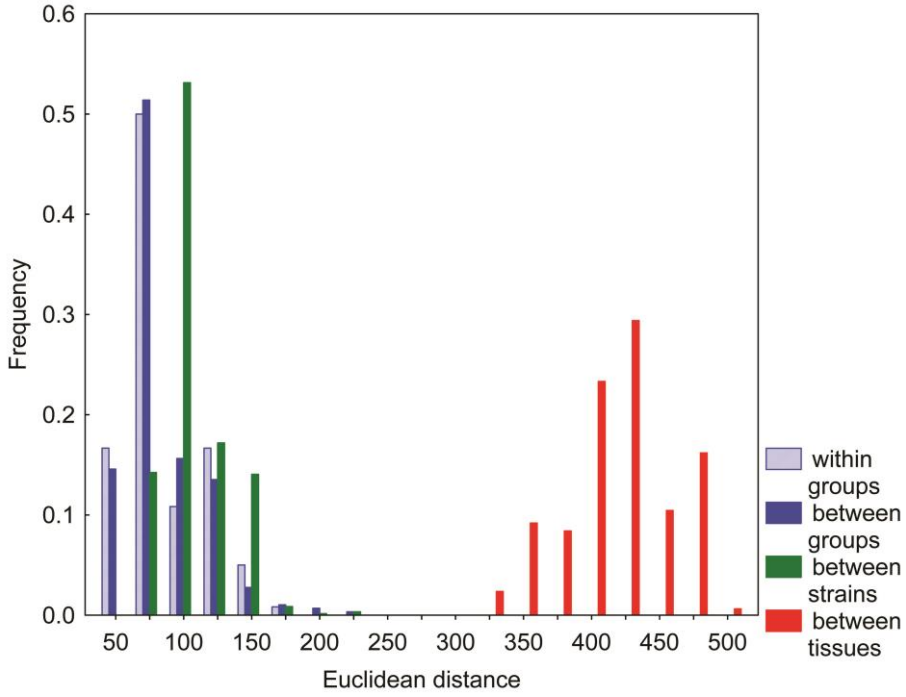

Figure 4.12: Distribution of Euclidean distances for all RNA samples (the Kruskal-Wallis test, χ^2 = 2524.08 (df = 2), $P < 10^{-16}$). Both within and between groups share identical distributions of Euclidean distances (see **section 4.3.1**) and were combined for this test. The Kruskal-Wallis test confirmed that the distributions of Euclidean distances from within strains, between strains and between tissues had significantly different location of distances.

4.3.4. Hierarchical clustering of all RNA samples using Euclidean distance

Unsupervised hierarchical clustering was used for the analysis of Euclidean distance data. This technique is exploratory and works by letting the data organise itself before finding biological meaning. This type of clustering involves grouping large sets of data into clusters (or, groups) of small and related sets of data and does not rely on any previous knowledge of the dataset (i.e. "unsupervised" learning) (Haines and

Pericak-Vance, 2006, p.201 [book]; Kimmel and Oliver, 2006, pp.198-199 [book]). The Statistica v7.1 software was used for clustering analysis.

Agglomerative hierarchical clustering is the most commonly used algorithm for identifying groups of closely related genes or tissues. Initially, each object (or sample) represents its own cluster but links between these objects can be formed by joining objects with the shortest distance. A distance metric and a linkage rule are used to find the best pair of objects for merging. Once pair(s) of objects have been formed the linkage rule forms higher order linkages between an object and a cluster or between clusters. The final product of linking all clusters is a tree structure called a dendrogram (Haines and Pericak-Vance, 2006, pp.201-202 [book]).

Complete linkage is a linkage method in which the distances between clusters are determined by the greatest distance between any two objects in different clusters (Kimmel and Oliver, 2006, p.201 [book]). Using agglomerative hierarchical clustering with a complete linkage method, all samples were clustered according to their similarity as measured by the Euclidean distance (**Figure 4.13**).

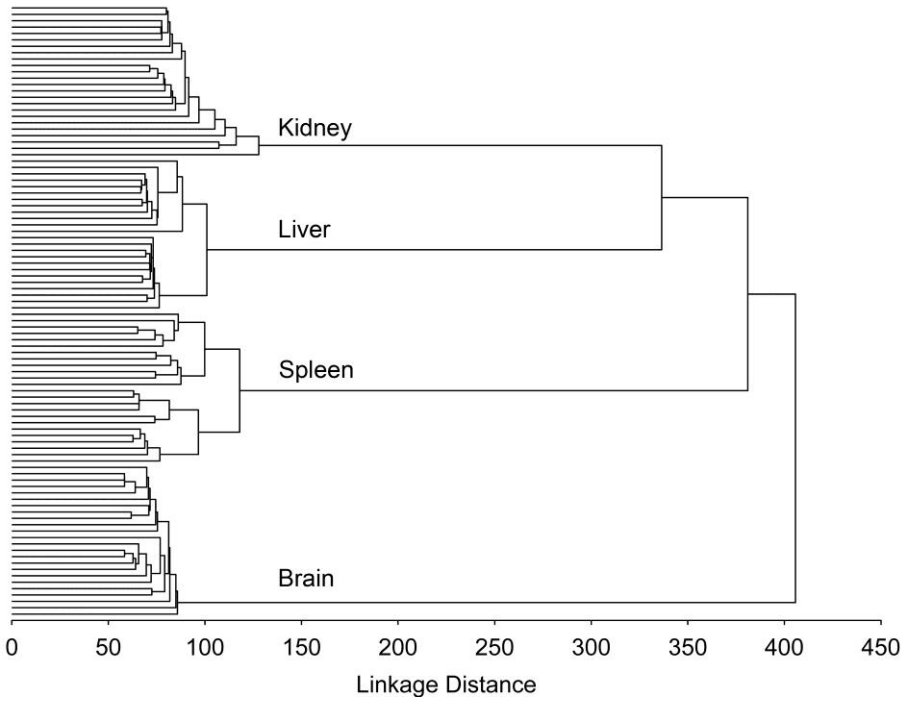

Figure 4.13: Unsupervised agglomerative hierarchical clustering analysis of the expression of 42,575 genes in all tissues. The data for all 96 animals are shown. Samples sharing the smallest distance were grouped into clusters according to tissue, almost completely according to strain, but not by experimental group.

After averaging technical and biological replicates, together, for each of the 16 different groups of animals, the resulting averaged gene expression profiles were subjected to hierarchical clustering (**Figure 4.14**). This second dendrogram fully confirmed all the above-mentioned results, namely: (1) the global dissimilarities in the expression profiles from offspring of irradiated and offspring of control mice is negligible; (2) within each tissue, the type of strain contributes substantially towards variation in the expression profiles; and, (3) the maximum distance corresponds to the global difference in transcriptional activity detected between tissues.

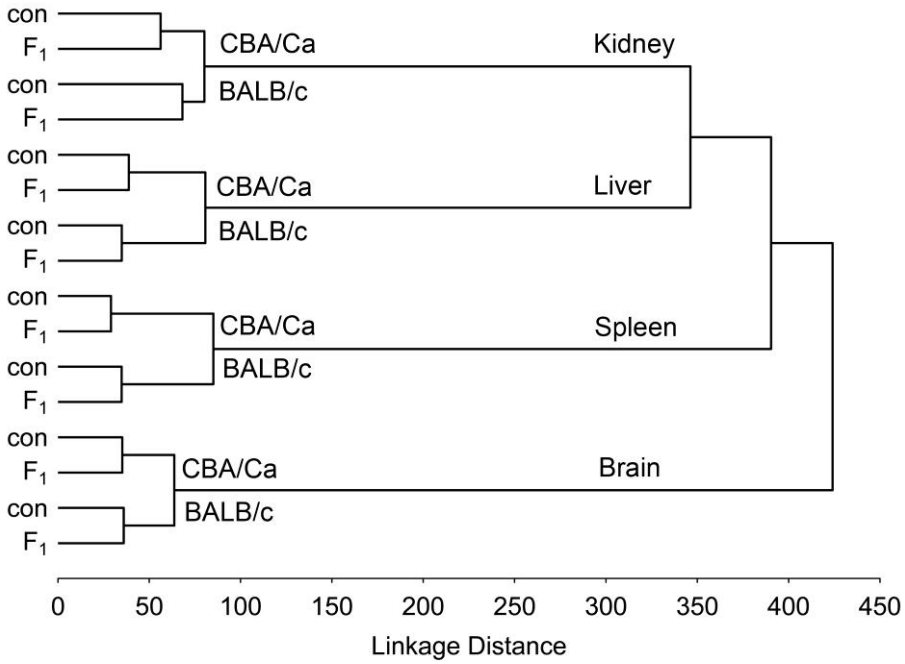

Figure 4.14: Unsupervised agglomerative hierarchical clustering analysis of the expression of 42,575 genes in all tissues. The mean values of gene expression were estimated for 16 groups of animals and the groups clustered according to their respective strain and tissue combination by hierarchical clustering.

4.3.5. Conclusion

The microarray technique that was validated using the Pearson's coefficient of correlation was validated again using the Euclidean distance metric. Correlation studies established very highly positive correlations in within technical replicates. Since no inverse correlations were present, the data set was regarded as optimal for Euclidean distance measurement of RNA profiles.

The biggest distances between gene expression profiles corresponded to RNA samples from different tissues. Inter-strain variation accounted for the second biggest group of distances. The most important observation was that the expression profiles from mice belonging to different experimental groups were indistinguishable. This

suggests that the contribution of paternal irradiation towards the variability of expression profiles, on a global level, is negligible.

4.4. The transgenerational effects of paternal irradiation on the pattern of gene expression

The purpose of this study was to identify any genes that are differentially expressed in the non-exposed offspring of irradiated mice. The differences, if present, would provide insight into the molecular processes underlining the transgenerational effects of paternal irradiation, including genomic instability observed in the offspring of exposed males. Given the very large differences between the transcriptional profiles of the four tissues analysed in this project, as well as a substantial inter-strain variation within each tissue, the microarray data were analysed using a number of different statistical techniques.

4.4.1. The effects of paternal irradiation on the pattern of gene expression in each tissue and strain

For each tissue, strain and group combination, probe values were averaged for each transcript and $\log_2$-transformed. Samples had $\log_2$-transformed transcript values (interval data) with a distribution close to the normal distribution; equal variances between samples; and, independent observations across samples, as described in **section 4.2.1** and tested for in **section 4.1**. Thus, a pair-wise statistical comparison between the experimental groups can be performed using the Student's *t*-test (*t*-test). For each tissue and strain, a two-tailed and two-sample *t*-test with $2(n-1)$ degrees of freedom was used on every transcript (Drăghici, 2003, pp.118-119 [book]). n is the number of biological and technical replicates in each group ($n = 6$). The H_0 of the *t*-test is that the

mean expression for a transcript following paternal irradiation is the same as the mean in the control population, and the H_1 is that it is not (Sokal and Rohlf, 2012, pp.223-226 [book]). For each tissue and strain, the test provided a *t*-statistic and *P*-value for the difference in transcript expression between the control and F_1 groups.

The *t*-test individually tested hypotheses for multiple, simultaneous comparisons of (42,575 transcripts) expression values across the two groups and is, therefore, prone to incorrectly rejecting the H_0 at one or more comparisons. An expression value that is incorrectly considered different between the two groups would be called a false positive and the test would have committed a type I error for rejecting the H_0 when it was true (Kimmel and Oliver, 2006, p.228 [book]). With the probability of detecting a false positive set to 5% ($\propto = 0.05$), a statistical method known as the false discovery rate (FDR) was used to adjust the test *P*-values, following the Benjamini and Hochberg method (Benjamini and Hochberg, 1995). Following this method, test *P*-values were ranked from smallest to largest, multiplied by the total number of transcripts and divided by the corresponding rank. If the adjusted *P*-value remained inferior to 0.05, the transcript was significantly differentially expressed between the two groups. A Volcano plot was used to visualise the magnitude and significance of effects of paternal irradiation on gene expression for each tissue and strain (**Figure 4.15**) (Wolfinger *et al.*, 2001). The x axis represents the fold change between the two groups on a log scale so that up and down regulation appear symmetric. The y axis represents the *P*-values for the *t*-test on a negative log scale so that smaller *P*-values appear at the top.

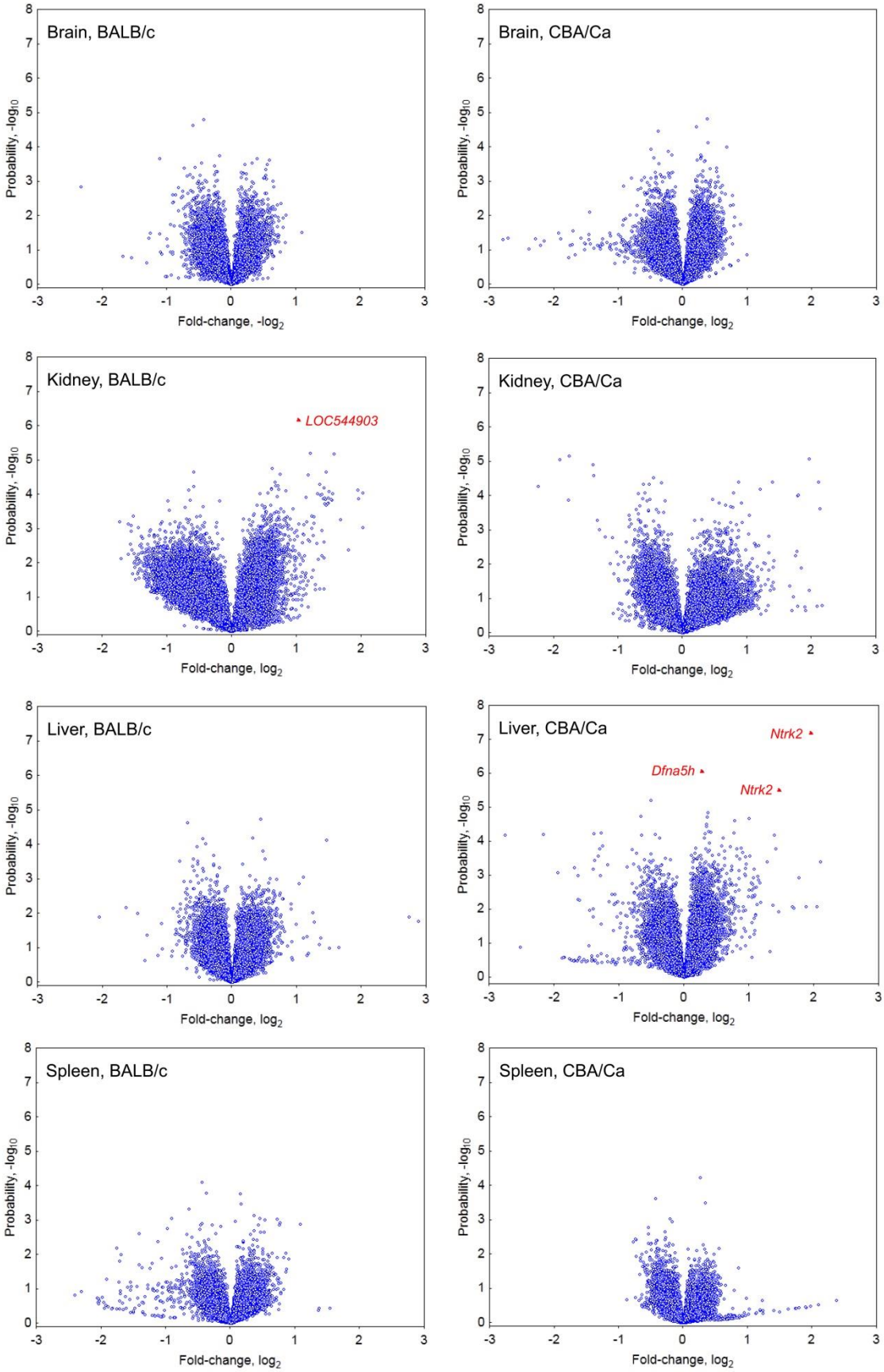
Brain, BALB/c
Brain, CBA/Ca
Kidney, BALB/c
LOC544903
Kidney, CBA/Ca
Liver, BALB/c
Liver, CBA/Ca
Ntrk2
Dfna5h
Ntrk2
Spleen, BALB/c
Spleen, CBA/Ca
Probability, -log10
Fold-change, -log2
Fold-change, log2

Figure 4.15: Volcano plots of significance against magnitude of effect of paternal irradiation on transcription for each tissue and strain. Significance is represented as negative $\log_{10}$ *P*-values from the t-test on the y axis, and magnitude of differential expression is represented as $\log_2$ fold-change on the x axis. Transcripts that are up-regulated in the F_1 group have positive fold-change while transcripts that are down-regulated in the F_1 group have negative fold-change. There seems to be a large proportion of differentially expressed transcripts, particular in the kidney of both strains. Despite the high probabilities of differential expression only 4 transcripts (in red) have passed the FDR < 0.05 representing 2 known genes in the liver of CBA/Ca and 1 unknown gene in the kidney of BALB/c.

At FDR < 0.05, only 4 transcripts are shown to be significantly differentially expressed: NM_008745 (*Ntrk2*), AK155135 (*Dfna5h*) and BC052014 (*Ntrk2*) for liver in the CBA/Ca strain; and, the unknown transcript BC106162 (*LOC544903*) for kidney in the BALB/c strain. Finally, the pair-wise analysis of transcription profiles detected a small number of significant differentially expressed transcripts between the offspring of irradiated and non-exposed males.

The effects of paternal irradiation on transcription are seemingly non-existent but gene expression variation caused by other factors could contribute towards masking the effects of paternal irradiation from being detected with a *t*-test between groups, for each tissue and strain. A two-sample *t*-test is mathematically equivalent to a Model I one-way analysis of variance (one-way ANOVA) for two groups since squaring the *t*-statistic value (t^2) should make it identical to the *F*-statistic of the ANOVA (Sokal and Rohlf, 2012, pp.223-225 [book]). Assuming equal variance between samples, both tests depend on estimates of variance, within samples (*t*-test) or within and among groups

(ANOVA), in order to analyse the significance for differences between means for two samples or groups (Sokal and Rohlf, 2012, pp.221, 224 [book]). Actually, in a Model I ANOVA, variances are called mean squares as they do not estimate a population variance (Sokal and Rohlf, 2012, p.198 [book]). Unlike the two-sample *t*-test, ANOVA can compare more than two groups and/or analyse multiple factors, all simultaneously, without increasing the probability of a type I error (Stekel, 2003, pp.134-135 [book]). Model I ANOVA is so called because it involves treatment effects (fixed factors) that are deliberately planned or fixed into the groups by the researcher. Model II ANOVA have added effects for each group that are not fixed treatments, are random and can only be partly controlled (random factors) (Sokal and Rohlf, 2012, p.195 [book]). Random factors, such as stochastic and individual variation, contribute towards gene expression variation but were reduced by having multiple probes per transcript in the same array, standardizing microarray experimentation, etc., (stochastic variation) or by extracting tissues at a fixed time of the day (individual variation) (Stekel, 2003, pp.100-102, 135 [book]; Kimmel and Oliver, 2006, pp.215-218 [book]). Tissue and strain (but not paternal irradiation) contributed more than random factors towards gene expression variation (**section 4.3.4**) and, therefore, should be controlled during statistical analysis. These three fixed factors can be analysed simultaneously for their individual and combined contribution to gene expression variation using a, Model I, three-way ANOVA (Sokal and Rohlf, 2012, p.354-365 [book]).

4.4.2. The effects of paternal irradiation on the pattern of gene expression in all tissues and strains

To establish the effects of paternal irradiation on the pattern of gene expression in all tissues, a three-way ANOVA was used. This statistical technique allows

estimating independent effects of three factors (tissue, strain and paternal treatment) on the value of gene expression for each of 42,575 transcripts. The data set submitted to the two-sample *t*-test met parametric assumptions and was used again for the 3-way ANOVA. The 3-way ANOVA was used, individually, for every transcript across the 16 animal groups. For each animal group, every transcript had 6 values corresponding to biological and technical replicates. For the effects of paternal irradiation, the H_0 and H_1 for the 3-way ANOVA are exactly the same as those for the two-sample *t*-test (see **section 4.4.1**). Because a Model I ANOVA was used, the presence of paternal irradiation effects was tested by the ratio of its mean square over the mean square for all 3 factors, which represents the error variance, on the assumption that the interaction between the three factors is zero (Sokal and Rohlf, 2012, p.362 [book]). For the effects of paternal irradiation, the values of *F*-statistics together with the corresponding *P*-values for each transcript were estimated. Similarly to the results presented in **section 4.4.1**, the statistical significance of all *P*-values was ascertained at the FDR < 0.05 following the method by Benjamini and Hochberg (1995). The *F*-statistic describes the combined effects of paternal exposure on the pattern of gene expression in all tissues. The 3-way ANOVA *P*-values and the fold-change in expression was estimated for each transcript and represented in a Volcano plot (**Figure 4.16**).

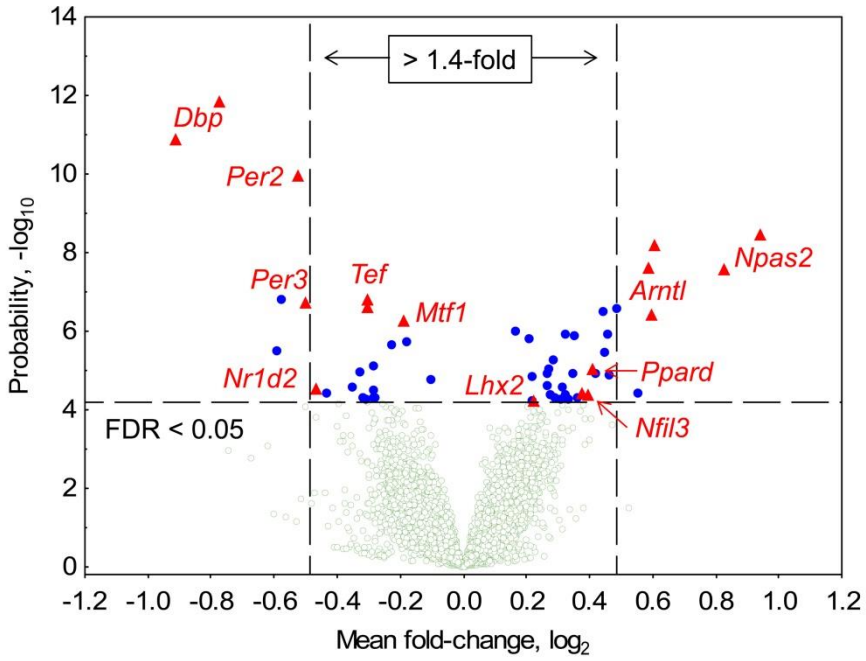

Figure 4.16: Volcano plot of significance against magnitude of effect of paternal irradiation on transcription for all tissues and strains. The mean fold change represents the transcript fold-change across groups averaged for all strains and tissues. Genes that have passed the FDR < 0.05 cut point are in blue and those genes with additional interest are in red. An arbitrary cut-off point of 1.4 fold-change, in both directions, was applied for representative purposes and shows that the most significantly differentially expressed transcripts belong to circadian genes.

Table 4.7 describes the results for the gene ontology (GO) class analysis. Transcripts from the three most significantly compromised GO classes in the offspring of irradiated male mice were listed with the probabilities for differential expression considering all tissues alone (*t*-test; **Tables 4.9.1-4.9.4**) or together (ANOVA; **Table 4.8**).

4.4.3. Grouping differentially expressed genes into functional classes

BioProfiling.de is a comprehensive analytical toolkit capable of associating current and novel biological principles to group genes into functional classes or to

associate genes into a global gene network. This resource is freely available at http://www.BioProfiling.de/gene_list. All 39 genes, with transcript(s) passing the selection threshold FDR < 0.05, were submitted to the Bioprofiling.de interface (input list). This interface contains a statistical framework, referred to as ProfCom, which deals with the grouping of genes into functional classes (GO classes) (Antonov *et al.*, 2008). Genes on the input list were assigned into one or multiple GO-classes. For each GO-class, the number of genes assigned into the same class belonging to the input list or the reference list (containing all genes from the *Mus musculus* genome), was counted and compared (Antonov, 2011). The H_0 that genes from the input list are independent of having a certain GO attribute was tested using hypergeometric, binomial or χ^2-tests retrieving a *P*-value for the probability of a certain GO-class occurring x times by chance in the input list (Antonov *et al.*, 2008). Bonferroni correction was used to adjust the *P*-values for multiple testing (Antonov, 2011). ProfCom is also capable of extending this enrichment scheme by forming gene-based combinations of GO-classes ("complex classes") at the pairwise, triplet and quadruplet level using Boolean combination of available GO-classes and a greedy heuristics search algorithm to reduce computational complexity (Antonov, 2011). For complex class enrichments, ProfCom retrieved *P*-values using the tests mentioned above and adjusted *P*-values for multiple testing using the Monte-Carlo simulation (described by Antonov *et al.*, 2010).

Genes with FDR <0.05 were submitted to GO analysis. Single class enrichment detects all GO-classes which are over-represented in the input list (**Table 4.7**), while the complex class enrichment provides a more specific biological function for a subset of genes in the input list (Antonov *et al.*, 2008). The complex class enrichment, mainly at the pairwise and triplet levels, provided some complex classes with odds ratio >459.63 and a *P*-value <1 x 10^{-7} that happened to combine the top 6 GO-classes on **Table 4.7**.

Whether using single or complex GO-class enrichments, it is clear that the majority of the genes in the most enriched classes are either components of the circadian clock or circadian output genes with roles in the regulation or initiation of transcription. Indeed, these genes have been found to be very significantly enriched into functional classes involved in Rhythmic Processes ($P = 1.25 \times 10^{-9}$), Circadian Rhythm ($P = 1.52 \times 10^{-7}$) and DNA-dependent Regulation of Transcription ($P = 1.62 \times 10^{-6}$).

Table 4.7: Single GO-class enrichment of genes with an FDR < 0.05.

