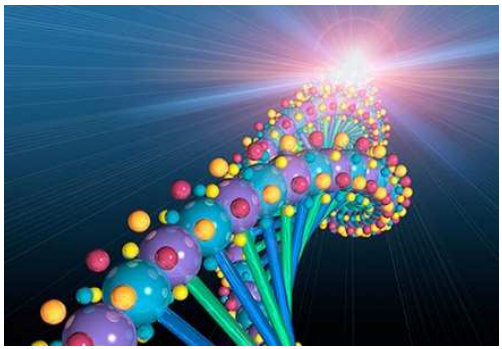
Sapienza Università di Roma

"THE GENE TARGETING APPROACH OF SMALL FRAGMENT HOMOLOGOUS REPLACEMENT: INFLUENCE OF EPIGENETICS, DNA REPAIR AND CELL CYCLE"



Tutore
Prof. Marco Lucarelli

Docente guida
Prof. Marco Lucarelli

Coordinatore
Prof. Marco Tripodi

Dottoranda
Silvia Pierandrei

Dottorato di Ricerca in Scienze Pasteuriane
XXV CICLO

INDEX

1. AIMS

Gene targeting strategies are aimed to the correction of a mutation *in situ*, allowing the recovery of a normal gene function. They have significant therapeutic and safety advantages over traditional gene therapy approaches (based on transgene delivery) as the mutant genetic instructions are directly repaired in a site-specific, long-term and genetically inheritable manner, in their native sequence context. Consequently, the targeted gene is regulated by its endogenous machinery, maintaining the physiologic expression pattern. Several gene targeting approaches are thought to be based on homologous recombination (HR). In mitotic cells, this is a basic mechanism to repair DNA damage, in particular DNA double-strand breaks. The low frequency of HR and the possible random (non-homologous) integration constitute the main barriers to the easy gene targeting in vertebrate cells. Within gene targeting protocols, the Small Fragment Homologous Replacement (SFHR) is aimed to stably modify genomic sequences, restoring the wild-type gene function, using small DNA fragments (SDFs) used as template for modification. Molecular mechanisms of SFHR are poorly understood, although DNA repair pathways, in particular HR, seem to be involved. The potential of this approach is currently

limited by a low and variable correction frequency due to several factors influencing its efficiency.

The general aim of my thesis was to contribute to the clarification of molecular mechanisms underlying the cell invasion by exogenous DNA and its processing, as well as to the comprehension of mechanisms underlying the SFHR gene therapy approach. To this purpose, this thesis focused on main pathways involved in epigenetic silencing, host defence, genome integrity maintenance and cell cycle control, after treatment with an exogenous SDF. The scarcely characterized relationships between SFHR and these main cellular pathways was studied at whole pathway level and at single gene level. In particular, the effect of SFHR on DNA methylation of targeted locus and on the expression of genes involved in the DNA damage repair and cell cycle progression were studied. The experimental evidences obtained by our studies, confirm the involvement and interconnection of the biochemical pathways analyzed in the cellular response to cell invasion by exogenous DNA. The selection of specific targets, at cellular and single gene level, is crucial to better clarify the exact mechanisms and networks involved. In turn, this better comprehension of the basic mechanisms, may allow to ameliorate the efficiency of SFHR in both differentiated and embryonic stem (ES) cells, primary targets of gene therapy protocols.

ABBREVIATIONS:

5-Aza-dC= 5-Aza-2'-Deoxycytidine	mutEGFP= mutated EGFP gene
BER= base excision repair	NER= nucleotide excision repair
CFTR= cystic fibrosis transmembrane conductance regulator gene;	NHEJ= non homologous end joining
DMD= dystrophin gene (muscular dystrophy, Duchenne and Becker types)	qRT-PCR arrays= real time expression arrays
DSB= double strand break	SFHR: small fragment homologous replacement
dsDNA= double strand DNA;	SDFs: small DNA fragments
EGFP= enhanced green fluorescent protein	SDF-DIG-WT= double strand plasmid-digested small DNA fragment
HR= homologous recombination;	SDF-PCR-WT= double strand PCR amplified small DNA fragment
MMR= mismatch repair ;	ssDNA= single strand DNA
MEF= mouse embryonic fibroblast	wtEGFP= wild type EGFP gene
MEF-mutEGFP= mouse embryonic fibroblast with integrated mutated EGFP gene	
MEF-wtEGFP= mouse embryonic fibroblast with integrated wild type EGFP gene	

2. INTRODUCTION AND BACKGROUND

2.1 Gene Therapy

The primary objective of gene therapy is to correct a gene, restoring the wild type phenotype. This can be achieved using DNA molecules as "drugs" to correct a permanent alteration of the genome or to express functional copies of a "therapeutic gene".

Virtually all cells are potential targets for gene therapy. However, the cells can be divided into two major categories: somatic cells or germline cells (eggs or sperm). Gene therapy using germ line cells results in permanent changes that passed down to subsequent generations. The appeal of germ line gene therapy is its potential for offering a permanent therapeutic effect for every person who inherit the target gene. Successful germ line therapies introduce the possibility of eliminating some diseases from a particular family and, ultimately, from the population. However, this also raises controversy. Somatic cells are non-reproductive. Somatic cell therapy is viewed as a more conservative, safer approach because it affects only the targeted cells in the patient, and is not passed on to future generations. In other words, the therapeutic effect is limited at the individual who received the therapy. However, this type of therapy

presents some problems. Often the effects of somatic cell therapy are short-lived. Because the cells of most tissues ultimately die and are replaced by new cells, repeated treatments over the course of the individual's life span are required to maintain the therapeutic effect. Delivering the gene to the target cells or tissue is also problematic. Regardless of these difficulties, however, somatic cell gene therapy is appropriate and acceptable for many disorders, including cystic fibrosis, muscular dystrophy, cancer, and certain infectious diseases like acquired immunodeficiency virus (AIDS) (Cavazzana-Calvo et al., 2000; Nemunaitis et al., 2001). In summary, to date, all gene therapy on humans has been directed at somatic cells; therefore, by these approaches any changes to the genes of an individual will only impact his cells and are not inherited by future generations. Depending on the cell target, the pathology and the delivery system, gene therapy protocols can be applied *in vivo*, *in vitro* or *ex vivo*.

The *in vivo* gene therapy is a strategy of gene transfer that theoretically can be used for the treatment of most genetic disorders. By this protocol the therapeutic DNA is inserted directly into cells and/or tissues of the patient (usually through viral vectors). In this case it is essential that the transfer of the therapeutic gene takes place in a sufficient number of affected cells, in order to obtain a clear change at the phenotypic level.

In vivo applications have been carried out in several diseases, including cystic fibrosis (Crystal et al., 1994; Flotte et al., 1996), in which an adeno-associated virus (AAV) containing the cystic fibrosis transmembrane conductance regulator gene (CFTR), was administered to patients in the form of aerosols.

About the *in vitro* protocols, the DNA is transferred into cultured cells to correct the mutated sequences. This procedure is used primarily for research purposes and as a preliminary step for *ex vivo* protocols.

In the *ex vivo* protocols, the DNA is transferred into the cells previously isolated from the patient; subsequently the corrected cells are reintroduced into the same patient (*ex vivo* gene therapy).

This indirect procedure, although longer than *in vivo* approach, allows to monitor and to evaluate the transfer efficiency; moreover it allows to isolate, in culture, the cells in which correction has been successful. These cells, once transplanted, are able to synthesize the proteins of interest at physiological levels.

Gene therapy involves two different therapeutic approaches depending on whether the endogenous defective gene is intended to be corrected permanently *in situ* (gene targeting approach), or to be over-expressed into the cell (gene augmentation approach). These two procedures are clearly distinct one from the other by specific characteristics, thus leading to advantages and

disadvantages in their application. Although very promising, till now the gene therapy has shown limited clinical successes; for this reason a vast effort worldwide is undergoing to allow its eventual therapeutic use.

2.1.1 Gene Targeting

The gene targeting is a methodology by which a specific sequence of the cellular genome, defined target sequence, is modified directly *in situ* (Yanez et al., 1998). The technique allows to stably modify the mutated genotype providing to the target cells the exact genetic information and restoring the normal structure and proper functioning of the protein (Capecchi, 1989). The modification is obtained introducing into the cells a sequence of exogenous DNA that specifically recognizes and modifies the target sequence. This possibly occurs exploiting the natural events of Homologous Recombination (HR) that take place in cells between two specific sequences of DNA (Smithies et al., 1985; Thomas et al., 1986). The gene targeting offers several advantages over the cDNA-based gene therapy strategies (Yanez et al., 1998). First of all, the technique allows a precise correction of the defect directly on the gene locus. It can also be applied in the treatment of both dominant and recessive monogenic diseases. The size of the gene to be corrected does not constitute a limiting factor because the targeting sequence corrects only the portion of the

altered DNA. Furthermore, the reversion of the genetic code is permanent and the integrity of the gene is respected. This is the most relevant advantage over the gene augmentation because the relationship between the coding and regulatory elements of the target gene is maintained. Only by keeping this level of genomic integrity one can be assured that the targeted gene will be appropriately expressed and translated in a specific cell physiological context. In fact, although eukaryotic cells have multiple levels of regulation of gene expression that are still poorly known, it is becoming increasingly clear the crucial role of flanking and intron sequences, as well as of polymorphisms. Even epigenetic factors can affect indirectly the physiological expression of a gene sequence. Consequently, it is arbitrary to consider gene therapy detached from the specific cellular and genomic environment. So far, the validity of this approach has been highlighted in human and murine cells, primary and immortalized, and in stem cells. However, there are two main barriers to easy gene targeting in vertebrate cells: the low frequency of HR, generally occurring at one event per 10^5 to 10^7 treated cells, and the high rate of random (non-homologous) integration, which occurs in about one per 10^2 to 10^4 treated cells (Capecchi, 1989; Capecchi, 1994). Many efforts are aimed to increase the efficiency of recombination, which depends mainly on the cell type target. In addition, the efficiency of targeting seems to be influenced by the

differentiation status of the cells, the degree of chromatin condensation and the phases of the cell cycle. As it is most likely that the process of gene targeting exploits the physiologic mechanisms of DNA damage repair, also the mechanisms of DNA repair appear to affect the efficiency of gene targeting.

2.1.2 Small Fragments Homologous Replacement

Several techniques of gene targeting have been developed as Small Fragments Homologous Replacement (SFHR), Chimeric RNA-DNA Oligonucleotide (RDO), Single Strand Oligonucleotide (SSO), Triple Helix-forming Oligonucleotide (TFO), Zinc Fingers Nucleases (ZFNs); within these, the SFHR is a gene repair strategy that involves the introduction of small DNA fragments (SDFs) (up to 1 kb) into the cells. The fragments are essentially homologous to the sequence of the mutant gene except that they code for the normal, rather than the mutant sequence. After entering the cells, these SDFs involve exchange between their sequence and the endogenous one through an, as yet, undefined mechanism (Gruenert, 1998, 1999; Yanez et al., 1998) that probably involves the HR and / or other pathways of DNA repair (Goncz, 2000). The SFHR has already been used to modify different kinds of genomic mutations *in vitro* and *in vivo* in both human and mouse cells, as well as animal model, demonstrating its potentiality for the treatment of several disease-associated genes. These include CFTR

(responsible for cystic fibrosis) (Kunzelmann et al., 1996; Colosimo et al., 2001; Sangiuolo et al., 2002; Maurisse et al., 2006; Sangiuolo et al., 2008), dystrophin (DMD, responsible for muscular dystrophies) (Kapsa et al. 2001; Kapsa et al., 2002; Todaro et al., 2007), survival motor neuron (SMN, responsible for spinal muscular atrophy) (Sangiuolo et al., 2005; Spitalieri et al. 2009), hypoxanthine phosphoribosyltransferase 1 (HPRT1, responsible for Lesch-Nyhan syndrome) (Bedayat et al., 2009), DNA - dependent protein kinase catalytic subunit (DNA-PKcs, responsible for severe combined immune deficiency) (Zayed et al., 2006) and β-globin (*HBB,* responsible for β-thalassemia and sickle cell disease) (Goncz et al., 2002; Goncz et al., 2006; Colosimo et al., 2007). In 2008, Sangiuolo's research group has applied SFHR to stably introduce a 3-bp deletion (F508del, the most frequent cystic fibrosis mutation) into the CFTR locus in the mouse embryonic stem (ES) cell genome. After transfection of F508del-SDF into murine ES cells, SFHR-mediated modification was evaluated at molecular levels on both DNA and mRNA obtained from transfected ES cells. The results indicate that, in spite of its low efficiency, the SFHR technique can be used to effectively target and modify genomic sequences in ES cells.

SFHR was successfully employed by Kapsa and collaborators to correct a nonsense mutation in the DMD locus of the *mdx* mouse model of Duchenne muscolar dystrophy both *in vitro* and in *vivo*. A 603-bp wild-type PCR product

was used to repair the exon 23 C to T *mdx* nonsense transition at dystrophin locus in cultured myoblasts and in affected muscles from male *mdx* mice. Conversion was observed at both the DNA and RNA levels. The level of conversion of *mdx* to wild-type sequence *in vitro* was about 15 %, although there was no detection of normal Dmd protein. *In vivo* the correction efficiency was up to 0.1 % in the tibialis anterior of male *mdx* mice, but again there was no evidence of gene expression at either the transcript or protein level. It was suggested that the disparity between the genomic repair and protein expression was possibly due to toxicity of the transfected agent on myoblasts or a delay in protein expression. The genetic correction of *mdx* myoblasts was shown to persist for up to 28 days in culture and for at least 3 weeks in muscles.

Zayed and collaborators described genotypic and functional correction of a point mutation in the gene encoding the DNA-PKcs that causes severe combined immune deficiency in mice. Using a T cell thymoma line they have shown that SDFs (621 bp) can provide genotypic and functional correction of these cells. Corrected cells were selected on the basis of protection from radiation hypersensitivity that occurs as a consequence of the SCID mutation. These results have indicated that SDFs can correct point mutations by HR with

the possibility of using ionizing radiation as a selection method to eliminate non corrected cells and enrich for corrected SCID radioresistant cells.

To determine if SFHR can be used as a gene therapy option for *ex vivo* treatment of sickle cell anemia, Goncz and collaborators carried out a study in which SFHR-mediated modification of the endogenous β-globin gene locus was analyzed in human chronic myelogenous leukemia lymphoblasts (K562), as well as in hematopoietic progenitor/stem cells (CD38- / lin- or CD34+). The DNA fragments (559 bp) used for transfection in these experiments are exactly homologous to normal β-globin (β^A) except for an A>T conversion located in the sixth codon of the gene (*i.e.* sickle or β^S genotype) and a silent mutation that introduces a unique Afl II site. The conversion of endogenous β^A globin to β^S globin was successful in both K562 cells and in hematopoietic progenitor / stem cells. This study showed that the endogenous β-globin locus can be modified by SFHR in clinically relevant cells. Specifically, the endogenous β^A globin allele of both normal lymphoblast and CD34+ and / or CD38-lin- cells can be converted to a β^S globin. The potential of SFHR is currently limited by a low and variable frequency of correction, ranging from 0.01 % to 5 % (Goncz et al., 1998, 2000 and 2001; Colosimo et al., 2001; Kapsa et al., 2002; Thorpe et al., 2002). Earliest works suffered from the lack of a suitable cellular experimental system for the systematic study of SFHR at DNA, RNA and protein level.

Considering these difficulties, functional assays scoring phenotypic changes upon gene correction have become fundamental to measure the actual frequencies of SFHR-mediated modification in different target cells. It is also important to highlight that the lack of a detailed knowledge of the mechanism of action of the SDF at the molecular level, sets a limit to the reproducibility of this strategy. The accessibility of the SDF to the target locus depends on several variables as, for example, an efficient nuclear uptake, the half-life of the SDF, the chromatin compaction and the level of recombination activities. Any factor that increases the dynamic nature of the chromatin, relax nucleoprotein conformation and enhance DNA exchange, may prove to be useful to increase target gene repair frequencies. The epigenetic changes and chromatin remodeling occurring during the different phases of the cell cycle, as well as the host defence response involved in the maintenance of genome structural integrity, are likely to influence the targeting mechanism. They may represent key factors underlying the low and variable SFHR efficiency. The clarification of basic mechanisms and networks involved is crucial to ameliorate the efficiency of SFHR, allowing its future use in clinical protocols.

2.1.3 Gene Augmentation

The approach of gene augmentation (Flotte et al., 2001) involves the use of either the coding cDNA sequence of a gene or its complete genomic sequence. In first case, the cDNA sequence, containing the complete coding sequence of the gene, is usually flanked by regulatory sequences, such as a viral promoter capable of ensuring high levels of expression. After the transfer, the introduced genes can integrate in the cell genome or remain in the form of extrachromosomal genetic elements. The integration in the genome allows the replication of the gene and its transfer to daughter cells that arise from cell division. In this way, it is possible to obtain long-term stable expression of the product that allows to maintain the effectiveness of therapy over time. However, the gene integration on chromosome also presents disadvantages because sometimes the insertion of the gene occurs in a random way; in fact, in some cases, the gene cannot be expressed because it is integrated in a region of heterochromatin that prevents the transcription. In other cases, instead, if the event of insertion takes place in a critical region for cell survival, the host cell can undergo death due to inactivation of genes responsible for cell survival. There is also the possibility that the recombination event can modify the expression of genes that control cell division and cell proliferation, as the

inactivation of a tumor suppressor or the activation of an oncogene (Remus et al., 1999; Muller et. al., 2001). In the second case large vectors, such as yeast artificial chromosomes (YACs) or bacterial artificial chromosomes (BACs), spanning all introns and regulatory regions of the gene of interest, may be used. Their main advantage in gene therapy are the physiological levels of expression and the long-term maintenance. This is achieved by using large genomic fragments, mostly behaving as cellular chromosomes, that contain all of the long-range controlling elements of the gene of interest, that will allow tissue-specific gene expression to physiological level (Conese et al., 2007). The main disadvantage of this kind of vectors is the handling and delivery difficulties. Due to the relevance for gene therapy of this kind of approach, in the first part of my phD, I also contributed to the structural and functional characterization of a BAC vector containing the CFTR locus (Auriche et al., 2010).

2.2 Delivery Systems

The current gene therapy techniques mainly utilize transport systems to transfer the genetic material into the nucleus. The success of gene therapy is largely dependent on the development of vectors that can selectively and efficiently deliver a gene to target cells with minimal toxicity. Gene delivery

systems generally fall into two categories: viral and non viral. The first one utilizes viral vectors; their expression do not require a delivery vehicle for DNA transfer because viruses infect cells as part of their natural biological function. However, the mode of entry for each virus is different. The second one utilizes either chemical or physical methods to introduce gene of interest into target cells. Non viral expression vectors usually show low transfection efficiency. Consequently, there are many different transfer methods based on transfection vehicles. These vehicles can be classified as either chemical (phosphate precipitation, cationic polymers, liposomes, molecular conjugates and others) or physical (electroporation, biolistic and microinjection). The choice of vehicle protocol depends, in part, on the cell type and number of cells to be transfected, and on the desired transfection efficiency. The general choice of the vector depends on several factors including the type of organ or the target cell type and the size of the gene to be transferred (Orkin, 1986). Another key to effective transfection is the growth state of the transfected cells. The best results are achieved when the cells are in log phase growth. Another factor to consider in transfection is that uncut, supercoiled DNA is optimal for transient expression systems, while linear DNA tends to be more recombinogenic and will facilitate stable transfection (McNally et al., 1988). Finally, the choice of the transfection vehicle will be critical when defining the appropriate system for *in*

vitro versus *in vivo* DNA transfer. To date, several efforts are focused to improve the techniques for the transfer of genetic material into cells and / or tissues (Sullenger, 2003) and to develop effective and safe protocols.