Go term	Description	Odds ratio	Prob	Genes
GO:0048511	Rhythmic process	158.49	1.25×10^{-9}	*Dbp, Nfil3, Npas2, Per2, Per3, Tef*
GO:0007623	Circadian rhythm	147.32	1.52×10^{-7}	*Arntl, Dbp, Nfil3, Per2, Per3*
GO:0006355	Regulation of transcription, DNA-dependent	11.27	1.62×10^{-6}	*Arntl, Dbp, Lhx2, Mtf1, Nfil3, Npas2, Per2, Per3, Ppard, Tef, Nr1d2*
GO:0003700	Sequence-specific DNA binding transcription factor activity	9.93	0.0002	*Arntl, Dbp, Lhx2, Mtf1, Nfil3, Npas2, Ppard, Tef, Nr1d2*
GO:0006350	Transcription	5.38	0.0028	*Arntl, Dbp, Lhx2, Mtf1, Nfil3, Npas2, Per2, Per3, Ppard, Tef, Nr1d2*
GO:0045449	Regulation of transcription	5.23	0.0036	*Arntl, Dbp, Lhx2, Mtf1, Nfil3, Npas2, Per2, Per3, Ppard, Tef, Nr1d2*
GO:0005634	Nucleus	3.05	0.0074	*Arntl, Cdkn1a, Cirbp, Dbp, Mtf1, Nfil3, Npas2, Per2, Per3, Ppard, Ppp2cb, Rbm3, Tef, Park7, Dnajb1, Nr1d2*
GO:0003677	DNA binding	4.28	0.0100	*Arntl, Dbp, Lhx2, Mtf1, Nfil3, Npas2, Ppard, Tef, Nr1d2*
GO:0007165	Signal transduction	3.66	0.0100	*Arntl, Cntfr, Grm8, Npas2, Ntrk2, Per2, Per3, Ppard, Nr1d2*
GO:0043565	Sequence-specific DNA binding	9.29	0.0100	*Dbp, Lhx2, Nfil3, Ppard, Tef, Nr1d2*
GO:0005515	Protein binding	2.22	0.0300	*Arntl, Cdkn1a, Cntfr, Edn1, Ntrk2, Per2, Per3, Ppard, Ppp2cb, Tef, Dtx4*

GO:0005737	Cytoplasm	1.94	0.0800	*Cdkn1a, Cirbp, Hsd17b1, Per2, Per3, Ppp2cb, Rbm3, Park7, Dnajb1, Dtx4*
GO:0046872	Metal ion binding	2.29	0.0900	*Cdkn1a, Cyp27b1, Lhx2,Mtf1, Ppard, Ppp2cb, Dtx4, Nr1d2*
GO:0008270	Zink ion binding	2.88	0.2500	*Lhx2, Mtf1, Ppard, Dtx4,Nr1d2*
GO:0004871	Signal transducer activity	2.48	0.3300	*Arntl, Grm8, Npas2, Per2,Per3*
GO:0004872	Receptor activity	1.89	0.3600	*Arntl, Cntfr, Grm8, Ntrk2, Ppard, Nr1d2*

The 6 genes enriched for the Rhythmic process class (*Dbp, Nfil3, Npas2, Per2, Per3* and *Tef*) have an odds ratio of 158.49. So, these genes are 158.49 times more likely to belong to the Rhythmic processes class than any other GO-class, with a probability of false positive of just 1.25×10^{-9} (calculated by BioProfiling.de, as described above), when submitted for enrichment in this list of 39 genes.

The genes from the top three most significantly enriched GO-classes were listed (**Table 4.8**). The transcript mean fold-change (detected by microarray) and the probability of effect by paternal irradiation for all individuals (calculated by 3-way ANOVA), were provided for each gene transcript. Only gene transcripts with significant probabilities were listed.

Table 4.8: The list of down- and up-regulated genes from the first three GO categories.

Name, symbol, GeneBank accession	Prob*	Mean fold-change
Down-regulated		
D site albumin promoter binding protein, *Dpb*		
NM_016974	1.38×10^{-12}	1.71↓
BC018323	1.28×10^{-11}	1.88↓
Period homolog 2, *Per2*		
NM_011066	1.11×10^{-10}	1.44↓
Period homolog 3, *Per3*		
NM_011067	1.85×10^{-7}	1.42↓
Nuclear receptor subfamily 1, group D, member 2, *Nr1d2*		
BC096461	2.99×10^{-5}	1.38↓
Thyrotroph embryonic factor, *Tef*		
AK042146	1.57×10^{-7}	1.23↓
NM_017376	2.53×10^{-7}	1.23↓
Metal regulatory transcription factor 1, *Mtf1*		
BC017679	5.51×10^{-7}	1.14↓
Up-regulated		
Neuronal PAS domain protein 2, *Npas2*		
NM_008719	3.35×10^{-9}	1.92↑
BC109166	2.78×10^{-8}	1.77↑
Aryl hydrocarbon receptor nuclear translocator-like, *Arntl*		
AB012601	6.45×10^{-9}	1.51↑
NM_007489	2.37×10^{-8}	1.50↑
AY690602	3.72×10^{-7}	1.52↑
Peroxisome proliferator-activated receptor, *Ppard*		
NM_011145	9.33×10^{-6}	1.33↑
Nuclear factor, interleukin 3 regulated, *Nfil3*		
NM_017373	4.00×10^{-5}	1.32↑
BC100384	3.97×10^{-5}	1.30↑
LIM/homeobox protein, *Lhx2*		
NM_010710	6.09×10^{-5}	1.17↑

*Probability for the effects of paternal irradiation, three-way ANOVA, all tissues combined.

Of the list of 39 genes submitted for GO analysis, 11 genes were found to play roles in circadian processes or were found to be transcription factors and included: *Arntl, Dbp, Lhx2, Mtf1, Nfil3, Npas2, Nr1d2, Per2, Per3, Ppard* and *Tef.*

Arntl and *Npas2* form a heterodimer driving the primary loop of the circadian clock and were both found to be significantly up-regulated. The direct targets of this

heterodimer involved genes with E-box elements including, from **Table 4.8**, *Dbp* and *Tef* which are transcription factors in charge of increasing the output of circadian expression (Zuber *et al.*, 2009). Curiously, both *Dbp* and *Tef* are down-regulated suggesting that there may be a phase shift or a change in the amplitude of expression in some or all of these genes. *Per2* and *Nr1d2* are also targets of the heterodimer and were also found to be down-regulated in offspring of irradiated males. *Per2* functions to prevent activity of the heterodimer by binding and assisting to its degradation. *Nr1d2* (*Rev-erbβ*) is a repressor at the RRE element sites in *Arntl*, *Npas2*, *Clock* and *Nfil3* and should reduce or prevent their transcription. *Per3* was found to be down-regulated which is consistent with the down-regulation of its transcription activator, *Dbp*, and with the up-regulation of *Nfil3* (*E4bp4*), a repressor at D-box elements which is present at the *Rev-erbα/β*, *RORα/β*, *Per1/2/3* and *Cry1* genes (see **Figure 1.6**). *Per3* is in control of the period and phase of rhythms in many tissues including liver and is important for endogenous timekeeping in tissues (Pendergast *et al.*, 2012). *Mtf1* is a metal-responsive factor that plays key roles in the liver as a regulator of transcription in response to stress conditions including heavy metal load and oxidative stress (Wimmer *et al.*, 2005). Oxidative stress can induce inflammation by induction of cyclo-oxygenase-2 and inducible nitric oxide synthase (iNOS) aberrant expression of inflammatory cytokines (TNF), interleukins 1 and 6 (IL-1, IL-6) and chemokines (IL-8, CXCR4) and alterations in the expression of specific microRNAs (Hussain and Harris, 2007). Oxidative stress and inflammation are usually associated following irradiation (Narayanan *et al.*, 1999; Giles *et al.*, 2000; Gregory, 2000 [review]). Oxidative stress causes RIBE (Shao *et al.*, 2004) while inflammation could cause RIGI (Neriishi *et al.*, 2001). A sustained inflammatory and oxidative environment causes the affected cellular region to enter a vicious cycle of cellular damage (Federico *et al.*, 2007). Additionally, ROS can reduce

the expression and activity of the DNA mismatch repair genes mutS homologs 2 and 6 and increase the expression of DNA methyltransferases causing global hypermethylation of the genome (Schetter *et al.,* 2010). This causes promoter silencing of several important genes including cyclin-dependent kinase inhibitor-2, breast cancer susceptibility gene 1, retinoblastoma protein, murine double minute 2 (MDM2) and the human DNA mismatch repair gene mutL homolog 1 (Fleisher *et al.,* 2000; Das and Singal, 2004 [review]). If sustained oxidative stress and inflammation environments were transmitted to the zygote and lasted until the DNA reprogramming occurring during the embryonic day 11 (van der Heijden *et al.*, 2006), then some essential regions could, potentially, suffer permanent promoter silencing by DNA hypermethylation to generate GI and TI. *Mtf1* was found to be down-regulated in offspring of irradiated males, when compared to controls. The observed *Mtf1* down-regulation is contradictory with the idea that the GI in the paternally irradiated offspring is caused by sustained high levels of oxidative stress or non-specific inflammation inherited from the irradiated parent. *Lhx2* is a transcriptional activator expressed cyclically and is essential for generation and regeneration of hair (Törnqvist *et al.,* 2010) and for the control of differentiation in developing lymphoid and neuronal cells. *Ppard* (also known as *PPARβ/PPARδ*) is an inhibitor of ligand-induced transcription activity by *PPARα* and *PPARγ*. *PPARα* is, in turn, found to activate *Arntl* expression through the PPAR response element (PPRE) on the *Arntl* gene (Froy, 2011 [review]). *Ppard* is up-regulated which could cause a down-regulation of *Arntl*, however, it does not. In fact, *Ppard* and *Arntl* have been reported to have a positively correlated expression during gestational diabetes mellitus (GDM) in pregnant women which manifested deregulated expression of other clock genes. This circadian clock deregulation was thought to play a pathogenic role in GDM (Pappa *et al.*, 2013). Finally, the RRE element is established to

be the main element controlling the *Arntl* cyclic expression and allows a transcription factor complex to form including factors with protein deacetylation and metabolite sensing functions (e.g. Sirt1) and others with ROR activating functions (e.g. PGC-1α) (Liu *et al.*, 2008; Froy, 2011 [review]).

The general observation of differential expression of these 11 genes is that they are not necessarily deregulated. Possibly, a simple phase shift or change in oscillation of one key gene or the whole clock could explain these observations. However, that would not exclude the possibility that this change could be an effect of paternal irradiation. Although phase shifting is part of the entrainment event, if it is both large and frequent over a long period of time (as with shift workers) it may lead to cancer onset (Schernhammer *et al.*, 2001). It is not possible to conclude if the observed phase shift in the paternally irradiated offspring has occurred only once, which would probably be inconsequential, or if there is deregulation of the environmental cue recognition and entrainment processes.

This network of genes and the direction of differential expression in all offspring of irradiated males (**Table 4.8**) are perfectly matched in liver samples of the CBA/Ca paternally irradiated offspring (**Table 4.9.3**). In the kidney samples of the same group of mice (**Table 4.9.2**) only Nr1d2 is unchanged while the other genes are differentially expressed in the same direction as those in liver of CBA/Ca. In the brain of paternally irradiated CBA/Ca offspring the number of significantly differentially expressed genes is smaller as only *Arntl, Dbp, Mtf1, Nr1d2, Per2* and *Per3* show differential expression (**Table 4.9.1**). Spleen samples of CBA/Ca offspring of irradiated males (**Table 4.9.4**) show differential expression of only *Npas2, Nr1d2* and *Dbp*. In the paternally irradiated BALB/c offspring *Arntl, Npas2, Nr1d2, Mtf1* and *Per2* were shown to be significantly differentially expressed in kidney samples only.

When significant, the direction of expression change for each gene transcript is the same in each tissue analysed for the offspring of CBA/Ca and BALB/c paternally irradiated mice. It is clear that paternal irradiation produces a similar cellular effect in all tissues and in both strains of offspring of irradiated males. What is more surprising is that this effect is not observed with the same intensity in all tissues, as liver and kidney are the most susceptible and spleen the most resistant. The increased sensitivity of the CBA/Ca strain in comparison to BALB/c, to these newly observed effects of paternal irradiation, is in disagreement with previous publications (Barber *et al.*, 2002; Barber *et al.*, 2006). These publications show, in offspring from both strains, similar ESTR and *hprt* mutation rate ratio/fold-change (to control mice from the same strain) in germline or somatic tissues but with paternal CBA/Ca mice having double X-ray exposure when compared to the paternal BALB/c mice. To sum up, it seems paternal irradiation can affect the circadian clock in the non-exposed offspring but that effect, which must be epigenetic, is not evenly kept in all tissues throughout the life of the non-exposed offspring.

When considering all paternally irradiated offspring per experimental group it is obvious that both liver and kidney of paternally irradiated CBA/Ca offspring have very large changes in gene expression (**Tables 4.9.2** and **4.9.3**). The size of expression change was able to compensate for unchanged expression, for each gene, in other tissues and/or strain when considering all paternally irradiated offspring together in **Table 4.8**.

More importantly, the level of significant differential expression in the liver and kidney tissues of paternally irradiated CBA/Ca offspring is impressive. In the kidney, *Dbp* (3.76↓), *Per2* (3.38↓), *Per3* (2.12↓), *Arntl* (3.88↑) and *Npas2* (3.48↑) have substantial fold-changes in expression. While in liver, a massive down-regulation in

Dbp for NM_016974 (6.75↓ at $P = 6.59 \times 10^{-5}$) and BC018323 (17.48↓ at $P = 8.94 \times 10^{-5}$) transcripts was observed.

Lastly, *Arntl, Npas2, Per2, Per3, Mtf1* and *Nr1d2* are very significantly differentially expressed in three of the four tissues analysed but *Dbp* is the only gene that is significantly down-regulated in all tissues of the paternally irradiated CBA/Ca offspring.

Table 4.9.1: The pattern of gene expression of the genes from the first three GO categories across brain.

		CBA/Ca				BALB/c			
		Mean ± s.e.m, log_2				Mean ± s.e.m, log_2			
Locus	Transcript	Control	F_1	Fold-change	Prob*	Control	F_1	Fold-change	Prob*
Brain									
Down-regulated									
Dbp	NM_016974	11.27±0.10	10.67±0.07	1.52↓	**0.0008**	11.36±0.08	11.32± .06	1.03↓	0.6902
Dbp	BC018323	12.12±0.04	11.57±0.13	1.47↓	**0.0022**	11.95±0.07	12.03±0.10	1.06↓	0.5166
Per2	NM_011066	10.58±0.07	10.17±0.05	1.32↓	**0.0012**	11.17±0.07	11.04±0.10	1.10↓	0.3051
Per3	NM_011067	10.05±0.04	9.86±0.07	1.14↓	**0.0423**	9.67 ± 0.06	9.64 ± 0.03	1.02↓	0.6444
Nr1d2	BC096461	11.35±0.07	10.92±0.07	1.35↓	**0.0011**	11.60±0.12	11.34±0.06	1.20↓	0.0773
Tef	AK042146	12.92±0.02	12.79±0.06	1.09↓	0.0843	12.93±0.03	12.87±0.03	1.04↓	0.2082
Tef	NM_017376	12.90±0.02	12.83±0.04	1.05↓	0.1694	12.99±0.03	12.95±0.04	1.03↓	0.4306
Mtf1	BC017679	11.72±0.03	11.44±0.08	1.22↓	**0.0076**	11.71±0.06	11.63±0.07	1.05↓	0.4333
Up-regulated									
Npas2	NM_008719	12.05±0.05	12.29±0.13	1.18↑	0.1191	11.64±0.09	11.90±0.10	1.20↑	0.0755
Npas2	BC109166	11.99±0.13	12.32±0.14	1.26↑	0.1231	11.43±0.12	11.65±0.11	1.17↑	0.2072
Arntl	AB012601	12.17±0.09	12.46±0.06	1.23↑	**0.0203**	11.97±0.03	12.01±0.06	1.03↑	0.5827
Arntl	NM_007489	11.74±0.03	12.03±0.08	1.22↑	**0.0099**	11.53±0.07	11.51±0.05	1.01↑	0.8629
Arntl	AY690602	13.03±0.08	13.10±0.06	1.05↑	0.5319	12.69±0.05	12.72±0.05	1.02↑	0.6693
Ppard	NM_011145	10.79±0.13	11.06±0.12	1.21↑	0.1563	10.58±0.16	10.73±0.07	1.11↑	0.3911
Nfil3	BC100384	10.90±0.12	11.08±0.02	1.19↑	0.1766	10.87±0.04	10.81±0.08	1.11↑	0.4976
Nfil3	NM_017373	10.22±0.14	10.47±0.13	1.14↑	0.2283	10.30±0.05	10.15±0.05	1.04↑	0.0613
Lhx2	NM_010710	13.26±0.09	13.48±0.08	1.16↑	0.0985	13.15±0.06	13.27± 0.08	1.09↑	0.2345

Table 4.9.2: The pattern of gene expression of the genes from the first three GO categories across kidney.

		CBA/Ca				BALB/c			
		Mean ± s.e.m, $\log_2$				Mean ± s.e.m, $\log_2$			
Locus	Transcript	Control	F_1	Fold-change	Prob*	Control	F_1	Fold-change	Prob*
Kidney									
Down-regulated									
Dbp	NM_016974	12.84±0.15	10.93±0.17	3.76↓	**9.00×10^{-6}**	12.68±0.22	12.53±0.21	1.11↓	0.6210
Dbp	BC018323	13.68±0.13	11.91±0.26	3.41↓	**0.0001**	13.60±0.23	13.39±0.19	1.16↓	0.4925
Per2	NM_011066	11.02±0.13	9.26±0.16	3.38↓	**6.93×10^{-6}**	11.91±0.18	11.26±0.08	1.58↓	**0.0068**
Per3	NM_011067	9.66±0.13	8.57±0.22	2.12↓	**0.0016**	9.42±0.15	9.33±0.14	1.06↓	0.6736
Nr1d2	BC096461	9.80±0.52	9.62±0.41	1.13↓	0.7907	10.57±0.27	9.42±0.13	2.21↓	**0.0033**
Tef	AK042146	13.74±0.06	13.18±0.07	1.47↓	**0.0001**	13.53±0.11	13.51±0.05	1.01↓	0.8968
Tef	NM_017376	13.63±0.07	13.09±0.06	1.45↓	**0.0003**	13.38±0.15	13.42±0.08	1.02↑	0.8492
Mtf1	BC017679	11.72±0.03	11.45±0.07	1.21↓	**0.0060**	11.80±0.09	11.32±0.07	1.40↓	**0.0015**
Up-regulated									
Npas2	NM_008719	10.07±0.19	11.85±0.21	3.48↑	**0.0001**	8.83±0.27	10.21±0.30	2.68↑	**0.0060**
Npas2	BC109166	10.06±0.17	11.86±0.23	3.43↑	**9.68×10^{-5}**	9.09±0.33	10.51±0.25	2.61↑	**0.0069**
Arntl	AB012601	9.27±0.27	11.38±0.13	4.31↑	**3.94×10^{-5}**	8.71±0.26	9.83±0.19	2.18↑	**0.0060**
Arntl	NM_007489	8.90±0.16	10.85±0.17	3.88↑	**8.23×10^{-6}**	7.99±0.20	9.09±0.27	2.14↑	**0.0085**
Arntl	AY690602	9.86±0.35	11.99±0.15	4.38↑	**0.0002**	9.20±0.32	10.46±0.29	2.39↑	**0.0158**
Ppard	NM_011145	12.41±0.17	13.09±0.22	1.60↑	**0.0323**	11.97±0.22	12.53±0.18	1.47↑	0.0761
Nfil3	BC100384	10.14±0.24	11.48±0.28	2.66↑	**0.0047**	10.69±0.33	11.03±0.23	1.27↑	0.4212
Nfil3	NM_017373	9.35±0.27	10.75±0.32	2.52↑	**0.0077**	9.90±0.31	10.24±0.19	1.27↑	0.3686
Lhx2	NM_010710	7.35±0.18	8.38±0.09	2.04↑	**0.0004**	6.42±0.07	6.59±0.13	1.12↑	0.2715

Table 4.9.3: The pattern of gene expression of the genes from the first three GO categories across liver.

		CBA/Ca				BALB/c			
		Mean ± s.e.m, $\log_2$				Mean ± s.e.m, $\log_2$			
Locus	Transcript	Control	F_1	Fold-change	Prob*	Control	F_1	Fold-change	Prob*
Liver									
Down-regulated									
Dbp	NM_016974	10.83±0.34	8.08±0.25	6.75↓	**6.59×10^{-5}**	11.34±0.18	11.27±0.31	1.05↓	0.8383
Dbp	BC018323	11.50±0.40	7.38±0.52	17.48↓	**8.94×10^{-5}**	12.17±0.12	12.22±0.31	1.04↑	0.8840
Per2	NM_011066	10.49±0.23	9.66±0.12	1.78↓	**0.0099**	11.03±0.23	10.82±0.12	1.16↓	0.4371
Per3	NM_011067	8.35±0.32	6.18±0.06	4.49↓	**6.18×10^{-5}**	8.41±0.16	8.34±0.36	1.05↓	0.8696
Nr1d2	BC096461	11.01±0.13	9.98±0.26	2.04↓	**0.0058**	11.22±0.06	11.14±0.14	1.06↓	0.5876
Tef	AK042146	11.67±0.16	10.41±0.14	2.39↓	**0.0001**	12.10±0.07	12.06±0.18	1.03↓	0.8532
Tef	NM_017376	11.70±0.16	10.31±0.14	2.61↓	**5.96×10^{-5}**	12.08±0.06	12.08±0.18	1.00↓	0.9776
Mtf1	BC017679	11.74±0.07	11.55±0.04	1.14↓	**0.0443**	11.80±0.10	11.62±0.04	1.14↓	0.1138
Up-regulated									
Npas2	NM_008719	9.39±0.61	11.44±0.12	4.13↑	**0.0083**	7.41±0.51	7.79±0.41	1.30↑	0.5820
Npas2	BC109166	9.70±0.49	11.37±0.14	3.18↑	**0.0082**	7.96±0.35	8.03±0.33	1.04↑	0.9010
Arntl	AB012601	11.42±0.29	12.49±0.04	2.11↑	**0.0042**	10.46±0.35	10.52±0.28	1.04↑	0.9060
Arntl	NM_007489	10.89±0.34	12.07±0.04	2.27↑	**0.0063**	9.88±0.38	9.98±0.30	1.07↑	0.8443
Arntl	AY690602	12.47±0.28	13.52±0.05	2.06↑	**0.0045**	11.76±0.31	11.61±0.31	1.11↓	0.7456
Ppard	NM_011145	9.74±0.32	11.10±0.25	2.57↑	**0.0075**	9.20±0.09	9.36±0.12	1.12↑	0.2934
Nfil3	BC100384	12.73±0.10	13.50±0.05	1.75↑	**5.63×10^{-5}**	12.24±0.24	12.54±0.14	1.26↑	0.2952
Nfil3	NM_017373	12.15±0.12	12.96±0.07	1.70↑	**0.0002**	11.52±0.25	11.85±0.14	1.24↑	0.2804
Lhx2	NM_010710	10.21±0.07	10.31±0.07	1.07↑	**0.3214**	10.58±0.12	10.56±0.04	1.01↓	0.8720

Table 4.9.4: The pattern of gene expression of the genes from the first three GO categories across spleen.

		CBA/Ca				BALB/c			
		Mean ± s.e.m, log_2				Mean ± s.e.m, log_2			
Locus	Transcript	Control	F_1	Fold-change	Prob*	Control	F_1	Fold-change	Prob*
Spleen									
Down-regulated									
Dbp	NM_016974	9.96±0.05	9.54±0.14	1.33↓	**0.0212**	9.93±0.09	9.70±0.18	1.17↓	0.2849
Dbp	BC018323	10.56±0.11	9.93±0.21	1.55↓	**0.0229**	10.36±0.09	10.24±0.18	1.08↓	0.5832
Per2	NM_011066	8.70±0.12	8.61±0.16	1.06↓	0.6860	9.51±0.11	9.39±0.17	1.09↓	0.5540
Per3	NM_011067	7.69±0.15	7.42±0.17	1.21↓	0.2630	7.90±0.15	7.79±0.19	1.08↓	0.6599
Nr1d2	BC096461	9.48±0.10	8.98±0.10	1.41↓	**0.0047**	9.82±0.08	9.73±0.14	1.07↓	0.5742
Tef	AK042146	10.35±0.07	10.00±0.17	1.27↓	0.0828	10.59±0.07	10.55±0.17	1.03↓	0.8361
Tef	NM_017376	10.33±0.08	9.99±0.15	1.27↓	0.0644	10.49±0.09	10.41±0.17	1.06↓	0.6703
Mtf1	BC017679	11.91±0.07	11.86±0.07	1.04↓	0.6196	12.00±0.09	12.02±0.08	1.02↓	0.8298
Up-regulated									
Npas2	NM_008719	7.24±0.28	8.10±0.17	1.81↑	**0.0244**	5.88±0.14	6.41±0.20	1.44↑	0.0552
Npas2	BC109166	8.27±0.40	9.05±0.21	1.72↑	0.1142	6.81±0.32	7.19±0.21	1.30↑	0.3485
Arntl	AB012601	11.42±0.16	11.61±0.06	1.15↑	0.2852	10.98±0.11	10.95±0.13	1.02↓	0.8756
Arntl	NM_007489	10.90±0.14	11.08±0.06	1.13↑	0.2735	10.50±0.07	10.42±0.11	1.06↓	0.5491
Arntl	AY690602	12.34±0.18	12.68±0.09	1.27↑	0.1229	11.94±0.19	12.00±0.21	1.04↑	0.8451
Ppard	NM_011145	10.56±0.19	10.52±0.14	1.03↓	0.8439	10.15±0.10	10.29±0.10	1.10↑	0.3538
Nfil3	BC100384	9.78±0.17	9.85±0.09	1.13↑	0.7184	9.25±0.12	9.31±0.11	1.00↓	0.6959
Nfil3	NM_017373	8.95±0.17	9.13±0.15	1.05↑	0.4532	8.49±0.08	8.49±0.11	1.05↑	0.9796
Lhx2	NM_010710	6.62±0.17	6.75±0.10	1.09↑	0.5282	6.18±0.10	6.23±0.15	1.03↑	0.8010

*Probability for difference between controls and F1, 3-way ANOVA; significant values are given in bold.

4.4.4. Conclusion

Differences in gene expression were observed in the non-exposed offspring of irradiated and control mice. Student's t-test detected a significant difference in the mean value of transcript expression between the control and F_1 groups for all tissue and strain combinations. FDR was used to correct false positives. Volcano plots showed that the effects of paternal irradiation on the pattern of gene expression for each tissue and strain were not strong enough to be detected by pair-wise analysis.

Consequently, 3-way ANOVA was used allowing the estimation of the independent effects of tissue, strain and paternal treatment. After applying an FDR < 0.05 and an arbitrary fold-change cut point of 1.4, the most significantly differentially expressed transcripts were shown to belong to circadian genes. This observation was confirmed after functional enrichment of the significantly differentially expressed genes into functional classes using ProfCom which showed enrichment for Rhythmic process and Circadian Rhythm GO-classes.

The fold-changes in expression and probabilities calculated by 3-way ANOVA for the respective transcripts of circadian genes in the top 3 GO-classes were presented for each tissue. It was observed that paternal irradiation affected the same genes in the same fashion but had different intensity of effect depending on the strain and tissue (i.e. higher sensitivity for CBA/Ca mice and kidney tissues). This observation informed that the signal of TI is susceptible to modulation and suggests a permissive mechanism of inheritance.

Chapter 5:

The pattern of expression of circadian genes

5.1. Validation of microarray data by qPCR

The microarray data was validated using qPCR which involves reverse transcription and PCR to quantify gene expression levels in cells or tissues. qPCR is based on normal PCR but is able to quantify the cDNA at all stages of amplification by using a fluorescent dye able to intercalate double-stranded DNA (Pfaffl, 2001). This approach was used in this study to measure the level of expression of the *Per2* and *Arntl* genes in RNA samples extracted from the offspring of irradiated and offspring of control male mice. These two genes showed the highest levels of deregulation in the paternally irradiated offspring, along with *Dbp* and *Npas2*. *Per2* and *Arntl* were selected, over *Dbp* and *Npas2*, since their loss of function has more severe consequences for circadian rhythmicity (Lopez-Molina *et al.*, 1997; Bunger *et al.*, 2000; Bae *et al.*, 2001; Dudley *et al.*, 2003) and for physiological health (Franken *et al.*, 2000; Fu *et al.*, 2002; Zhu *et al.*, 2007; Taniguchi *et al.*, 2009).