2.2.1 Viral gene transfer systems

Viral vector systems were developed to facilitate efficient delivery of DNA expression vectors into cells. The gene of interest is delivered in the form of a defective virus encapsulated within a virus envelope that increases the efficiency of gene transfer. The recipient cells are infected with these defective viruses. The viruses have been modified so that they are no longer able to replicate autonomously and require passage through a helper cell line that contains the genetic instructions for coating or packaging the recombinant viral vector. Each viral vector system is characterized by an inherent set of properties that affect its suitability for specific gene therapy applications so the choice of a given viral delivery system will depend on the endpoints desired. There are now numerous recombinant viral delivery systems.

Retroviruses were the first vectors used in gene therapy (Boris-Lawrie et al., 1994). Retroviruses are RNA viruses with a single stranded genome ranging from 7 to 11 kilobases. Following entry into target cells, the RNA genome is retro-transcribed into linear doublestranded DNA and integrated into the cell

chromatin. This family of viruses includes several varieties being exploited for gene therapy: the mammalian and avian C-type retroviruses, lentiviruses (such as HIV and other immunodeficiency viruses) and spumaviruses. They tend to establish chronic infection that is usually well tolerated by the host but may also cause latent diseases ranging from malignancy to immunodeficiency. A useful property of retroviral vectors is the ability to integrate efficiently into the chromatin of target cells. Thanks to integration, this system achieves stable, long-term expression of therapeutic gene. The disruption of the nuclear membrane is required to gain access to the chromatin to allow the complex integration. The productive transduction by retroviral vectors is strictly dependent on target cell mitosis shortly after entry. Because only a fraction of cells pass through mitosis at any given time, this severely limits the applications of retroviral vectors in gene therapy since the differentiated cellular systems are excluded (Scott-Taylor et al., 1998).

Lentiviruses (for example HIV, human immunodeficiency virus) have a morphology and a replicative cycle very similar to retroviruses. Unlike retroviruses, they rely on active transport of the preintegration complex through the nucleopore by the nuclear import machinery of the target cell (Bukrinsky et al., 1999). The lentiviral strategy for nuclear targeting enables infection of non-dividing cells, an attractive attribute for a gene therapy vector.

The lentivirus ensure stable long-term expression and have a high transfer efficiency, without causing immune responses (Quinonez et al., 2002). The use of lentiviruses in gene therapy is a widely used experimental approach. The attraction of this technique depends on the unique combination of a highly manipulable genome, with the capacity to carry portions of very large DNA (7-8 kb).

Adenoviruses (double-stranded DNA viruses), have achieved greater success (Einfeld et al., 2002) because they are able to infect rapidly different cell types and to produce high concentrations of the therapeutic protein (Silman et al., 2000). These viruses do not integrate into the genome of host cells, therefore do not damage the genetic information of the host; on the other hand, for this reason, this approach requires periodic treatments (Silman et al., 2000). This causes an immune response against the capsid proteins of the vector and limits the effectiveness of the treatment to short periods (Yang et al., 1996).

In conclusion, viruses are excellent carriers for the transport of therapeutic genes within the host cell as they are provided with natural mechanisms that allow it to bind to specific cell types and to transfer the genetic information with efficiency. Their use, however, involves a series of problems difficult to control. In fact, they can damage the DNA of target cells and, once inside the cell, the attenuated virus might regain its pathogenic activity.

2.2.2 Non viral gene transfer systems

Although viral vectors such as retrovirus, lentivirus and adenovirus are potentially efficient, non viral vectors have some advantages: they are less toxic, less immunogenic and easier to prepare. So far, several non viral gene transfer systems have been developed to deliver DNA directly including naked DNA injection, various chemical agents and physical methods. However, there are drawbacks with each of these non viral methods, including lower efficiency compared with viral vectors and transient gene expression. For efficient *in vivo* gene transfer, non viral vectors should overcome the delivery barriers existing in the process of their biodistribution, cellular uptake, and intracellular routing (Nishikawa et al., 2001). The simplest system for DNA delivery is the injection of naked plasmid DNA. The plasmids containing the therapeutic molecule penetrate into the cell nucleus, but a small fraction does not integrate into the genome and it remains in episomal form. This causes a limited expression of the transgene in both proliferating and non-proliferating cells. The low efficiency of the transgene transfection and its transient and low levels of expression limit this approach (Ogris, 2003). For systemic administration, the plasmid DNA needs to be protected from degradation before reaching to target cells. Therefore, when administered systemically, plasmid DNA requires a

delivery system, such as cationic liposomes or gene gun that protects it from *in vivo* degradation. Liposomes are chemical agents widely used in *in vitro* transfection to introduce DNA into the cells. These are artificial membrane vesicles that either fuse with the cell membrane and deliver their contents into the cell for expression (Matsui et al., 1997) or enter the cell through an endocytic pathway (Zabner et al., 1993; Friend et al., 1996). Liposome-mediated gene transfer offers several advantages over other chemical transfection systems. Among these are a relatively high efficiency of gene delivery, ability to transfect various cell types and successful delivery of a wide range of DNA (Colosimo et al., 2000). Disadvantages include the cytotoxicity of the liposome formulation and the need to optimize various parameters for each formulation, such as the DNA-to-liposome charge ratio, the amount of DNA, cell density and the transfection period (Gao et al., 1995). Other critical factors that must be optimized for effective DNA delivery are the size and homogeneity of lipid - DNA complexes (Colosimo et al., 2000). Cationic liposomes have been successfully used for the transfer of genes both *in vitro* and *in vivo* (Colosimo et al., 2000). They have already been tested in clinical trials for *in vivo* gene therapy, mostly of cancer and cystic fibrosis (Colosimo et al., 2000). However, these studies have been inconsistent in demonstrating long-term transgene expression and effective delivery of plasmid DNA. Regarding physical methods

for DNA delivery there is the biolistic particle bombardment (or gene gun) technique (Klein et al., 1993; Yang et al., 1995). The biolistic technology involves accelerated delivery of plasmid DNA (coated with high-density metal particles) into cells using one of several acceleration instruments. Gold particles are preferred to tungsten because of their size uniformity and spherical shape. Compared to other physical methods, like the electroporation, smaller amounts of DNA and fewer target cells are required for the DNA transfer. The major advantages of this gene delivery strategy include the relatively low level of cell damage and its potential application for *in vivo* transfection (Sanford et al., 1993; Zelein et al., 1993). However, *in vivo* gene gun application typically results in short-term and low-level expression of the gene product. Others drawbacks include poor tissue penetration, the manipulations required for microparticle preparation and the initial cost of the acceleration instrument. Another approach that has been used for high efficiency gene transfer is intranuclear microinjection (Brinster et al., 1985; De Pamphilis et al., 1988). This approach has been used to transfer genes into mammalian oocytes and preimplantation embryos (Shen et al., 1982; De Pamphilis et al., 1988), ES cells (Thomas et al., 1986; De Pamphilis et al., 1988), cultured fibroblasts (Daicumakos, 1973), airway epithelial cells (Kunzelmann et al., 1996; Goncz et al., 1999) and blood stem cells (Davis et al., 2000). While this approach is

tedious, it is highly efficient and directly delivers the exogenous DNA to the nucleus. It is possible to control both the site of DNA delivery and the amount of DNA that enters the nucleus. Thus, microinjection provides a mechanism for minimizing nuclease degradation of the extracellular DNA in intracellular vesicles and enhancing transfection efficiency. However, while this technique has the advantage of circumventing exposure to intracellular compartments that will degrade the DNA, it cannot be readily used to introduce DNA into large numbers of cells. It is therefore most useful for generating cell lines that carry transgenes of interest. The electroporation is one of the most widely and efficient method that can be used both *in vitro* and *in vivo* to transfer physiologically the gene into the cell without complicated preparations. This technique is based on the application of external electric fields that results in transient, reversible permeabilization of the cell membrane in localized areas. During the process, the membrane becomes extremely conductive and, as a consequence of the current passage inside the cell, the formation of pores is achieved, through which small molecules are transferred into the cytosol. Developments over the past decades have led to sophistication of equipment and optimization of protocols. Among the different successful innovations about electroporation, the most interesting one is the new Amaxa Nucleofector technology, able to introduce the genetic material directly into the nucleus. This

reduces assay time to a couple of hours versus 24 to 48 h for standard transfection methods that use liposome mediated delivery. Despite the transfer system results invasive, it is very efficient. Moreover, this system is not restricted to DNA only, but it can also be used to transfer siRNA (small interference RNA) into the cells, especially in primary lines (Brummelkamp et al., 2002). A disadvantage is the impossibility to standardize the protocol that must be adapted to each specific cell type used. The Nucleofector method was used to transfect different cell lines as embryonic stem cells (ES) (Lakshmipathy et al., 2004), T lymphocytes and dendritic cells (Lenz et al., 2003), demonstrating that the frequency of transfection changes depending on the cell type. For example, a transfection frequency higher than 50 % in nature killer (NK) cells (Trompeter et al., 2003; Maasho et al., 2004) and 60 % in dendritic (DC) cells (Lenz P et al., 2003) have been shown.

2.3 DNA Methylation

The epigenetics is defined as the study of the heritable modifications of specific genes or gene-associated proteins, not involving changes in the DNA sequence. Epigenetic modifications contribute to define how the information in genes is expressed and used by cells. An epigenetic mark should be heritable, self-

perpetuating and reversible. Although the topic is debated, epigenetic regulatory systems include DNA methylation, post-translational modification of histone proteins, nucleosome location and noncoding RNA (Woodcock, 2006). Methylation of cytosine residues in the DNA is a well-known epigenetic mechanism that controls many functions: it is associated with gene transcription modulation and silencing and was reported to be essential for embryonic development (Okano et al., 1999), genomic imprinting (Li et al., 1993), X-inactivation in mammals (Goto et al., 1998) and silencing of potential harmful DNA elements like transposons or endogenous retroviruses (Kass et al., 1997). In addition, aberrant DNA methylation has been linked to abnormal developmental processes and cancer formation (Plass et al., 2002). From the biochemical point of view, this modification, performed by the enzymes DNA methyltransferases (DNMTs), consists in the replacement of the hydrogen linked to carbon in position 5 of the cytosine with a methyl group, with formation of 5-methylcytosine. In mammals, this modification is mainly assigned to cytosines followed by a guanine (G), the so called CpG dinucleotides (CpGs). About 80 % of the CpG dinucleotides in the mammalian genome are methylated (Cheung et al., 2005), while the remaining 20 % are unmethylated and mainly located with high density ("CpG Island") in the promoter regions of constitutive or inducible genes. The methylation pattern is the result of the

combination of three processes: the maintenance methylation, the *de novo* methylation and the demethylation. The maintenance methylation, that occurs immediately after DNA replication, allows the heritability of the CpG dinucleotide modification. When the cytosine is methylated on one of the two strands of DNA also the symmetric base on the complementary strand is methylated. After the duplication of the nucleic acid a specific DNMT (DNMT1) recreates the methylation profile of the parental DNA in the neo-synthesized filament. The d*e novo* methylation, in which are involved DNMT3A and DNMT3B, acts during development to establish new methylation patterns. In the early phases of development the eukaryotic genome is demethylated; in the blastula stage a new methylation profile is established by a process called epigenetic reprogramming. A hypomethylation, as well as a remodulation of the methylation pattern, may originate from both the absence of methylating activity following the DNA replication (passive demethylation) and active mechanisms to remove methylated moieties (active demethylation) that do not require DNA replication. These mechanisms are to the basis of the developmental patterns in the early stages and also of pathological events. DNA methylation is generally associated to inhibition of gene expression, which can originate either directly through inhibition of transcription factors that bind DNA or indirectly influencing the structure of chromatin. In fact, the methylated

cytosine represents the binding site for several proteins that modify the chromatin conformation. For example, the methyl-CpG-binding protein 2 (MeCP2) and, in general, the methyl-CpG-binding domain proteins (MBDs) bind to methylated CpG and repress transcription by interacting with the enzyme histone deacetylase (HDAC) and other chromatin remodeling proteins. They modify histones, thereby forming compact, inactive chromatin, termed heterochromatin, that is less accessible to the transcriptional machinery with consequent repression of gene expression (Jaenisch et al., 2003). The link between DNA methylation and chromatin structure is relevant in both health and disease (Robertson et al., 2000). For example, the loss of MeCP2 has been implicated in Rett syndrome and methyl-CpG-binding domain protein 2 (MBD2) mediates the transcriptional silencing of hypermethylated genes in cancer.

A response of the DNA methylation machinery to gene therapy interventions appears likely. The DNA methylation also acts as a genome-defence system to silence expression and suppress expansion of parasitic DNA elements (such as retrotrasposons, LINE and SINE elements) limiting their spread through the genome. It protects genome integrity, stabilising genomes that contain large amounts of repetitive DNA by inhibiting HR between such repeats. Moreover, in human epithelial cells has been specifically demonstrated that micro and small

interfering RNAs are able to target endogenous and inserted genes, producing modification of DNA strands by DNA methylation Finally, epigenetic changes were observed during stem cells *in vitro* manipulation, in particular after gene targeting and in mouse embryonic fibroblasts (MEF), as demonstrated by our research group, after gene targeting by SFHR (Luchetti et al., 2012).

2.4 DNA Repair Mechanisms

During their evolution higher eukaryotic cells acquired a large genome and increasingly complex molecular machinery to preserve chromosome integrity. DNA repair enzymes continuously monitor chromosomes to correct replication errors and damaged nucleotide residues generated by exposure to carcinogens and cytotoxic compounds. The damage may be consequence of environmental agents such as ultraviolet (UV) light from the sun, inhaled cigarette smoke or incompletely defined dietary factors. However, a large proportion of DNA alterations are caused unavoidably by endogenous weak mutagens including water, reactive oxygen species, and metabolites that can act as alkylating agents. A very slow turnover of DNA consequently occurs even in cells that do not proliferate. Genome instability caused by the great variety of DNA-damaging agents would be an overwhelming problem for cells and organisms

without the intervention of DNA repair mechanisms. Maintaining the continuity and stability of each nuclear DNA molecule is fundamental for preventing chromosomal rearrangements that can lead, for example, to cancer through altered gene expression (Venkitaraman, 2009; Nagasawa et al., 2010). Unrepaired DNA double-strand breaks (DSBs) may also contribute to cell senescence (d'Adda di Fagagna, 2008) and diseases.

DNA-repair genes can be sub-grouped into genes associated with distinct signaling and regulation of DNA repair pathway, according with distinct repair mechanisms such as base excision repair (BER), nucleotide excision repair (NER), mismatch repair (MMR), direct damage reversal and DSB repair. There are at least two DSB repair pathways, the HR and the Non-Homologous End Joining (NHEJ), which are error-free and error-prone respectively. **BER** is responsible for removing DNA-damaged bases, which can be recognised by specific enzymes, the DNA glycosylases. The main lesions subjected to BER are oxidised DNA bases, arising spontaneously within the cell, during inflammatory responses, or from exposure to exogenous agents, including ionising radiation and long-wave UV light. Another main source of lesions repaired by BER is DNA alkylation induced by endogenous alkylating species and exogenous carcinogens such as nitrosamines. DNA glycosylases initiate this process by releasing the modified base. This is followed by cleavage of the sugar-

phosphate chain, excision of the abasic residue, and local DNA synthesis and ligation. **NER** mainly removes bulky adducts caused by environmental agents, such as UV-light-induced photolesions intrastrand cross-links, large chemical adducts generated from exposure to aflatoxine, benzo(a)pyrene and other genotoxic agents. In NER about 30 proteins are involved. In E. coli, the three polypeptides UvrA, UvrB, and UvrC can locate a lesion and incise on either side of it to remove a segment of nucleotides containing the damage. Eukaryotes, including yeast and human cells, do not have direct UvrABC homologs but use a more elaborate assembly of gene products to carry out NER (Lindahl et al., 1999). Cells defective in NER belong to different complementation groups and UV-hypersensitive disorders such as xeroderma pigmentosum (XP), Cockayne's syndrome (CS), trichothiodystrophy (TTD), UV-sensitive syndrome (UVSS) and a variety of UV-hypersensitive rodent lines, in which the defect can be complemented by human genes belonging to the excision repair cross-complementing group (ERCC) (Vermeulen et al., 1997). NER consists of two distinct pathways termed global genomic repair (GGR) and transcription-coupled repair (TCR). GGR is thought to be largely transcription-independent and removes lesions from the non-transcribed regions of the genome and from the non-transcribed strand of transcribed regions. TCR removes different RNA-polymerase-blocking lesions from the transcribed strand of active genes. The

MMR system is responsible for removal of base mismatches caused by spontaneous and induced base deamination, oxidation, methylation and replication errors (Modrich et al., 1996; Umar et al., 1996). The main targets of MMR are base mismatches such as G / T (arising from deamination of 5-methylcytosine), G / G, A / C and C / C (Fang et al., 1993). A few unusual enzymes reverse rather than excise DNA damage. In human cells, O6-alkylation lesions can be repaired in a single-step reaction by **O6-methylguanine–DNA methyltransferase (MGMT)**, which is the homologue of the Ada (or the constitutively expressed OGT) gene product of E. coli. The repair protein transfers the methyl or chloroethyl group from the alkylated guanosine in a one-step reaction onto an internal cystein residue in its active centre (Pegg et al., 1995). This alkyl group transfer leads to irreversible inactivation of the MGMT protein, and targets it for ubiquitination and proteasome-mediated degradation.

DSBs are highly potent inducers of genotoxic effects (chromosomal breaks and exchanges) and cell death (Dikomey et al., 1998; Pfeiffer et al., 2000; Lips et al., 2001). In higher eukaryotes a single non-repaired DSB inactivating an essential gene can be sufficient for inducing cell death via apoptosis (Rich et al., 2000). In simple eukaryotes like yeast, HR is the main pathway, whereas in mammals the NHEJ pathway predominates (Cromie et al., 2001; Haber, 2000). In fact, in

mammals the frequency of HR mechanism is low and generally occurring at one event per 10^5 to 10^7 treated cells while the random integration occurs in about one per 10^2 to 10^4 treated cells (Capecchi, 1989; Capecchi, 1994). The use of NHEJ rather than HR also depends on the phase of the cell cycle. NHEJ occurs mainly in G0 / G1, whereas HR occurs during the late S and G2 phases (Johnson, 2000; Takata et al., 1998).

Repair via HR uses homologous DNA sequence as a template for repair synthesis, whereas NHEJ often simply ligates the broken DSB ends. If the ends are not perfectly matched, resection of the broken ends usually occurs before the ligation. The first step in **NHEJ** is the binding of a heterodimeric complex consisting of the proteins Ku70 (alias XRCC6; Reeves et al., 1989) and Ku80 (alias XRCC5; Jeggo et al., 1992) to the damaged DNA, thus protecting the DNA from exonuclease digestion. Following DNA binding, the Ku heterodimer associates with the catalytic subunit of DNA–PK (XRCC7, DNA–PKcs) thereby forming the active DNA–PK holoenzyme. DNA–PKcs is activated by interaction with a single-strand DNA at the site of DSB (Hammarsten et al., 2000; Martensson et al., 2002). One of the targets of DNA–PKcs is XRCC4, which forms a stable complex with DNA ligase IV. The XRCC4–ligase IV complex binds to the ends of DNA molecules and links together duplex DNA molecules with complementary but non-ligatable ends (Lee et al., 2003).