5.2. Relative quantification of circadian gene expression using REST 2009

The relative amount of target mRNA is quantified by comparing the target gene expression to the expression of one or more endogenous control genes. The genes used for normalisation in relative quantification are often referred to as reference or endogenous control genes and need to be stable in cells/tissues following the treatment

regime under analysis (Pfaffl *et al.*, 2002). For this study, *Aprt* and *Actb* loci demonstrated constant expression in all tissues making them suitable for normalization procedures in this study.

The REST 2009 software uses the Pfaffl method, geometric mean of multiple internal controls and randomisation and bootstrap techniques, to accurately and reliably measure relative expression of a target gene (Qiagen, 2009b). Relative quantification requires a method for normalisation against a reference gene. The Pfaffl method considers both C_t values and the qPCR reaction efficiency of target and reference genes in the control and test offspring making this method the most powerful for the analysis of relative expression (Filion, 2012, p.57 [book]). The Pfaffl method is represented by the equation below.

$$Relative\ expression = \frac{(Efficiency_{target})^{[\Delta C_t,\ target(control-test)]}}{(Efficiency_{reference})^{[\Delta C_t,\ reference(control-test)]}} \quad \textbf{(5.1)}$$

In which the denominator represents the efficiency of the reference gene to the power of the average C_t of the reference gene in the control samples minus the average C_t of the reference gene in the test samples (Pfaffl, 2001). The denominator can be called the concentration of the reference gene while the numerator can be called the concentration of the target gene (Qiagen, 2009b). The expression of endogenous control genes can, occasionally, vary considerably in a given cell type or experimental condition. Using multiple reference genes for normalisation improves the reliability of the results by compensating for unstable reference genes (Pfaffl *et al.*, 2002). When using multiple reference genes, the geometric mean of the concentration of all reference genes can be used to calculate the relative expression of a target gene as described in the equation below where n represents the number of reference genes (Vandesompele *et al.*, 2002).

$$Relative\ expression = \frac{concentration\ of\ target\ gene}{\sqrt[n]{conc.of\ reference\ gene\ 1 \times (\ldots) \times conc.of\ reference\ gene\ n}} \quad \textbf{(5.2)}$$

Relative quantification provides a ratio value which in a ratio distribution would not have a standard deviation. Thus, by itself, ratio data cannot be analysed statistically. REST 2009 overcomes this limitation using randomisation and bootstrapping methods which involve, respectively, random allocation and random reallocation of C_t values, corresponding to the same gene, into the qPCR efficiency correction equation (**equation 5.1**) from a set of control and test samples (Qiagen, 2009b). After the random allocation of C_t values into **equation 5.1** for 1 target gene and 2 reference genes (see **equation 5.2**), the relative expression was calculated. This randomisation was repeated 10000 times using the REST 2009 software to generate a distribution of relative expression values. By ranking these ratio values in order of size, a confidence interval (C.I.) can be calculated for the relative expression distribution. A non-parametric test is performed by the software to determine the probability that the difference between the control and test samples is only due to chance. This is done by generating a second relative expression distribution from the same data set but, in this case, C_t values are allocated irrespective of whether they belong to control or test samples. The software performs 10000 random reallocations of C_t values between the control and test samples, and counts the number of times the relative expression calculated from the randomly assigned samples is greater than those calculated by the properly assigned samples (Qiagen, 2009b).

The REST 2009 software was used to estimate the altered gene expression levels for *Per2* and *Arntl* genes in the offspring of irradiated mice (**Table 5.1**).

Table 5.1: Validation of the differential expression of *Per2* and *Arntl* genes by qPCR.

				qPCR data				Arrays	
Gene	ID	Tissue	Strain	Fold-change	Std. Error	95% C.I.	Prob	Fold-change	Prob
Per2	NM_011066	Kidney	CBA/Ca	0.221	0.106 - 0.545	0.008 - 0.906	**$>10^{-3}$**	0.318	**$6.93x10^{-6}$**
			BALB/c	0.507	0.260 - 1.017	0.137 - 1.898	**$>10^{-3}$**	0.621	**0.0068**
		Liver	CBA/Ca	0.355	0.113 - 0.943	0.044 - 2.804	**$>10^{-3}$**	0.554	**0.0099**
			BALB/c	0.682	0.246 - 1.886	0.081 - 4.208	**0.033**	0.819	0.4371
Arntl	NM_007489	Kidney	CBA/Ca	2.556	1.063 - 6.190	0.468 - 15.400	**$>10^{-3}$**	3.387	**$8.23x10^{-6}$**
	AB012601							3.930	**$3.94x10^{-5}$**
	AY690602							3.845	**0.0002**
	NM_007489		BALB/c	1.128	0.474 - 2.935	0.126 - 4.523	0.434	2.253	**0.0085**
	AB012601							2.120	**0.0060**
	AY690602							2.228	**0.0158**
	NM_007489	Liver	CBA/Ca	3.803	1.494 - 10.478	0.894 - 33.315	**$>10^{-3}$**	1.982	**0.0063**
	AB012601							1.885	**0.0042**
	AY690602							1.874	**0.0045**
	NM_007489		BALB/c	1.656	0.617 - 4.574	0.235 - 10.456	**0.004**	1.011	0.8443
	AB012601							0.982	0.9060
	AY690602							0.919	0.7456

The qPCR validation was performed for the manually annotated and reviewed mRNAs of *Per2* (NM_011066) and *Arntl* (NM_007489) present in NCBI RefSeq (http://www.ncbi.nlm.nih.gov/refseq/). Microarrays measured significant down- and up-regulation of *Per2* and *Arntl*, respectively, in the kidney tissues of paternally irradiated CBA/Ca and BALB/c and in the liver tissues of paternally irradiated CBA/Ca. qPCR validated *Per2* down-regulation in the kidney tissues of paternally irradiated CBA/Ca and BALB/c and in the liver tissues of paternally irradiated CBA/Ca.

Unlike microarrays, qPCR detected significant down-regulation of *Per2* in the liver tissues of paternally irradiated BALB/c. qPCR also validated *Arntl* up-regulation in kidney and liver tissues of paternally irradiated CBA/Ca. *Arntl* up-regulation was not validated in the kidney but was detected in the liver tissues of paternally irradiated BALB/c, contradicting microarray observations. *Arntl* up-regulation was generally found to be stronger than *Per2* down-regulation, with the exceptions being the strong *Per2* down-regulation in kidney tissues of both strains (qPCR data) and the absent *Arntl* up-regulation in the liver tissues of paternally irradiated BALB/c (microarray data). More importantly, *Per2* down-regulation and *Arntl* up-regulation were stronger in kidney tissues from both strains and in the tissues from paternally irradiated CBA/Ca mice.

5.3. Conclusion

The qPCR data validates and supports the results obtained from the NimbleGen 12x135K microarrays. Both microarrays and qPCR measured significant *Per2* down-regulation and *Arntl* up-regulation, with few exceptions, in tissues from paternally irradiated BALB/c. Interestingly, there was consistency between microarray and qPCR measurements on the following point: the most significant, and strongest, *Per2* down-regulation and *Arntl* up-regulation was observed preferentially in the kidney tissue over the liver and in the paternally irradiated offspring from CBA/Ca mice over those from the BALB/c mice.

qPCR further confirmed that paternal irradiation results in transgenerational changes to expression of the *Per2* and *Arntl* genes. According to literature the deregulation of a single core circadian gene can affect the whole circadian system and cause the organism to develop diseases such as cancer. This is the case for *Arntl* (Tokunaga *et al.*, 2008; Taniguchi *et al.*, 2009), *Clock* (Sahar and Sassone-Corsi, 2007 [review]; Dai *et al.*, 2011), *Npas2* (Zhu *et al.*, 2007; Zhu *et al.*, 2009), *Per1* (Shih *et al.*, 2006; Winter *et al.*, 2007; Cao *et al.*, 2009; Oda *et al.*, 2009; Xia *et al.*, 2010), *Per2* (Fu *et al.*, 2002; Shih *et al.*, 2006; Winter *et al.*, 2007; Oda *et al.*, 2009; Xia *et al.*, 2010) and *Per3* (Yang *et al.*, 2006; Yang *et al.*, 2011). Both microarray and qPCR detected significant changes in the expression of *Per2* and *Arntl* in tissues of paternally irradiated offspring questioning the mechanism underlying these effects. Differential expression of core circadian genes could result from an alteration to the DNA methylation pattern of their transcription controlling regions. This hypothesis was tested using bisulfite sequencing to analyse DNA methylation, mainly at the promoter, of genes involved in the primary and negative feedback loops of circadian transcription (most of which were, in turn, significantly differentially expressed).

Chapter 6:

DNA methylation profiling using bisulfite sequencing

6.1. DNA methylation analysis by direct bisulfite sequencing

In all mammals, DNA methylation occurs almost exclusively on the 5th carbon of the pyrimidine ring of deoxycytidine bases forming 5-methydeoxycytidine (5-me C) at CG or CpG dinucleotides. DNA methylation can cause transcriptional silencing in promoter regions particularly if there is a high CpG density region, called a CpG island, in the promoter. Low CpG density regions distant from promoters may also control gene expression (Kennedy and Oswald, 2011, p.168 [book]).

As stated in **section 1.8.4**, X-ray irradiation causes genome-wide DNA hypomethylation and promoter DNA hypermethylation that affects gene expression in tissues of directly irradiated mice/rats. There is support for the transmission of these altered DNA methylation patterns to the non-exposed offspring (Barton *et al.*, 2005; Tamminga *et al.*, 2008a; Tamminga *et al.*, 2008b; Filkowski *et al.*, 2010). This hypothesis was tested by the analysis of DNA methylation in somatic and germline tissues from paternally irradiated offspring mice using bisulfite sequencing.

6.1.1. Bisulfite conversion reaction

The best methods for determining the methylation status of a DNA sequence involve an initial incubation of the DNA with bisulfite salts. Bisulfite initiated the conversion of non-methylated cytosines (C) into uracils (U) while leaving the 5-me C unmodified. This chemical conversion of C into U is called deamination and is achieved

in vitro, in a three-step reaction, by bisulfite (i.e. sulphonation, hydrolytic deamination and desulphonation) (Qiagen, 2009a). The sulphonation reaction is carried at low pH and adds a bisulfite ion to the carbon 5-6 double bond of cytosine. Hydrolytic deamination of the resulting cytosine sulphonate requires an acidic pH and gives uracil sulphonate. Finally, an alkali treatment removes the sulphonate group to give uracil (alkali desulphonation reaction) (Craig and Wong, 2011, p.295 [book]). At the end of the treatment the converted DNA is single-stranded and of significantly lower complexity as it is composed mostly by 3-bases: adenine, guanine and adenine complementary bases (thymine and U). Bisulfite treatment provides different sequences for two identical DNA sequences if one is methylated and the other is non-methylated. The sequence from methylated DNA will still have Cs only at CpGs, while that obtained from non-methylated DNA will not have any Cs (Neumann, 2007, pp.92, 95 [book]). These differences can be detected and/or quantified using techniques such as conventional and quantitative PCR, bisulfite sequencing, methylation specific PCR (MSP), combined bisulfite restriction enzyme analysis (COBRA) and Pyrosequencing (Qiagen, 2009a).

6.1.2. Prevention of DNA degradation during bisulfite conversion

Each of the techniques on **section 6.1.1** relies on the information that a U in the converted DNA represents a non-methylated C while a C represents a 5-me C. For this reason, complete conversion is necessary in order to produce meaningful results. Unfortunately, this means incubating the DNA with high concentrations of bisulfite at high temperatures and for a long time which is estimated to cause the degradation of around 90% of the incubated DNA (Grunau *et al.*, 2001). DNA degradation occurs during deamination and is usually due to depurination (Ehrich *et al.*, 2007). Excessive

alkylation may prevent complete desulfonation of the pyrimidine residues and cause inhibition of DNA polymerases during PCR (Fraga and Esteller, 2002). For this reason, the bisulfite reaction is initiated at a slightly acidic pH5 and monitored by a pH-indicator dye (Qiagen, 2009a). The acidic condition does, however, contribute to fragment the DNA (Craig and Wong, 2011, p.298 [book]). Also, increased DNA degradation at high temperatures can be minimized by cycling the denaturation and incubation temperatures (Neumann, 2007, p.111 [book]). The cycling of the conversion temperatures was also coupled with a reduction of the denaturation time (see **section 3.10.1**).

6.1.3. Nested PCR and SM-PCR provide detection and quantification of template

Nested PCR was used to amplify the fragile and rare converted DNA template to levels detectable in a sequencing reaction (as described in **section 3.10.3**).

Coupling the nested PCR to bisulfite sequencing ensures the detection of the methylated DNA but is not able to quantify the proportion of methylation in the individual CpGs in a DNA sample. Also, conventional PCR of bisulfite converted DNA has a bias towards amplification of non-methylated DNA. The bisulfite converted methylated DNA is a template with a higher (G + C) content and this may increase the melting temperature and contribute to the formation of secondary structures, which reduces PCR efficiency. Thus, a sample containing a mixture of methylated and non-methylated bisulfite-treated DNA should show a higher proportion of non-methylated DNA after PCR (Warnecke *et al.*, 1997).

During the mapping of methylation patterns in DNA, the PCR-bias can be avoided by diluting the bisulfite converted DNA to the point where SM-PCR can be

used to accurately determine the proportion of variants of the same template (Chhibber and Schroeder, 2008). At SM level, there is no preferential amplification of an "easy" template as a single variant of a template, containing a single collection of SNPs, is present. The PCR-bias can happen due to: (1) the polymerase slowing down when reaching the "difficult" SNP or polymorphic region (e.g. G/C-rich region with secondary structure potential); and, (2) due to template change in which the polymerase switches in the middle of the amplification to an easier template. Both events contribute to the underestimation of the "difficult" template when compared to the "easy" template and are avoided in SM-PCR (Kraytsberg and Khrapko, 2005 [review]).

6.1.4. Assessing the success of SM-PCR using the Poisson distribution

The success of PCR in amplifying a DNA template can be determined by detecting a band of DNA on a gel. In SM-PCR there is a series of PCR reactions for one sample and the success of the procedure depends on the number of reactions with positive and negative PCR product. To be more precise, what is actually being measured is the number of DNA templates at the start of the SM-PCR. Negative PCR reactions are expected to have no templates while positive PCR reaction can have ≥1 template(s).

Performing PCR on the samples containing on average one amplifiable molecule results either in successful amplification (≥1 template) or a complete absence of PCR products. The distribution of the number of positive and negative PCR reactions can be approximated by the Poisson distribution (Sokal and Rohlf, 2012, pp.78-80 [book]).

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \qquad \textbf{(6.1)}$$

where λ is the mean number of amplifiable molecules per reaction. Using the frequency of negative PCR reactions, $P_0 = P(X = 0) = e^{-\lambda}$, the mean number of amplifiable molecules per reaction was estimated as:

$$\lambda = -\ln\left(\frac{total\ number\ of\ negative\ PCR\ reactions}{total\ number\ of\ PCR\ reactions}\right) \quad \textbf{(6.2)}$$

For example, if among the 24 SM-PCR reactions 9 negatives were recorded, the mean number of amplifiable molecules per reaction is $\approx -\ln\left(\frac{9}{24}\right) \approx 0.98$.

6.1.5. Sanger dye-terminator sequencing

Direct bisulfite sequencing involves bisulfite treatment of DNA, amplification using a primer specific to a single-stranded bisulfite treated template (avoiding annealing to bisulfite treated CpG cytosines to prevent incomplete annealing) and automated sequencing. The Sanger sequencing method used involved dye-terminator sequencing by random incorporation of four chain-terminating dideoxynucleotides (ddNTPs) by the DNA polymerase in a single DNA replication reaction (Sanger *et al.*, 1977). Each of the four ddNTPs was labelled with fluorescent dyes that emitted light at different wavelengths. The sequencing end-reaction was submitted to a Sequencing machine which performed capillary electrophoresis separating the labelled DNA according to size before a fluorescence detector could detect each ddNTP and assemble a chromatogram representing the sequenced DNA (Karger and Guttman, 2009).

6.2. The *Per2* gene is an ideal candidate for DNA methylation analysis

The knockout of *Per2* has been associated with increased cell proliferation, reduced apoptosis, increased onset of cancer, and reduced life expectancy in mice (Fu *et al*., 2002). The down-regulation of *Per2* has also been associated with the development of cancer (Shih *et al.*, 2006; Winter *et al.*, 2007; Oda *et al.*, 2009; Xia *et al.*, 2010).

In this study the *Per2* gene was found to be significantly down-regulated in the male offspring of irradiated male mice. Other circadian genes were also found to be significantly differentially expressed and the direction of their expression change supports the view that *Per2* deregulation, as a consequence of paternal irradiation, may not be a coincidence.

Epigenetic changes affecting the pattern of DNA methylation have been suggested as a plausible mechanism underlying the transgenerational effects of paternal irradiation (Dubrova, 2003 [review]; Barber and Dubrova, 2006 [review]). To verify whether this may be the case for the observed changes in expression of this locus, the pattern of *Per2* methylation was established in DNA samples extracted from kidney of the offspring of irradiated and control males (**Figure 6.1**).

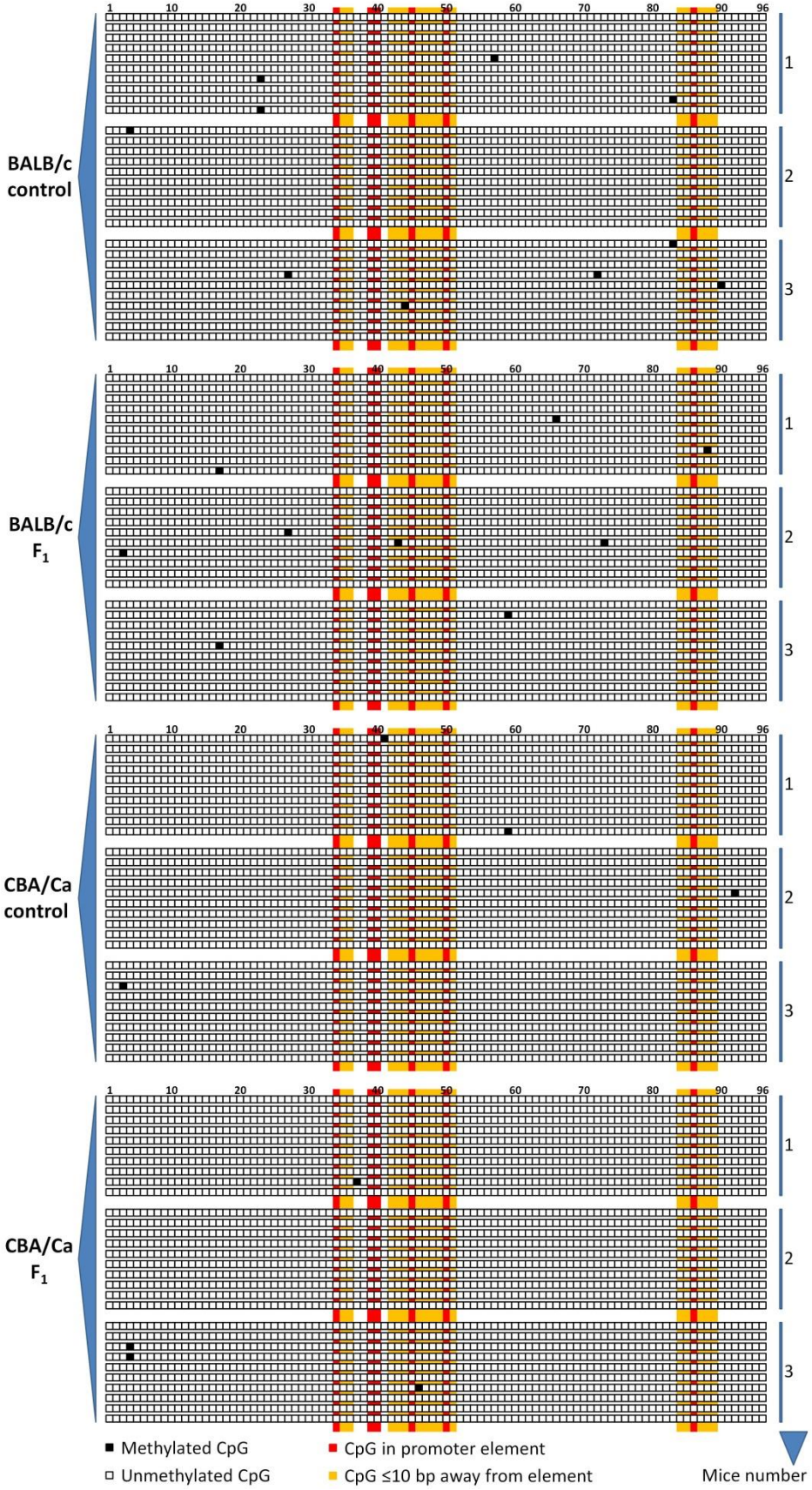

Figure 6.1: The pattern of CpG island methylation of the *Per2* promoter in DNA samples extracted from kidney. Three mice were profiled for control mice and offspring

of paternally irradiated mice belonging to the BALB/c and CBA/Ca strains. For each mouse, ten DNA samples were bisulfite treated and sequenced. The five CpG promoter elements Sp1/CG box 1, 2, and E'-box 2, 3, and D-box 2 are shown in red. This analysis failed to detect any changes in the pattern of DNA methylation among the CBA/Ca and BALB/c offspring of irradiated male mice.

In both strains, the pattern of *Per2* methylation in DNA samples extracted from kidney did not significantly differ between F_1 and controls. Random methylation of a single CpG in the amplicons of *Per2* was often detected in all groups of mice and may be attributed to an incomplete deamination during the bisulfite conversion. However, the deamination efficiency was around 99.2% (8 incompletely deaminated cytosines per 1000 cytosines) which is higher than the acceptable ≥98% (Darst *et al.*, 2010). Moreover, the ten repetitions of methylation analysis in *Per2* for each mouse tissue studied allowed the incompletely deaminated and non-methylated CpGs to be easily detected and excluded from further analysis.

According to the hypothesis that heritable changes in DNA-methylation patterns underlie the transgenerational effects, the erroneous methylation pattern should be established in the germline DNA of the irradiated parents and be transmitted to their offspring. To verify this, the pattern of DNA methylation of the *Per2* gene was also analysed in sperm samples taken from directly irradiated BALB/c and CBA/Ca male mice, as well as from the first-generation offspring of paternally irradiated and offspring of control mice (**Figures 6.2**, **6.3**). The sperm samples collected from the non-exposed F_1 offspring of control mice were used as the reference for all comparisons. For this analysis, sperm was taken from the caudal epididymis.

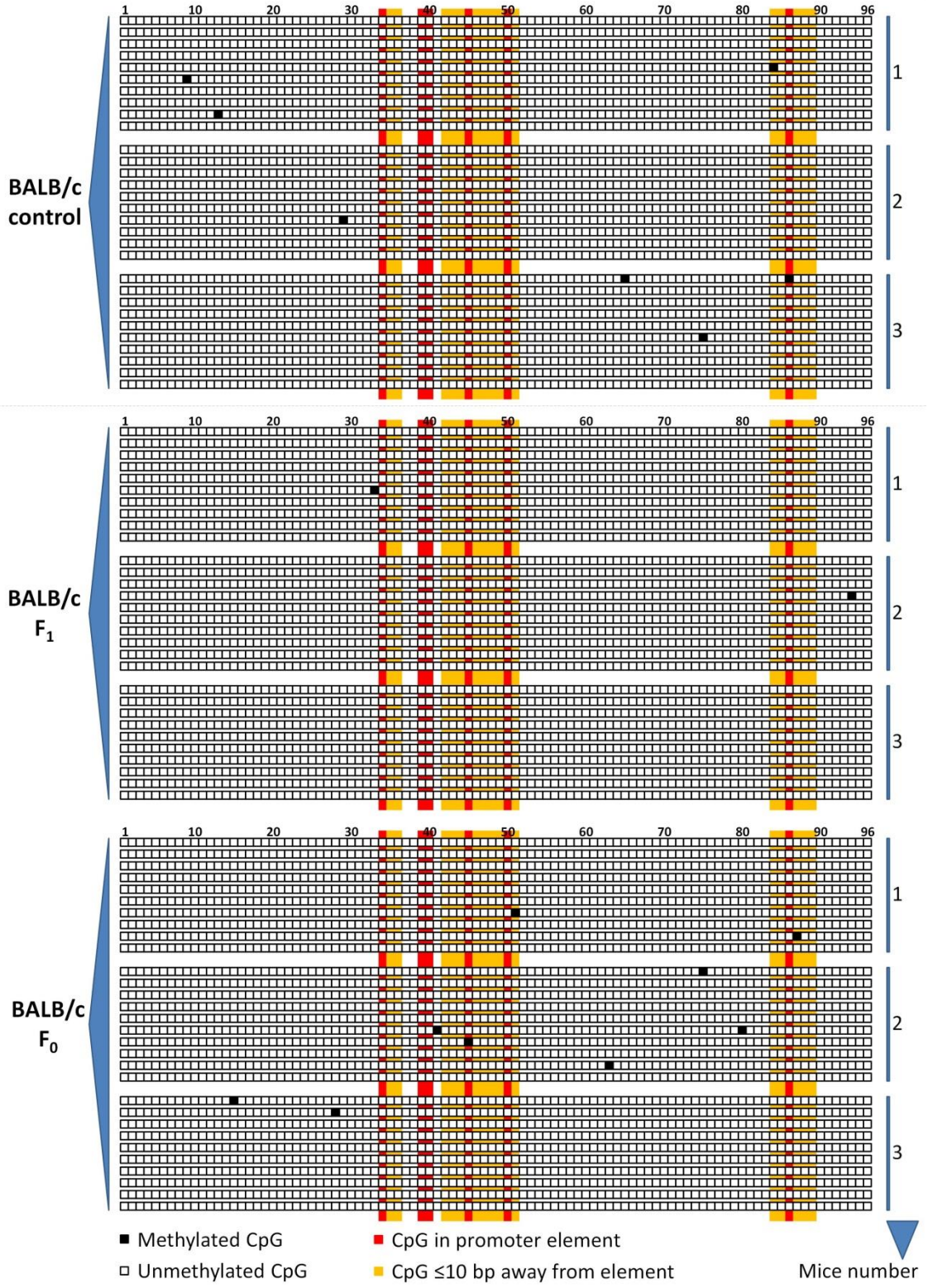

Figure 6.2: Pattern of CpG island methylation of *Per2* promoter in sperm DNA from directly irradiated, offspring of irradiated and offspring of control BALB/c males. Three mice were profiled in each group. For each mouse, ten DNA samples were bisulfite treated and sequenced. The five CpG promoter elements Sp1/CG box 1, 2, and E'-box

2, 3, and D-box 2 are shown in red. This analysis failed to detect any changes in the pattern of DNA methylation in the germline of irradiated male mice and their male offspring.

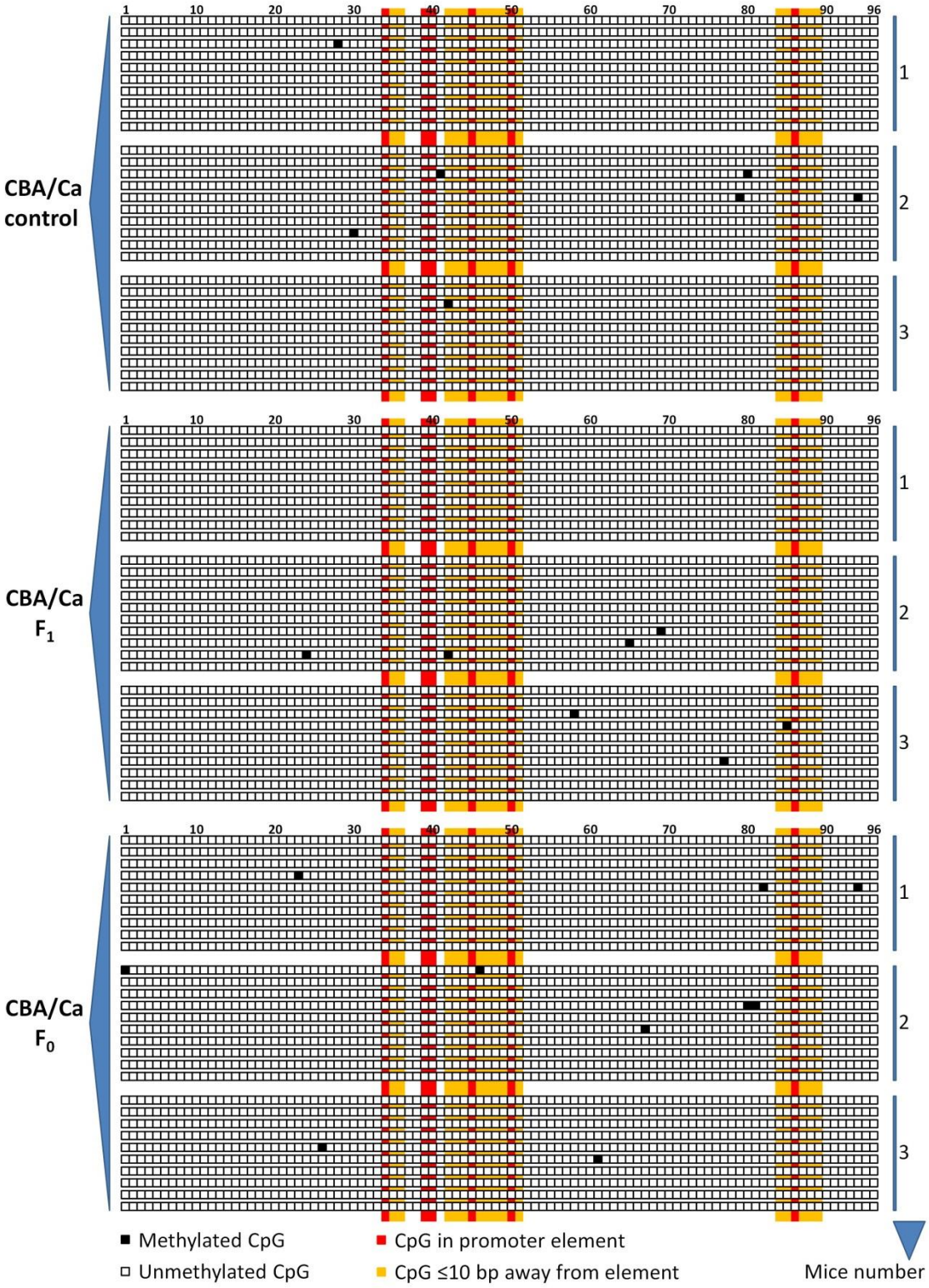

Figure 6.3: Pattern of CpG island methylation of *Per2* promoter in sperm DNA from directly irradiated, offspring of irradiated and offspring of control CBA/Ca males. Three mice were profiled in each group. For each mouse, ten DNA samples were bisulfite treated and sequenced. The five CpG promoter elements Sp1/CG box 1, 2, and E'-box

2, 3, and D-box 2 are shown in red. This analysis failed to detect any changes in the pattern of DNA methylation in the germline of irradiated male mice and their male offspring.

The *Per2* gene showed no differential methylation in sperm samples from directly irradiated males, or their non-exposed offspring, and the control mice belonging to the BALB/c and CBA/Ca strains.

It does not seem that the differential expression of *Per2* in offspring of irradiated males can be explained by differential DNA methylation in the *Per2* gene. However, there is a possibility that a disruptive DNA methylation pattern is present in other circadian genes and that the down-regulation of *Per2* is a consequence of their differential expression.

6.2.1. Conventional bisulfite sequencing of circadian genes

Because SM bisulfite sequencing is a lengthy and costly procedure, a preliminary analysis was performed using (quick and cheap) conventional bisulfite sequencing of the circadian genes in kidney and sperm of F_1 and control groups of BALB/c mice. The sperm of the F_0 group of BALB/c mice was also included in this analysis. The sequencing using amplicons from normal PCR will be less reliable than those from SM-PCR as there may be a bias of up to 40% towards amplification of non-methylated strands during PCR (Warnecke *et al.*, 1997). However, the sequencer greatly exaggerates the size of the rare C peaks on a converted template. So, as long as the percentage of methylation on circadian genes is higher than that of the bias, conventional bisulfite sequencing may be enough to detect methylation.

This analysis was meant to be used as an exploratory tool and repeated using the more reliable SM bisulfite sequencing but this was not possible due to time limitations.

6.2.1.1. CpG methylation in *Per1* gene

The *Per1* gene is located on chromosome 1, has a length of 11,007 bp and consists of 23 exons. It has a proximal promoter of about 1.9 kb ranging [-1803, +40] (Ishida *et al.*, 2008; Ji *et al.*, 2010) but the full promoter is much larger and ranges [-5602, +1646] (Yamaguchi *et al.*, 2000; Li *et al.*, 2004). The full promoter of *Per1* contains 5 E-boxes (one of which forms an EE-element), 2 D-boxes, 4 CRE elements, 3 Sp1 elements and 1 CTF/NF-1 element (Hida *et al.*, 2000; Yamamoto *et al.*, 2004; Kumaki *et al.*, 2008; Ukai-Tadenuma *et al.*, 2008; Yamajuku *et al.*, 2010; Ye *et al.*, 2011)[2]. In humans, the 510 bp upstream of exon 1 (mouse homolog in sequenced region 4, **Figure 6.4-6**) contains the core promoter of *Per1* and E-boxes 4 and 5 (Taruscio *et al.*, 2000). Methylation of these two E-boxes plus E-box 3, upstream of the core promoter, is sufficient to completely inhibit *Per1* expression (Ji *et al.*, 2010). The position of the CpGs located in promoter elements are highlighted in red on **Figure 6.4** and are present in the following elements: D-box 1, E-box 1, CRE 2 and E-box 2 (sequenced region 1); Sp1, CRE 3 and E-box 3 (sequenced region 3); E-box 4 and 5 and EE-element (sequenced region 4); CRE 4 and Sp1 (sequenced region 5).

Bisulfite sequencing of *Per1*, promoter and genic regions, detected CpG methylation in the D-box 1 element for sperm samples of test mice (F_0 and F_1) and no methylation in the control BALB/c mice. This differential CpG methylation could explain a possible *Per1* down-regulation but that was not observed for any of the

[2] Some elements in *Per1/2/3*, *Arntl* and *Npas2* genes may result from predictions made by the referenced authors. Yamamoto *et al.* (2004) and Yamajuku *et al.* (2010) are in agreement with most of the literature and state that *Per1* has 5 E-boxes and 1 D-box, *Per2* has 2 E-boxes and 1 D-box, *Per3* has 2 D-boxes and both *Arntl* and *Npas2* have 2 RRE elements each.

strains. Also, this methylation does not occur in the same manner in somatic DNA of the F_1 and control group samples (**Figure 6.4-3**). There is strong CpG methylation in exons 20 and 21 and associated introns of *Per1* for all samples analysed (**Figure 6.4-8**).

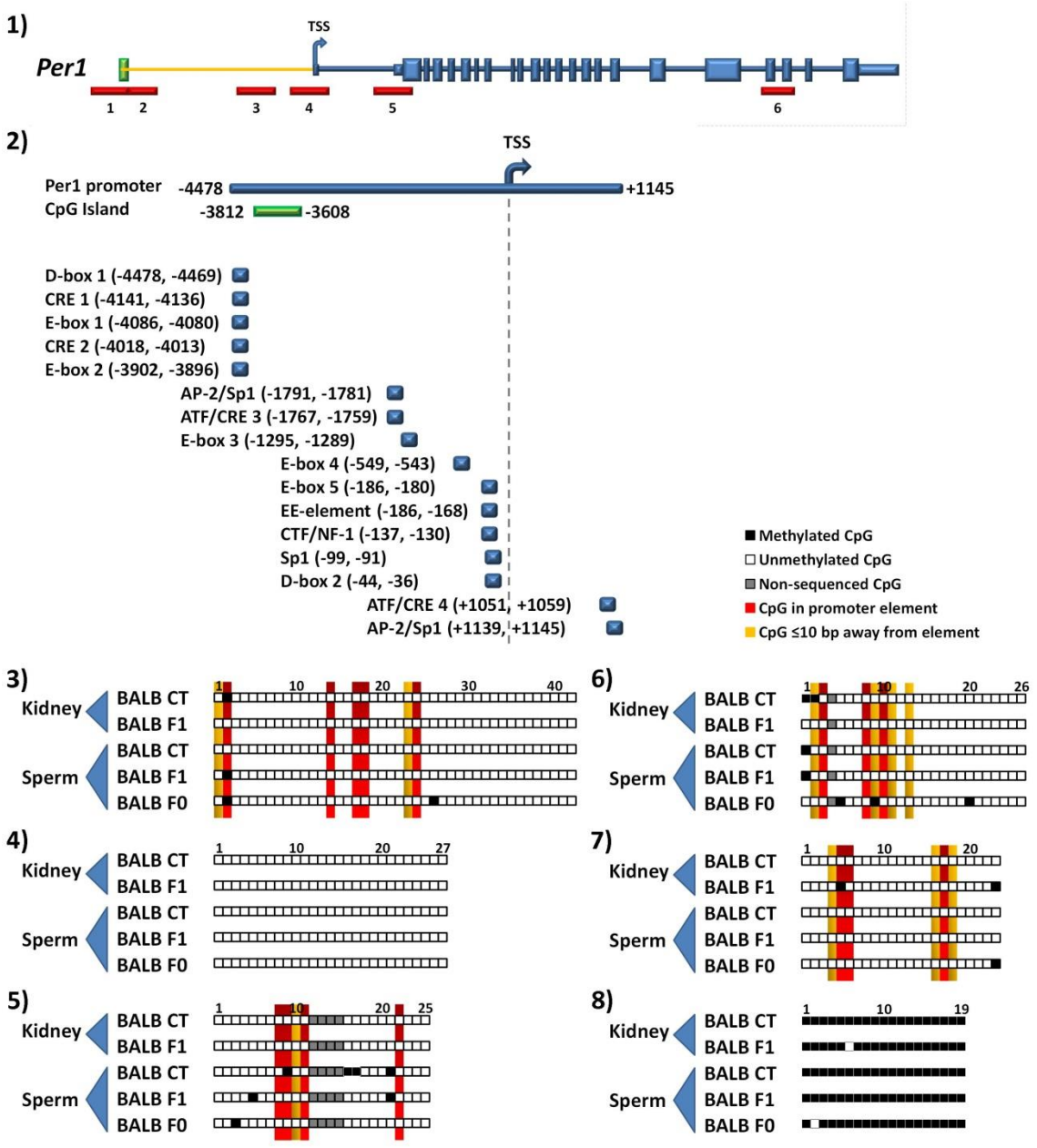

Figure 6.4: CpG methylation analysis of the *Per1* gene in kidney and sperm of F_1 and CT groups and in the sperm of F_0 group in BALB/c mice. **1)** Representation of the forwardly oriented *Per1* gene highlighting the six regions submitted for CpG methylation analysis (red), the CpG island (green), the untranscribed (yellow) and

transcribed (blue) regions and the exonic untranslated (small blue rectangle) and translated (large blue rectangle) regions. **2)** *Per1* promoter region containing a 204 bp-long CpG island and 16 different promoter elements (blue rectangles). **3)** Sequenced region 1 presents CpG methylation in the test mice (F_0 and F_1) and no methylation in the control mouse. This differential CpG methylation rests on top of the D-box 1 of the *Per1* gene in germline DNA but unfortunately is not consistent in the test and control samples of somatic DNA. **4)** Sequenced region 2, no CpG methylation. **5)**, **6)** and **7)** Sequenced regions 3, 4 and 5, respectively, present no relevant CpG methylation to explain effects of treatment. **8)** Sequenced region 6 contains exons 20 and 21 and some intronic regions of *Per1* and present high DNA methylation including almost all CpGs from test and control samples.

6.2.1.2. CpG methylation in *Per2* gene

The *Per2* gene is located on chromosome 1, has a length of 43,313 bp and consists of 23 exons. The proximal and core promoters of *Per2* are within the range [-181, +1]. The full promoter of *Per2* contains: a distant CRE element; a region ranging [-254, +41] that is conserved between mice and humans and contains 3 E-boxes (with two of them forming an EE-element), 1 D-box and 2 Sp1 elements; and, a second D-box downstream of the TSS (Etchegaray *et al.*, 2003; Ohno *et al.*, 2007; Yamajuku *et al.*, 2010; Koyanagi *et al.*, 2011). In **Figure 6.5**, the CpGs highlighted in red sit on promoter elements. The CRE element is present in sequenced region 1 while the remaining elements are in sequenced region 2.

Bisulfite sequencing of *Per2*, promoter and genic regions, detected CpG methylation in the CRE element in kidney and sperm of the F_1 mouse group and in the kidney of the control mouse group. Sequenced region 3 presents methylation in 2-3

CpGs belonging to sperm samples from all groups. There was strong CpG methylation in *Per2* exons 2, 10, 11 and 21, and surrounding intronic regions, for all samples analysed (**Figure 6.5**). No CpG methylation, able to explain the down-regulation of *Per2* in the offspring of irradiated males, was observed for the BALB/c mice.

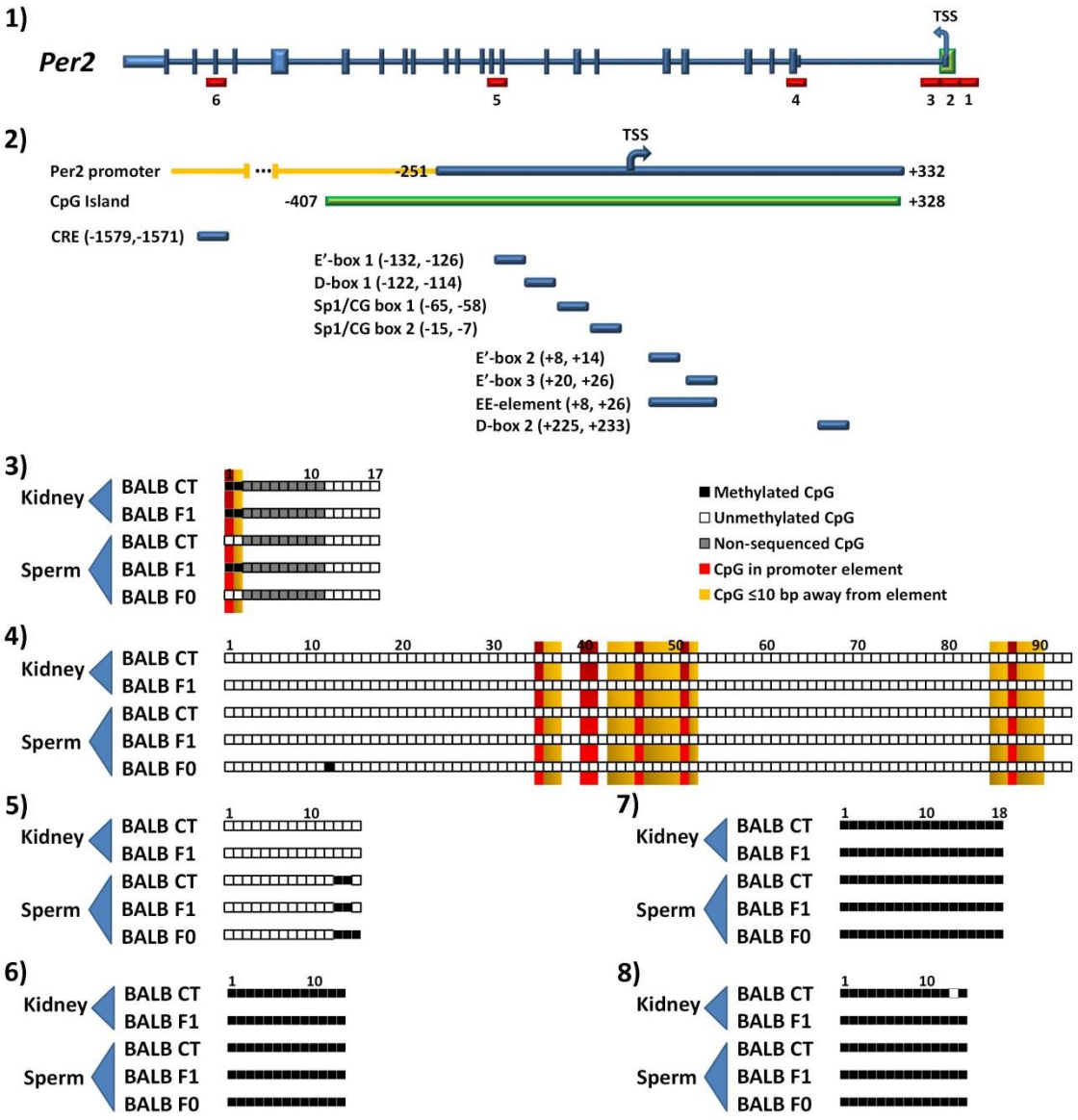

Figure 6.5: CpG methylation analysis of the *Per2* gene in kidney and sperm of F_1 and CT groups and in the sperm of F_0 group in BALB/c mice. **1)** Representation of the reversely oriented *Per2* gene highlighting the six regions submitted for CpG methylation analysis (red), the CpG island (green), the transcribed (blue) region and the

exonic untranslated (small blue rectangle) and translated (large blue rectangle) regions. **2)** *Per2* core promoter region (blue) containing a 735 bp-long CpG island (green) 7 different proximal promoter elements plus one distant promoter element (blue rectangles). **3)** Sequenced region 1 presents CpG methylation in the CRE promoter element in kidney and sperm of the F_1 mouse group and in the kidney of the control mouse group. **4)** Sequenced region 2 includes almost every CpG in the CpG island and presents no significant methylation. **5)** Sequenced region 3 presents methylation in 2-3 CpGs belonging to sperm samples from all groups. **6)** Sequenced region 4 contains exon 2 plus introns. **7)** Sequenced region 5 contains exons 10 and 11 plus introns. **8)** Sequenced region 6 contains exon 21 plus introns. Sequenced regions 4, 5 and 6 contain virtually complete CpG methylation in test and control samples.

6.2.1.3. CpG methylation in *Per3* gene

The *Per3* gene is located on chromosome 1, has a length of 39,878 bp and consists of 21 exons. The promoter of *Per3* contains 2 D-boxes and one EE-element composed of two E-boxes (Nakahata *et al.*, 2008; Yamajuku *et al.*, 2010). In humans, the promoter region of *Per3* is within the range [-874, +1] (Archer *et al.*, 2010).

Bisulfite sequencing of the *Per3* promoter and CpG island detected no CpG methylation in kidney and sperm samples from test (F_0 and F_1) and control BALB/c mice (**Figure 6.6**). No CpG methylation, able to explain the down-regulation of *Per3* in the offspring of irradiated males, was observed for the BALB/c mice.

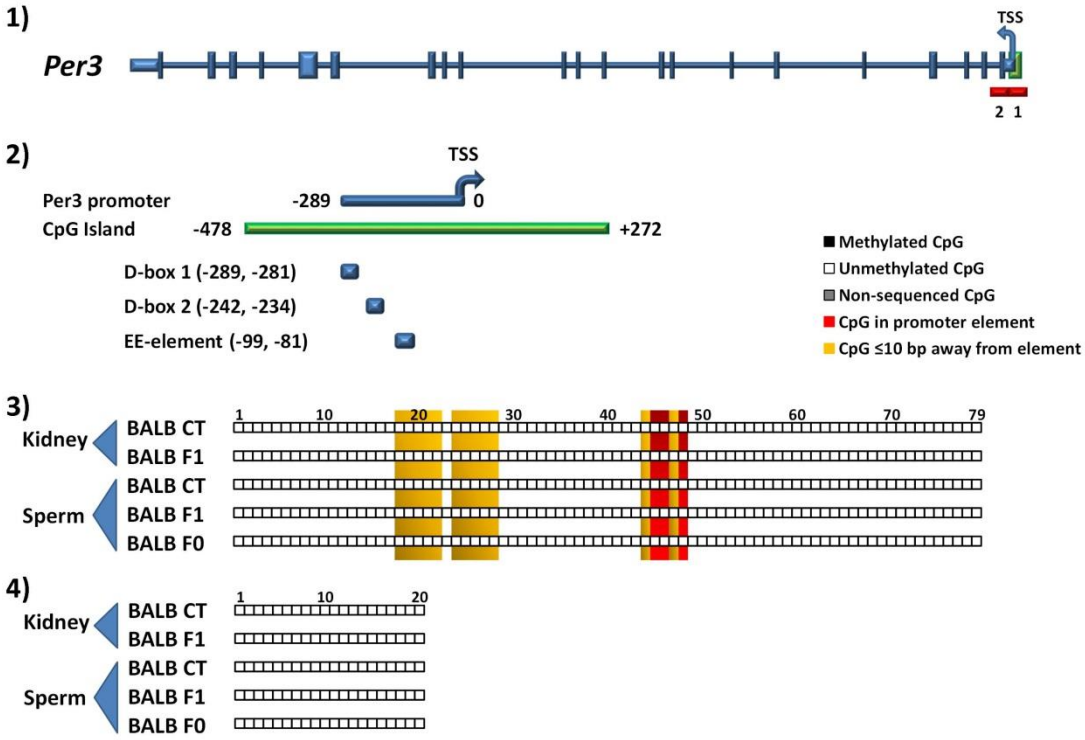

Figure 6.6: CpG methylation analysis of the *Per3* gene in kidney and sperm of F_1 and CT groups and in the sperm of F_0 group in BALB/c mice. **1)** Representation of the reversely oriented *Per3* gene highlighting the two regions submitted for CpG methylation analysis (red), the CpG island (green), the transcribed (blue) region and the exonic untranslated (small blue rectangle) and translated (large blue rectangle) regions. **2)** *Per3* core promoter region (blue) is within a 750 bp-long CpG island (green) and contains 3 different promoter elements (blue rectangles). **3)** Sequenced region 1 includes every CpG in the CpG island and presents no methylation in test and control samples. **4)** Sequenced region 2 contains exon 1 plus introns and also presents no methylation in test and control samples.

6.2.1.4. CpG methylation in *Arntl* gene

The *Arntl* gene is located on chromosome 7, has a length of 106,658 bp and consists of 20 exons. The promoter of *Arntl* ranges [-422, +108] contains 2 RRE

elements and can be silenced by hypermethylation on the region ranging [-290, +49] (Preitner *et al.*, 2002; Sato *et al.*, 2004; Taniguchi *et al.*, 2009). In **Figure 6.7**, the CpGs highlighted in yellow sit ≤10 bp away from the promoter elements RRE1 or RRE2.

Bisulfite sequencing of *Arntl* detected no CpG methylation in the promoter and CpG island (sequenced regions 1, 2 and 3) for test and control BALB/c mice. There was strong CpG methylation in exon 5 and associated introns of *Arntl* for all mouse groups (**Figure 6.7**). There was no *Arntl* promoter hypermethylation for normal BALB/c mice reported in literature or observed in this analysis. Thus, CpG methylation results for *Arntl* alone are unable to explain *Arntl* up-regulation in the offspring of irradiated BALB/c male mice.

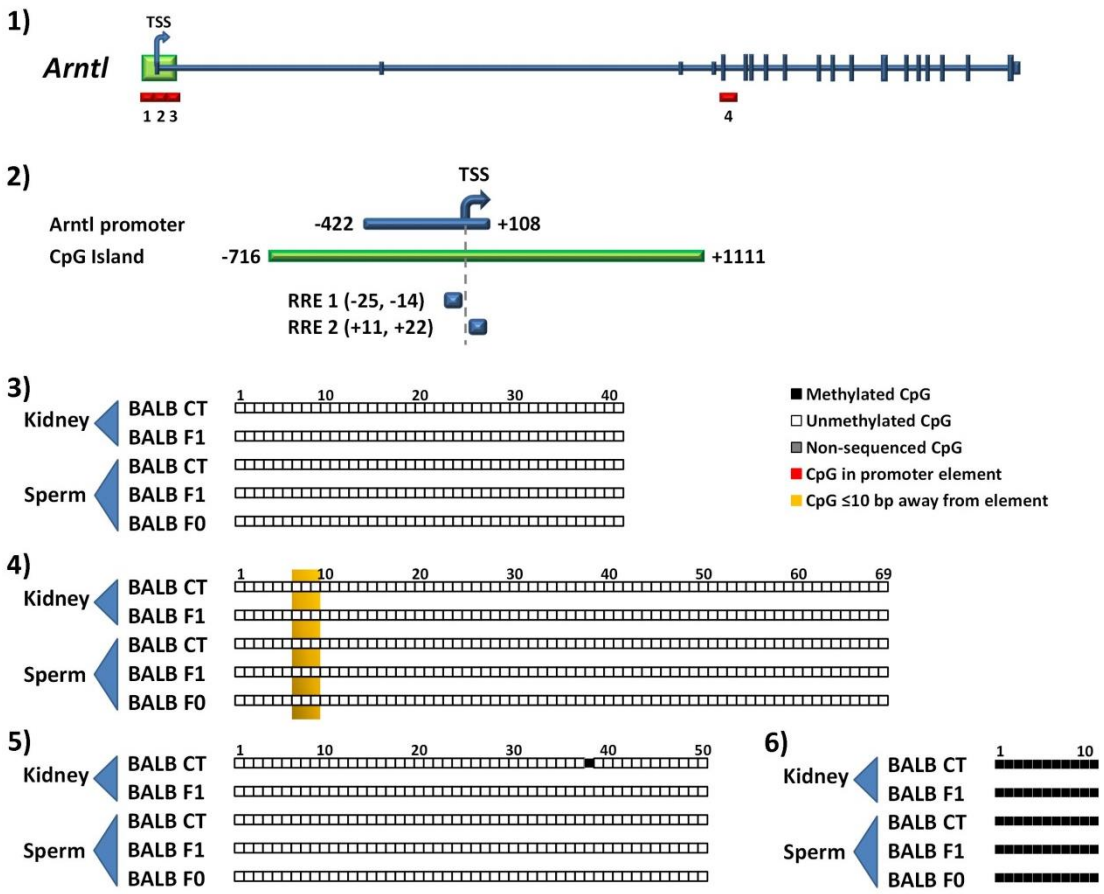

Figure 6.7: CpG methylation analysis of the *Arntl* gene in kidney and sperm tissues of F_1 and CT groups and in the sperm of F_0 group in BALB/c mice. **1)** Representation of

the forwardly oriented *Arntl* gene highlighting the four regions submitted for CpG methylation analysis (red), the CpG island (green), the transcribed (blue) regions and the exonic untranslated (small blue rectangle) and translated (large blue rectangle) regions. **2)** *Arntl* core promoter region is within a 1,827 bp-long CpG island and contains 2 different promoter elements (blue rectangles). **3)** Sequenced region 1 includes the CpGs from the 5' third portion of the CpG island and part of the core promoter. **4)** Sequenced region 2 includes the CpGs from the centre of the CpG island and those next to the promoter elements. **5)** Sequenced region 3 includes the CpGs from the 3' third portion of the CpG island. Sequenced regions 1, 2 and 3 include all the CpGs present in the promoter and CpG island of *Arntl* but present virtually no methylation in test and control samples. **6)** Sequenced region 4 contains exon 5 and some intronic regions and presents high DNA methylation including almost all CpGs from test and control samples.