During **HR**, the damaged chromosome enters into physical contact with an undamaged DNA molecule with which it shares sequence homology and which is used as template for repair (Sonoda et al., 2001). HR is initiated by a nucleolytic resection of the DSB in the 5′–>3′ direction by the MRE11–Rad50–NBS1 complex (Fig. 1). The resulting 3′ single-stranded DNA is thereafter bound by a heptameric ring complex formed by Rad52 proteins (Stasiak et al., 2000), which protects against exonucleolytic digestion. Rad52 competes with the Ku complex for the binding to DNA ends. This may determine whether the DSB is repaired via the HR or the NHEJ pathway (Van Dyck et al., 1999). Rad52 interacts with Rad51 and RPA, stimulating DNA strand exchange activity of Rad51 (New et al., 1998). The human Rad51 protein is the homologue of the E. coli recombinase RecA. It forms nucleofilaments, binds single- and double-stranded DNA and promotes ATP-dependent and RPA-stimulated interaction with a homologous region on an undamaged DNA molecule. Thereafter Rad51 catalyzes strand-exchange events with the complementary strand in which the damaged DNA molecule invades the undamaged DNA duplex, displacing one strand as D-loop. The assembly of the Rad51 nucleoprotein filament is facilitated by five different paralogues of Rad51 (Rad51B, C and D; and XRCC2 and XRCC3) that could play a role during pre-synapsis. Another important protein that interacts with Rad 51 is RPA. It is supposed that RPA stabilizes

Rad51-mediated DNA pairing by binding to the displaced DNA strand (Eggler et al., 2002). After DSB recognition and strand exchange performed by Rad proteins, the resulting structures are resolved according to the classical model of Holliday (Holliday, 1964; Constantinou et al., 2001).

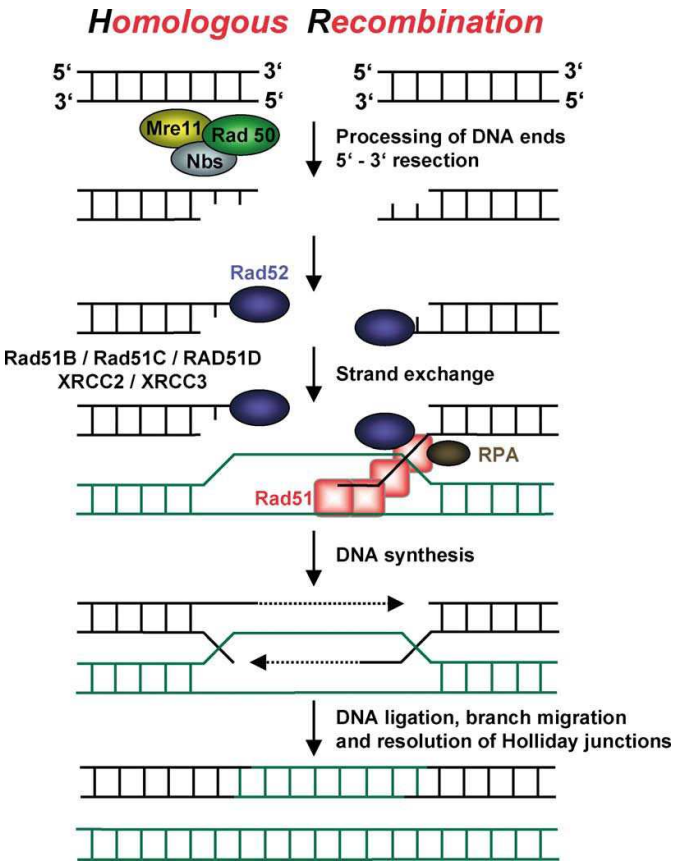


Figure 1 - **Mechanism of HR.** HR starts with nucleolytic resection of the DSB in the 5′ → 3′ direction by the MRE11–Rad50–NBS1 complex, forming a 3′ single-stranded DNA fragment to which Rad52 binds. Rad52 interacts with Rad51, provoking a DNA strand exchange with the undamaged, homologous DNA molecule. Assembly of the Rad51 nucleoprotein filament is facilitated by different Rad51 paralogues (such as Rad51B, Rad51C and Rad51D, XRCC2 and XRCC3). After DNA synthesis, ligation and branch migration, the resulting structure is resolved (Christmann et al., 2003) .

To efficiently repair DNA damages, cells have to activate several interacting and functionally cooperating pathways, including DNA repair and cell cycle checkpoints.

2.5 Cell cycle control, checkpoints and the interconnection with DNA repair

During the cell cycle, an ordered sequence of events allows a cell to divide into 2 daughter cells. For this to happen, DNA has to be duplicated. A complex regulatory network controls DNA duplication and cell cycle on multiple levels (Warmerdam et al., 2010). The so-called cell cycle checkpoints ensure a correct progression into the next cell cycle phase, by regulating Cyclin - CDK (Cyclin - dependent kinase) activity. In this pathway, slow processes (within hours) control gene expression, whereas fast responses (within minutes) depend upon direct protein - protein interactions and post-translational modifications. A main controlling mechanism is the phosphorylation of CDK, which reduces the activity of the Cyclin - CDK complex and, in turn, temporarily halts the cell cycle progression (Fig. 2). During the G1 phase, the Cyclin D – CDK4 / 6 complex is responsible for cell cycle progression. Then, the Cyclin E – CDK2 complex come up, leading to further progression into S-phase and duplication of the DNA.

Cyclin A – CDK2 levels rise during S and G2 phase and, finally, Cyclin A / B – CDK1 drives cells into mitosis. The G1 / S and G2 / M borders, together with intra-S-phase, are the 3 main checkpoints which preserve correct progression.

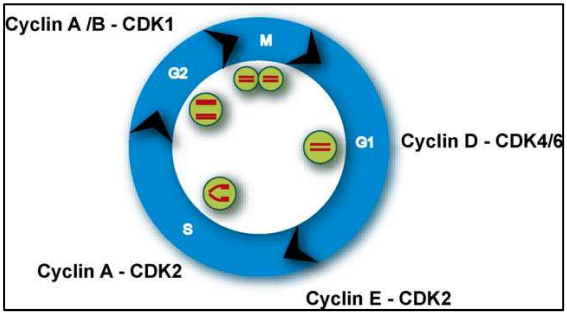


Figure 2 - Cyclin–CDK regulation throughout the different phases of the cell cycle. (Warmerdam et al., 2010).

Throughout the cell cycle the DNA is constantly monitored for damages possibly caused by errors during replication, products of metabolism, toxic drugs or ionizing radiations. As previously shown, genotoxic insult results in the activation of an interwoven network of DNA damage sensors and DNA repair pathways (Warmerdam et al., 2010). Due to the complex chemical nature of DNA, the different kinds of DNA lesions that can occur is vast. Repair of such

a huge set of possible DNA alterations, requires proportionally complex and different systems of damage recognition. The processing of DNA lesions by structure -specific nucleases in DNA - protein intermediates, forms the basis for checkpoint activation and DNA repair. However, checkpoints are less diverse. They do not directly act on the DNA lesion, but respond to the common DNA – protein complex built up at a lesion by the specific DNA repair pathway. In turn, cell cycle - dependent regulation additionally plays a key role in the regulation of both DNA repair and checkpoint activation. An initial description of events may be done from the point of view of the kind of DNA damage recognition. For example, in response to double-strand breaks, ATM activation is necessary for cell cycle checkpoint activation whereas, after single-strand DNA gaps, ATR signaling is initiated. However, as the DNA repair events take place in the cellular context, for a more realistic description of the processing of DNA lesions and cell cycle progression an integrated vision of DNA repair pathways, checkpoints and cell cycle phases have to be considered. In the G1 phase of the cell cycle only one copy of each chromosome is present: HR can therefore not be efficient. In effect, the major DNA repair pathway for DSBs during G1, activating G1 / S checkpoint, is NHEJ (Fig. 3). Since many cells in the human body are post-replicative and thus non-cycling, error-free DNA repair is often not essential for cell viability. In effect, core proteins involved in HR, like Rad51

and BRCA2 do not form DSB-induced foci in G1, suggesting HR is not active during this phase. During the S-phase (when sister chromatids originate) the response to DNA damage is quite different, with a predominant DSBs repair via HR (Fig. 3).

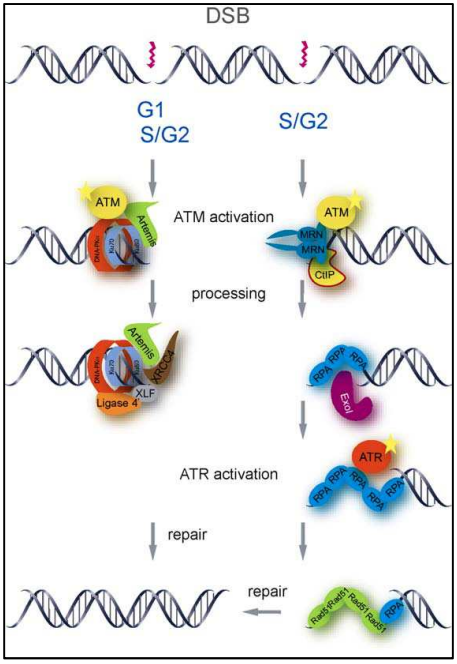


Figure 3 - Cell cycle-dependent processing of DSBs and subsequent checkpoint activation. DSBs generated in G1 and S / G2 can be repaired by NHEJ. The Ku70/80 complex binds the broken DNA ends first. ATM becomes activated, inducing a checkpoint response. During S and G2 phase, ATM becomes activated via the MRN complex eliciting an early checkpoint response. DSBs are processed first by the MRN complex and CtIP, creating a small stretch of single strand DNA (ssDNA) bound by RPA. Secondary processing is performed by ExoI (or Sgs1 / Dna2), creating longer stretches of RPA coated ssDNA. This initiates an additional checkpoint response through ATR. Finally RPA is exchanged for Rad51 that initiates repair through HR. (Warmerdam et al., 2010).

In this case, after DNA damage, CDK - dependent phosphorylation of BRCA2 goes down immediately, stimulating the interaction between BRCA2 and Rad51

and repair via HR. Finally, before the cell goes into mitosis it must go through the G2 / M checkpoint, which shows its own distinct set of effectors. When cells enter mitosis with DSBs, this can lead to gross chromosomal rearrangements and an uneven separation of chromosomes between the two daughter cells. Therefore the G2 / M checkpoint must work effectively to turn off the cell cycle progression.

It is not easy and, probably, not useful a functional separation between sensors and effectors of DNA repair and cell cycle regulators. Cell cycle - dependencies influence the DNA damage response extensively. On the other hand several protein are multifunctional with direct involvement in both pathways, as for example: Claspin, Atm, Atr, Sens2, Rad9, Ppm1d, Brca1 and Brca2. This makes the situation more complicated, also considering that DNA damage repair and checkpoint pathways are not static protein cascades but dynamic and highly regulated mechanisms. It is advisable to study the cell cycle - dependent regulation of checkpoint and repair pathways onto specific types of DNA damage, using and developing different and better ways of inducing DNA damage.

3. MATERIALS AND METHODS

3.1 Experimental system set-up (before the start of this thesis)

The experimental system used in this thesis was carried out by the group of the "Department of Biopathology and Diagnostic Imaging", University of Rome Tor Vergata, with which my research group collaborates. For clarity, in this preliminary section 3.1 are described the set-up phases of the experimental system (Luchetti et al. 2012; see also Results section 4.1), performed before I started the experiments of my thesis.

Cells. MEF were isolated from mice Knock-out SMA1 generated by prof. Arthur Burghes (Ohio State University) and subsequently immortalized with the SV40 virus. After the infection, the cells were cultured in DMEM (Euroclone, Milan, Italy) supplemented with 10 % FBS (Euroclone, Milan, Italy), 1 % L-glutamine (Euroclone, Milan, Italy), 1 % Non Essential Amino Acids (NEAA) (Euroclone, Milan, Italy), 0.01 mM of 2-Mercaptoethanol (Gibco, Life Technologies, Foster City, CA, USA), 20 mM Hepes (Euroclone, Milan, Italy) and incubated at 37° C under 5 % CO_2. 48 h after infection cells have been trypsinized and plated adding 400 μg of G418 (Euroclone, Milan, Italy) to select the infected cell.

Construction of the stably integrated mutant and wild type EGFP cells line. The wild type enhanced green fluorescent protein gene (EGFP) was obtained from the pEGFP-N1 vector (Clonetech Lab. Inc., USA) by *Xho I* and *Hind III* restriction (New England Biolabs, Ipswich, MA, USA), and cloned in vector pCR-2.1 (Invitrogen, CA, USA). The wild type sequence of the gene was mutated by QuikChange Site-Directed Mutagenesis Kit (Agilent Technologies Inc, Santa Clara, CA, USA) using two specific mutagenesis primers that insert a nonsense mutation creating a stop codon at codon 70 (**C**AG > **T**AG) of EGFP protein. This causes the loss of the restriction site for the enzyme *Bts I*, unique within the gene, and allows the screening of the corrected clones. The mutated and the wild-type gene were extracted from pCR-2.1 vector by restriction with *Xho I* and *Hind III* and cloned inside the pCEP4 vector (Invitrogen, CA, USA), between pCMV promoter and SV40-pA, using the same restriction enzymes. To create stable cell clones, the MEF cells were transfected by electroporation and were plated in fresh medium containing 200 μg / ml of hygromycin (Sigma-Aldrich, Milan, Italy) for stable vector integration selection. The cells were cultured for two weeks in selective medium using 500 μg / ml of hygromycin. After selection, several single cell clones were isolated by serial dilution in 96 well plate and screened by PCR and FISH analyses to check genomic plasmid integration. By TaqMan qPCR the pCEP4 / EGFP copy number was determined

for each clone isolated. Among selected clones, D1 and C2 clones were chosen for subsequent experiments as integrating, respectively, 1 copy and 13 copies of the mutated enhanced green fluorescent protein (mutEGFP) gene sequence; the C1 clone was chosen as control, because containing wild type enhanced green fluorescent protein (wtEGFP) gene. After transfection, to confirm the genomic modification and its persistence over time several molecular analysis were performed in D1 corrected cells such as Southern Blotting to exclude any detectable random integration of SDF within genomic DNA, RFLP analysis to check whether the restriction site for *Bts I* enzyme was recovered as result of successful SDF replacement and, finally, direct sequencing to confirm the presence of wt nucleotide (cytosine) in D1 corrected cells.

In order to enhance the efficiency of gene modification and to optimize the experimental system, different parameters were tested in mouse embryonic fibroblast with integrated the mutated EGFP (MEF-mutEGFP) gene, such as the amount and the nature of the SDF to transfect, the different phases of the cell cycle and the conditions of transfection.

Nature of SDF. Three different experimental protocols for SDFs synthesis was evaluated. A 876 bp double strand DNA (dsDNA) SDF homologous to EGFP wild type sequence (named SDF-PCR-WT) was obtained by amplification of the region cloned in pCR-2.1 vector (Invitrogen, CA, USA), using HYGRO 1F (5'-

ACTCATCAATGTATCTTATCAT-3') and HYGRO 1R primers (5'-AGGTCTATATAAGCAGAGCT-3'). The PCR product was purified from 1 % agarose gel by QIAquick Gel Extracion Kit (Qiagen, Manchester, U. K.). Another kind of dsDNA SDF (752 bp), homologous to EGFP wild type sequence, was obtained by *Hind III* and *Xho I* restriction of the pCR-2.1 vector (named SDF-DIG-WT) (Fig. 6 A). The sequence of the SDFs was checked by DNA sequencing. The single strand SDF-PCR-WT was obtained by heat denaturation, incubating 10 min at 100° C and soon after placed on ice. All fragments were dosed by spectrophotometer (ND-1000, Nanodrop, USA).

Amount of SDF. 1.7 x 10^6 unsynchronized cells were transfected with increasing amounts of SDF ranging from 5 μg (6 x 10^6 SDF / cell) up to 30 μg (18 x 10^6 SDF / cell). Targeted correction rates was evaluated three days later by flow cytometry and expressed as the ratio between the number of living cells EGFP positive and the number of cells analyzed for each experiment (approximately 300000).

Cell Synchronization. To investigate SFHR-mediated gene repair at various phases of the cell cycle, different concentrations of mimosine, thymidine and vinblastine were tested to synchronize cells in G0 / G1, S or G2 / M phase, respectively. For the mimosine (Sigma-Aldrich, Milan, Italy) cells were grown for 12 h at a concentration ranging from 250 μM to 750 μM. For the thymidine

(Sigma-Aldrich, Milan, Italy) cells were grown for 15 h at a concentration ranging from 0.5 mM to 4 mM. Synchronization in G2 / M phase was obtained growing cells for 14 h at a concentration ranging from 25 nM to 200 nM of vinblastine (Sigma-Aldrich, Milan, Italy). Once 60 % of confluence was reached, cells were treated, washed in PBS, fixed in 70 % ethanol, stained with 0.05 mg / ml propidium iodide and then analyzed by FACS-Calibur Flow Cytometer (Becton Dickinson) to determine DNA content. Sync-Wizard Model was used to model the cell cycle.

Nucleofection parameters. The Amaxa Nucleofection System (Lonza, Cologne, Germany) was used to perform electroporation. Appropriate Nucleofection program was evaluated in order to have an optimal transfection efficiency and cell viability. Two different electroporation programs A-23 and T-20 and two different transfection solutions MEF-1 and MEF-2 was tested using a 21 bp fluorescent oligonucleotide. The combination of program T-20 and solution MEF-2 was chosen.

3.2 Cells and culture conditions

From here and below, the experiments I performed for this thesis are described. D1 and C2 MEF-mutEGFP clones were chosen to perform all my

experiments and C1 mouse embryonic fibroblast with integrated the wild type EGFP (MEF-wtEGFP) clone was used as a control. The cells were cultured in DMEM (Euroclone, Milan, Italy) with 10 % FBS (Euroclone, Milan, Italy), 1 % L-glutammine (Euroclone, Milan, Italy), 1 % Penicillin / Streptomycin (Euroclone, Milan, Italy), 1 % Non-Essential AmminoAcids (Euroclone, Milan, Italy), 20mM Hepes (Euroclone, Milan, Italy), 0.01 mM 2-β-Mercaptoethanol (Gibco, Life Technologies, Foster City, CA, USA), 120 ng/μL of G418 ((Euroclone, Milan, Italy) and 200 ng/μL of Hygromicin (Sigma-Aldrich, Milan, Italy), according to the protocol previously optimized (Luchetti et al., 2012), and incubated at 37° C under 5 % CO_2.

3.3 Methylation analyses of SDF

The methylation patterns of SDF-DIG-WT and SDF-PCR-WT (obtained as described above) were tested by treating the targets with specific methylation-sensitive restriction endonucleases (*Hpa II* or *Msp I*, *PspG I* or *BstN I*, *Mbo I* and *Sau3A I*) able to recognize eukaryotic, Dcm or Dam methylation. PCR amplification with flanking primers was performed after restriction. Specifically, 300 ng of DNA were digested at 37 °C with 3 units of enzyme for 12 h in a total volume of 20 μl. PCR was performed in a volume of 15 μl containing:

6 pmol of each primer, 6 ng of SDF DNA, 175 µM dNTPs, 0.5 unit of Yieldace DNA Polymerase (Agilent Technologies), and 1X Yieldace reaction buffer. The reaction was a multiplex PCR of both the specific target and an internal standard (Table 1). PCR cycle was 2' at 92 °C, 45'' at 94°C, 1'30'' at specific Ta (Table 1) and 2'30" at 72 °C for 20 cycles.

Table 1 - Primers used to perform methylation analysis of SDF. See Fig. 7 for the scheme of amplified fragments.