6.2.1.5. CpG methylation in *Npas2* gene

The *Npas2* gene is located on chromosome 1, has a length of 167,975 bp and consists of 21 exons. The *Npas2* promoter ranges [-1534, +81] contains 3 RRE elements and most of the CpG island (Yamamoto *et al.*, 2004; Takeda *et al.*, 2011). In **Figure 6.8**, the CpGs highlighted in yellow sit ≤10 bp away from the promoter elements RRE1 or RRE2.

Bisulfite sequencing of *Npas2* detected no CpG methylation in the promoter and CpG island (sequenced regions 1, 2 and 3) for test and control BALB/c mice. There was strong CpG methylation in exon 2 and associated introns of *Npas2* for all mouse groups (**Figure 6.8**). There was no *Npas2* promoter hypermethylation for normal BALB/c mice reported in literature or observed in this analysis. Thus, CpG methylation results for

Npas2 alone are unable to explain *Npas2* up-regulation in the offspring of irradiated BALB/c male mice.

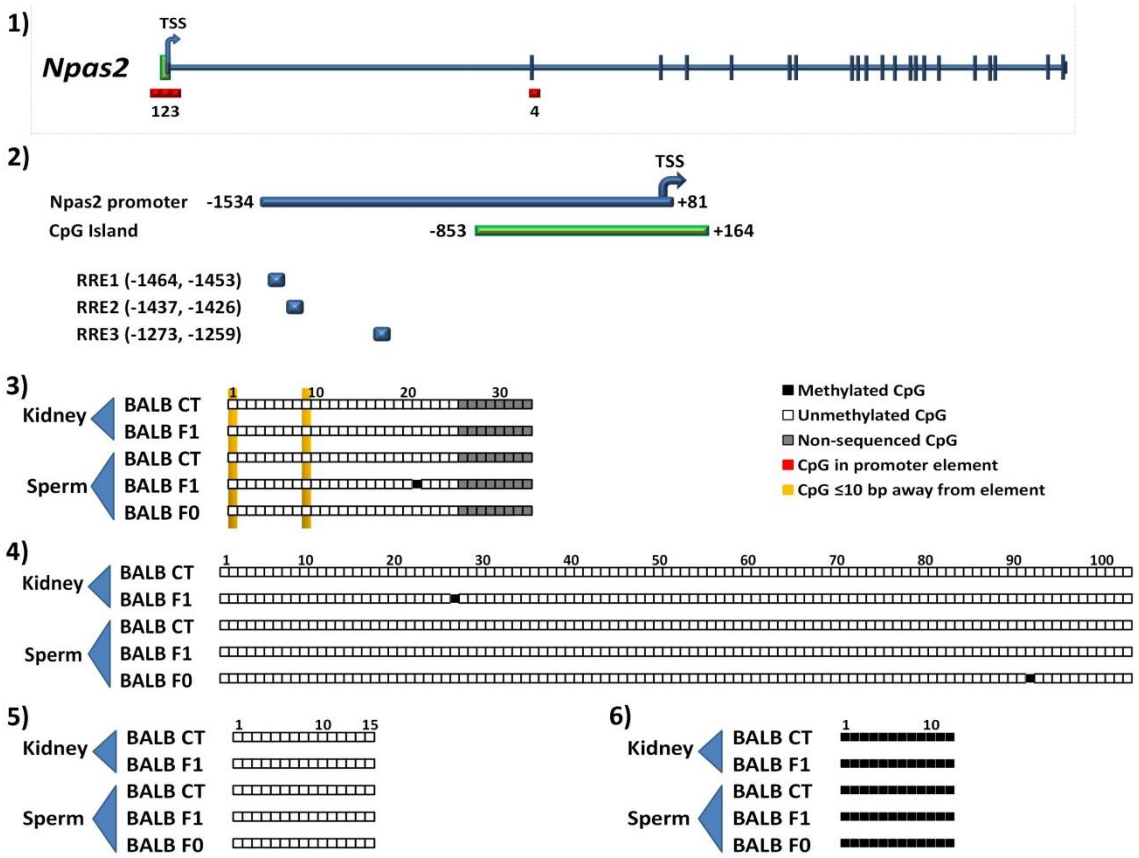

Figure 6.8: CpG methylation analysis of the *Npas2* gene in kidney and sperm tissues of F_1 and CT groups and in the sperm of F_0 group in BALB/c mice. **1)** Representation of the forwardly oriented *Npas2* gene highlighting the four regions submitted for CpG methylation analysis (red), the CpG island (green), the transcribed (blue) regions and the exonic untranslated (small blue rectangle) and translated (large blue rectangle) regions. **2)** *Npas2* promoter region includes most of a 1,017 bp-long CpG island and contains 3 different promoter elements (blue rectangles). **3)** Sequenced region 1 includes the CpGs near the promoter elements RRE 1 and 2. **4)** Sequenced region 2 includes the CpGs near the promoter element RRE 3 and from most of the CpG island. **5)** Sequenced region 3 includes the CpGs from the 3' remaining portion of the CpG

island. Sequenced regions 1, 2 and 3 include all the CpGs present in the promoter and CpG island of *Npas2* but present no methylation in test and control samples. **6)** Sequenced region 4 contains exon 2 and some intronic regions and presents high DNA methylation including all CpGs from test and control samples.

6.2.2. Conclusion

Although the use of conventional bisulfite sequencing was meant as a preliminary analysis before the use of SM bisulfite sequencing, a high PCR-bias could prevent detection of low ratios of methylation over non-methylation at each CpG in a sample of DNA template. There were methods that if coupled to the conventional bisulfite sequencing would have significantly reduced PCR-bias: (1) reducing the number of PCR cycles to < 30 or starting with very small amounts of template (e.g. 6800 copies) (Kennedy and Oswald, 2011, pp.178-179 [book]); and, (2) using primers designed to bind to 1-2 methylated CpGs in the template or increasing the annealing temperature during PCR (Moskalev *et al.*, 2011). The best method is to measure PCR-bias and relies on running parallel PCR reactions with templates of known methylation percentage (both high and low) which are used post-PCR to correct data obtained from samples of interest, as described in literature (Warnecke *et al.*, 1997; Moskalev *et al.*, 2011). This latter method is faster and less laborious than the previous alternatives (Moskalev *et al.*, 2011).

A more serious misjudgement was made by giving priority to the BALB/c strain for methylation analysis when it was the CBA/Ca strain that presented the strongest changes in circadian gene expression in both kidney and liver tissues from paternally irradiated offspring.

Despite detecting a significant change in expression of the circadian genes in microarray and qPCR analyses, no relevant CpG methylation was found in the CpG islands or promoters of the affected genes that could explain the changes observed.

SM bisulfite sequencing established that there is no methylation change in *Per2* regulatory elements in sperm and kidney tissues of paternally irradiated BALB/c and CBA/Ca offspring. Thus, the hypothesis that *Per2* DNA methylation changes were responsible for the *Per2* down-regulation observed in the affected offspring was rejected. SM bisulfite sequencing also determined that there are no radiation-induced *Per2* DNA methylation changes in the sperm tissues of directly exposed BALB/c and CBA/Ca paternal mice. Thus, the hypothesis that an abnormal *Per2* DNA methylation pattern was established in the paternal genome, before fertilisation, and caused TI was rejected (but only for the *Per2* gene). Another hypothesis, that was not tested, was that the *Per2* DNA methylation pattern was established in the paternal genome, after fertilisation, but removed during embryonic reprogramming in the offspring and that repressing histone marks remained to down-regulate *Per2* in the offspring. This hypothesis is based on the observation that gene silencing by DNA methylation is usually preceded by formation of repressing PTMs on histones (Miranda and Jones, 2007 [review]).

Because the methylation analyses carried out for other circadian genes were not completely conclusive, there is still a chance that DNA methylation in other circadian genes could be responsible for the observed differential gene expression in the paternally irradiated offspring. If, on the other hand, the results obtained by conventional bisulfite sequencing are accurate, then it is likely that CpG methylation is not the main repressing vehicle for the circadian genes. Other epigenetic mechanisms (non-coding RNA and changes in histones) could be involved in the control of gene

expression. RNA interference of *Per2* transcription could potentially cause a chain-reaction affecting the expression of other circadian genes, similar to that observed in the paternally irradiated offspring. However, between these two mechanisms, abnormal histone marks induced by radiation are the most likely event for reasons listed in **section 1.8.1**. Additionally, a recent study shows that circadian RNA polymerase II binding induces circadian histone modifications and circadian transcription initiation (Koike *et al.*, 2012; more on this topic in the discussion).

Chapter 7:

Discussion

7.1. Discussion

7.1.1. Study of transgenerational transmission of GI in mice

IR has the potential to cause varied cellular damage (**section 1.1**) and indirect damage by non-targeted effects such as GI and TI (Barber *et al.*, 2002; Barber *et al.*, 2006). Both *in vivo* and *in vitro* studies have measured changes in gene expression in directly irradiated organisms/cells and most of them detect changes in genes involved in DNA replication/repair, proliferation/apoptosis, cell cycle control or RNA processing/protein turnover (Burns and El-Deiry, 2003; Snyder and Morgan, 2004a; Snyder and Morgan, 2004b; Dai *et al.*, 2006). Both *in vivo* and *in vitro*, the direct gene expression responses to IR are heterogeneous meaning that the number and type of genes, as well as the orientation of expression change, varies significantly between studies (Burns and El-Deiry, 2003; Snyder and Morgan, 2004a; Dai *et al.*, 2006; Zhao *et al.*, 2006). In fact, even in different tissues extracted from the same irradiated wild-type mice, the radiation-induced change in expression profiles could share a proportion of similarity as low as 3% (for the type of affected genes and orientation of change) (Zhao *et al.*, 2006). Additionally, the immediate and long-term transcriptional responses to IR were found to be different in cell cultures (Snyder and Morgan, 2004b). Thus, if there is an altered gene expression component to the GI phenotype it would be difficult to find it in directly irradiated cells or animals given the number of cellular responses activated

by IR. For this reason, this study focuses on analysing the transcription of paternally irradiated offspring which inherit GI in a process called TI and in whom the increased DNA damage and mutation rate is caused almost exclusively by GI (Barber *et al.*, 2006). This is because pre-mutational lesions in sperm DNA can be effectively repaired in the egg a few hours after fertilisation (Generoso *et al.*, 1979; Brandriff and Petersen, 1981).

7.1.1.1. Summary of DNA microarray data validation

The gene expression study was initiated by extracting RNA samples from kidney, liver, spleen and brain tissues of paternally irradiated F_1 offspring and control offspring. These samples were then hybridised to the NimbleGen 12x135K multiplex arrays to measure the expression of 42,575 known mouse transcripts. Technical replicates of each sample were performed to measure reproducibility. Reproducibility of the microarray data was evaluated by similarity comparison of gene expression profiles among samples using the Pearson's correlation coefficient and the Euclidean distance. The correlation coefficient measured the influence of stochastic variation and strain on gene expression profiles and confirmed a positive correlation between all samples (necessary for the use of Euclidean distance). The Euclidean distance measured the influence of paternal irradiation and strain on gene expression profiles. Pairwise comparison of samples were divided into classes which were, in turn, statistically compared in pairs using Mann-Whitney and Kolmogorov-Smirnov tests or in triplets using the Kruskal-Wallis test.

For stochastic variation, the Mann-Whitney and Kolmogorov-Smirnov tests confirmed, respectively, significantly different location and distribution of correlation coefficient values for within and between replicates when the data for each tissue was

combined (**Table 4.1**; **Figure 4.1**). These results validated the microarray technique and confirmed low stochastic variation. For strain variation, the same statistical tests showed significantly different location or distribution of correlation coefficient values (**Table 4.2**; **Figures 4.2-4.5**) or of Euclidean distance values (**Table 4.5**; **Figure 4.8-4.11**) for within and between strains when the data for each tissue was considered separately or combined. Additionally, the Kruskal-Wallis test showed different location of correlation coefficient values (**Figure 4.6**), or of Euclidean distance values (**Figure 4.12**), for within strains, between strains and between tissues classes. These results showed that correlation coefficient and Euclidean distance provided, respectively, stronger correlations and shorter distances the closer the biological conditions between samples and, therefore, that the microarray technique was accurate and reliable.

7.1.1.2. Paternal irradiation has negligible effects on the global expression profile of male offspring

Mann-Whitney and Kolmogorov-Smirnov tests did not detect significant differences between Euclidean distances from within and between experimental groups for data from each or all tissues (**Table 4.4**; **Figure 4.7**). This means that the two classes were globally similar. This view was illustrated by an agglomerative hierarchical clustering of all samples using a complete linkage method and Euclidean distance as the metric. This analysis was able to cluster samples according to tissue, almost completely according to strain, but not by experimental group (**Figure 4.13**). After averaging the samples belonging to the same animal group, the resulting 16 expression profiles were re-submitted to hierarchical clustering and all 16 animal groups were clustered according to their respective strain and tissue combination (**Figure 4.14**). These results

support the view that the effects of paternal irradiation on the expression profiles are negligible.

7.1.1.3. Gene expression analysis detected significant differential expression of circadian genes

In order to detect which genes were being affected in the first generation male offspring of directly irradiated male mice, irrespective of strain or tissue effects, gene expression data was $\log_2$ transformed for each sample and submitted to analyses with parametric tests. Student's *t*-test was used to test the null hypothesis that, within the same strain and tissue combination, expression values for a transcript across paternally irradiated and control groups were the same (**Figure 4.15**). Since the *t*-test individually tested hypotheses for multiple, simultaneous comparisons of transcript expression values across the two groups, the FDR was used to adjust the test *P*-values (Benjamini and Hochberg, 1995). At FDR < 0.05, only 4 transcripts were shown to be significantly differentially expressed: two *Ntrk2* and one *Dfna5h* transcripts for liver in the CBA/Ca strain and one unknown transcript of *LOC544903* for kidney in the BALB/c strain. No transcripts were identified as significantly differentially expressed between experimental groups in all strains and tissues studied. This result prompted the hypothesis that the effects of paternal irradiation on the gene expression profiles were being concealed by the combined effect of random factors (stochastic and individual variation) and other fixed factors (tissue and strain). Strain and tissue were shown to contribute the most towards gene expression variation. Thus, tissue, strain and paternal treatment were analysed simultaneously for their individual contribution to gene expression variation using a 3-way ANOVA (**Figure 4.16**). The 3-way ANOVA tested the same hypothesis as the Student's *t*-test and also had to be corrected for multiple

testing. A total of 39 different genes had transcripts passing the FDR < 0.05 threshold and were submitted to GO-analysis. Using the Bioprofiling.de website analysis the genes were grouped into GO-classes in order to inform on their cellular function and facilitate analysis of the biological effects of paternal irradiation. Single GO-class enrichment demonstrated that the majority of the genes in the most enriched classes were either components of the circadian clock or circadian output genes with roles in the regulation or initiation of transcription (**Table 4.7**). The *Arntl* and *Per2* genes had essential roles on the circadian clock (Bunger *et al.*, 2000; Fu *et al.*, 2002) and were amongst the most significantly differentially expressed circadian genes (**Tables 4.8-4.9.4**). Both *Arntl* and *Per2* expression were validated by qPCR (**Table 5.1**).

7.1.1.4. DNA methylation analysis of circadian genes

For both CBA/Ca and BALB/c strains, the patterns of DNA methylation were studied at the *Per2* promoter in the kidney and sperm tissues of the offspring of paternally irradiated mice and in the sperm tissues of directly irradiated male mice using SM-bisulfite sequencing (**Figures 6.1-6.3**). SM-bisulfite sequencing determined that altered DNA methylation patterns are not present in the *Per2* promoter of irradiated mice and their offspring and that the observed *Per2* down-regulation is not caused by *Per2* DNA methylation. A preliminary analysis using conventional bisulfite sequencing allowed DNA methylation patterns to be mapped at promoter or regulatory elements elsewhere in *Arntl*, *Npas2*, *Per1*, *Per2* and *Per3* genes. These analyses took place in kidney and sperm tissues of offspring of paternally irradiated and offspring of control mice and in the sperm tissue of the directly irradiated mice (**Figures 6.4-6.8**). Unfortunately, this analysis was carried only in the BALB/c strain which had actually shown weaker differential expression of circadian genes when compared to the CBA/Ca

strain (**Tables 4.9.1-4.9.4**). Additionally, no changes in DNA methylation at *Arntl*, *Npas2*, *Per1*, *Per2* and *Per3* were detected in any of these analyses. Although not as reliable as SM-bisulfite sequencing, the results with conventional bisulfite sequencing in this study suggest that DNA methylation is not the main repressing vehicle for circadian gene expression.

7.1.1.5. Circadian gene expression at a single time point was accurately analysed

At the time this project was designed it was not expected that any observed differentially expressed gene would have circadian expression or belong to the circadian clock. However, because it was known that the transcriptome is not static and changes throughout the day, all mice were light-entrained (**section 2.6**) with paternally irradiated offspring being culled and tissues extracted at a fixed daytime period (**section 3.2**). This design does not provide information regarding the pattern or direction of gene expression during the daytime period used for culling. However, gene expression levels were accurately compared between the offspring of paternally irradiated and offspring of control mice. The reason this occurred was because there is homogeneity of variance between gene expression profiles of animals belonging to the same animal group (**section 4.1**) meaning their gene expression profile does not seem to vary significantly within the selected daytime period. The animals culled for each animal group are not littermates and were, therefore, culled at very different points of the daytime period.

7.1.2. ESTR instability is a reporter of GI and requires cell proliferation

ESTRs are the most effective system for measuring radiation-induced germline mutation (Dubrova *et al.*, 1998; Dubrova *et al.*, 2000). Mutation at ESTR loci was shown to be independent of meiotic recombination in the germline (Barber *et al.*, 2000) and dependent on the ESTR size (Bois *et al.*, 2001). Mutation at ESTR loci is currently thought to be caused by replication fork stalling, followed by polymerase slippage (Yauk *et al.*, 2002; Barber *et al.*, 2004). In normal cells, replication fork stalling is probably caused by hairpin or quadruplex structures within the structurally homogeneous arrays (Fukuda *et al.*, 2002; Mirkin, 2006 [review]). In irradiated mice and their genomically unstable progeny, additional replication fork stalling could be caused by increased DNA damage formation and recognition, triggering cell cycle arrest in a cell proliferation-dependent event (Yauk *et al.*, 2002; Barber *et al.*, 2004). This hypothesis is supported by the observation that non-proliferating brain tissue only manifests increased ESTR mutations if male and female mice are irradiated *in utero* (Barber *et al.*, 2009). Male BALB/c mice irradiated 1 week prior to culling showed no elevation of ESTR mutations in their sperm tissues (Barber *et al.*, 2009). Conversely, male BALB/c mice irradiated *in utero* (at 12 dpc) or irradiated 8 weeks prior to culling presented elevated frequency of ESTR mutations in their sperm tissues (Barber *et al.*, 2009). Thus, radiation-induced ESTR mutation only takes place if the somatic/germ cells are able to perform cell division and, consequently, undergo cell-cycle arrest due to DNA damage. The offspring of maternally or paternally irradiated mice present increased frequency of ESTR mutation in brain and sperm tissues (Barber *et al.*, 2009). This is because GI is present during gestation, when cell division is active in all tissues.

7.1.3. Evidence against possible mechanisms of TI

Significantly increased frequency of ESTR mutation is observed in the somatic and germline tissues of male and female F_1 offspring of irradiated males (Dubrova *et al.*, 2000; Niwa and Kominami, 2001; Barber *et al.*, 2002; Barber *et al.*, 2006). This transgenerational effect was observed in multiple strains and affected both the irradiated paternal allele and the non-exposed maternal allele suggesting a genome-wide destabilisation occurring after fertilisation, in the zygote or embryo (Dubrova *et al.*, 2000; Niwa and Kominami, 2001; Barber *et al.*, 2002; Barber *et al.*, 2006). Germline ESTR mutation frequency was equally elevated in offspring of F_1 males or F_1 females meaning that GI is equally transmitted by male and female germlines. Although germline ESTR mutation rates were significantly different between strains, they remained similarly elevated between F_1 and F_2 generations within a strain (Barber *et al.*, 2002). This event is partly explained by increased mutational mosaicism in the germline of F_1 mice that causes the F_2 mice to share more than 50% of *de novo* mutations between two or more littermates (Dubrova *et al.*, 2000). Nevertheless, mutations present in only one littermate were more frequent in F_2 mice than in controls suggesting that GI is still present in the F_2 generation (Dubrova *et al.*, 2000). The mutation frequency of the protein-coding *hprt* gene was also elevated 3.7 and 3.3-fold in the spleen tissue of CBA/Ca and BALB/c paternally irradiated offspring, respectively (Barber *et al.*, 2006). A significant increase in SSB (measured by Comet tails) and DSB (measured by γ-H2AX foci) and a surprising absence of oxidised DNA bases (measured by FPG Comet assay) was detected in bone marrow cells of CBA/Ca and BALB/c paternally irradiated F_1 offspring (Barber *et al.*, 2006). Alkaline Comet assays measured DNA repair efficiency in irradiated *ex vivo* bone marrow cells. No difference in DNA repair was detected between paternally irradiated and control mice belonging to the CBA/Ca and

BALB/c strains (Barber *et al.*, 2006). The observed lack of oxidative stress damage in the paternally irradiated offspring suggests that, unlike its parent, the offspring do not present an increased free radical production (Brune *et al.*, 1996; Gregory, 2000 [review]). Cytokine production is possible, but is unlikely to accumulate in sufficient quantities in the very small cytoplasmic component of the spermatozoa and then be transported to the zygote to initiate TI (Rakyan *et al.*, 2001; Dubrova, 2003 [review]).

Exposure to IR during the pre and post-meiotic stages of germline development causes an increase in ESTR mutation rate in the germline of directly exposed males (Barber *et al.*, 2002). During the late post-meiotic stage of germline development, corresponding to condensed spermatids and spermatozoa, transcription is inactive (Leduc *et al.*, 2008). So, if non-coding RNA is responsible for TI then mice formed from irradiated mature sperm should be stable. TI can manifest itself through increased frequency of ESTR mutations in somatic and germline tissues of F_1 offspring conceived from irradiated spermatogonia (Dubrova *et al.*, 2000; Barber *et al.*, 2002; Barber *et al.*, 2006; Barber *et al.*, 2009), round spermatids (Barber *et al.*, 2002) or transcriptionally-inactive and mitotically-unable condensed spermatids and spermatozoa (Niwa and Kominami, 2001; Hatch *et al.*, 2007; Barber *et al.*, 2009). There is also increased p^{un} mutation rates in somatic tissues of F_1 offspring generated from irradiated spermatozoa (Shiraishi *et al.*, 2002). These results exclude the hypothesis that non-coding RNA is the epigenetic mechanism of TI responsible for carrying a GI-signal from the mature sperm into the egg. Nevertheless, these results do not exclude the possibility that non-coding RNA could contribute towards GI, after fertilisation, once transcription is started.

7.1.4. Evidence supporting an epigenetic mechanism of TI

TI cannot be explained by Mendelian segregation of a mutant allele because: (1) almost all offspring (~100%) of irradiated parents manifest genomic instability; (2) unexposed F1 offspring present a similar amount of damage on both ESTR alleles from the irradiated father and unexposed mother; and, (3) transgenerational genomic instability equally manifests in the F1 and F2 offspring of irradiated parents (Barber *et al.*, 2002; Barber *et al.*, 2006). Therefore, the transmission of GI to non-irradiated generations must be dependent on an epigenetic mechanism. Taking into account the literature discussed above, this epigenetic mechanism could be based on either DNA methylation, histone PTM, or both.

7.1.5. Retrotransposable elements could explain most end-points of TI in paternally irradiated offspring

Epigenetic reprogramming in primordial germ cells (PGC) involves genome-wide decreased H3K9me2 and increased H3K27me3 (Seki *et al.*, 2005; Seki *et al.*, 2007) and DNA demethylation (Hajkova *et al.*, 2002). Between 11.5 and 12.5 dpc in PGC development, there is active DNA demethylation at imprinted and non-imprinted loci and IAP and LINE1 retrotransposable elements show incomplete DNA demethylation (Hajkova *et al.*, 2002). Individual IAP and LINE1 elements resist DNA demethylation to different extents (Hajkova *et al.*, 2002) but no transcription of IAP retrotransposons were identified during PGC development (Walsh *et al.*, 1998). Indeed DNA methylation is essential for the silencing of these elements as was demonstrated by the knockout of DNMT3L in mice (Bourc'his and Bestor, 2004). DNMT3L provides essential assistance to DNMT3a/3b during *de novo* methylation and its absence causes IAP and LINE1 retrotransposons to be highly transcribed in spermatogonia and

spermatocytes causing SSB, DSB, meiotic failure, non-homologous chromosomal synapsis and germ cell loss (Bourc'his and Bestor, 2004). Epigenetic inheritance has been show to occur through both the male and female germlines at the *Axin*^{Fu} allele expressing a mutated transcript manifesting a kinky tail phenotype (Rakyan *et al.*, 2003). Expression of aberrant transcripts at this allele is controlled by an IAP element located at an intron of the *Axin*^{Fu} allele and was shown to be inversely correlated with the level of DNA methylation at the IAP 5' LTR (Rakyan *et al.*, 2003). Thus, retrotransposons could potentially be activated by DNA hypomethylation to cause the SSB and DSB observed in the paternally irradiated offspring which could, in turn, lead to chromosomal aberrations and cell cycle arrest causing microsatellite instability (Yauk *et al.*, 2002; Barber *et al.*, 2004). There is support for the transmission of altered DNA methylation patterns to the non-exposed offspring (Barton *et al.*, 2005; Tamminga *et al.*, 2008a; Tamminga *et al.*, 2008b; Filkowski *et al.*, 2010). In fact, there has been a single report of radiation-induced miRNA expression resulting in DNA hypomethylation at LINE1 and SINE B2 elements in the germline of directly irradiated mice and in the thymus of non-exposed offspring (Filkowski *et al.*, 2010). However, there have been no reports of substantially increased retrotransposon expression in the germline of directly irradiated mice and in tissues of paternally irradiated offspring (Prof. Yuri E. Dubrova, *personal communication*, May 29, 2013). The failure to detect expression of LINE1 and SINE B2 hypomethylated elements could suggest a protective role is played by the chromatin in which they occur (Filkowski *et al.*, 2010). More studies should address the direct and transgenerational effects of radiation on retrotransposon DNA methylation, chromatin status and expression.

7.1.6. Genome-wide control of transcription by the circadian clock involves a very dynamic chromatin driven by changes in histone marks

RNA polymerase II (RNApII) was shown to be controlled by the circadian clock since it has a circadian rhythm of C-terminal domain (CTD) hypophosphorylation, when recruited genome-wide to pre-initiation complexes, and a circadian rhythm of CTD phosphorylation, when transcription is initiated (Koike *et al.*, 2012). The transcription coactivators p300 and CBP also showed circadian rhythms of genome-wide occupancy (Koike *et al.*, 2012). Both p300 and CBP are histone acetyltransferases (HATs) (Etchegaray *et al.*, 2003; Curtis *et al.*, 2004). p300 interacts with Clock to facilitate transcription from the *Per1* promoter (Etchegaray *et al.*, 2003). CBP was found to be able to function as a coactivator or a corepressor by binding Per2 (Koike *et al.*, 2012). The H3K4me3, H3K9ac and H3K27ac histone PTMs were mostly located at promoters and, similarly to RNApII, showed a circadian rhythm of occupancy at transcription start sites of expressed genes (Koike *et al.*, 2012). Bmal1, Clock, Npas2, Per1, Per2, Cry1 and Cry2 bind to thousands of binding sites in the genome including intergenic, promoter and intronic regions. Cry1 bound the most, a total of 16,506 sites, but there was major overlap between the activators and repressors for the same binding sites (Koike *et al.*, 2012). All of these core circadian proteins were found to bind at three sites in *Dbp* with circadian rhythmicity, one promoter and two intronic regions (Koike *et al.*, 2012). *Dbp* has H3K4me and H3K9ac marks, formed in phase with transcriptional activity and associated with RNApII, that are reversed as the RNA is elongated (Eissenberg and Shilatifard, 2006). Thus, RNApII has circadian recruitment, probably by circadian transcription factors, bringing circadian histone modification and transcription initiation to clock-controlled and clock core genes on a genome-wide scale

(Koike *et al.*, 2012). Only 22% of cycling mRNA is synthesised *de novo* so, the circadian clock modulates post-transcriptional processes like RNA splicing, polyadenylation or mRNA stability to provide circadian patterns to mRNA levels (Koike *et al.*, 2012). There is also marked control at the post-translational level, involving protein phosphorylation, meaning that the circadian cycles of mRNA and protein levels are often different (Reddy *et al.*, 2006). These results suggest a genome-wide control of transcription by the circadian clock.

7.1.7. Changes in histone marks and DNA methylation are the most likely candidates for the epigenetic mechanism of TI

The circadian genes that were shown to be differentially expressed in the paternally irradiated offspring from the CBA/Ca and BALB/c strain, varied in number but showed the same direction of expression across the tissues studied (e.g. *Dbp* was always down-regulated while *Npas2* was always up-regulated across all four tissues studied). Only *Dbp* was significantly differentially expressed in all tissues while *Per2* and *Arntl* were significant in brain, kidney and liver and *Npas2* was significant in kidney, liver and spleen of the paternally irradiated CBA/Ca offspring. The liver of paternally irradiated CBA/Ca offspring showed the highest number of significantly differentially expressed circadian genes while the kidney of BALB/c offspring was the only tissue showing differential expression of circadian genes (i.e. *Arntl*, *Mtf1*, *Npas2*, *Nr1d2* and *Per2*). These results suggest that not only different strains but also different tissues have distinct resistance to the effects of paternal irradiation on circadian gene expression.

Since IR could cause transgenerational effects on patterns of chromatin (Baulch *et al.*, 2007) and DNA methylation (Tamminga *et al.*, 2008a; Tamminga *et al.*, 2008b;

Filkowski *et al.*, 2010), the genome-wide initiation of transcription by the circadian clock (Koike *et al.*, 2012) could potentially be affected. This could then result in sub-optimal function of tissues, such as the kidney and liver, relying on the predictive function of the circadian clock on transcription to play essential physiological roles in a timely manner (Lamia *et al.*, 2008; Zuber *et al.*, 2009; Eckel-Mahan *et al.*, 2012). The circadian control over cell cycle progression (Canaple *et al.*, 2003 [review]) could play a role in the accumulation of mutations at ESTR loci (Barber *et al.*, 2004) or other DNA damage resultant from TI (Barber *et al.*, 2006) due to disruption caused by paternal irradiation.

Single molecule bisulfite sequencing of the *Per2* promoter and additional promoter elements, introns and exons did not show changes in the pattern of methylation able to explain this effect of paternal irradiation. Equally, no differences in the DNA methylation pattern in sperm were found between controls, the directly irradiated fathers and their non-exposed offspring. However, because most clock components regulate most of the other clock components (Baggs *et al.*, 2009) the possibility that DNA methylation occurs at other circadian genes must be considered and tested. Indeed, circadian gene silencing by DNA hypermethylation in hematologic malignancies has been reported (Taniguchi *et al.*, 2009). Importantly, histone PTMs are reported to control circadian gene activity particularly at the *Dbp* gene (Eissenberg and Shilatifard, 2006) which was the gene most significantly affected by paternal irradiation. So, the role of histone PTMs in the observed effect of paternal irradiation on circadian gene expression must be considered in subsequent experiments.

Appendix I

Table S1: Primers for qPCR.

Gene	Type of primer	Sequence (5' to 3')	Melting temperature (T_m)	cDNA amplicon (bp)	gDNA amplicon (bp)
Aprt	Forward	GTCATTGTGGATGACCTCC	60.4	213	341
	Reverse	CCACCAAGCAGTTCCTG	60.5		
ActB	Forward	GCTGTATTCCCCTCCATCGTG	67.6	265	264 (ch.X); 265 (ch.17); 352 (ch.5)
	Reverse	CACGGTTGGCCTTAGGGTTCAG	70.5		
Per2	Forward	CGAGCGGCTGCAGTAGTG	66.1	224	N/A
	Reverse	CACACTGCAGGGCTGGC	67.3		
Arntl	Forward	GATCAAGAATGCAAGGGAGG	63.5	184	3580
	Reverse	GTTGGTGGCACCTCTCAAAG	64.7		

Guidelines for primer design were described on Methods (**section 3.10.1**). Primers were designed to easily distinguish amplifications between mRNA and gDNA templates. *Aprt* contained a single splicing variant but *ActB* contained 8 splicing variants of its transcript as shown by the Ensembl database (http://www.ensembl.org/index.html). The *ActB* cDNA amplified an exon 4-intron-exon 5 region, overlapping three *ActB* protein coding splicing variants plus the longest splicing variant with the consensus coding DNA sequence. As mentioned in **section 5.2**, *ActB* showed constant expression in all tissues despite having multiple splicing variants.

Table S2: Primers for PCR and sequencing of gDNA.

Gene	Type of primer	Sequence (5' to 3')	Melting temperature (T_m)	gDNA amplicon (bp)
Per2	Forward	GTTTTGTGCTTTGGCCCATCT	67.1	1135
	Reverse	ACCCTCTGGACCCATCCTAG	64.2	
Arntl	Forward	CAGGGGAAGTGGGAAGTTAC	62.3	2863
	Reverse	TCAAGGGGTCAGAGGGTCAA	67.0	
	Forward	AGGACAAGCCAGGGGTTTCA	68.0	372
	Reverse	CAGGGCTGGTTTACCTGCTA	64.0	
	Forward	GGGAAGGCAGAAAGTAGGTC	61.6	628
	Reverse	GGAGAAGTGTAGAAGGGTTGA	59.8	
Npas2	Forward	TGCTGGAATGTGAGATGTAGAA	62.2	3163
	Reverse	AAGAGTAGACTGCTGGGAAAC	59.0	
	Forward	CTGGGGTTACAAGGGGACAT	64.8	2249
	Reverse	TGAAGAGAGAGACAGGAAGGT	59.6	
	Forward	AGGGGGACAAGACAATCACAT	64.5	1408
	Reverse	TACTTGGGGAGGGGGCACA	69.9	

These primers were used to amplify the CpG island gDNA sequence of *Per2, Arntl* or *Npas2* genes and for sequencing. Although amplicon sizes are given for primer pairs, different combinations of forward and reverse primers may have been used for each gene. Sequencing gDNA allowed the mapping of the CpG Island and detection of single nucleotide polymorphisms (SNPs) that could have been present at the priming sites of bisulfite treated sequences. This exercise was performed since SNPs are still being abundantly discovered for fully sequenced genomes and no genome browser, at the moment, contains a complete collection of all SNPs for the mice strains used in this study. It was concluded less wasteful to abandon this approach and return to it if no product was generated by bisulfite SM-PCR primers.

Table S3: Primers for 1° PCR in nested bisulfite conventional-PCR or SM-PCR.

Gene	Primer name	Type of primer	Sequence (5' to 3')	Melting temperature (T_m)	cDNA amplicon (bp)
Per1	per1_SM 1_F	Forward	TTTTAGGTTTAATGTT TTTATAGG	54.5	1081
	per1_SM 1_R	Reverse	ATAATCTCTTTCTCCA CAACC	56.3	
	per1_SM 9_F	Forward	GTTTTAGGAGTGAAG AAATTTG	56.2	1013
	per1_SM 9_R	Reverse	ATCTACTTCATTCTCA ACTTCC	55.6	
	per1_SM 3_F	Forward	AATTGTTTTTTTAGGA ATTTTTGG	60.1	869
	per1_SM 4a_R	Reverse	AAAACAACATACCAA TCTAAATC	55.5	
	per1_SM 5a_F	Forward	TTTTGGAGATTTTTTT TTTGATTG	61.8	969
	per1_SM 5_R	Reverse	ATCTCTTAAATTCAAA ATAACTAC	51.5	
	per1_SM 7_F	Forward	GTTAGTATGAGGGTGT GTTG	54.4	986
	per1_SM 7_R	Reverse	CTTACCATTACCAAAA AAAAAAC	57.4	
	per1_SM 11_F	Forward	TTGGAAAGGGTTTAGT ATTATG	56.4	790
	per1_SM 11_R	Reverse	TACAAATAAACTAAT CTCTCTAC	48.4	
Per2	per2_SM 10_F	Forward	TTATTGGTGGAAATAG AAAAGG	58.6	488
	per2_SM 10_R	Reverse	ACTAAACTACTAATAT ACACAAC	45.9	
	per2_SM 14_F	Forward	TTAGGGTTTTTTTGGA AAGGG	62.7	1043
	per2_SM 14_R	Reverse	ACCTACAACAACATA ACCAAC	55.5	
	per2_S M6_F	**Forward**	**GTTTGTTTTAGGGTT ATTTAGTT**	**54.2**	**1105**
	per2_S M2_R	**Reverse**	**AACTCCTTAAAAAAC CCTCTAA**	**56.8**	
	per2_SM 22_F	Forward	TAGTTGGAGTTGGGGT TTTG	61.4	827
	per2_SM 22_R	Reverse	AAACTACCAAACCTA ATTTTTAC	54.2	
	per2_SM 16_F	Forward	TTAGGTATTTTAGGGG TTTTTG	57.7	959
	per2_SM 16_R	Reverse	ATACTACCATCCTACA AATCC	53.7	

	per2_SM18_F	Forward	GGAGTTTGTGTTTGTTTTTGG	60.9	988
	per2_SM24_R	Reverse	TTACCCATCCATATAACCAAC	57.7	
	per2_SM20_F	Forward	GGGTATAAGTTAGGTTAATAAG	50.2	673
	per2_SM20_R	Reverse	CTCCATTACCAAATCTAACC	55.1	
Per3	per3_SM1_F	Forward	GGTTGGATTTTATTTGGGTTTA	60.3	1086
	per3_SM1_R	Reverse	TCACCTATAACTACTATCCAC	49.9	
	per3_SM3_F	Forward	GTAGAGTTTGTGGAAATGTTG	56.2	985
	per3_SM3_R	Reverse	ACTAACTTTTCAAACTTAAAAAAC	53.8	
Arntl	Arntl_SM4_F	Forward	TTTGTTTAGGTTTTATTTTGTGG	58.6	930
	Arntl_SM4_R	Reverse	TCCCTAACCTACTTTCTACC	54.0	
	Arntl_SM11_F	Forward	GTTTTGTTATTGGTTAGAGGTT	55.5	1112
	Arntl_SM11_R	Reverse	ATACCTACTCCCAAACCCAAC	60.5	
	Arntl_SM7_F	Forward	GTTTGTTTGGTTAATTTTTTTATAT	55.5	649
	Arntl_SM7_R	Reverse	AAAAAAACCAAATCTCAAAAAATC	57.9	
	Arntl_SM9_F	Forward	TTATTATAGAAGAGTAGATTGAG	48.5	920
	Arntl_SM9_R	Reverse	TTCTACTCCTTCCTTAAAAAC	55.1	
Npas2	Npas2_SM4_F	Forward	GTTAGTAGAGGGTTGTTGTG	53.6	1129
	Npas2_SM4a_R	Reverse	CTTCCTCTTAAACCTCCAAAC	58.4	
	Npas2_SM7_F	Forward	GTGATAATTTTAGTAGGATTTTG	53.2	1137
	Npas2_SM7a_R	Reverse	CCAAACAACCACACTCTCC	60.6	
	Npas2_SM9_F	Forward	GTTAAGAGAGAGGATTAGGTA	51.0	882
	Npas2_SM9_R	Reverse	AACTACTTCAAAACTAACTACC	50.5	
	Npas2_SM11_F	Forward	AAATTTGTAAGAAAGATTTAGAAG	53.8	927
	Npas2_SM11_R	Reverse	ACTAACCTTTCATATTTCTCTC	52.6	

Primers were designed (as described in **section 3.10.2**) to amplify regions of circadian genes in a 1°PCR of nested conventional PCR and SM-PCR. The primer pair used in the amplification of the *Per2* promoter is in bold and was the only primer pair used in SM-PCR.

Table S4: Primers for 2° PCR in nested bisulfite conventional PCR or SM-PCR.

Gene	Primer name	Type of primer	Sequence (5' to 3')	Melting temperature (T_m)	cDNA amplicon (bp)
Per1	per1_SM2_F	Forward	ATTAAGTTATTTTTATGGTGTTAG	52.3	1024
	per1_SM2_R	Reverse	AAACATTTATTCTACAACTAATTC	52.2	
	per1_SM10_F	Forward	TAGGAGTGAAGAAATTTGGAG	57.5	998
	per1_SM10_R	Reverse	TCTCAACTTCCTAATAATAAAAC	52.9	
	per1_SM4a_F	Forward	TTTTTTAGGAATTTTTGGTTTTTG	60.9	882
	per1_SM3_R	Reverse	AAAATCCCCTTCCCTAAACC	61.8	
	per1_SM5_F	Forward	TGATTGGTTAGTGAAGTATTTG	55.9	948
	per1_SM6a_R	Reverse	TCTTAAATTCAAAATAACTACTCT	52.1	
	per1_SM8_F	Forward	GTTAGTATAGGATTGGGTTTG	54.4	901
	per1_SM8_R	Reverse	TAAACTCATTACCACTTAAACC	54.2	
	per1_SM12_F	Forward	TTTTAGTTTGGTTTTGTGTTATG	57.4	708
	per1_SM12_R	Reverse	AAATAAAAACTCTCCAACTCAC	55.7	
Per2	per2_SM11_F	Forward	TTAAATTTAGGGAGGAATTGTG	58.6	395
	per2_SM11_R	Reverse	CTACCTAACAAAACAAACATAC	52.4	
	per2_SM15_F	Forward	AGGGTTGTTGAAATGAATTTAG	58.6	996
	per2_SM15_R	Reverse	CCAAAAAACCAAAAATACTCTC	58.2	
	per2_SM3_F	**Forward**	**GGGAAGGTGTTTAGTAGTAG**	**52.0**	**1018**
	per2_S	**Reverse**	**CCTTAAACAACACCA**	**58.1**	

	M1_R		**TAAAAATC**		
	per2_SM23_F	Forward	TTGGGGTTTTGAGTTAGGAG	60.5	991
	per2_SM23_R	Reverse	ATACAACTTACCAACCTATAAC	50.9	
	per2_SM17_F	Forward	GATTTAGGGTAAGTTTGAGTTT	54.8	926
	per2_SM17_R	Reverse	TCCTACAAATCCTAATACTTAC	51.2	
	per2_SM19_F	Forward	TGTGTTTGTTTTTGGATGAGG	62.3	905
	per2_SM18_R	Reverse	CAAAACCCCCTAAAACCAAATA	62.4	
	per2_SM21_F	Forward	TTTTTAAAGGTGGATGGATTTG	61.4	606
	per2_SM21_R	Reverse	CCAATATTCCAAAATATCCTAC	55.8	
Per3	per3_SM2_F	Forward	GGGAGAAAGTAGTAAGTTTTG	53.9	1011
	per3_SM2_R	Reverse	AAAAAACCCTCCTAACCAAAC	59.8	
	per3_SM4_F	Forward	TGAGGAGAAAAGTGTTTTTTAG	56.8	927
	per3_SM4_R	Reverse	CATCAACTTTAAACTTACTTATC	51.1	
Arntl	Arntl_SM5_F	Forward	ATGTTTTTGTTTTAGTTTTTTGAG	57.2	876
	Arntl_SM5a_R	Reverse	AACCTACTTTCTACCTTCCC	55.4	
	Arntl_SM6_F	Forward	GGTTAATGGGAAGAGGTAGG	58.9	972
	Arntl_SM6a_R	Reverse	AAAAAAACACATCCTCCAAAAC	60.7	
	Arntl_SM7a_F	Forward	GTTTGGTTAATTTTTTTATATTTTTT	56.1	595
	Arntl_SM8_R	Reverse	AAAAAAAAACTAAAAAAAATACACC	56.9	
	Arntl_SM10_F	Forward	AAAGTATTTAAAAAGAAGTAAGTG	51.9	520
	Arntl_SM10_R	Reverse	AAAAAAAAACTAAATAAAAACTCTC	54.5	
Npas2	Npas2_SM5_F	Forward	GGGGTTAGTAATTTTTTTATTATT	54.7	1068
	Npas2_SM4_R	Reverse	CTCTTAAACCTCCAAACTCC	56.9	
	Npas2_SM8_F	Forward	GTATTAAGTATAGTTTTTTTTTTTA	51.1	1063
	Npas2_SM7_R	Reverse	AACAACCACACTCTCCAAAC	59.1	

	Npas2_SM10_F	Forward	GAGAGGATTAGGTATTTTAAATA	52.1	847
	Npas2_SM10_R	Reverse	ATCCAAATAAAAAATCCCTCCT	60.4	
	Npas2_SM12_F	Forward	TTTAGAAGTAGATTATTATTATTAG	47.4	877
	Npas2_SM12_R	Reverse	ATCCCTTATACCCCAATAATC	57.2	

Primers were designed (as described in **section 3.10.2**) to amplify regions of circadian genes in a 2°PCR of nested conventional PCR and SM-PCR. The primer pair used in the amplification of the *Per2* promoter is in bold and was the only primer pair used in SM-PCR.

Table S5: Primers used for sequencing amplicons from conventional PCR and SM-PCR.

Gene	Primer name	Type of primer	Sequence (5' to 3')	Melting temperature (T_m)
Per1	per1_SM2a_F	Forward	AAGTTATTTTTATGGTGTTAGAG	52.9
	per1_SM10a_R	Reverse	CCTAATAATAAAACCCAAAATTC	56.5
	per1_SM4_F	Forward	GAATTTTTGGTTTTTGTATAGGA	57.7
	per1_SM4_R	Reverse	TTCTCAAACTTCCCTACAAAC	57.4
	per1_SM6_F	Forward	GTTTTTTTTTATAGTAGTTATTGG	52.1
	per1_SM6_R	Reverse	TAAACACCTAAAACCCAAAAAC	58.5
	per1_SM8a_R	Reverse	ACAACATCTACTAACTACACC	49.7
	per1_SM12a_F	Forward	GTTTGGTTTTGTGTTATGGTG	59.3
Per2	per2_SM11a_R	Reverse	CTAACAAAACAAACATACAATTC	54.7
	per2_SM15_R	Reverse	CCAAAAAACCAAAAATACTCTC	58.2
	per2_SM1_R	**Reverse**	**CCTTAAACAACACCATAAAAATC**	**58.1**
	per2_SM23a_R	Reverse	ACTTACCAACCTATAACTATAC	48.0
	per2_SM	Reverse	TCCTACAAATCCTAATACTTAC	51.2

	17_R			
	per2_SM19_R	Reverse	AAACCCTAAAAACACTCACAC	57.2
	per2_SM21a_F	Forward	TGGATGGATTTGTTTTTAGAGT	59.0
Per3	per3_SM2a_F	Forward	AGTAGTAAGTTTTGGAAAGATG	53.5
	per3_SM4a_R	Reverse	TTCACTAAATAAAAAAAAAACAAAC	56.5
Arntl	Arntl_SM5_R	Reverse	CTTTCTACCTTCCCTAAATATC	54.5
	Arntl_SM2_F	Forward	GAAGGGGGGTTGGGTATAG	61.9
	Arntl_SM8_F	Forward	TTTGTATTGGTATGTTTTTATGGT	57.9
	Arntl_SM10a_R	Reverse	AAAAACTAAATAAAAACTCTCCC	55.6
Npas2	Npas2_SM6a_F	Forward	TAGGATTTTTTTTTTTATTTGGTG	59.0
	Npas2_SM6_F	Forward	TTTTTTATAGGAAAAAAAAAAAAGG	58.8
	Npas2_SM6_R	Reverse	AACACCAAATAAAAAAAAAAATCC	59.9
	Npas2_SM8_R	Reverse	AACAATCTTAAAAACTTCTATCC	54.2
	Npas2_SM10a_R	Reverse	TCCTCTACTTAAAAAAAAAACAC	55.2
	Npas2_SM12a_R	Reverse	CCCTTATACCCCAATAATCTA	56.2

Sequencing of some PCR amplicons produced high background noise and/or truncated sequence read with one or both of the PCR primer pairs. Homopolymeric sequences of ≥10 nt were observed to cause the DNA polymerase to slip, producing a truncated sequence. In these situations, and if the opposing primer could not be used, new primers were designed specifically for sequencing. The new primers were positioned 2-3 nt away from the template extremity or designed to exclude the homopolymeric region from sequencing.

Table S6: Groupings of primer pairs involved in 1° PCR of multiple regions.

Genes	Primer name	Group
Per3	per3_SM1_F	1
	per3_SM1_R	
Npas2	Npas2_SM4_F	
	Npas2_SM4_R	
Per1	per1_SM3_F	2
	per1_SM4a_R	
Arntl	Arntl_SM4_F	
	Arntl_SM4_R	
Npas2	Npas2_SM9_F	
	Npas2_SM9_R	
Per1	per1_SM7_F	3
	per1_SM7_R	
Per2	per2_SM6_F	
	per2_SM2_R	
Npas2	Npas2_SM7_F	
	Npas2_SM7a_R	
Per2	per2_SM16_F	4
	per2_SM16_R	
Per2	per2_SM20_F	
	per2_SM20_R	
Per2	per2_SM14_F	5
	per2_SM14_R	
Per2	per2_SM18_F	
	per2_SM24_R	

List of primer pairs able to assist amplification of a target region while in the presence of other distant amplification reactions. Primer details are available on Table S3 of this Appendix. Other primer pairs, listed on Table S3, were either unable to amplify in a multiple target PCR or were excluded from testing.

References

Antonov, A.V., Schmidt, T., Wang, Y. and Mewes, H.W. (2008). ProfCom: a web tool for profiling the complex functionality of gene groups identified from high-throughput data. *Nucleic Acids Res*, **36**(Web Server issue), W347-W351. (http://dx.doi.org/10.1093/nar/gkn239)

Antonov, A.V., Schmidt, E.E., Dietmann, S., Krestyaninova, M. and Hermjakob, H. (2010). R spider: a network-based analysis of gene lists by combining signaling and metabolic pathways from Reactome and KEGG databases. *Nucleic Acids Res*, **38**(Web Server issue), W78-W83. (http://dx.doi.org/10.1093/nar/gkq482)

Antonov, A.V. (2011). BioProfiling.de: analytical web portal for high-throughput cell biology. *Nucleic Acids Res*, **39**(Web Server issue), W323–W327. (http://dx.doi.org/10.1093/nar/gkr372)

Aravin, A., and Bourc'his, D. (2008). Small RNA guides for de novo DNA methylation in mammalian germ cells. *Genes & Development*, **22**(8), 970-975.

Archer, S.N., Carpen, J.D., Gibson, M., Lim, G.H., Johnston, J.D., Skene, D.J. and von Schantz, M. (2010). Polymorphism in the PER3 promoter associates with diurnal preference and delayed sleep phase disorder. *Sleep*, **33**(5), 695-701.

Atwood, A., DeConde, R., Wang, S.S., Mockler, T.C., Sabir, J.S., Ideker, T., and Kay, S.A. (2011). Cell-autonomous circadian clock of hepatocytes drives rhythms in transcription and polyamine synthesis. *Proc Natl Acad Sci U S A*, **108**(45), 18560-18565.

Azzam, E.I., de Toledo, S.M. and Little, J.B. (2001). Direct evidence for the participation of gap junction-mediated intercellular communication in the transmission of damage signals from alpha -particle irradiated to nonirradiated cells. *Proc Natl Acad Sci U S A*, **98**(2), 473-478.

Azzam, E.I., De Toledo, S.M., Spitz, D.R. and Little, J.B. (2002). Oxidative metabolism modulates signal transduction and micronucleus formation in bystander cells from alpha-particle-irradiated normal human fibroblast cultures. *Cancer Res*, **62**(19), 5436-5442.

Azzam, E.I., de Toledo, S.M., and Little, J.B. (2003). Oxidative metabolism, gap junctions and the ionizing radiation-induced bystander effect. *Oncogene*, **22**(45), 7050-7057.

Bae, K., Jin, X., Maywood, E.S., Hastings, M.H., Reppert, S.M. and Weaver, D.R. (2001). Differential functions of mPer1, mPer2, and mPer3 in the SCN circadian clock. *Neuron*, **30**(2), 525-536.

Baggs, J.E., Price, T.S., DiTacchio, L., Panda, S., Fitzgerald, G.A. and Hogenesch, J.B. (2009). Network features of the mammalian circadian clock. *PLoS Biol*, **7**(3). (http://dx.doi.org/10.1371/journal.pbio.1000052)

Barbas, C.F. 3rd, Burton, D.R., Scott, J.K., and Silverman, G.J. (2007). Quantitation of DNA and RNA. *CSH Protoc*. (http://dx.doi.org/10.1101/pdb.ip47)

Barber, R., Plumb, M., Smith, A.G., Cesar, C.E., Boulton, E., Jeffreys, A.J. and Dubrova, Y.E. (2000). No correlation between germline mutation at repeat DNA and meiotic crossover in male mice exposed to X-rays or cisplatin. *Mutat Res*, **457**(1-2), 79-91.

Barber, R.C., Plumb, M.A., Boulton, E., Roux, I. and Dubrova, Y.E. (2002). Elevated mutation rates in the germ line of first- and second-generation offspring of irradiated male mice. *Proc Natl Acad Sci U S A*, **99**(10), 6877-6882.