Name	Sequence 5'-3'	T_a	Amplicon	Fragment
Hygro1f	5'-ACTCATCAATGTATCTTATCAT-3'	55°C	876 bp	SDF-PCR-WT
Hygro1r	5'- AGGTCTATATAAGCAGAGCT-3'			
Pr GFP F-2f	5'-AAGGGCGAGGAGCTGTTCA-3'	62°C	505 bp	b
Pr GFP R- 2r	5'-GTTGTGGCGGATCTTGAAGT-3'			
GFP St1- 3f	5'-ACCACATGAAGCAGCACGAC-3'	62°C	246 bp	St1
GFP St1- 3r	5'-TCTGCTTGTCGGCCATGATAT-3'			
GFP St2- 4f	5'-ATCATGGCCGACAAGCAGAA-3'	62°C	251 bp	St2
GFP St2- 4r	5'-TCGTCCATGCCGAGAGTGAT-3'			

3.4 Production and i*n vitro* methylation of SDFs

To produce the SDF-PCR-WT, homologous to EGFP wild type sequence, an amplification of the region cloned in pCR-2.1 vector (Invitrogen, CA, USA) was performed following the protocol previously optimized (Luchetti et al., 2012).
The SDF-PCR-WT was used as target for *in vitro* methylation. For *Dam* and *Sss I* methylation, 1 µg of SDF-PCR-WT was incubated with 2 units of corresponding methyltransferase (New England Biolabs, Ipswich, MA, USA) in a final reaction volume of 20 µl. To create a fragment with both *Dam* and *Sss I* methylation patterns, a second *Sss I* methylation step on previously *Dam* methylated samples was performed. The extent of methylation was checked by incubating overnight treated samples with respective methylation-sensitive restriction endonucleases followed by PCR amplification.

3.5 General protocol of cell transfection, growth evaluation and FACS analysis

Electroporation was carried out using the Amaxa Nucleofection System (Lonza, Cologne, Germany). 1.7 x 10^6 cells of specific clones were treated with different amount of SDF, depending on the experimental plan. In particular: 3 x 10^6 SDFs

(corresponding to 5 µg SDF) / cell or 12 x 10^6 SDFs (corresponding to 20 µg SDF) / cell of C2 clone were used in real time expression array (qRT-PCR array) experiments to test, respectively, the effect of low and high SDF dosage; 12 x 10^6 SDFs (corresponding to 20 µg SDF) / cell of D1 clone were used in all other experiments. After transfection, cells were suspended in 100 µl of supplemented Nucleofection Buffer MEF-2. At the same time, 1.7 x 10^6 C2 or D1 clone cells (depending on the experiment) were treated only with the Nucleofection Buffer MEF-2, without adding SDF. Nucleofection program T-20 was used to transfect the cells. Then the cells was incubated at 37° C under 5 % CO_2. C2 or D1 clone cells (depending on the experiment) untrasfected control were also plated. Every experimental condition was tested in triplicate (3 independent biological samples from 3 independent experiments). The overall experimental times used were: 8 h, 24 h and 72 h after transfection (or plating for untransfected control); depending on the experiment, all or only some experimental times were included.

The cells were trypsinized and the cellular growth after transfection was evaluated for every experimental condition: 10 µl of cell suspension were added to 10 µl of 2X Trypan blue to exclude from the analysis dead cells. Ten µl of this mix were placed in a glass slide to perform cell counts. The cellular growth is expressed as the ratio between the number of total living cells and

the number of cells that initially underwent transfection. The remaining trypsinized cells were pelleted and used to perform DNA and RNA extraction. Cells were FACS analyzed using nucleic acid dye Topro-3 (0.1 mM; Invitrogen, CA, USA) to exclude dead cells. Data from 300000 living cells were analyzed by the BD-ARIA-DIVA software, to obtain the percentage of EGFP positive cells. To gate EGFP positive cells, parental wt C1 clone was used.

3.6 KU-55933, 1,5-Isoquinolinediol, α-Amanitin and 5-Aza-2'-Deoxycytidine treatments

Cells were treated by three different inhibitors of specific proteins. KU-55933 (Tocris, Bristol, U.K.), a potent, selective and competitive ATM kinase inhibitor, was used at 10 μM 1 h prior transfection. 1,5-Isoquinolinediol (Sigma- Aldrich, Milan, Italy), an inhibitor of Poly-(ADP-ribose) synthetase-1, was used soon after transfection at 0.622 mM for 24 h. α-Amanitin (Sigma-Aldrich, Milan, Italy), an inhibitor of eukaryotic RNA polymerase II and III, was used 24 h after transfection at 1 mM for 24 h. Treated cells, previously synchronized in G2 / M phase with 25 nM vinblastine (Sigma-Aldrich, Milan, Italy), were then transfected. To verify a relationship between methylation status of EGFP locus and time-dependent EGFP expression one day after transfection, cells were

incubated with 0.5 μM 5-Aza-2'-Deoxycytidine (5-Aza-dC) (Sigma-Aldrich, Milan, Italy) for 48 h and then FACS analyzed. Every experiment was performed in triplicate as previously described.

3.7 Fluorescence microscopy, RNA extraction and real time PCR for EGFP gene expression analysis

C1 and D1 fluorescent cells (corrected, after sorting) were plated in a 100 mm^2 plate 24 h prior to 0.5 μM 5-Aza-dC (Sigma-Aldrich, Milan, Italy) treatment and incubated for 48 h. After treatment, cells were analyzed for EGFP expression. Total RNA was extracted according to Trizol protocol (Invitrogen, CA, USA) and dosed by spectrophotometer (ND-1000, Nanodrop, USA) to determine the concentration and the A260 / A280 ratio. The quality of the RNA extracted was evaluated with electrophoresis in agarose denaturing gel that allows to distinguish the bands related to the portion 28S, 18S and 5S. DNase treatment was carried out to remove residual amounts of contaminating genomic DNA: 6 μg of total RNA was incubated with 2.2 units of *DNase I* (New England Biolabs, Ipswich, MA, USA) at 37°C for 10' in a final volume of 20 μl. Subsequently, the sample was treated with 1X EDTA (50 mM, pH = 8) at 75° C for 10', to deactivate the *DNase I*. After digestion, 1.5 μg of RNA were reverse transcribed

to cDNA according to High-Capacity cDNA Archive kit protocol (Applied Biosystems, Foster City, CA). Real time PCR was performed using a 7500 Fast Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) using the primers reported in Table 2. Bright field and fluorescence microscopy for EGFP expression detection were performed by common procedures.

Table 2 - Primers used to perform Real-Time PCR for expression study of EGFP gene in C1 and D1 positive cells.

Name	Sequence 5'-3'	T_a	Amplicon
EGFP f	5'-CTGCTGCCCGACAACCA-3'	60° C	74 bp
EGFP r	5'-ATGTGATCGCGCTTCTCGTT-3'		
EGFP probe	6FAM-CCAGTCCGCCCTGAGCAAAG-TAMRA		

A commercially available endogenous gene, Glyceraldehyde 3-phosphate dehydrogenase gene (GAPDH: Mm99999915g1 Applied Biosystems) was used as reference for the TaqMan assay. All PCR reactions were performed in triplicate.

3.8 DNA extraction and DNA methylation analysis of EGFP locus

DNA extraction was performed with the QIAamp DNA Mini Kit (Qiagen, Manchester, U.K.). In order to characterize the relationship between methylation status of EGFP locus and time-dependent EGFP expression, studies by multiplex *Hpa II* / PCR or *Aci I* / PCR were performed on genomic DNA of C1 wt and D1 corrected clones, treated by *Hpa II* or *Aci I* methylation-sensitive restriction endonucleases (New England Biolabs, Ipswich, MA, USA). The methylation status of different sites of the locus were analyzed in 3 regions (see also Fig. 10 A): the upstream / SFHR target region (indicated as *c* amplicon), the SFHR target / downstream region (indicated as *d* amplicon) and the downstream region (indicated as *e* amplicon). Two regions which possesses no *Hpa II* or *Aci I* recognition sites, were used as internal standards for *Hpa II* / PCR and *Aci I* / PCR (indicated respectively as St3 and St4 amplicon). Primer pairs were designed (Table 2) with at least one primer located outside the SFHR targeted region avoiding the amplification of non integrated SDF.

Table 3 - Primers used to perform methylation studies on genomic DNA of C1 and D1 clones. See Fig. 10 for the scheme of amplified fragments.

Name	Sequence 5'-3'	T_a	Amplicon	Fragment
Hygro5f	5'-CGTAAGTTATGTAACGCGGAACTC-3'	64°C	167 bp	St3
Hygro5r	5'-GGCCATTGCATACGTTGTATC-3'			
Hygro6f	5'-GTAGGTCAGGGTGGTCACG-3'	64°C	699 bp	c
Hygro6r	5'-GTAACGCCAATAGGGACTTTCC-3'			
Hygro7f	5'-AACTTGTTTATTGCAGCTTATAATGG-3'	64°C	383 bp	d
Hygro7r	5'-AGATCCGCCACAACATCG-3'			
Hygro8f	5'-AAAACCTCTACAAATGTGGTATGG-3'	64°C	503 bp	e
Hygro8r	5'-CATAAAGGCAATGTTGTGTTGC-3'			
Hygro8f1	5'-GGGGACCAAACACAAAGG-3'	64°C	138 bp	St4
Hygro8r	5'-CATAAAGGCAATGTTGTGTTGC-3'			

300 ng of genomic DNA were digested at 37° C with 3 units of each enzyme for 12 h in a final volume of 20 µl. Only after the treatment with methylation-sensitive restriction endonucleases, that cut the target genomic DNA only if

unmethylated, every single region was amplified together with the internal standard in a multiplex touchdown PCR. The touchdown PCR cycle was performed as follow using a PTC100 Thermal Cycler (Bio-Rad): 2' of at 92° C, 40 cycles (45" at 94° C, 1'30" at 64° C - 0.2° C per cycle, and 4' at 72° C) and a final extension of 7' at 72° C. Gel electrophoresis run was scanned by a CCD camera (VisiDoc-It, UVP) and acquired on the VisionWorks LS software version 6.7.3 (UVP) for densitometry. For a semi-quantitative evaluation of methylation patterns of integrated EGFP construct, the densitometric value of each target amplicon was normalized using the value of the corresponding control amplicon. The final result is the percentage of methylation of the examined region in respect to the corresponding uncut control. All PCR reactions and densitometric analyses were performed at least in triplicate.

3.9 RNA extraction and retrotranscription for qRT-PCR arrays

Total RNA was extracted from transfected C2 cells according to RNeasy mini kit protocol (Qiagen, Manchester, U.K.). The quantification, the purity and the DNase treatment were performed as described above.

The reverse transcription was performed with the RT^2 First Strand kit (Qiagen, Manchester, U.K.): 2.4 µg of RNA (300 ng / µl) was further treated with a Buffer

GE (5X) and incubated at 42° C for 5' to perform a further step of DNase treatment. According to Qiagen protocol, the reverse-transcription mix was prepared in a final volume of 10 µl and subsequently added to 10 µl of each preparation of RNA treated with genomic DNA elimination mix. The mix was incubated at 42° C for 15', then immediately stopped by incubating at 95° C for 5'. The reaction was placed on ice for 5' then was added with 91 µl of RNase free water to proceed with the qRT-PCR array protocol.

3.10 Quantitative expression study by qRT-PCR arrays of both 84 DNA repair genes and 84 cell cycle genes

The quantitative analysis of the expression of 84 genes involved in different pathways of DNA repair (Fig. 4) and cell cycle (Fig. 5) was conducted using the RT^2 Profiler PCR Array (SAbioscience, Qiagen, Manchesters, U.K.). RT^2 Profiler PCR Arrays contain primer assays for 84 pathway - or disease-focused genes and 5 housekeeping genes. In addition, one well contains a genomic DNA control, 3 wells contain reverse-transcription controls, and 3 wells contain positive PCR controls. Assays for 5 housekeeping genes included in the arrays enable normalization of data. The genomic DNA control (GDC) is an assay that specifically detects genomic DNA contamination with a high level of sensitivity.

The reverse - transcription control (RTC) is an assay that tests the efficiency of the reverse - transcription reaction performed with the RT^2 First Strand Kit by detecting template synthesized from the kit's built-in external RNA control. The positive PCR control (PPC) consists of a predispensed artificial DNA sequence and the assay that detects it, for the test of PCR efficiency. Controls provided in replicates can be used to test for inter-well, intra-plate consistency.

Real-time PCR was performed using RT^2 Profiler PCR Arrays in combination with RT^2 SYBR Green Mastermixes. The PCR array mastermix was composed by 1350 µl of 2X RT^2 SYBR Green Mastermix, 102 µl of cDNA synthesis reaction and 1248 µl of RNase free water in a final volume of 2700 µl. Twenty-five µl of the PCR array mastermix were dispensed into each well of the RT^2 Profiler PCR Array using an 8-channel pipettor. The plate was run on Applied Biosystems 7500 Fast real-Time PCR System (Applied Biosystems, Foster city, CA, USA) according to the cycle: 95° C for 10' and 40 cycles at 95° C for 15", 60° C for 1'. As control of the primers specificity, the derivative-melting curves of PCR products generated at the end of the amplification cycle were used. All PCR products were always specific.

Once acquired the threshold cycles (C_T) of the individual genes, the analysis was performed using the Excel analysis template developed by Qiagen. ΔC_T were calculated using the difference between the C_T values of the individual

genes of interest (GOI) and the average C_T of housekeeping genes (HKG). All the HKG genes analyzed remain stable in both the control and the experiment (with a difference <1.5 C_T); for this reason the average C_T of housekeeping genes was calculated considering all values (with no need of HKG exclusion). The $\Delta\Delta C_T$ was calculated by the difference between the ΔC_T of the gene in the experiment and the ΔC_T of the gene in the control. Fold-Change was calculated as $2^{(-\Delta\Delta CT)}$. Average values were calculated using the results of three independent experiments.

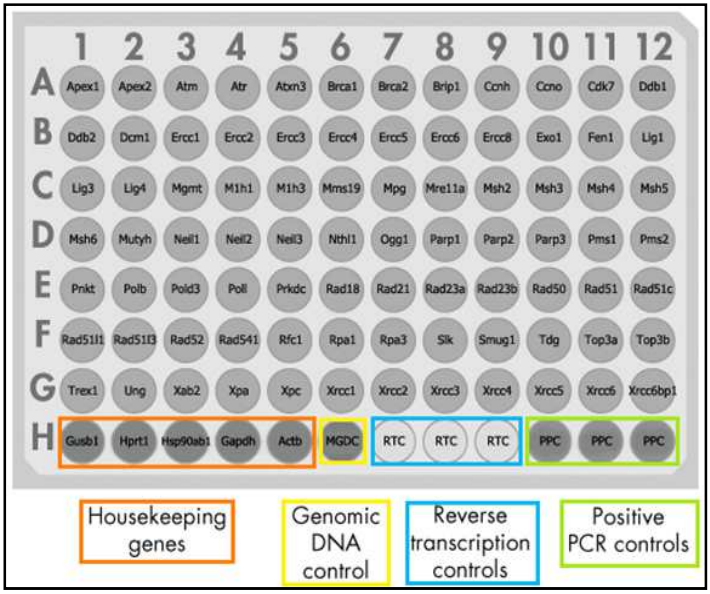


Figure 4 - DNA repair RT² Profiler PCR Array format C layout. (See text for explanations).
BER: Apex1, Apex2, Ccno, Lig3, Mpg, Mutyh, Neil1, Neil2, Neil3, Nthl1, Ogg1, Parp1, Parp2, Parp3, Polb, Smug1, Tdg, Ung, Xrcc1 (19 genes).
NER: Atr, Atxn3, Brip1, Ccnh, Cdk7, Ddb1, Ddb2, Ercc1, Ercc2, Ercc3, Ercc4, Ercc5, Ercc6, Ercc8, Lig1, Mms19, Pnkp, Poll, Rad23a, Rad23b, Rfc1, Rpa1, Rpa3, Slk, Xab2, Xpa, Xpc, (27 genes).
MMR: Exo1, Mlh1, Mlh3, Msh2, Msh3, Msh4, Msh5, Msh6, Pms1, Pms2, Pold3, Trex1 (12 genes).
HR: Atm, Brca1, Brca2, Dmc1, Mre11a, Rad18, Rad21, Rad50, Rad51, Rad51c, Rad51l1, Rad51l3, Rad52, Rad54l, Xrcc2, Xrcc3, Top3a, Top3b (18 genes).
NHEJ: Fen1, Lig4, Prkdc, Xrcc4, Xrcc5, Xrcc6, Xrcc6bp1 (7 genes).
Other gene involved in DNA repair: Mgmt (1 gene).

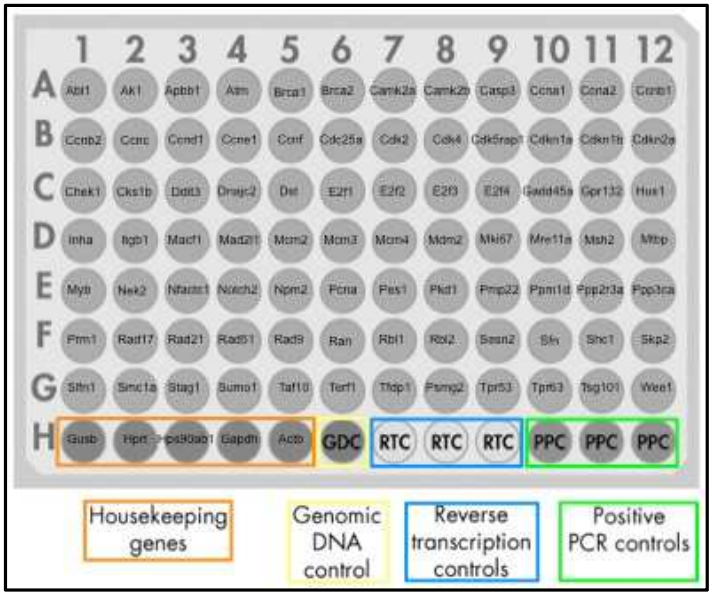


Figure 5 - Cell cycle RT² PCR Array format C layout. (See text for explanation).
Checkpoint and Arrest: Ak1, Apbb1, Brca2, Casp3, Cdk5rap1, Cdkn1a, Cdkn1b, Cdkn2a, Chek1, Cks1b, Ddit3, Dst, Gadd45a, Hus1, Inha, Macf1, Mad2l1, Mdm2, Msh2, Notch2, Pkd1, Pmp22, Ppm1d, Rad9, Sesn2, Sfn, Slfn1, Smc1a, Tsg101 (29 genes).
Positive Regulation: Abl1, Ccna1, Ccna2, Ccnb, Ccnb2, Ccnc, Ccnd1, Ccne1, Ccnf, Cdk4, E2f1, E2f2, E2f3, E2f4, Psmg2, Ran, Shc1, Skp2, Tfdp1 (19 genes).
Negative Regulation: Atm, Brca1, Itgb1, Rbl1, Rbl2, Trp53, Trp63 (7 genes).
G1 and G1/S transition: Camk2a, Camk2b, Gpr132, Mtbp, Myb, Nfatc1, Ppp2r3a, Ppp3ca, Taf10 (9 genes).
S and replication: Dnajc2, Mcm2, Mcm3, Mcm4, Mki67, Mre11a, Pcna, Rad17, Rad51, Sumo1 (10 genes).
M phase: Cdc25a, Cdk2, Nek2, Npm2, Pes1, Prm1, Rad21, Stag1, Terf1, Wee1 (10 genes).

3.11 Statistical analysis

In qRT-PCR array experiments, for comparison of proportions of modulated genes at different experimental times, the χ^2 or exact Fisher's test (depending on the absolute frequencies of contingency tables) were used. To perform qRT-PCR array statistical analysis aimed to the selection of genes with expression modification statistically significant in respect to control at least at one experimental time, a two-tailed Student's t-test with Bonferroni's correction for multiple tests was used. In this case a $p<0.0006$ was considered statistically significant. Only for these pre-selected genes, a two-tailed Student's t-test was further used to statistically evaluate both expression differences in respect to control at any experimental time and a dose effect according to the quantity of SDF administered. In this case a value of $p<0.05$ was considered statistically significant. The densitometric data about EGFP locus methylation were analyzed using ANOVA, with a $p<0.05$ level of significance. All other experimental data were analyzed by a two-tailed Student's t-test with a level of significance of $p<0.05$.

4. RESULTS

4.1 Construction, optimization and characterization of the experimental system (before the start of this thesis)

For a better comprehension of the outcomes of my thesis, in this preliminary section 4.1 I briefly report the results obtained from the research group of the "Department of Biopathology and Diagnostic Imaging", University of Rome Tor Vergata, during the construction, optimization and characterization of the experimental system used (Luchetti et al., 2012; see also Materials and Methods, section 3.1) before the start of this thesis. SV-40 immortalized MEF line was modified stably integrating within its genome either the wtEGFP gene or a mutEGFP. In the wtEGFP gene, a non-sense mutation was introduced by *in vitro* mutagenesis carried out on pCEP4 residue 210 located in the coding region of the gene. The glutamine (CAG) to stop codon (TAG) transition causes, at the same time, a fluorescence switch off and a Bts I restriction site disruption (Fig. 6 A). In MEF-mutEGFP, a SDF homologous to wtEGFP sequence, after integration within genomic DNA, can restore the wild type sequence and EGFP fluorescence. Clonal dilution and hygromycin selection were performed to obtain homogeneous transgenic cell lines, stably integrating wild type or

mutated copies of EGFP gene, as demonstrated by sequencing (Fig. 6 B) and FACS analyses (Fig. 6 C).

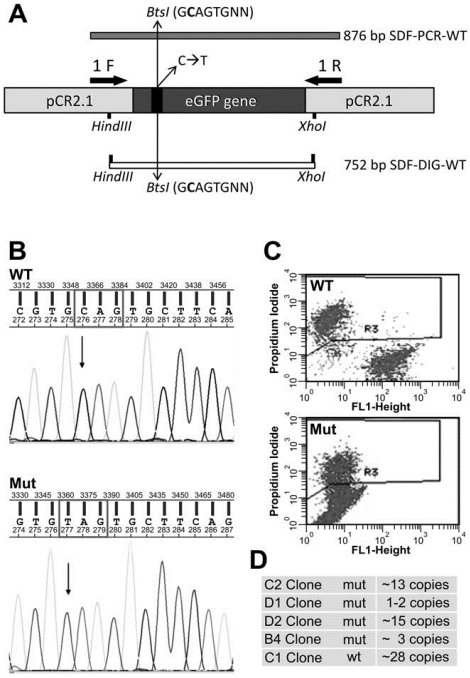


Figure 6 - Experimental design for SDF and cell clone generation. A) SDF sequence is homologous to the entire wild type EGFP coding sequence. SDF-PCR-WT, 876 bp long was generated by PCR amplification, SDF-DIG-WT, 752 bp long, was obtained by HindIII and XhoI digestion of pCR-2.1 vector. B) Sequencing analysis showing wild type (WT; top panel) and mutated (Mut; bottom panel) pCEP4-EGFP in C1 and D1 cell clones, respectively. C) FACS density plot of C1 (WT; top) and D1 (Mut; bottom) respectively. D) pCEP4-EGFP copy number determination for each cell clone. (Luchetti et al., 2012).

For each clone pCEP4 / EGFP copy number was determined by Taqman qPCR (Fig. 6 D). Among four mutated cell clones, D1 was employed for all the experiments, but those by qRT-PCR arrays, because containing only one copy of the transgene. For qRT-PCR array experiments, the EGFP C2 multy-copy clone was used because of its better experimental performance to this kind of experiments. The MEF-wtEGFP and MEF-mutEGFP constitute a reporter system by which also low frequency of phenotypic changes can be easily detected and quantified by FACS. In this assay system different parameters were tested and the best experimental conditions in terms of cell viability, transfection and correction efficiency were established such as, for example, the optimal quantity and structural characteristics of the SDF, the optimal cellular and transfection conditions, as well as the best electroporation program on the Amaxa nucleofection system. In unsynchronized MEF-mutEGFP, the best correction efficiency was obtained using 12 x 10^6 molecules of SDF / cell (20 μg). This amount has been used for all further transfections. The best correction frequency of 0.05 % was detected by FACS analysis when SDF-PCR-WT, resulting five-folds higher than both SDF-DIG-WT and single strand and PCR amplified (both 0.01 %, Student's t-test $p<0.07$). The SDF-PCR-WT was used in all further experiments. To determine the cell cycle influence, gene targeting was evaluated in cell populations enriched in G0 / G1, G1 / S and G2 / M phases.

In G2 / M the SFHR resulted favored, in fact G2 / M synchronized cells showed an increased correction efficiency to 0.5 % (Student's t-test $p<0.0001$ in respect to unsynchronized cells).

4.2 SDF methylation and modification efficiency

From here and below, the experiments I performed for this thesis are described. Since the SDF-PCR-WT showed higher correction efficiency than SDF-DIG-WT after transfection in D1 unsynchronized cells, to test the hypothesis that specific methylation patterns of SDF can influence the correction process of SFHR, the methylation patterns of SDF-DIG-WT, SDF-PCR-WT and of pCR 2.1 plasmid containing SDF and used as control, were studied (Fig. 7). In particular, these samples were treated with methylation-sensitive or insensitive restriction endonucleases, and used as targets for subsequent PCR amplification. For each sample, a multiplex PCR with two amplicons was obtained: the smallest is referred to a zone with no restriction sites (the internal standard always amplified), the largest is the amplicon of the target zone. The successful amplification of the target amplicon demonstrates the presence of the specific methylation pattern. Both *Dcm* and *Dam* bacterial methylation patterns resulted present on pCR2.1 plasmid and on the SDF-DIG-

WT, as clearly evidenced by the presence of an amplified largest band in samples treated by the methylation sensitive *PspG I* and *Mbo I*, respectively. There is no eukaryotic HpaII methylation pattern on pCR2.1 and on SDF-DIG-WT, as demonstrated by the absence of the specific amplicon in Hpa II-treated samples. As expected, SDF-PCR-WT showed no methylation pattern, as evidenced by the absence of any specific amplicon from *Hpa II-*, *PspG I-* or *Mbo I*-treated samples.

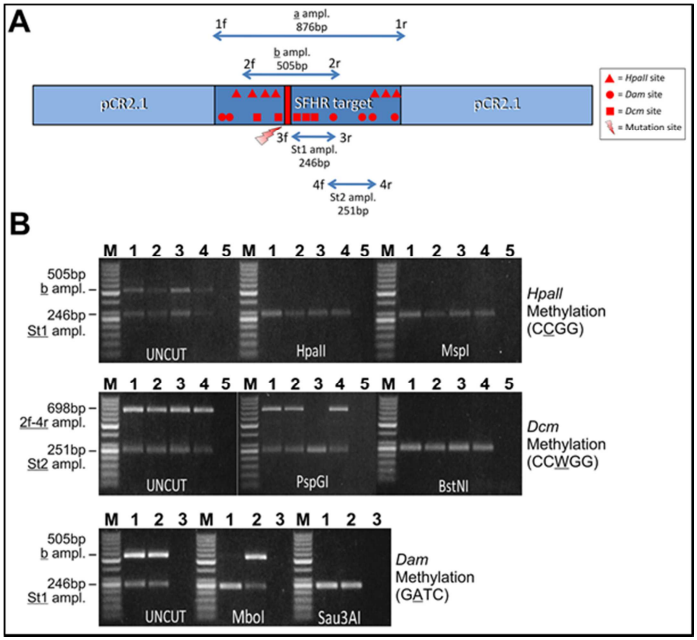


Figure 7 - Analysis of methylation patterns of SDF-PCR-WT and of SDF-DIG-WT. A) Analysis design. B) Amplification of SDF-PCR-WT and SDF-DIG-WT after treatment with methylation sensitive restriction enzymes for Hpa II, Dcm or Dam methylation (central panels: *Hpa II, PspG I, Mbo I*). Untreated samples, or samples treated with heat inactivated restriction enzymes (left panels) are shown as positive controls. Samples cut with methylation insensitive isoschizomers (right panels: *Msp I, BstN I, Sau3A I*) are shown as negative controls. In all panels the lower band is the internal control of amplification always amplified, while the upper band is the amplicon from the target sequence. The recognition sequence of each restriction enzyme is reported on the right. In every lane M a GeneRuler™ 50 bp DNA Ladder is shown. For Hpa II and Dcm methylations target samples are repeated in the same order, as follows: lanes 1 and 2 SDF-DIG-WT; lane 3, SDF-PCR-WT; lane 4, pCR 2.1 plasmid; lane 5, negative control. For Dam methylation: lane 1, SDF-PCR-WT; lane 2, pCR2.1 plasmid; lane 3 negative control. (Luchetti et al., 2012).

Subsequently, differently methylated SDFs were produced *in vitro* using *Sss I* or *Dam* or both DNA methyltransferases and transfected into G2 / M synchronized cells. The efficiency of SDF replacement (Fig. 8) was up to 75 % lower (Dam+ methylation = 0.12 %) than that obtained transfecting SDF-PCR-WT (= 0.48 %), where no methylation was present. A reduction of about 50 % was observed when SDF-DIG-WT (= 0.22 %) harboring prokaryotic methylation, was used.

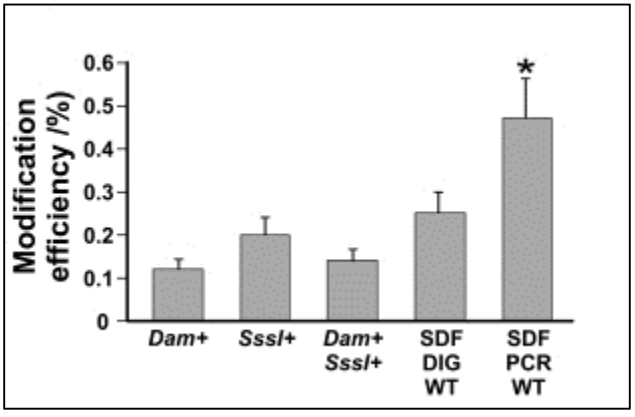


Figure 8 - Modification efficiencies obtained testing SDFs with different superimposed methylations patterns, on D1 cells. Differently *in vitro* methylated SDFs were tested to assess methylation involvement in gene modification efficiency. SDF-PCR-WT gave the highest efficiency of modification (*significant when compared to all treatments; specifically Student's t-test, $p<0.005$ respect to Dam+, $p<0.05$ respect to SssI+, $p<0.05$ respect to Dam+ / SssI+, and $p<0.05$ respect to SDF-DIG-WT). (Luchetti et al., 2012).

4.3 Genomic DNA methylation involvement in the inactivation of EGFP expression

By culturing D1 fluorescent cells after sorting, with increasing cell passages a gradual loss of EGFP expression was noticed (Fig. 9 A). A similar trend was observed in parental C1. Retro-mutation was excluded by both RFLP and sequencing analyses, confirming the presence of SDF-mediated C nucleotide, regardless to cell fluorescent phenotype (data not shown). Thus, the hypothesis was that DNA methylation may cause transcriptional inactivation of the corrected EGFP gene and consequently the fluorescence switch off in both D1 SDF-modified and in parental C1 clone. To assess DNA methylation involvement in the EGFP expression, D1 SDF-modified cells were resorted. EGFP sorted negative cells (but still carrying the correction) were treated for 24 and 48 h with 0.5 μM of 5-Aza-dC. EGFP expression, monitored by Real Time-PCR (Fig. 9B), showed more than a four-fold increase after 24 h (Student's t-test $p<0.005$). The expression further doubled after 48 h of treatment, when compared to untreated cells (Student's t-test $p<0.005$). Untreated cells usually showed a decreasing in relative EGFP expression according to fluorescence decrement (data not shown). These results demonstrated that the percentage of SDF-mediated modification in transfected cells was underestimated, because of a methylation-mediated silencing of EGFP expression.

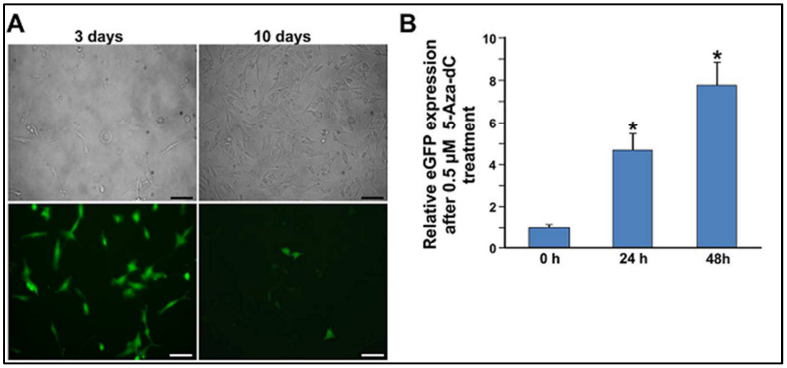


Figure 9 - EGFP expression increased after 5-Aza-dC treatment. A) Bright field (upper row) and fluorescent (bottom row) images of D1 sorted corrected cells at different experimental time. B) EGFP expression, analyzed by Real Time PCR, after 24 h and 48 h of treatment with 0.5 µM 5-Aza-dC respect to untreated cells (0 h) (Student's t-test, *$p < 0.005$). (Luchetti et al., 2012).

To investigate the correlation between the EGFP locus methylation status and its time-dependent expression, studies by multiplex *Hpa II* / PCR and *Aci I* / PCR analysis were performed using either C1 parental (divided in two sub-populations according to their fluorescence intensity) or D1 SDF-corrected cells (Fig. 10). Three different amplicons (*c*, *d* and *e*), spanning EGFP locus including CMV promoter, were analyzed (Fig. 10 A). Figures 10 B and 10 C showed the percentage of methylation obtained from densitometric analyses of electrophoretic restriction pattern of the three amplicons.

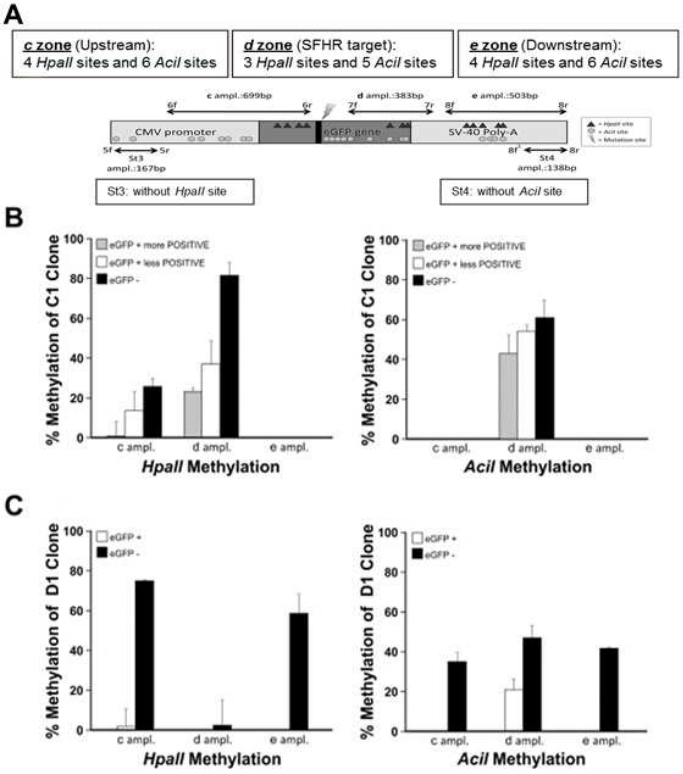


Figure 10 – *Hpa II* and *Aci I* methylation analyses of integrated EGFP in C1 and D1 clones. A) Experimental design. *Hpa II* and *Aci I* sites are indicated. B) Densitometric analyses of parental C1 clone methylation pattern on EGFP+ more positive, EGFP+ less positive and EGFP- cells. ANOVA test gave a statistical significance of $p<0.001$ and $p<0.005$ respectively for *Hpa II* and *Aci I* panels. C) Densitometric analysis of methylation pattern of D1 SFHR-modified clone on both fluorescent and non-fluorescent cells. ANOVA test gave a statistical significance of $p<0.001$ for each panel. (Luchetti et al., 2012).

In both C1 fluorescent and non-fluorescent sorted populations, the *d* amplicon (Fig. 10 B) resulted the most methylated, either for *Hpa II* or *Aci I*; on the contrary, neither *Hpa II* nor *Aci I* methylation were evidenced in the *e* amplicon. In the *c* amplicon the *Hpa II* methylation was lower than in *d* fragment, whereas the *Aci I* methylation resulted absent. Importantly, C1 non-fluorescent cells showed higher levels of *Hpa II* methylation than C1 fluorescent cells in *c* and *d* amplicons. A similar quantitative correlation was evidenced for *Aci I* methylation, although only for the *d* amplicon. So in the parental C1 clone there is a good correlation between EGFP inactivation and both *Hpa II* and *Aci I* DNA methylation patterns; in fact, non-fluorescent cells always showed the highest levels of both kind of methylation. The same analysis was carried out on SFHR-modified D1 cells, fluorescent after sorting but that lost their fluorescence after several passages in culture. As a result of re-sorting of phenotypically heterogeneous corrected D1 cells were possible to distinguish between EGFP + and EGFP - cells. In fluorescent modified D1 cells (Fig.10 C) all the analyzed zones showed no or very low levels of both *HpaII* or *AciI* methylation whereas non-fluorescent modified D1 cells showed considerable level of both *Hpa II* and *Aci I* methylation, with the exception of the *d* zone that resulted negative for *Hpa II* methylation. Also in D1 modified cells, methylation resulted to be directly correlated with EGFP expression.

4.4 1,5-Isoquinolinediol drug treatment increases correction efficiency only in presence of 5-Aza-dC

Finally, three main biological pathways involved in the repair of the DNA damage were investigated, to test their influence on the SFHR mechanism:

a) the DSB repair based on the ATM kinase pathway;

b) the transcription-coupled NER based on the RNA polymerase II activity;

c) the BER and NHEJ based on the poly-(ADP-ribose) polymerase pathway.

Specifically KU55933, 1,5-Isoquinolinediol (1,5-ISQ) and α-Amanitin, inhibitors respectively of ATM kinase, PARP-1 and RNA polymerase II (Bryant et al., 2006; Hickson et al. 2004; Shiloh, 2003; Semionov et al., 2003; Abraham, 2001; Semionov et al., 1999; Rotman et al., 1999; Selby et al., 1993;) were added to transfected cells. No statistically significant variations in modification efficiency were observed three days after transfection (Fig. 11, black bars) in respect to control sample in which no drugs were added.