Barber, R.C., Miccoli, L., van Buul, P.P., Burr, K.L., van Duyn-Goedhart, A., Angulo, J.F. and Dubrova, Y.E. (2004). Germline mutation rates at tandem repeat loci in DNA-repair deficient mice. *Mutat Res*, **554**(1-2), 287-295.

Barber, R.C., and Dubrova, Y.E. (2006). The offspring of irradiated parents, are they stable? *Mutat. Res*, **598**(1-2), 50-60.

Barber, R.C., Hickenbotham, P., Hatch, T., Kelly, D., Topchiy, N., Almeida, G.M., Jones, G.D.D., Johnson, G.E., Parry, J.M., Rothkamm, K. and Dubrova, Y.E. (2006).

Radiation-induced transgenerational alterations in genome stability and DNA damage. *Oncogene*, **25**(56), 7336-7342.

Barber, R.C., Hardwick, R.J., Shanks, M.E., Glen, C.D., Mughal, S.K., Voutounou, M. and Dubrova, Y.E. (2009). The effects of in utero irradiation on mutation induction and transgenerational instability in mice. *Mutat. Res*, **664**(1-2), 6-12.

Barnard, A.R. and Nolan, P.M. (2008). When Clocks Go Bad: Neurobehavioural Consequences of Disrupted Circadian Timing. *PLoS Genet*, **4**(5). (http://dx.doi.org/10.1371/journal.pgen.1000040)

Barton, T.S., Robaire, B. and Hales, B.F. (2005). Epigenetic programming in the preimplantation rat embryo is disrupted by chronic paternal cyclophosphamide exposure. *Proc Natl Acad Sci U S A*, **102**(22), 7865-7870.

Baulch, J.E., Li, M. and Raabe, O.G. (2007). Effect of ATM heterozygosity on heritable DNA damage in mice following paternal F0 germline irradiation. *Mutat. Res*, **616**(1-2), 34-45.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B.*, **57**(1), 289-300.

Berger, S.L., Kouzarides, T., Shiekhattar, R. and Shilatifard, A. (2009). An operational definition of epigenetics. *Genes Dev*, **23**(7), 781-783.

Bois, P.R., Southgate, L. and Jeffreys, A.J. (2001). Length of uninterrupted repeats determines instability at the unstable mouse expanded simple tandem repeat family MMS10 derived from independent SINE B1 elements. *Mamm Genome*, **12**(2), 104-111.

Bolstad, B.M., Irizarry, R.A., Astrand, M. and Speed, T.P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, **19**(2), 185-193.

Bourc'his, D. and Bestor, T.H. (2004). Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature*, **431**(7004), 96-99.

Brandriff, B. and Pedersen, R.A. (1981). Repair of the ultraviolet-irradiated male genome in fertilized mouse eggs. *Science*, **211**(4489), 1431-1433.

Brune, B., Golkel, C. and von Knethen, A. (1996). Cytokine and low-level nitric oxide prestimulation block p53 accumulation and apoptosis of RAW 264.7 macrophages. *Biochem Biophys Res Commun*, **229**(2), 396–401.

Buijs, R.M., Hermes, M.H. and Kalsbeek, A. (1998). The suprachiasmatic nucleus-paraventricular nucleus interactions: a bridge to the neuroendocrine and autonomic nervous system. *Prog Brain Res*, **119**, 365-382.

Bunger, M.K., Wilsbacher, L.D., Moran, S.M., Clendenin, C., Radcliffe, L.A., Hogenesch, J.B., Simon, M.C., Takahashi, J.S. and Bradfield, C.A. (2000). Mop3 is an

essential component of the master circadian pacemaker in mammals. *Cell*, **103**(7), 1009–1017.

Burns, T.F. and El-Deiry, W.S. (2003). Microarray analysis of p53 target gene expression patterns in the spleen and thymus in response to ionizing radiation. *Cancer Biol Ther*, **2**(4), 431-443.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J. and Wittwer, C.T. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem*, **55**(4), 611-622.

Canaple, L., Kakizawa, T. and Laudet, V. (2003). The days and nights of cancer cells. *Cancer Res*, **63**(22), 7545-7552.

Cao, Q., Gery, S., Dashti, A., Yin, D., Zhou, Y., Gu, J. and Koeffler, H.P. (2009). A role for the clock gene Per1 in prostate cancer. *Cancer Res*, **69**(19), 7619–7625.

Chhibber, A. and Schroeder, B.G. (2008). Single-molecule polymerase chain reaction reduces bias: application to DNA methylation analysis by bisulfite sequencing. *Anal Biochem*, **377**(1), 46-54.

Chial, H. (2008). Somatic mosaicism and chromosomal disorders. *Nature Education*, **1**(1). (http://www.nature.com/scitable/topicpage/somatic-mosaicism-and-chromosomal-disorders-867)

Clark, S.J. (2007). Action at a distance: epigenetic silencing of large chromosomal regions in carcinogenesis. *Hum Mol Genet*, **16**(1), R88-R95. (http://dx.doi.org/10.1093/hmg/ddm051)

Cropley, J.E., Suter, C.M., Beckman, K.B. and Martin, D.I. (2010). CpG methylation of a silent controlling element in the murine Avy allele is incomplete and unresponsive to methyl donor supplementation. *PLoS One*, **5**(2), e9055. (http://dx.doi.org/10.1371/journal.pone.0009055)

Curtis, A.M., Seo, S.B., Westgate, E.J., Rudic, R.D., Smyth, E.M., Chakravarti, D., FitzGerald, G.A. and McNamara, P. (2004). Histone acetyltransferase-dependent chromatin remodeling and the vascular clock. *J Biol Chem*, **279**(8), 7091-7097.

Dai, J.M., Sun, D.C., Lin, R.X., Yang, J., Lou, S. and Wang, S.Q. (2006). Microarray analysis of differentially expressed genes in mouse bone marrow tissues after ionizing radiation. *Int J Radiat Biol*, **82**(7), 511-521.

Dai, H., Zhang, L., Cao, M., Song, F., Zheng, H., Zhu, X., Wei, Q., Zhang, W. and Chen, K. (2011). The role of polymorphisms in circadian pathway genes in breast tumorigenesis. *Breast Cancer Res Treatment*, **127**(2), 531–540.

Darst, R.P., Pardo, C.E., Ai, L., Brown, K.D. and Kladde, M.P. (2010). Bisulfite sequencing of DNA. *Curr Protoc Mol Biol*, **7**(9), 1-17.

Das, P.M. and Singal, R. (2004). DNA methylation and cancer. *J Clin Oncol*, **22**(22), 4632–4642.

de Lange, T. (2002). Protection of mammalian telomeres. *Oncogene*, **21**(4), 532-540.

de Ménorval, M.A., Mir, L.M., Fernández, M.L. and Reigada, R. (2012). Effects of dimethyl sulfoxide in cholesterol-containing lipid membranes: a comparative study of experiments in silico and with cells. *PLoS One*, **7**(7). (http://dx.doi.org/10.1371/journal.pone.0041733)

DeBruyne, J.P., Weaver, D.R. and Reppert, S.M. (2007). CLOCK and NPAS2 have overlapping roles in the suprachiasmatic circadian clock. *Nat Neurosci*, **10**(5), 543-545.

DeForge, L.E., Preston, A.M., Takeuchi, E., Kenney, J., Boxer, L.A. and Remick, D.G. (1993). Regulation of interleukin 8 gene expression by oxidant stress. *J Biol Chem*, **268**(34), 25568-25576.

Doi, M., Hirayama, J. and Sassone-Corsi, P. (2006). Circadian regulator CLOCK is a histone acetyltransferase. *Cell*, **125**(3), 497-508.

Dubrova, Y.E., Plumb, M., Brown, J., Fennelly, J., Bois, P., Goodhead, D. and Jeffreys, A.J. (1998). Stage specificity, dose response, and doubling dose for mouse minisatellite germ-line mutation induced by acute radiation. *Proc Natl Acad Sci U S A*, **95**(11), 6251-6255.

Dubrova, Y.E., Plumb, M., Gutierrez, B., Boulton, E. and Jeffreys, A.J. (2000). Genome stability: Transgenerational mutation by radiation. *Nature*, **405**(6782), 37-37.

Dubrova, Y.E. (2003). Radiation-induced transgenerational instability. *Oncogene*, **22**(45), 7087-7093.

Dudley, C.A., Erbel-Sieler, C., Estill, S.J., Reick, M., Franken, P., Pitts, S. and McKnight, S.L. (2003). Altered patterns of sleep and behavioral adaptability in NPAS2-deficient mice. *Science*, **301**(5631), 379-383.

Eckel-Mahan, K.L., Patel, V.R., Mohney, R.P., Vignola, K.S., Baldi, P. and Sassone-Corsi, P. (2012). Coordination of the transcriptome and metabolome by the circadian clock. *Proc Natl Acad Sci*, **109**(14), 5541-5546.

Ehrich, M., Zoll, S., Sur, S. and van den Boom, D. (2007). A new method for accurate assessment of DNA quality after bisulfite treatment. *Nucleic Acids Res*, **35**(5), e29. (http://dx.doi.org/10.1093/nar/gkl1134)

Ehrlich, M. (2002). DNA methylation in cancer: too much, but also too little. *Oncogene*, **21**(35), 5400-5413.

Eissenberg, J.C. and Shilatifard, A. (2006). Leaving a mark: the many footprints of the elongating RNA polymerase II. *Curr Opin Genet Dev*, **16**(2), 184-190.

Etchegaray, J.P., Lee, C., Wade, P.A. and Reppert, S.M. (2003). Rhythmic histone acetylation underlies transcription in the mammalian circadian clock. *Nature*, **421**(6919), 177-182.

Federico, A., Morgillo, F., Tuccillo, C., Ciardiello, F. and Loguercio, C. (2007). Chronic inflammation and oxidative stress in human carcinogenesis. *Int J Cancer*, **121**(11), 2381–2386.

Filkowski, J.N., Ilnytskyy, Y., Tamminga, J., Koturbash, I., Golubov, A., Bagnyukova, T., Pogribny, I.P. and Kovalchuk, O. (2010). Hypomethylation and genome instability in the germline of exposed parents and their progeny is associated with altered miRNA expression. *Carcinogenesis*, **31**(6), 1110-1115.

Fleisher, A.S., Esteller, M., Harpaz, N., Leytin, A., Rashid, A., Xu, Y., Liang, J., Stine, O.C., Yin, J., Zou, T.T., Abraham, J.M., Kong, D., Wilson, K.T., James, S.P., Herman, J.G. and Meltzer, S.J. (2000). Microsatellite instability in inflammatory bowel disease-associated neoplastic lesions is associated with hypermethylation and diminished expression of the DNA mismatch repair gene, hMLH1. *Cancer Res*, **60**(17), 4864–4868.

Fraga, M.F. and Esteller, M. (2002). DNA methylation: a profile of methods and applications. *Biotechniques*, **33**(3), 632, 634, 636-49.

Franken, P., Lopez-Molina, L., Marcacci, L., Schibler, U. and Tafti, M. (2000). The transcription factor DBP affects circadian sleep consolidation and rhythmic EEG activity. *J Neurosci*, **20**(2), 617-625.

Froy, O. (2011). The circadian clock and metabolism. *Clinical Science,* **120**(2), 65–72.

Fu, L., Pelicano, H., Liu, J., Huang, P. and Lee, C. (2002). The circadian gene Period2 plays an important role in tumor suppression and DNA damage response in vivo. *Cell,* **111**(1), 41-50.

Fukuda, H., Katahira, M., Tsuchiya, N., Enokizono, Y., Sugimura, T., Nagao, M. and Nakagama, H. (2002). Unfolding of quadruplex structure in the G-rich strand of the minisatellite repeat by the binding protein UP1. *Proc Natl Acad Sci U S A,* **99**(20), 12685-12690.

Generoso, W.M., Cain, K.T., Krishna, M. and Huff, S.W. (1979). Genetic lesions induced by chemicals in spermatozoa and spermatids of mice are repaired in the egg. *Proc Natl Acad Sci U S A,* **76**(1), 435-437.

Generoso, W.M., Bishop, J.B., Gosslee, D.G., Newell, G.W., Sheu, C.J. and von Halle, E. (1980). Heritable translocation test in mice. *Mutat Res,* **76**(2), 191-215.

Geyfman, M., Kumar, V., Liu, Q., Ruiz, R., Gordon, W., Espitia, F., Cam, E., Millar, S.E., Smyth, P., Ihler, A., Takahashi, J.S. and Andersen, B. (2012). Brain and muscle Arnt-like protein-1 (BMAL1) controls circadian cell proliferation and susceptibility to UVB-induced DNA damage in the epidermis. *Proc Natl Acad Sci U S A,* **109**(29), 11758-11763.

Giles, K.M., Hart, S.P., Haslett, C., Rossi, A.G., and Dransfield, I. (2000). An appetite for apoptotic cells? Controversies and challenges. *Br. J. Haematol*, **109**(1), 1–12.

Goldberg, Z. (2003). Clinical implications of radiation-induced genomic instability. *Oncogene*, **22**(45), 7011-7017.

Goodhead, D.T. (1994). Initial events in the cellular effects of ionizing radiations: clustered damage in DNA. *Int J Radiat Biol*, **65**(1), 7-17.

Green, S., Lavappa, K.S., Manandhar, M., Sheu, C.J., Whorton, E. and Springer, J.A. (1987). A guide for mutagenicity testing using the dominant lethal assay. *Mutat Res*, **189**(2), 167-174.

Gregory, C.D. (2000). CD14-dependence clearance of apoptotic cells: relevance to the immune system. *Curr Opin Immunol*, **12**(1), 27–34.

Grunau, C., Clark, S.J. and Rosenthal, A. (2001). Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res*, **29**(13), e65-65. (http://dx.doi.org/10.1093/nar/29.13.e65)

Haaf, T. and Ward, D.C. (1995). Higher order nuclear structure in mammalian sperm revealed by in situ hybridization and extended chromatin fibers. *Exp Cell Res*, **219**(2), 604-611.

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J. and Surani, M.A. (2002). Epigenetic reprogramming in mouse primordial germ cells. *Mech Dev*, **117**(1-2), 15-23.

Hatch, T., Derijck, A.A., Black, P.D., van der Heijden, G.W., de Boer, P. and Dubrova, Y.E. (2007). Maternal effects of the scid mutation on radiation-induced transgenerational instability in mice. *Oncogene*, **26**(32), 4720-4724.

Hattar, S., Lucas, R.J., Mrosovsky, N., Thompson, S., Douglas, R.H., Hankins, M.W., Lem, J., Biel, M., Hofmann, F., Foster, R.G. and Yau, K.W. (2003). Melanopsin and rod-cone photoreceptive systems account for all major accessory visual functions in mice. *Nature*, **424**(6944), 76-81.

Hida, A., Koike, N., Hirose, M., Hattori, M., Sakaki, Y. and Tei, H. (2000). The human and mouse Period1 genes: five well-conserved E-boxes additively contribute to the enhancement of mPer1 transcription. *Genomics*, **65**(3), 224-233.

Huang, L., Snyder, A.R. and Morgan, W.F. (2003). Radiation-induced genomic instability and its implications for radiation carcinogenesis. *Oncogene*, **22**(37), 5848-5854.

Hughes, M.E., DiTacchio, L., Hayes, K.R., Vollmers, C., Pulivarthy, S., Baggs, J.E., Panda, S. and Hogenesch, J.B. (2009). Harmonics of circadian gene transcription in mammals. *PLoS Genet*, **5**(4), e1000442. (http://dx.doi.org/10.1371/journal.pgen.1000442)

Hussain, S.P. and Harris, C.C. (2007). Inflammation and cancer: an ancient link with novel potentials. *Int J Cancer*, **121**(11), 2373–2380.

Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B. and Speed, T.P. (2003). Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res*, **31**(4), e15. (http://dx.doi.org/10.1093/nar/gng015)

Ishida, M., Ueha, T. and Sagami, I. (2008). Effects of mutations in the heme domain on the transcriptional activity and DNA-binding activity of NPAS2. *Biochem Biophys Res Commun*, **368**(2), 292-297.

Jackson, A.L. and Loeb, L.A. (2001). The contribution of endogenous sources of DNA damage to the multiple mutations in cancer. *Mutat. Res*, **477**(1-2), 7-21.

Ji, Y., Qin, Y., Shu, H. and Li, X. (2010). Methylation analyses on promoters of mPer1, mPer2, and mCry1 during perinatal development. *Biochem Biophys Res Commun*, **391**(4), 1742-1747.

Jones, P.A. and Baylin, S.B. (2002). The fundamental role of epigenetic events in cancer. *Nat Rev Genet*, **3**(6), 415-428.

Karger, B.L. and Guttman, A. (2009). DNA Sequencing by Capillary Electrophoresis. *Electrophoresis*, **30**(Suppl 1), S196–S202. (http://dx.doi.org/10.1002/elps.200900218)

Keller, M., Mazuch, J., Abraham, U., Eom, G.D., Herzog, E.D., Volk, H.D., Kramer, A. and Maier, B. (2009). A circadian clock in macrophages controls inflammatory immune responses. *Proc Natl Acad Sci U S A*, **106**(50), 21407-21412.

Koike, N., Yoo, S.H., Huang, H.C., Kumar, V., Lee, C., Kim, T.K. and Takahashi, J.S. (2012). Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science*, **338**(6105), 349-354. (http://dx.doi.org/10.1126/science.1226339)

Korenčič, A., Bordyugov, G., Košir, R., Rozman, D., Goličnik, M. and Herzel, H. (2012). The interplay of cis-regulatory elements rules circadian rhythms in mouse liver. *PLoS One*, **7**(11), e46835. (http://dx.doi.org/10.1371/journal.pone.0046835)

Koturbash, I., Rugo, R.E., Hendricks, C.A., Loree, J., Thibault, B., Kutanzi, K., Pogribny, I., Yanch, J.C., Engelward, B.P. and Kovalchuk, O. (2006). Irradiation induces DNA damage and modulates epigenetic effectors in distant bystander tissue in vivo. *Oncogene*, **25**(31), 4267-4275.

Koturbash, I., Boyko, A., Rodriguez-Juarez, R., McDonald, R.J., Tryndyak, V.P., Kovalchuk, I., Pogribny, I.P. and Kovalchuk, O. (2007). Role of epigenetic effectors in maintenance of the long-term persistent bystander effect in spleen in vivo. *Carcinogenesis*, **28**(8), 1831-1838.

Kovalchuk, O., Burke, P., Besplug, J., Slovack, M., Filkowski, J. and Pogribny, I. (2004). Methylation changes in muscle and liver tissues of male and female mice exposed to acute and chronic low-dose X-ray-irradiation. *Mutat. Res*, **548**(1-2), 75-84.

Koyama, S., Kodama, S., Suzuki, K., Matsumoto, T., Miyazaki, T. and Watanabe, M. (1998). Radiation-induced long-lived radicals which cause mutation and transformation. *Mutat Res*, **421**(1), 45-54.

Koyanagi, S., Hamdan, A.M., Horiguchi, M., Kusunose, N., Okamoto, A., Matsunaga, N. and Ohdo, S. (2011). cAMP-response element (CRE)-mediated transcription by activating transcription factor-4 (ATF4) is essential for circadian expression of the Period2 gene. *J Biol Chem*, **286**(37), 32416-32423.

Kraytsberg, Y. and Khrapko, K. (2005). Single-molecule PCR: an artifact-free PCR approach for the analysis of somatic mutations. *Expert Rev Mol Diagn*, **5**(5), 809-815.

Kumaki, Y., Ukai-Tadenuma, M., Uno, K.D., Nishio, J., Masumoto, K.H., Nagano, M., Komori, T., Shigeyoshi, Y., Hogenesch, J.B. and Ueda, H.R. (2008). Analysis and synthesis of high-amplitude Cis-elements in the mammalian circadian clock. *Proc Natl Acad Sci U S A*, **105**(39), 14946-14951.

Lamia, K.A., Storch, K.F. and Weitz, C.J. (2008). Physiological significance of a peripheral tissue circadian clock. *Proc Natl Acad Sci U S A*, **105**(39), 15172-15177.

Lampe, P.D. and Lau, A.F. (2000). Regulation of gap junctions by phosphorylation of connexins. *Arch Biochem Biophys*, **384**(2), 205-215.

Leduc, F., Maquennehan, V., Nkoma, G.B. and Boissonneault, G. (2008). DNA damage response during chromatin remodeling in elongating spermatids of mice. *Biol Reprod*, **78**(2), 324-332.

Lehnert, B.E. and Goodwin, E.H. (1997). Extracellular Factor(s) following Exposure to α Particles Can Cause Sister Chromatid Exchanges in Normal Human Cells. *Cancer Research*, **57**(11), 2164-2171.

Li, Y., Song, X., Ma, Y., Liu, J., Yang, D. and Yan, B. (2004). DNA binding, but not interaction with Bmal1, is responsible for DEC1-mediated transcription regulation of the circadian gene mPer1. *Biochem J*, **382**(Pt 3), 895-904.

Limoli, C.L., Kaplan, M.I., Phillips, J.W., Adair, G.M. and Morgan, W.F. (1997). Differential induction of chromosomal instability by DNA strand-breaking agents. *Cancer Res*, **57**(18), 4048–4056.

Liu, A.C., Tran, H.G., Zhang, E.E., Priest, A.A., Welsh, D.K. and Kay, S.A. (2008). Redundant Function of REV-ERBa and b and Non-Essential Role for Bmal1 Cycling in Transcriptional Regulation of Intracellular Circadian Rhythms. *PLoS Genet*, **4**(2), e1000023. (http://dx.doi.org/10.1371/journal.pgen.1000023)

Lopez-Molina, L., Conquet, F., Dubois-Dauphin, M. and Schibler, U. (1997). The DBP gene is expressed according to a circadian rhythm in the suprachiasmatic nucleus and influences circadian behavior. *EMBO J*, **16**(22), 6762-6771.

Lorimore, S.A. and Wright, E.G. (2003). Radiation-induced genomic instability and bystander effects: related inflammatory-type responses to radiation-induced stress and injury? A review. *Int J Radiat Biol*, **79**(1), 15–25.

Lorimore, S.A., Coates, P.J. and Wright, E.G. (2003). Radiation-induced genomic instability and bystander effects: inter-related nontargeted effects of exposure to ionizing radiation. *Oncogene*, **22**(45), 7058-7069.

Loudon, A.S. (2012). Circadian biology: a 2.5 billion year old clock. *Curr Biol*, **22**(14), R570-R571. (http://dx.doi.org/10.1016/j.cub.2012.06.023)

Loukinov, D.I., Pugacheva, E., Vatolin, S., Pack, S.D., Moon, H., Chernukhin, I., Mannan, P., Larsson, E., Kanduri, C., Vostrov, A.A., Cui, H., Niemitz, E.L., Rasko, J.E., Docquier, F.M., Kistler, M., Breen, J.J., Zhuang, Z., Quitschke, W.W., Renkawitz, R., Klenova, E.M., Feinberg, A.P., Ohlsson, R., Morse, H.C. 3rd and Lobanenkov, V.V. (2002). BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma. *Proc Natl Acad Sci U S A*, **99**(10), 6806-6811.

Luning, K.G., Frolen, H. and Nilsson, A. (1976). Genetic effects of ^{239}Pu salt injections in male mice. *Mutat Res*, **34**(3), 539–542.

Marmur, J. and Doty, P. (1962). Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *J Mol Biol*, **5**, 109-118.

Miller, B.H., McDearmon, E.L., Panda, S., Hayes, K.R., Zhang, J., Andrews, J.L., Antoch, M.P., Walker, J.R., Esser, K.A., Hogenesch, J.B. and Takahashi, J.S. (2007). Circadian and CLOCK-controlled regulation of the mouse transcriptome and cell proliferation. *Proc Natl Acad Sci U S A*, **104**(9), 3342-3347.

Miller, D., Brinkworth, M. and Iles, D. (2009). Paternal DNA packaging in spermatozoa: more than the sum of its parts? DNA, histones, protamines and epigenetics. *Reproduction*, **139**(2), 287-301.

Miranda, T.B. and Jones, P.A. (2007). DNA Methylation: The Nuts and Bolts of Repression. *J Cell Physiol*, **213**(2), 384-390.

Mirkin, S.M. (2006). DNA structures, repeat expansions and human hereditary disorders. *Curr Opin Struct Biol*, **16**(3), 351-358.

Mograbi, B., Corcelle, E., Defamie, N., Samson, M., Nebout, M., Segretain, D., Fénichel, P. and Pointis, G. (2003). Aberrant Connexin 43 endocytosis by the carcinogen lindane involves activation of the ERK/mitogen-activated protein kinase pathway. *Carcinogenesis*, **24**(8), 1415-1423.

Morgan, H.D., Santos, F., Green, K., Dean, W. and Reik, W. (2005). Epigenetic reprogramming in mammals. *Hum Mol Genet*, **14**(1), 47-58.

Morgan, W.F. (2003). Is there a common mechanism underlying genomic instability, bystander effects and other nontargeted effects of exposure to ionizing radiation? *Oncogene*, **22**(45), 7094-7099.

Moskalev, E.A., Zavgorodnij, M.G., Majorova, S.P., Vorobjev, I.A., Jandaghi, P., Bure, I.V. and Hoheisel, J.D. (2011). Correction of PCR-bias in quantitative DNA methylation studies by means of cubic polynomial regression. *Nucleic Acids Res*, **39**(11), e77. (http://dx.doi.org/10.1093/nar/gkr213)

Mullenders, J., Fabius, A.W., Madiredjo, M., Bernards, R. and Beijersbergen, R.L. (2009). A large scale shRNA barcode screen identifies the circadian clock component ARNTL as putative regulator of the p53 tumor suppressor pathway. *PLoS One*, **4**(3), e4798. (http://dx.doi.org/10.1371/journal.pone.0004798)

Nagar, S., Smith, L.E. and Morgan, W.F. (2003a). Characterization of a novel epigenetic effect of ionizing radiation: the death inducing effect. *Cancer Res*, **63**(2), 324-328.

Nagar, S., Smith, L.E. and Morgan, W.F. (2003b). Mechanisms of cell death associated with death-inducing factors from genomically unstable cell lines. *Mutagenesis.*, **18**(6), 549-560.

Nagar, S. and Morgan, W. (2005). The death-inducing effect and genomic instability. *Radiat Res*, **163**(3), 316-323.

Nakahata, Y., Yoshida, M., Takano, A., Soma, H., Yamamoto, T., Yasuda, A., Nakatsu, T. and Takumi, T. (2008). A direct repeat of E-box-like elements is required for cell-autonomous circadian rhythm of clock genes. *BMC Mol Biol*, **9**, 1.

Narayanan, P.K., LaRue, K.E., Goodwin, E.H. and Lehnert, B.E. (1999). Alpha particles induce the production of interleukin-8 by human cells. *Radiat Res*, **152**(1), 57-63.

Nelson, R.J. and Zucker, I. (1981). Photoperiodic control of reproduction in olfactory-bulbectomized rats. *Neuroendocrinology*, **32**(5), 266-271.

Neriishi, K., Nakashima, E. and Delongchamp, R.R. (2001). Persistent subclinical inflammation among A-bomb survivors. *Int J Radiat Biol*, **77**(4), 475–482.