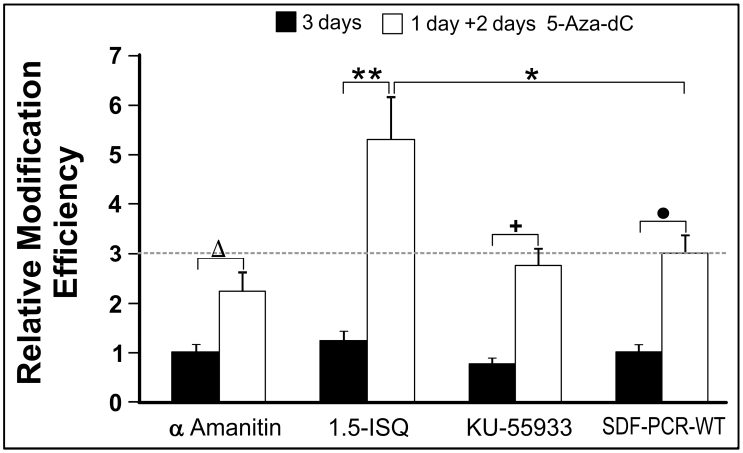


Figure 11: Relative modification efficiency in D1 cells transfected with SDF-PCR-WT and treated with α-Amanitin, 1,5-Isoquinolinediol and KU-55933. Transfected samples were analyzsed three days after transfection (3 days; black columns) or in parallel treated 24 h after transfection with 0.5 μM of 5-Aza-dC for 48 h (1 day +2 days 5-Aza-dC; white columns). No statistically significant differences were observed at 3 days (black bars) respect to untreated cells (SDF-PCR-WT). Demethylating effect of 5-Aza-dC increased EGFP detection in all samples (white columns) in a statistically significant manner (Student's t-test, Δ $p<0.05$; +$p<0.05$; •$p<0.05$). 5-Aza-dC addition also disclosed the effect of 1,5-Isoquinolinediol on SDF-mediated correction in a statically significant manner in respect either to cells not treated with 5-Aza-dC (Student's t-test **$p<0.001$) and to the cells transfected with SDF-PCR-WT in which no drug was added (Student's t-test *$p<0.005$). Dashed lines refers to modification efficiency observed in cells without addition of any drug but treated by 5-Aza-dC. (Luchetti et al., 2012).

To disclose methyl-hidden correction events, 5-Aza-dC was added to all samples 24 h after transfection (Fig. 11, white bars), that resulted in a statistically significant overall increase of fluorescence. When 5-Aza-dC is added to 1,5-ISQ treated cells a statistically significant increase in correction

efficiency was obtained in respect to both cells untreated with 5-Aza-dC (Fig. 11, 1,5 ISQ black bar, Student's t-test $p<0.001$) and control PCR (Fig.11, SDF-PCR-WT white bar, Student's t-test $p<0.005$). These data indicated a synergistic effect on SFHR correction efficiency of PARP pathway, which revealed a potential SFHR-efficiency modifier, and DNA methylation pathway.

4.5 Overall analysis of the effect of SFHR on DNA repair genes

To better understand the mechanisms of DNA repair possibly involved in SFHR, I extended my studies to a larger number of genes in order to identify new proteins involved in the gene repair system, essential for the integration of the therapeutic DNA. The experiments were conducted on 1700000 unsynchronized C2 MEF-mutEGFP cells that were transfected with different amounts of SDF-PCR-WT (5 μg and 20 μg). After transfection the cells were FACS analyzed to determine the correction efficiency and counted to assess the cellular growth (Fig.12). As expected, the gene modification efficiency (Fig.12 A) was enhanced when cells were treated with high dose (20 μg) of SDF-PCR-WT reaching the 0.05 % of correction (Student's t-test $p<0.001$) in respect to 0.01 % obtained with the low dose (5 μg) (Student's t-test $p<0.001$). Regarding the cell growth (Fig. 12 B), the negative effect on growth due to transfection is

highlighted. In fact, even the control transfected without SDF (indicated as No SDF) shows, in particular after 72h from transfection, growth levels significantly lower than the non-transfected control (Student's t-test $p<0.01$). This effect is further accentuated when the cells underwent transfection with the SDF (Student's t-test $p<0.01$). However, the effect appeared to be not dose-dependent, with similar growth values after administration of different amount of SDF.

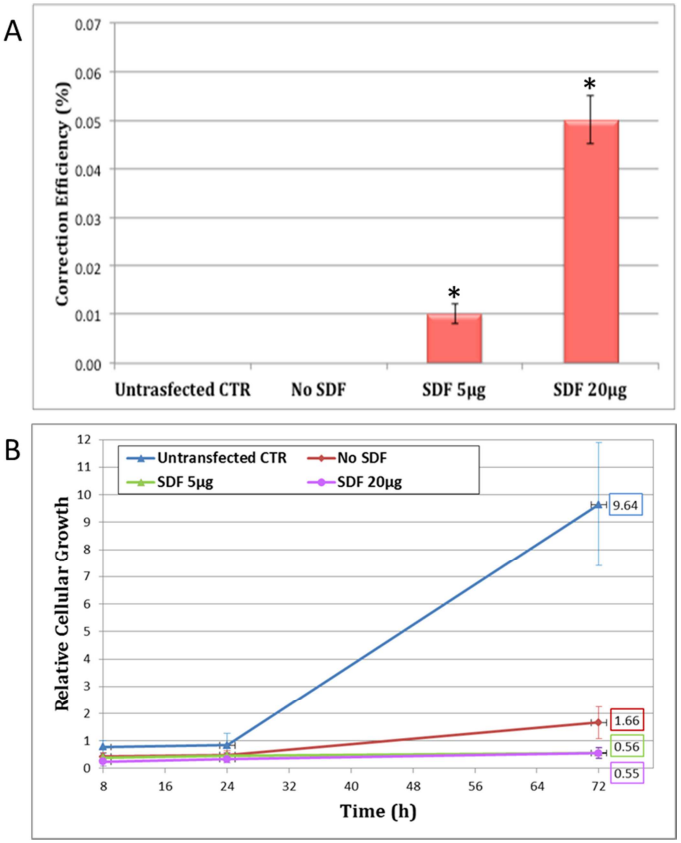


Figure 12 - A) Correction efficiencies after transfection of different amounts of SDF-PCR-WT in C2 cells. Student's t-test, *p<0.001 in respect to controls. **B) Relative cellular growth after transfection of different amounts of SDF-PCR-WT in C2 cells.** The values of growth after 72 h from transfection are indicated in the corresponding colored boxes.
Untrasfected CTR = cells that did not undergo transfection; No SDF = cells that underwent transfection without SDF; SDF 5 µg = cells transfected with low SDF dosage; SDF 20 µg = cells transfected with low SDF dosage.

After RNA extraction, the quantitative expression of the 84 genes involved in the response to several kinds of DNA damage, using qRT-PCR arrays, was investigated in MEF-mutEGFP. These genes were classified as follows: 18 related to HR, 7 to the NHEJ, 12 to the MMR, 19 to the BER, 27 to the NER and 1 with an interconnected and regulatory role within several repair pathways. The basal expression levels of DNA repair genes in untreated MEF-mutEGFP were heterogeneous (Fig. 13), with some high expressed and several low expressed genes, and resulted to change according to experimental time (8 h, 24 h or 72 h).

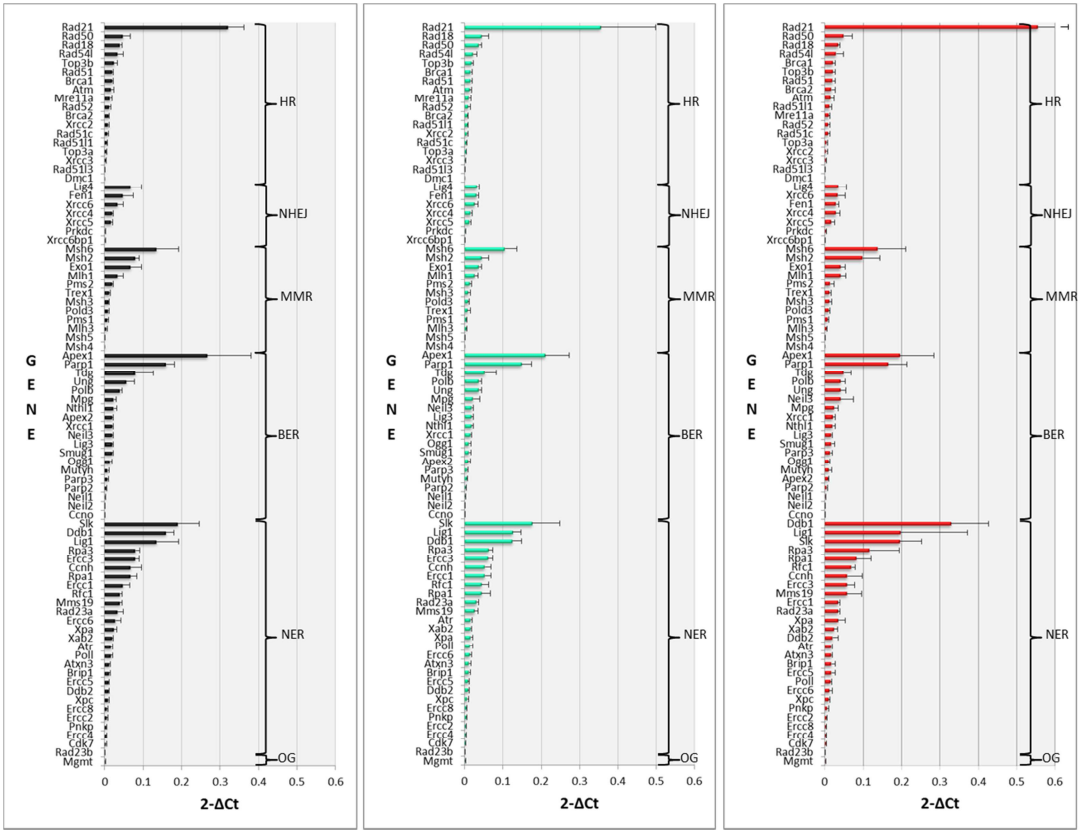


Figure 13: Basal expression of DNA repair genes in untransfected MEF-mutEGFP at 8 h (black bars), 24 h (green bars) and 72 h (red bars) from plating (simultaneously to nucleofected experimental lines). On the right are indicated the grouping of the analyzed genes according to their main role in the DNA repair pathway .

For a preliminary analysis, we used the fold change of expression in experiments as compared to respective controls to select 3 different classes of expression modulation. We considered upregulated those genes with a fold change in respect to the control greater than 1.2, downregulated those genes with a fold change lower than 0.8, and with an expression similar to the control those genes with a fold change within 0.8 and 1.2 (extremes included). After this preliminary analysis, a more stringent statistical analysis was performed (see below). To firstly evaluate the overall effect of both the SDF and the nucleofection protocol, we compared cells transfected by the SDF at low (5 μg) and high (20 μg) dosage with untransfected controls at the corresponding experimental time. At low dosage of SDF (5 μg, Fig. 14 A) and early experimental time (8 h after nucleofection) an upregulation of 56 (66.7 %) DNA repair genes was observed, with only 2 (2.4 %) downregulated genes. At intermediate experimental time (24 h after treatment) 42 (50.0 %) overexpressed genes left, with 1 (1.2 %) downregulated genes. Finally (72 h after treatment), only 4 (4.8 %) upregulated genes were evidenced, with 50 (59.5 %) genes reaching a lower expression level than untreated control and 30 genes with expression similar to the untransfected control.

Figure 14 - Overall analysis of the fold change of studied genes of the DNA repair pathway in MEF-mutEGFP at 8 h, 24 h and 72 h after treatment, according to different experimental conditions (from A to F). The average fold change is reported with error bars indicating s.d.

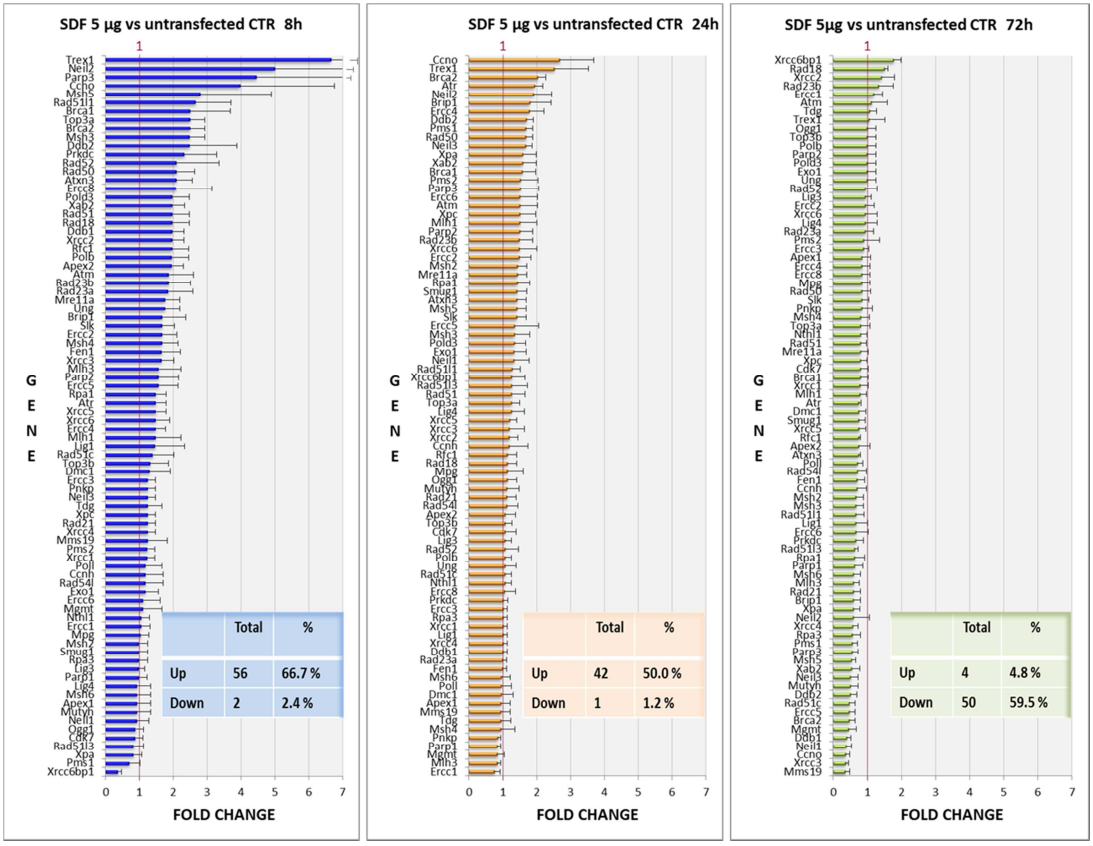


Figure 14 - A) cells treated with 5 µg of SDF as compared to untransfected cells (CTR, set to 1).

At high dosage of SDF (20 μg, Fig. 14 B) and early experimental time (8 h after nucleofection) an upregulation of 58 (69.0 %) DNA repair genes was observed, with only 5 (6.0 %) downregulated genes. At intermediate experimental time (24 h after treatment) an increment of overexpressed genes to 70 (83.3 %) could be evidence, with no downregulated gene. Finally (72 h after treatment), the number of upregulated genes decreased to 21 (25.0 %), with 35 (41.7 %) genes reaching a lower expression level than untreated control and 28 genes with expression similar to the untransfected control.

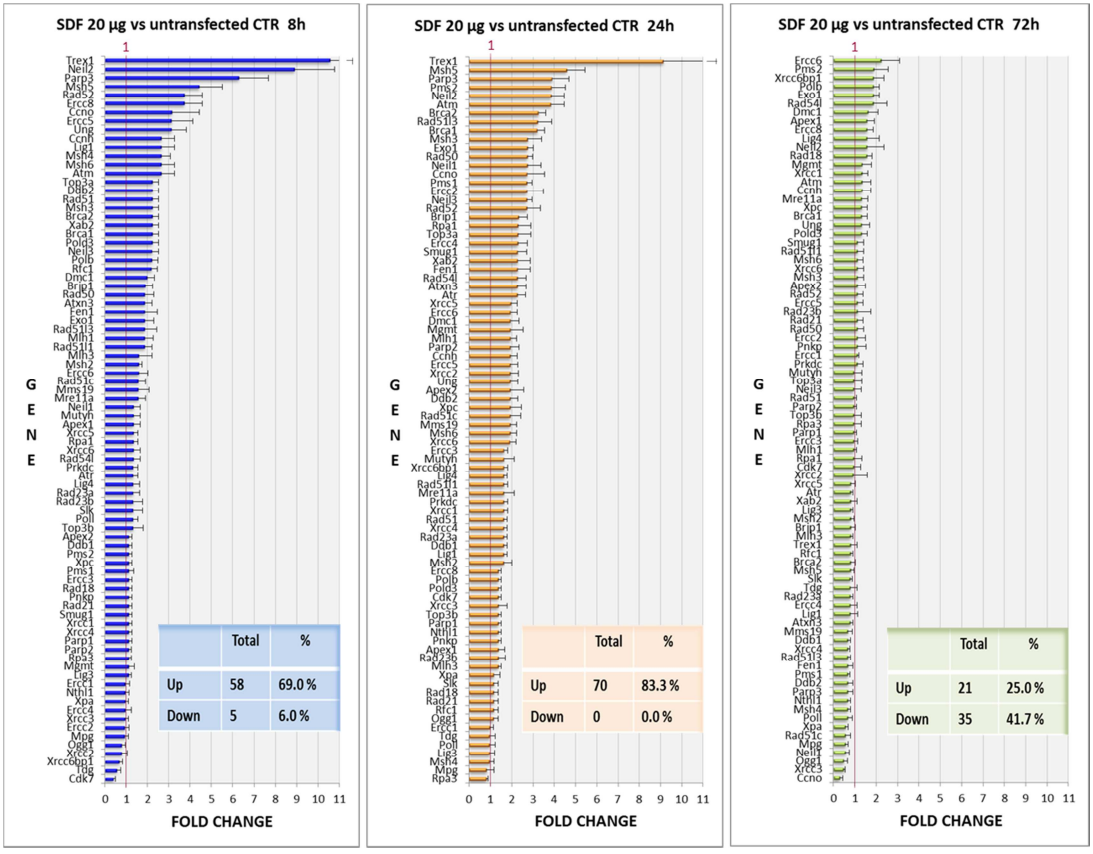


Figure 14 - B) cells treated with 20 µg of SDF as compared to untransfected cells (CTR, set to 1).

It appears that after an early phase of upregulation there is progressively dowregulation of DNA repair genes, most of them reaching, in the late phase, expression levels lower than control (set to 1). The proportion of early

upregulated genes seems to be similar at low (5 μg) and high (20 μg) SDF dosage. However, the early phase of upregulation seems to be shorter after treatment with a low dosage than with a high dosage. In the latter condition, 24 h after the treatment the number of upregulated genes is still growing, whereas, in the former condition, the number of upregulated genes is already decreasing. At 72 h, the number of upregulated genes is still considerably higher at high dosage than low dosage and, in addition, in the latter condition a higher number of downregulated genes could be evidenced. Also from the fold change point of view, the treatment with high dosage of SDF appeared to produce a greater effect than the treatment with low dosage, mainly at 8 h and 24 h. The overall dose effect that produced the prolonged and quantitatively greater consequence, on the DNA repair gene expression, of SDF at high dosage can also be clearly seen by a direct comparison between cells treated with 20 μg of SDF with those treated with 5 μg (experimental condition in this case set to 1) (Fig. 14 C).

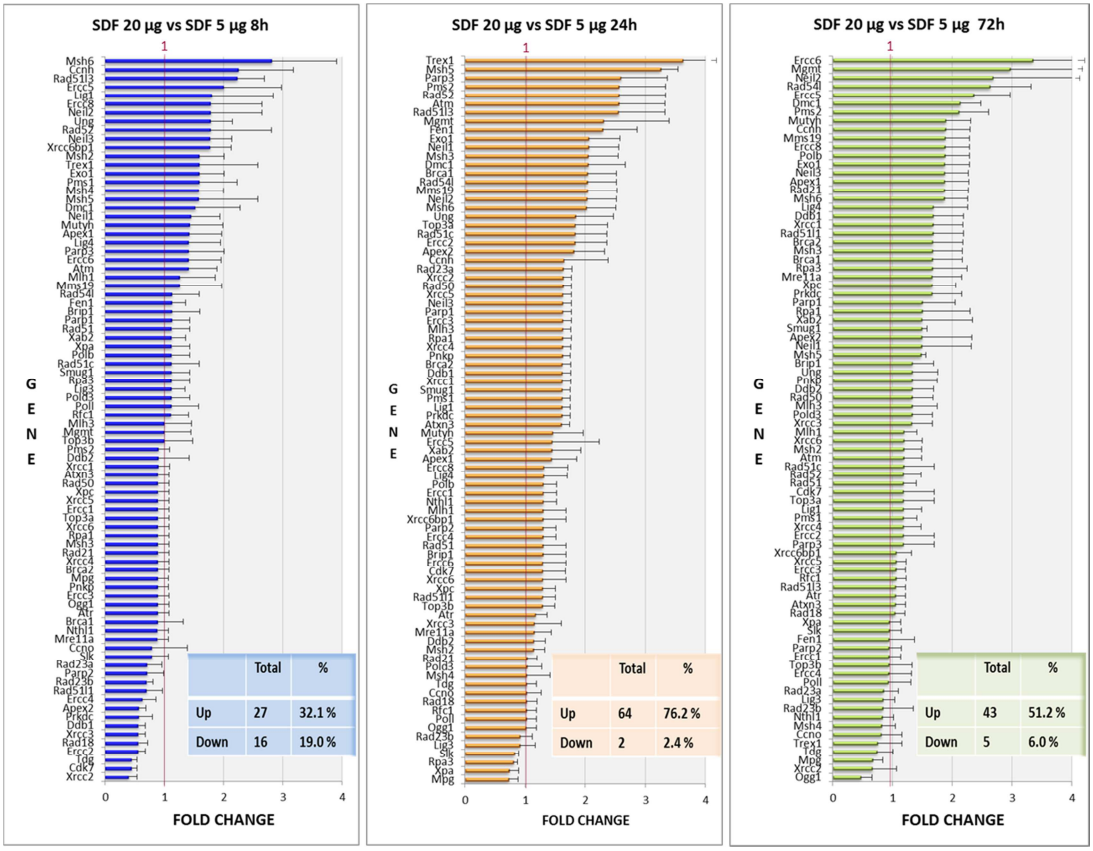


Figure 14 - C) cells treated with 20 µg of SDF as compared to those treated with 5 µg of SDF (set to 1).

To evaluate a possible effect of the nucleofection protocol on the expression of DNA repair genes, we analyzed cells that underwent the transfection protocol

without SDF (we used the same nucleofection protocol and buffer with no SDF inside) (Fig. 14 D).

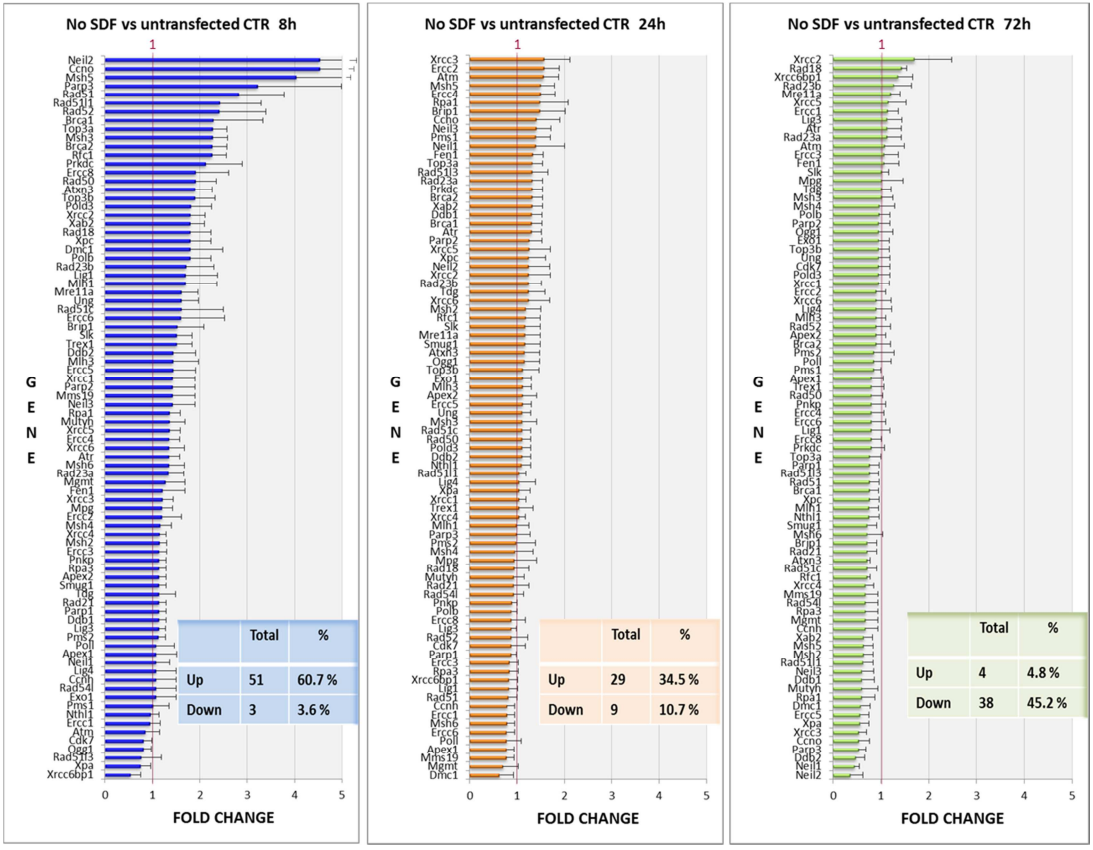


Figure 14 - D) cells treated (transfected) without SDF (indicated in the figure as No SDF) as compared to untransfected cells (CTR, set to 1).

The expression temporal pattern after nucleofection without SDF was very similar to the general expression temporal pattern, with an early upregulation followed by a progressive downregulation. However, this effect resulted reduced if compared with that induced by the SDF (also at low dosage) in terms of both the proportion of involved genes and fold change. In addition, the genes modulated by the transfection protocol appears, at least partially, different from those modulated by the SDF.

To subtract the effect of the nucleofection protocol, which allows to distinguish the specific effect of SDF, we compared cells nucleofected with or without SDF, either at low dosage or at high dosage. After the subtraction of the effect of the nucleofection protocol, the specific effect of 5 μg of SDF alone (Fig. 14 E) appeared quantitatively reduced (in terms of fold change) in respect to the additive effects of 5 μg of SDF and nucleofection protocol (Fig. 14 A).

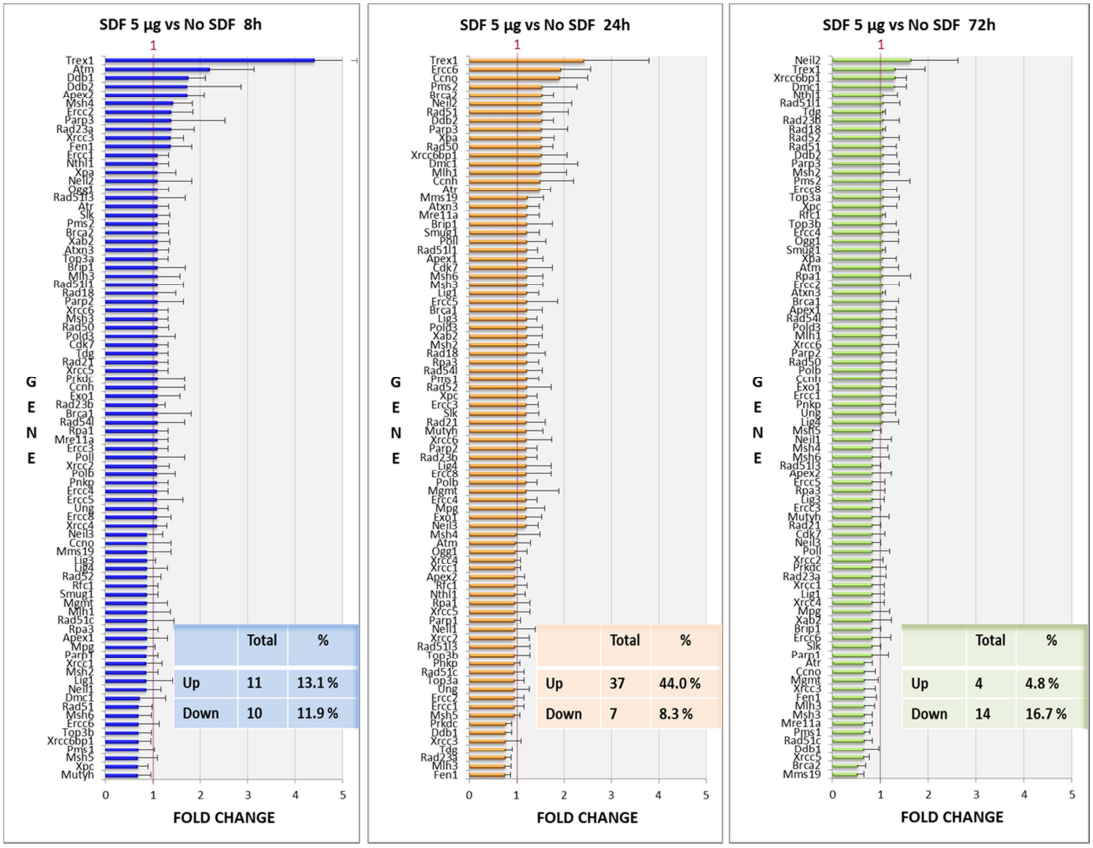


Figure 14 - E) cells treated with 5 µg of SDF as compared to cells treated (transfected) without SDF (indicated in the figure as H_2O, set to 1).

The temporal pattern specifically induced by the SDF (Fig. 14 E) is quite similar to that cumulative of SDF and nucleofection protocol (Fig. 14 A) at 24 h and 72 h. However, the large number of upregulated genes at 8 h evidenced after

treatment with SDF and nucleofection (Fig. 14 A) seems to depend more from the nucleofection protocol (Fig. 14 D) than from a specific effect of the SDF itself (Fig. 14 E). Similar considerations may be done about the specific effect exerted by 20 μg of SDF, after subtraction of the effect of the nucleofection protocol (Fig. 14 F), although with a quantitatively greater response (in terms of fold change) specifically depending from the SDF itself. Also a later persistence (up to 72 h) of a greater number of upregulated genes (53.6 %) appeared to be specifically induced by high dosage of SDF (Fig. 14 F as compared with Fig. 14 D).

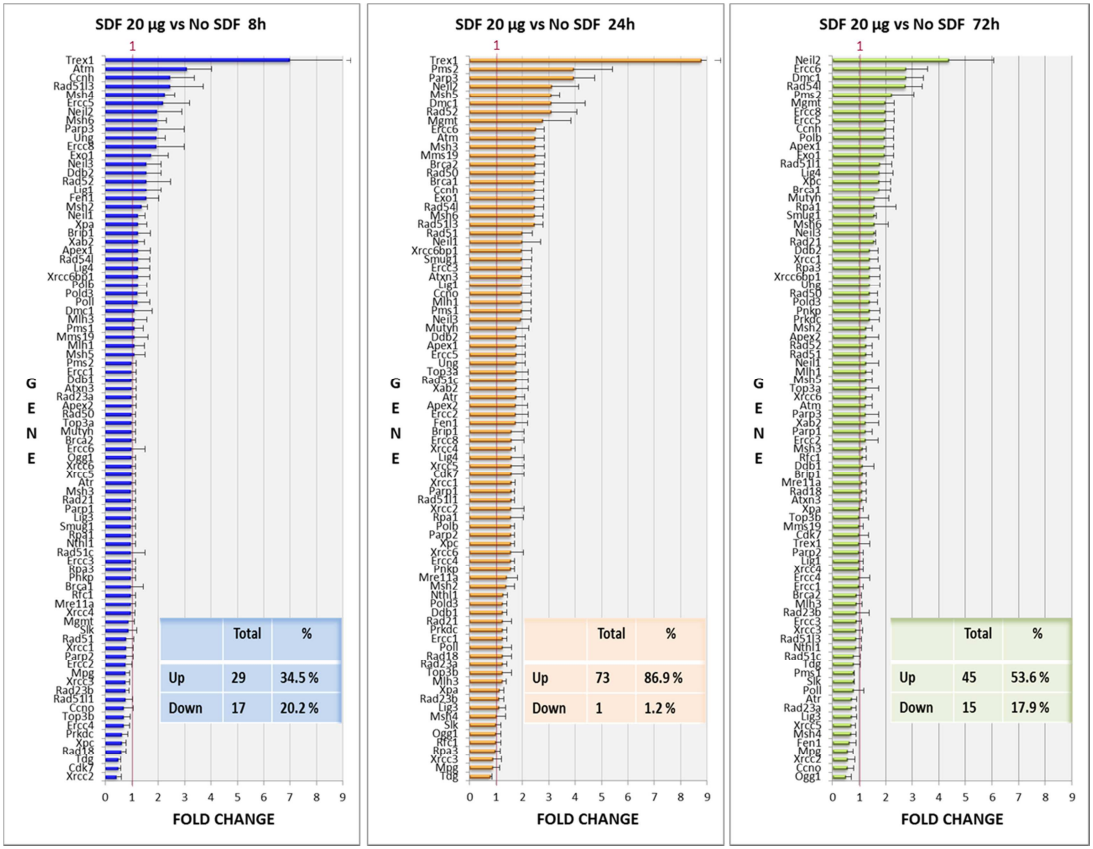


Figure 14 – F) cells treated with 20 µg of SDF as compared to cells treated (transfected) without SDF (indicated in the figure as No SDF, set to 1).

A synthetic view of these results about DNA repair gene expression is reported in Fig. 15.

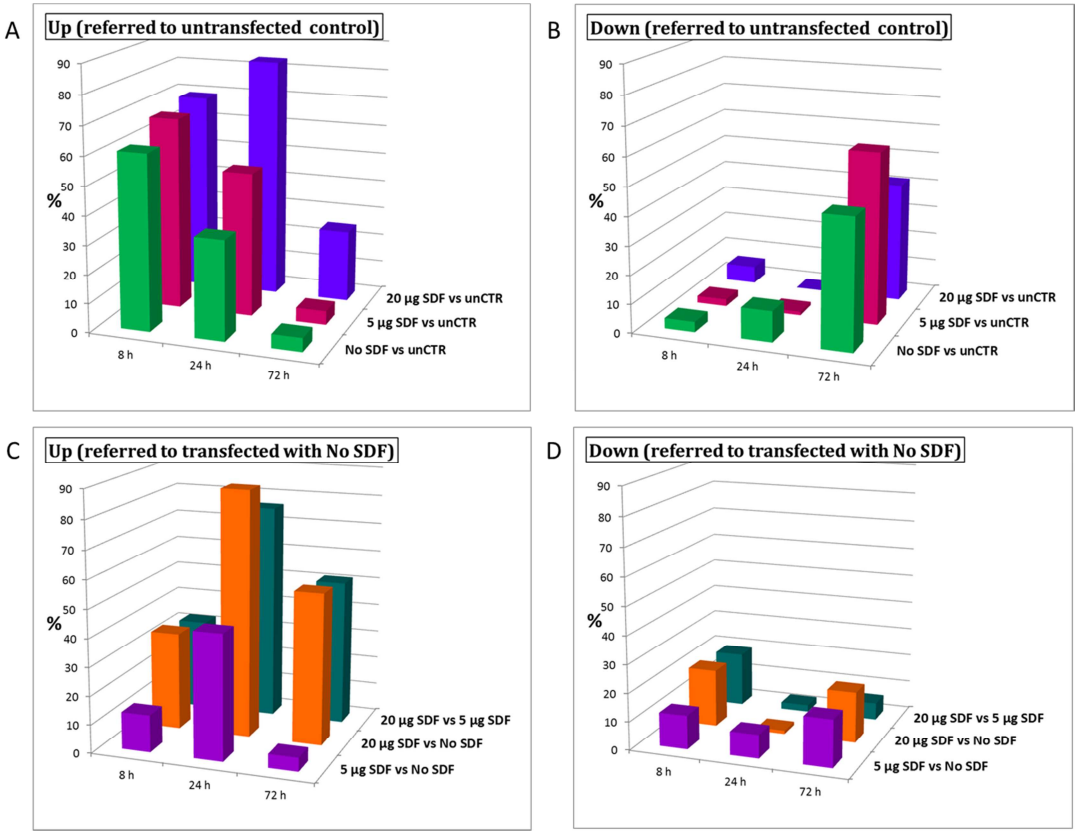
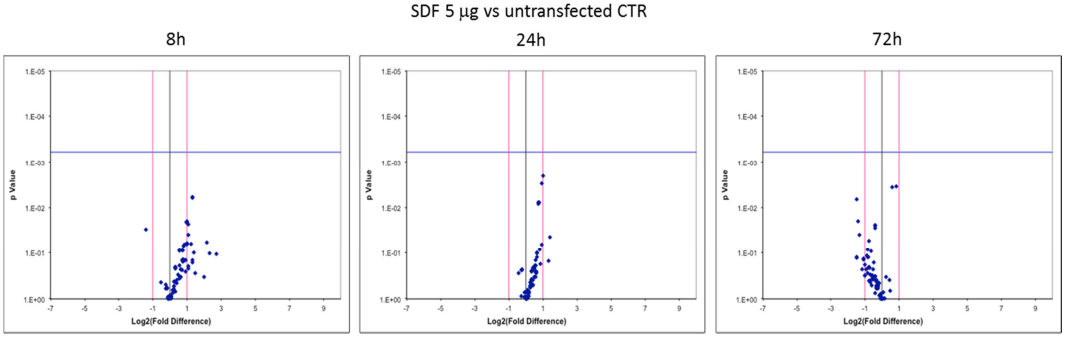


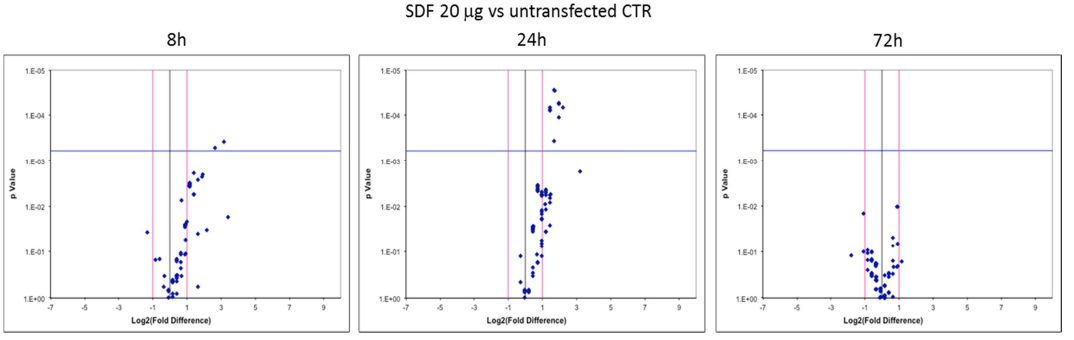
Figure 15 - Synthesis of results about the proportion of modulated DNA repair genes. A) Upregulated and downregulated (B) genes after transfection with no SDF (No SDF), low (5 µg SDF) or high (20 µg SDF) dosage of SDF, as compared with untransfected control (unCTR). C) Upregulated and D) downregulated genes after treatment with low (5 µg SDF) or high (20 µg SDF) dosage of SDF as compared with transfected control with no SDF (No SDF); a comparison for dosage effect evaluation (20 µg SDF vs 5 µg SDF) is also reported. Differences in the proportion of both upregulated and downregulated genes between the different experimental times (8 h, 24 h and 72 h) were statistically significant (every χ^2 or exact Fisher's tests, $p<0.001$). See text for the criteria used for the definition of upregulation and downregulation.