Niwa, O. and Kominami, R. (2001). Untargeted mutation of the maternally derived mouse hypervariable minisatellite allele in F1 mice born to irradiated spermatozoa. *Proc Natl Acad Sci U S A*, **98**(4), 1705-1710.

Niwa, O. (2003). Induced genomic instability in irradiated germ cells and in the offspring; reconciling discrepancies among the human and animal studies. *Oncogene*, **22**(45), 7078-7086.

Nomura, T. (1982). Parental exposure to x rays and chemicals induces heritable tumours and anomalies in mice. *Nature*, **296**(5857), 575-577.

Nomura, T. (1990). Of mice and men? *Nature*. **345**(6277), 671.

Oda, A., Katayose, Y., Yabuuchi, S., Yamamoto, K., Mizuma, M., Shirasou, S., Onogawa, T., Ohtsuka, H., Yoshida, H., Hayashi, H., Rikiyama, T., Kim, H., Choe, Y., Kim, K., Son, H., Motoi, F., Egawa, S. and Unno, M. (2009). Clock gene mouse period2 overexpression inhibits growth of human pancreatic cancer cells and has synergistic effect with cisplatin. *Anticancer Res*, **29**(4), 1201–1209.

Ohno, T., Onishi, Y. and Ishida, N. (2007). A novel E4BP4 element drives circadian expression of mPeriod2. A novel E4BP4 element drives circadian expression of mPeriod2. *Nucleic Acids Res*, **35**(2), 648-655.

Palmer, D.K., O'Day, K. and Margolis, RL. (1990). The centromere specific histone CENP-A is selectively retained in discrete foci in mammalian sperm nuclei. *Chromosoma*, **100**(1), 32-36.

Panda, S., Provencio, I., Tu, D.C., Pires, S.S., Rollag, M.D., Castrucci, A.M., Pletcher, M.T., Sato, T.K., Wiltshire, T., Andahazy, M., Kay, S.A., Van Gelder, R.N. and Hogenesch, J.B. (2003). Melanopsin is required for non-image-forming photic responses in blind mice. *Science*, **301**(5632), 525-527.

Pappa, K.I., Gazouli, M., Anastasiou, E., Iliodromiti, Z., Antsaklis, A. and Anagnou, N.P. (2013). Circadian clock gene expression is impaired in gestational diabetes mellitus. *Gynecol Endocrinol*, **29**(4), 331-335.

Parsons, W.B. Jr., Watkins, C.H., Pease, G.L. and Childs, D.S. Jr. (1954). Changes in sternal marrow following roentgen-ray therapy to the spleen in chronic granulocytic leukemia. *Cancer*, **7**(1), 179-189.

Pendergast, J.S., Niswender, K.D. and Yamazaki, S. (2012). Tissue-specific function of Period3 in circadian rhythmicity. *PLoS One*, **7**(1), e30254. (http://dx.doi.org/10.1371/journal.pone.0030254)

Peters, A.H., O'Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schöfer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A., Opravil, S., Doyle, M., Sibilia, M. and Jenuwein, T. (2001). Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell*, **107**(3), 323-337.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*, **29**(9), e45. (http://dx.doi.org/10.1093/nar/29.9.e45)

Pfaffl, M.W., Horgan, G.W. and Dempfle, L. (2002). Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res*, **30**(9), e36. (http://dx.doi.org/10.1093/nar/30.9.e36)

Pittoggi, C., Renzi, L., Zaccagnini, G., Cimini, D., Degrassi, F., Giordano, R., Magnano, A.R., Lorenzini, R., Lavia, P. and Spadafora, C. (1999). A fraction of mouse sperm chromatin is organized in nucleosomal hypersensitive domains enriched in retroposon DNA. *J Cell Sci*, **112**(12), 3537-3548.

Pogribny, I., Raiche, J., Slovack, M. and Kovalchuk, O. (2004). Dose-dependence, sex- and tissue-specificity, and persistence of radiation-induced genomic DNA methylation changes. *Biochemical and Biophysical Research Communications*, **320**(4), 1253-1261.

Preitner, N., Damiola, F., Lopez-Molina, L., Zakany, J., Duboule, D., Albrecht, U. and Schibler, U. (2002). The orphan nuclear receptor REV-ERBalpha controls circadian transcription within the positive limb of the mammalian circadian oscillator. *Cell*, **110**(2), 251-260.

Pugacheva, E.M., Suzuki, T., Pack, S.D., Kosaka-Suzuki, N., Yoon, J., Vostrov, A.A., Barsov, E., Strunnikov, A.V., Morse, H.C. 3rd, Loukinov, D. and Lobanenkov, V. (2010). The structural complexity of the human BORIS gene in gametogenesis and cancer. *PLoS One*, **5**(11), e13872. (http://dx.doi.org/10.1371/journal.pone.0013872)

Qiagen (2009a). EpiTect Bisulfite Handbook – For complete bisulfite conversion and cleanup of DNA for methylation analysis. (http://www.qiagen.com/Products/Catalog/Assay-Technologies/Epigenetics/EpiTect-Bisulfite-Kits#resources)

Qiagen (2009b). REST 2009 Software User Guide – for gene expression analysis using real-time PCR data from the Rotor-Gene Q and other cyclers. REST 2009 Software. (http://www.REST.de.com)

Raiche, J., Rodriguez-Juarez, R., Pogribny, I. and Kovalchuk, O. (2004). Sex- and tissue-specific expression of maintenance and de novo DNA methyltransferases upon low dose X-irradiation in mice. *Biochemical and Biophysical Research Communications*, **325**(1), 39-47.

Rakyan, V.K., Preis, J., Morgan, H.D. and Whitelaw, E. (2001). The marks, mechanisms and memory of epigenetic states in mammals. *Biochem. J*, **356**(Pt 1), 1-10.

Rakyan, V.K., Chong, S., Champ, M.E., Cuthbert, P.C., Morgan, H.D., Luu, K.V. and Whitelaw, E. (2003). Transgenerational inheritance of epigenetic states at the murine Axin(Fu) allele occurs after maternal and paternal transmission. *Proc Natl Acad Sci U S A*, **100**(5), 2538-2543.

Raman, T., O'Connor, T.P., Hackett, N.R., Wang, W., Harvey, B.G., Attiyeh, M.A., Dang, D.T., Teater, M. and Crystal, R.G. (2009). Quality control in microarray assessment of gene expression in human airway epithelium. *BMC Genomics*, **10**(493). (http://dx.doi.org/10.1186/1471-2164-10-493)

Rassoulzadegan, M., Grandjean, V., Gounon, P., Vincent, S., Gillot, I. and Cuzin, F. (2006). RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature*, **441**(7092), 469-474.

Reick, M., Garcia, J.A., Dudley, C. and McKnight, S.L. (2001). NPAS2: an analog of clock operative in the mammalian forebrain. *Science*, **293**(5529), 506-509.

Rensing, L. and Ruoff, P. (2002). Temperature effect on entrainment, phase shifting, and amplitude of circadian clocks and its molecular bases. *Chronobiol Int*, **19**(5), 807-864.

Reppert, S.M. and Weaver, D.R. (2002). Coordination of circadian timing in mammals. *Nature*, **418**(6901), 935-941.

Roche NimbleGen (2008). NimbleScan v2.5 Software – User's Guide. NimbleScan v2.5 Software. (http://www.nimblegen.com/support/dna-microarray-support.html)

Roche NimbleGen (2009). NimbleGen Arrays User's Guide – Gene Expression Analysis v3.2. (http://chromatin.bio.fsu.edu/docs/EXP_userguide.pdf)

Roche NimbleGen (2011). NimbleGen Software – Guide to Interpreting the Experimental metrics report v3.0. (http://www.nimblegen.com/downloads/support/NG_GuideExpMetricsReport_v3p0.pdf)

Roemer, I., Reik, W., Dean, W. and Klose, J. (1997). Epigenetic inheritance in the mouse. *Curr Biol*, **7**(4), 277-280.

Routledge, F.S., McFetridge-Durdle, J.A. and Dean, C.R. (2007). Night-time blood pressure patterns and target organ damage: a review. *Can J Cardiol*, **23**(2), 132-138.

Russell, L.B. and Russell, W.L. (1992). Frequency and nature of specific-locus mutations induced in female mice by radiations and chemicals: a review. *Mutat Res*, **296**(1-2), 107-127.

Russell, W.L. (1951). X-ray inducted mutation in mice. *Cold Spring Harb Symp Quant Biol*, **16**, 327-336.

Russell, W.L., Bangham, J.W. and Russell, L.B. (1998). Differential response of mouse male germ-cell stages to radiation-induced specific-locus and dominant mutations. *Genetics*, **148**(4), 1567-1578.

Sahar, S. and Sassone-Corsi, P. (2007). Circadian clock and breast cancer: a molecular link. *Cell Cycle*, **6**(11), 1329–1331.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*, **74**(12), 5463-5467.

Sato, F., Kawamoto, T., Fujimoto, K., Noshiro, M., Honda, K.K., Honma, S., Honma, K. and Kato, Y. (2004). Functional analysis of the basic helix-loop-helix transcription factor DEC1 in circadian regulation. Interaction with BMAL1. *Eur J Biochem*, **271**(22), 4409-4419.

Schernhammer, E.S., Laden, F., Speizer, F.E., Willett, W.C., Hunter, D.J., Kawachi, I. and Colditz, G.A. (2001). Rotating night shifts and risk of breast cancer in women participating in the nurses' health study. *J Natl Cancer Inst*, **93**(20), 1563-1568.

Schetter, A.J., Heegaard, N.H. and Harris, C.C. (2010). Inflammation and cancer: interweaving microRNA, free radical, cytokine and p53 pathways. *Carcinogenesis*, **31**(1), 37–49.

Schibler, U., Ripperger, J. and Brown, S.A. (2003). Peripheral circadian oscillators in mammals: time and food. *J Biol Rhythms*, **18**(3), 250–260.

Searle, A.G. (1974). Mutation Induction in Mice. *Advances in Radiation Biology*, **4**, 131-207.

Seki, Y., Hayashi, K., Itoh, K., Mizugaki, M., Saitou, M. and Matsui, Y. (2005). Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. *Dev Biol*, **278**(2), 440-458.

Seki, Y., Yamaji, M., Yabuta, Y., Sano, M., Shigeta, M., Matsui, Y., Saga, Y., Tachibana, M., Shinkai, Y. and Saitou, M. (2007). Cellular dynamics associated with the genome-wide epigenetic reprogramming in migrating primordial germ cells in mice. *Development*, **134**(14), 2627-2638.

Shao, C., Folkard, M., Michael, B.D. and Prise, K.M. (2004). Targeted cytoplasmic irradiation induces bystander responses. *Proc Natl Acad Sci U S A*, **101**(37), 13495-13500.

Shi, G., Xing, L., Liu, Z., Qu, Z., Wu, X., Dong, Z., Wang, X., Gao, X., Huang, M., Yan, J., Yang, L., Liu, Y., Ptácek, L.J. and Xu, Y. (2013). Dual roles of FBXL3 in the mammalian circadian feedback loops are important for period determination and robustness of the clock. *Proc Natl Acad Sci U S A*, **110**(12), 4750-4755.

Shih, M.C., Yeh, K.T., Tang, K.P., Chen, J.C. and Chang, J.G. (2006). Promoter methylation in circadian genes of endometrial cancers detected by methylation-specific PCR. *Mol Carcinog*, **45**(10), 732–740.

Shiraishi, K., Shimura, T., Taga, M., Uematsu, N., Gondo, Y., Ohtaki, M., Kominami, R. and Niwa, O. (2002). Persistent induction of somatic reversions of the pink-eyed unstable mutation in F1 mice born to fathers irradiated at the spermatozoa stage. *Radiat Res*, **157**(6), 661-667.

Smith, L.E., Nagar, S., Kim, G.J. and Morgan, W.F. (2003). Radiation-induced genomic instability: radiation quality and dose response. *Health Physics*, **85**(1), 23-29.

Snyder, A.R. and Morgan, W.F. (2004a). Gene expression profiling after irradiation: clues to understanding acute and persistent responses? *Cancer Metastasis Rev*, **23**(3-4), 259-268.

Snyder, A.R. and Morgan, W.F. (2004b). Radiation-induced chromosomal instability and gene expression profiling: searching for clues to initiation and perpetuation. *Mutat Res*, **568**(1), 89-96.

Stow, L.R. and Gumz, M.L. (2011). The circadian clock in the kidney. *J Am Soc Nephrol*, **22**(4), 598-604.

Suzuki, T., Kosaka-Suzuki, N., Pack, S., Shin, D.M., Yoon, J., Abdullaev, Z., Pugacheva, E., Morse, H.C. 3rd, Loukinov, D. and Lobanenkov, V. (2010). Expression of a testis-specific form of Gal3st1 (CST), a gene essential for spermatogenesis, is regulated by the CTCF paralogous gene BORIS. *Mol Cell Biol*, **30**(10), 2473-2484.

Takeda, Y., Kang, H.S., Angers, M. and Jetten, A.M. (2011). Retinoic acid-related orphan receptor γ directly regulates neuronal PAS domain protein 2 transcription in vivo. *Nucleic Acids Res*, **39**(11), 4769-4782.

Tamminga, J., Koturbash, I., Baker, M., Kutanzi, K., Kathiria, P., Pogribny, I., Sutherland, R. and Kovalchuk, O. (2008a). Paternal cranial irradiation induces distant bystander DNA damage in the germline and leads to epigenetic alterations in the offspring. *Cell Cycle*, **7**(9), 1238–1245.

Tamminga, J., Kathiria, P., Koturbash, I. and Kovalchuk, O. (2008b). DNA damage-induced upregulation of miR-709 in the germline downregulates BORIS to counteract aberrant DNA hypomethylation. *Cell Cycl*e, **7**(23), 3731-3736.

Tamminga, J. and Kovalchuk, O. (2011). Role of DNA damage and epigenetic DNA methylation changes in radiation-induced genomic instability and bystander effects in germline in vivo. *Curr Mol Pharmacol*, **4**(2), 115-125.

Taniguchi, H., Fernández, A.F., Setién, F., Ropero, S., Ballestar, E., Villanueva, A., Yamamoto, H., Imai, K., Shinomura, Y. and Esteller, M. (2009). Epigenetic inactivation of the circadian clock gene BMAL1 in hematologic malignancies. *Cancer Res*, **69**(21) 8447-8454.

Taruscio, D., Zoraqi, G.K., Falchi, M., Iosi, F., Paradisi, S., Di Fiore, B., Lavia, P. and Falbo, V. (2000). The human per1 gene: genomic organization and promoter analysis of the first human orthologue of the Drosophila period gene. *Gene*, **253**(2), 161-170.

Terman, J.S., Remé, C.E. and Terman, M. (1993). Rod outer segment disk shedding in rats with lesions of the suprachiasmatic nucleus. *Brain Res*, **605**(2), 256-264.

Tokunaga, H., Takebayashi, Y., Utsunomiya, H., Akahira, J., Higashimoto, M., Mashiko, M., Ito, K., Niikura, H., Takenoshita, S. and Yaegashi, N. (2008). Clinicopathological significance of circadian rhythm-related gene expression levels in patients with epithelial ovarian cancer. *Acta Obstet Gynecol Scand*, **87**(10), 1060–1070.

Törnqvist, G., Sandberg, A., Hägglund, A.C. and Carlsson, L. (2010). Cyclic expression of lhx2 regulates hair formation. *PLoS Genet*, **6**(4), e1000904. (http://dx.doi.org/10.1371/journal.pgen.1000904)

Tosini, G. and Fukuhara, C. (2002). The mammalian retina as a clock. *Cell Tissue Res*, **309**(1), 119-126.

Ueda, H.R., Chen, W., Adachi, A., Wakamatsu, H., Hayashi, S., Takasugi, T., Nagano, M., Nakahama, K., Suzuki, Y., Sugano, S., Iino, M., Shigeyoshi, Y. and Hashimoto, S. (2002). A transcription factor response element for gene expression during circadian night. *Nature*, **418**(6897), 534-539.

Ukai-Tadenuma, M., Kasukawa, T. and Ueda, H.R. (2008). Proof-by-synthesis of the transcriptional logic of mammalian circadian clocks. *Nat Cell Biol*, **10**(10), 1154-1163.

Ukai-Tadenuma, M., Yamada, R.G., Xu, H., Ripperger, J.A., Liu, A.C. and Ueda, H.R. (2011). Delay in feedback repression by cryptochrome 1 is required for circadian clock function. *Cell*, **144**(2), 268-281.

Ullrich, R.L. and Ponnaiya, B. (1998). Radiation-induced instability and its relation to radiation carcinogenesis. *Int J Radiat Biol*, **74**(6), 747–754.

UNSCEAR (2000). Sources and effects of ionizing radiation. Vol. 1. Sources. United Nations, New York.

UNSCEAR (2001). Hereditary Effects of Radiation. United Nations, New York.

UNSCEAR (2006). Effects of Ionizing Radiation. Vol. 2. Annex C – Non-targeted and delayed effects of exposure to ionizing radiation. United Nations, New York.

UNSCEAR (2008). Sources and effects of Ionizing radiation. United Nations, New York.

Valerie, K., Yacoub, A., Hagan, M.P., Curiel, D.T., Fisher, P.B., Grant, S. and Dent, P. (2007). Radiation-induced cell signaling: inside-out and outside-in. *Mol Cancer Ther*, **6**(3), 789-801.

van der Heijden, G.W., Derijck, A.A.H.A., Ramos, L., Giele, M., van der Vlag, J. and de Boer, P. (2006). Transmission of modified nucleosomes from the mouse male germline to the zygote and subsequent remodeling of paternal chromatin. *Developmental Biology*, **298**(2), 458-469.

Vance, M.M. and Wiley, L. (1999). Gap junction intercellular communication mediates the competitive cell proliferation disadvantage of irradiated mouse preimplantation embryos in aggregation chimeras. *Radiat Res*, **152**(5), 544-551.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*, **3**(7), 1-12.

Veenstra, R.D. (1996). Size and selectivity of gap junction channels formed from different connexins. *J Bioenerg Biomembr*, **28**(4), 327-337.

Vollmers, C., Gill, S., DiTacchio, L., Pulivarthy, S.R., Le, H.D. and Panda, S. (2009). Time of feeding and the intrinsic circadian clock drive rhythms in hepatic gene expression. *Proc Natl Acad Sci U S A*, **106**(50), 21453-21458.

Vorobtsova, I.E. (2000). Irradiation of male rats increases the chromosomal sensitivity of progeny to genotoxic agents. *Mutagenesis*, **15**(1), 33-38.

Walsh, C.P., Chaillet, J.R. and Bestor, T.H. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat Genet*, **20**(2), 116-117.

Warnecke, P.M., Stirzaker, C., Melki, J.R., Millar, D.S., Paul, C.L. and Clark, S.J. (1997). Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA. *Nucleic Acids Res*, **25**(21), 4422-4426.

Weissenborn, U. and Streffer, C. (1988). Analysis of structural and numerical chromosomal anomalies at the first, second, and third mitosis after irradiation of one-cell mouse embryos with X-rays or neutrons. *Int J Radiat Biol*, **54**(3), 381–394.

Widlund, H.R., Cao, H., Simonsson, S., Magnusson, E., Simonsson, T., Nielsen, P.E., Kahn, J.D., Crothers, D.M. and Kubista, M. (1997). Identification and characterization of genomic nucleosome-positioning sequences. *J Mol Biol*, **267**(4), 807-817.

Wiley, L.M., Baulch, J.E., Raabe, O.G. and Straume, T. (1997). Impaired cell proliferation in mice that persists across at least two generations after paternal irradiation. *Radiat Res*, **148**(2), 145-151.

Wimmer, U., Wang, Y., Georgiev, O. and Schaffner, W. (2005). Two major branches of anti-cadmium defense in the mouse: MTF-1/metallothioneins and glutathione. *Nucleic Acids Res*, **33**(18), 5715–5727.

Winter, S.L., Bosnoyan-Collins, L., Pinnaduwage, D. and Andrulis, I.L. (2007). Expression of the circadian clock genes Per1 and Per2 in sporadic and familial breast tumors. *Neoplasia*, **9**(10), 797–800.

Wolfinger, R. D., Gibson, G., Wolfinger, E. D., Bennett, L., Hamadeh, H., Bushel, P., Afshari, C. and Paules, R. S. (2001). Assessing gene significance from cDNA microarray expression data via mixed models. *J. Comp. Biol.*, **8**(6), 625–637.

Wouters-Tyrou, D., Martinage, A., Chevaillier, P. and Sautière, P. (1998). Nuclear basic proteins in spermiogenesis. *Biochimie*, **80**(2), 117-128.

Wu, L.J., Randers-Pehrson, G., Xu, A., Waldren, C.A., Geard, C.R., Yu, Z. and Hei, T.K. (1999). Targeted cytoplasmic irradiation with alpha particles induces mutations in mammalian cells. *Proc Natl Acad Sci U S A*, **96**(9), 4959–4964.

Xia, H.C., Niu, Z.F., Ma, H., Cao, S.Z., Hao, S.C., Liu, Z.T. and Wang, F. (2010). Deregulated expression of the Per1 and Per2 in human gliomas. *Can J Neurol Sci*, **37**(3), 365–370.

Yamaguchi, S., Mitsui, S., Yan, L., Yagita, K., Miyake, S. and Okamura, H. (2000). Role of DBP in the circadian oscillatory mechanism. *Mol Cell Biol*, **20**(13), 4773-4781.

Yamajuku, D., Shibata, Y., Kitazawa, M., Katakura, T., Urata, H., Kojima, T., Nakata, O. and Hashimoto, S. (2010). Identification of functional clock-controlled elements involved in differential timing of Per1 and Per2 transcription. *Nucleic Acids Res*, **38**(22), 7964-7973.

Yamamoto, T., Nakahata, Y., Soma, H., Akashi, M., Mamine, T. and Takumi, T. (2004). Transcriptional oscillation of canonical clock genes in mouse peripheral tissues. *BMC Mol Biol*, **5**, 18.

Yang, M.Y., Chang, J.G., Lin, P.M., Tang, K.P., Chen, Y.H., Lin, H.Y., Liu, T.C., Hsiao, H.H., Liu, Y.C. and Lin, S.F. (2006). Downregulation of circadian clock genes in chronic myeloid leukemia: alternative methylation pattern of hPER3. *Cancer Sci*, **97**(12), 1298–1307.

Yang, M.Y., Yang, W.C., Lin, P.M., Hsu, J.F., Hsiao, H.H., Liu, Y.C., Tsai, H.J., Chang, C.S. and Lin, S.F. (2011). Altered expression of circadian clock genes in human chronic myeloid leukemia. *J Biol Rhythms*, **26**(2), 136–148.

Yauk, C.L., Dubrova, Y.E., Grant, G.R. and Jeffreys, A.J. (2002). A novel single molecule analysis of spontaneous and radiation-induced mutation at a mouse tandem repeat locus. *Mutat Res*, **500**(1-2), 147-156.

Ye, R., Selby, C.P., Ozturk, N., Annayev, Y. and Sancar, A. (2011). Biochemical analysis of the canonical model for the mammalian circadian clock. *J Biol Chem*, **286**(29), 25891-25902.

Young, M.E. (2006). The circadian clock within the heart: potential influence on myocardial gene expression, metabolism, and function. *Am J Physiol Heart Circ Physiol*, **290**(1), H1-H16. (http://dx.doi.org/10.1152/ajpheart.00582.2005)

Zalenskaya, I.A., Bradbury, E.M. and Zalensky, A.O. (2000). Chromatin structure of telomere domain in human sperm. *Biochem Biophys Res Commun*, **279**(1), 213-218.

Zhao, W., Chuang, E.Y., Mishra, M., Awwad, R., Bisht, K., Sun, L., Nguyen, P., Pennington, J.D., Wang, T.J., Bradbury, C.M., Huang, L., Chen, Z., Bar-Sela, G., Robbins, M.E. and Gius, D. (2006). Distinct effects of ionizing radiation on in vivo murine kidney and brain normal tissue gene expression. *Clin Cancer Res*, **12**(12), 3823-3830.

Zhou, H., Randers-Pehrson, G., Waldren, C.A., Vannais, D., Hall, E.J. and Hei, T.K. (2000). Induction of a bystander mutagenic effect of alpha particles in mammalian cells. *Proc Natl Acad Sci U S A*, **97**(5), 2099-2104.

Zhu, Y., Leaderer, D., Guss, C., Brown, H.N., Zhang, Y., Boyle, P., Stevens, R.G., Hoffman, A., Qin, Q., Han, X. and Zheng, T. (2007). Ala394Thr polymorphism in the clock gene NPAS2: a circadian modifier for the risk of non-Hodgkin's lymphoma. *Int J Cancer*, **120**(2), 432–435.

Zhu, Y., Stevens, R.G., Hoffman, A.E., Fitzgerald, L.M., Kwon, E.M., Ostrander, E.A., Davis, S., Zheng, T. and Stanford, J.L. (2009). Testing the circadian gene hypothesis in prostate cancer: a population-based case-control study. *Cancer Res*, **69**(24), 9315–9322.

Zuber, A.M., Centeno, G., Pradervand, S., Nikolaeva, S., Maquelin, L., Cardinaux, L., Bonny, O. and Firsov, D. (2009). Molecular clock is involved in predictive circadian adjustment of renal function. *Proc Natl Acad Sci U S A*, **106**(38), 16523-16528.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*, **31**(13), 3406-3415.

Books

Craig, J.M. and Wong, N.C. eds. (2011). *Epigenetics – A Reference Manual*. Norfolk: Caister Academic Press.

Drăghici, S. (2003). *Data Analysis Tools for DNA Microarrays*. Washington, D.C.: Chapman & Hall/CRC.

Dubrova, Y.E. (2007). Radiation-Induced Genomic Instability in the Offspring of Irradiated Parents. In: C. Mothersill, I. Mosse, and C. Seymour, eds. (2007). *Multiple Stressors: A Challenge for the Future*. Dordrecht: Springer. pp.139-154.

Field, A., Miles, J. and Field, Z. (2012). *Discovering Statistics using R.* London: Sage Publications Ltd.

Filion, M. ed. (2012). *Quantitative Real-time PCR in Applied Microbiology*. Norfolk: Caister Academic Press.

Haines, J.L. and Pericak-Vance, M.A. (2006). *Genetic Analysis of Complex Disease*. 2nd ed. New Jersey: John Wiley & Sons, Inc.

Kennedy, S. and Oswald, N. eds. (2011). *PCR Troubleshooting and Optimization – The Essential Guide*. Norfolk: Caister Academic Press.

Kimmel, A. and Oliver, B. eds. (2006). *Methods in Enzymology, Volume 411: DNA Microarrays (Part B: Databases and Statistics)*. California: Elsevier, Inc.

Neumann, H.P. ed. (2007). *Progress in DNA methylation Research*. New York: Nova Science Publishers, Inc.

Sokal, R.R. and Rohlf, F.J. (2012). *Biometry*. 4th ed. New York: W.H. Freeman and Company.

Sprawls, P. Jr. (1995). *Physical principles of Medical Imaging.* 2nd ed. Madison: Medical Physics Publishing Corporation. (available at: http://www.sprawls.org/ppmi2/RADQU/)

Stekel, D. (2003). *Microarray Bioinformatics*. Cambridge: Cambridge University Press.