After the overall analysis reported above, that took under consideration all genes grouped into classes of expression level, a more statistically stringent analysis was performed, by using the Student's t test and the Bonferroni's correction for multiple comparisons. By this method, a difference was considered statistically significant if $p<0.0006$. The DNA repair genes with statistically significant expression differences, at least at one experimental time resulted to be 14. The statistically significant increase of expression of these genes was evidenced (using the $p<0.0006$ level) only after nucleofection with 20 μg of SDF (Fig. 16 B, C and F); this increase occurred for 2 genes at both 8 h and 24 h and for the remaining 12 genes at 24 h. No statistically significant effects (at $p<0.0006$ level) could be evidenced after nucleofection with 5 μg of SDF (Fig. 16 A and E) or after transfection with no SDF (Fig. 16 D). Each of these genes showed a quantitative expression value greater than 2. These results suggested a dose effect and a specific temporal pattern.

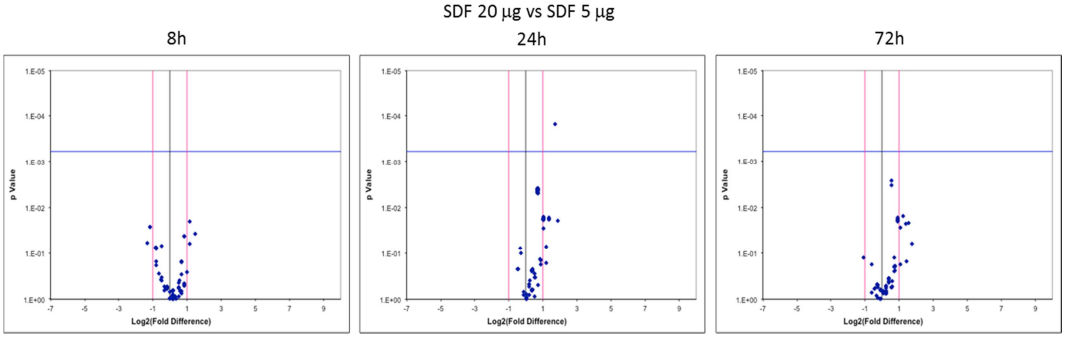
Figure 16 - Volcano plots of studied DNA repair genes in MEF-mutEGFP at 8 h, 24 h and 72 h after treatment, according to different experimental conditions (from A to F). The spots above the blue line mark those genes with statistically significant expression difference in respect to control (Student's t-test, after Bonferroni's correction for multiple comparisons, $p<0.0006$); the spots on the left and on the right of the pink lines mark those genes with expression level 2-fold, respectively, lower or greater than control.



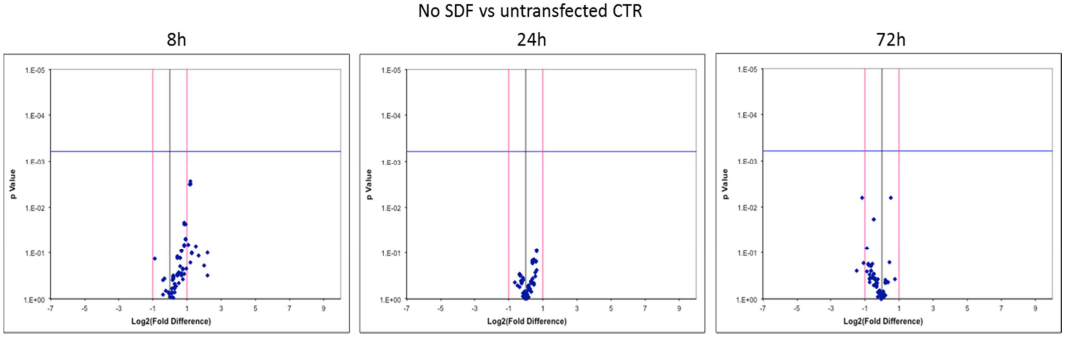
A) cells treated with 5 µg of SDF as compared to untransfected cells (CTR).



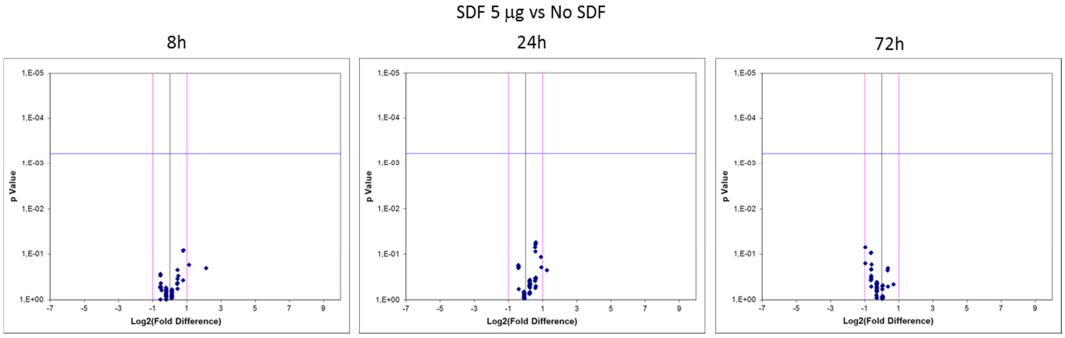
B) cells treated with 20 µg of SDF as compared to untransfected cells (CTR).



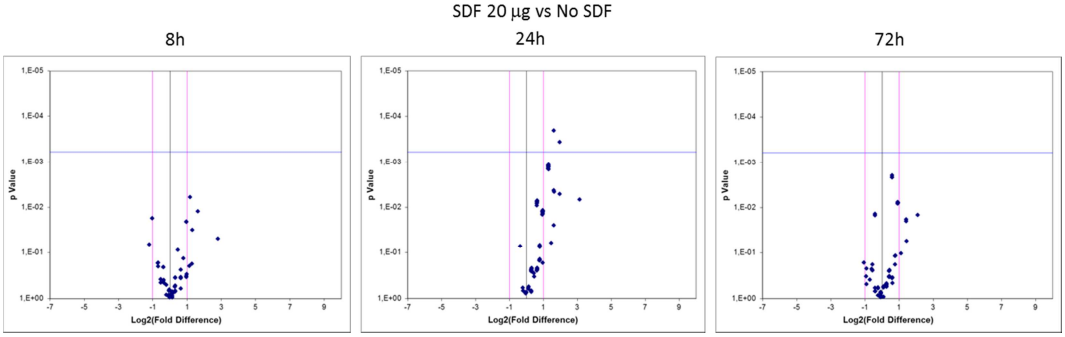
C) cells treated with 20 µg of SDF as compared to those treated with 5 µg of SDF.



D) cells treated (transfected) without SDF (indicated in the figure as No SDF) as compared to untransfected cells (CTR).



E) cells treated with 5 µg of SDF as compared to cells treated (transfected) without SDF (indicated in the figure as No SDF).



F) cells treated with 20 µg of SDF as compared to cells treated (transfected) without SDF (indicated in the figure as H_2O).

4.6 Overall analysis of the effect of SFHR on cell cycle regulatory genes

Using qRT-PCR arrays, also the quantitative expression of 84 genes involved in the cell cycle control was studied. These genes were classified as follows: 29 generally related to checkpoint and arrest, 9 specifically to G1 / G1-S transition, 10 specifically to S and replication phases, 10 to M phase, 19 to a general positive regulation and 7 to a general negative regulation of cell cycle. Also for cell cycle genes, the basal expression levels in untreated MEF-mutEGFP cells were heterogeneous (Fig. 17), with some high expressed and several low expressed genes, and variable depending on the experimental time (8 h, 24 h or 72 h). Also in this case, to firstly evaluate the overall effect of both the SDF and the nucleofection protocol, we compared the cells transfected by the SDF at low (5 μg) and high (20 μg) with untransfected controls at the corresponding experimental time.

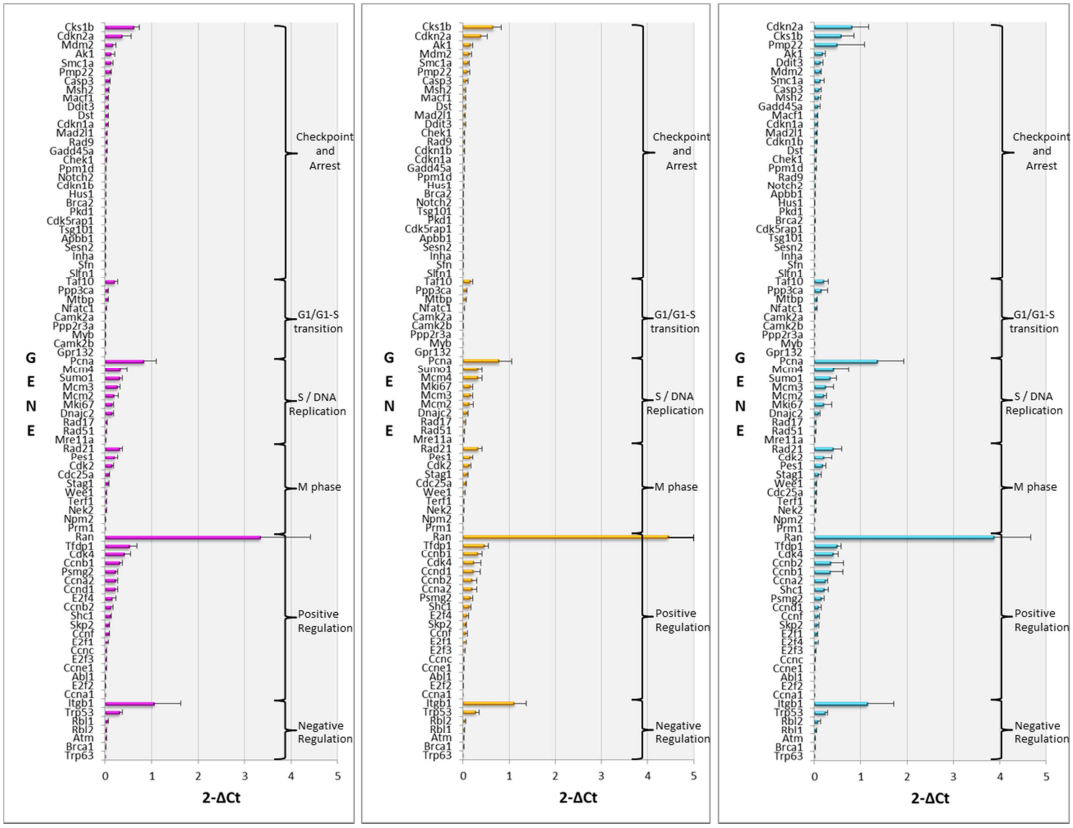
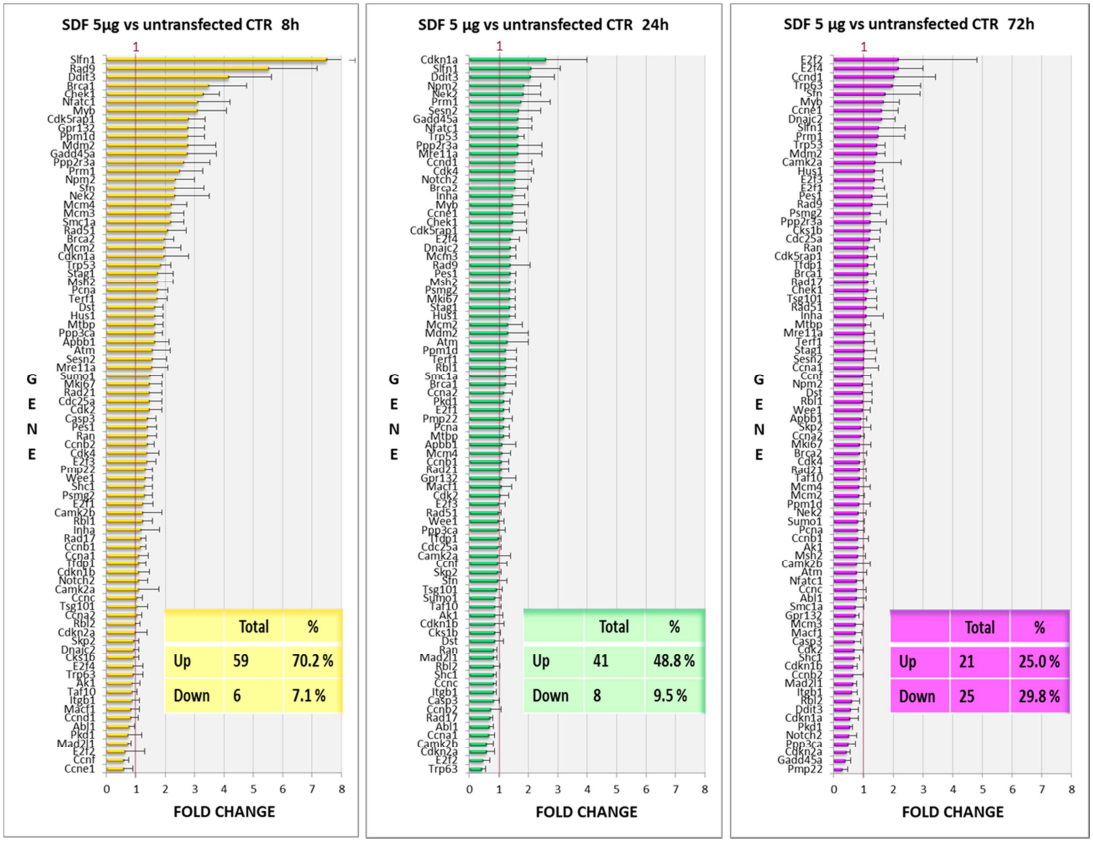


Figure 17 - Basal expression of cell cycle genes in untransfected MEF-mutEGFP at 8 h (purple bars), 24 h (yellow bars) and 72 h (light blue bars) from plating (simultaneously to nucleofected experimental lines). On the right are indicated the grouping of the analyzed genes according to their main role in the cell cycle pathway.

Applying the same thresholds than for the DNA repair gene analysis, at low dosage of SDF (5 μg, Fig. 18 A) and early experimental time (8 h after nucleofection) an upregulation of 59 (70.2 %) cell cycle genes and the downregulation of 6 (7.1 %) genes was observed. At intermediate experimental time (24 h after treatment), 41 (48.8 %) genes resulted upregulated and 8 (9.5 %) downregulated. Finally (72 h after treatment), 21 (25.0 %) genes resulted to be upregulated and 25 (29.8 %) genes were downregulated, with 38 similar to the control.

Figure 18 - Overall analysis of the fold change of studied genes of the cell cycle pathway in MEF-mutEGFP at 8 h, 24 h and 72 h after treatment, according to different experimental conditions (from A to F). The average fold change is reported with error bars indicating s.d.



A) cells treated with 5 µg of SDF as compared to untransfected cells (CTR, set to 1).

At high dosage of SDF (20 µg, Fig. 18 B) and early experimental time (8 h after nucleofection) an upregulation of 70 (83.3 %) cell cycle genes was observed, with only 6 (7.1 %) downregulated genes.

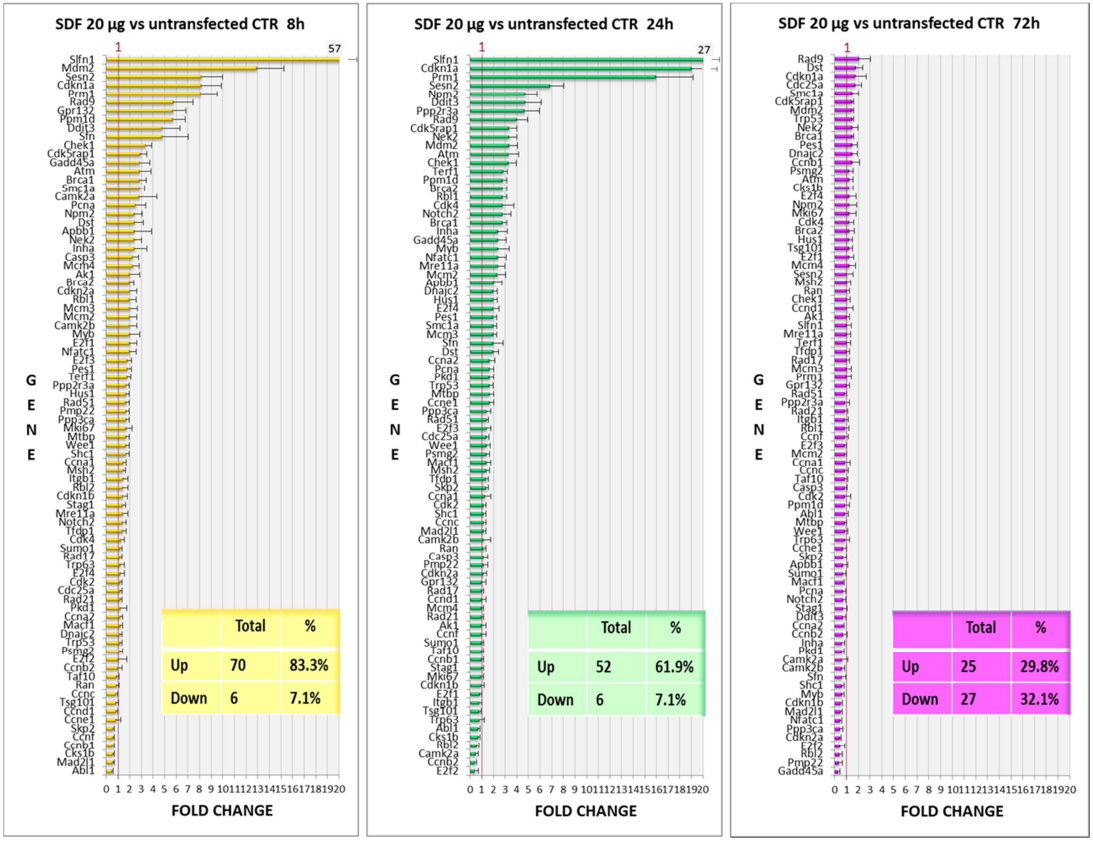


Figure 18 - B) cells treated with 20 µg of SDF as compared to untransfected cells (CTR, set to 1).

At intermediate experimental time (24 h after treatment) a decrement of overexpressed genes to 52 (61.9 %) could be evidence, with 6 (7.1 %) downregulated genes. Finally (72 h after treatment), the number of upregulated genes decreased to 25 (29.8 %), with 27 (32.1 %) genes reaching a lower expression level than untreated control, and 32 genes with expression similar to the untransfected control.

Also for cell cycle genes, there is an early phase of expression stimulation followed by a progressive phase of downregulation; most of the cell cycle genes reached, in the late phase (72 h), expression level lower than control (set to 1). As for the DNA repair genes, the treatment with high dosage of SDF appeared to produce a higher proportion of upregulated genes and a quantitative greater effect on the fold change than the treatment with low dosage. However, on the contrary to DNA repair genes, the treatment with either low or high dosage of SDF appeared to induce a similar temporal pattern, with a consistent upregulation still at 24 h, although with no further increase (at this experimental time) induced from the high dosage.

The overall dose effect that produced the quantitatively greater consequences, on the cell cycle genes, of SDF at high dosage can also be clearly seen by a direct comparison between cells treated with 20 μg of SDF with those treated with 5 μg (experimental condition in this case set to 1) (Fig. 18 C).

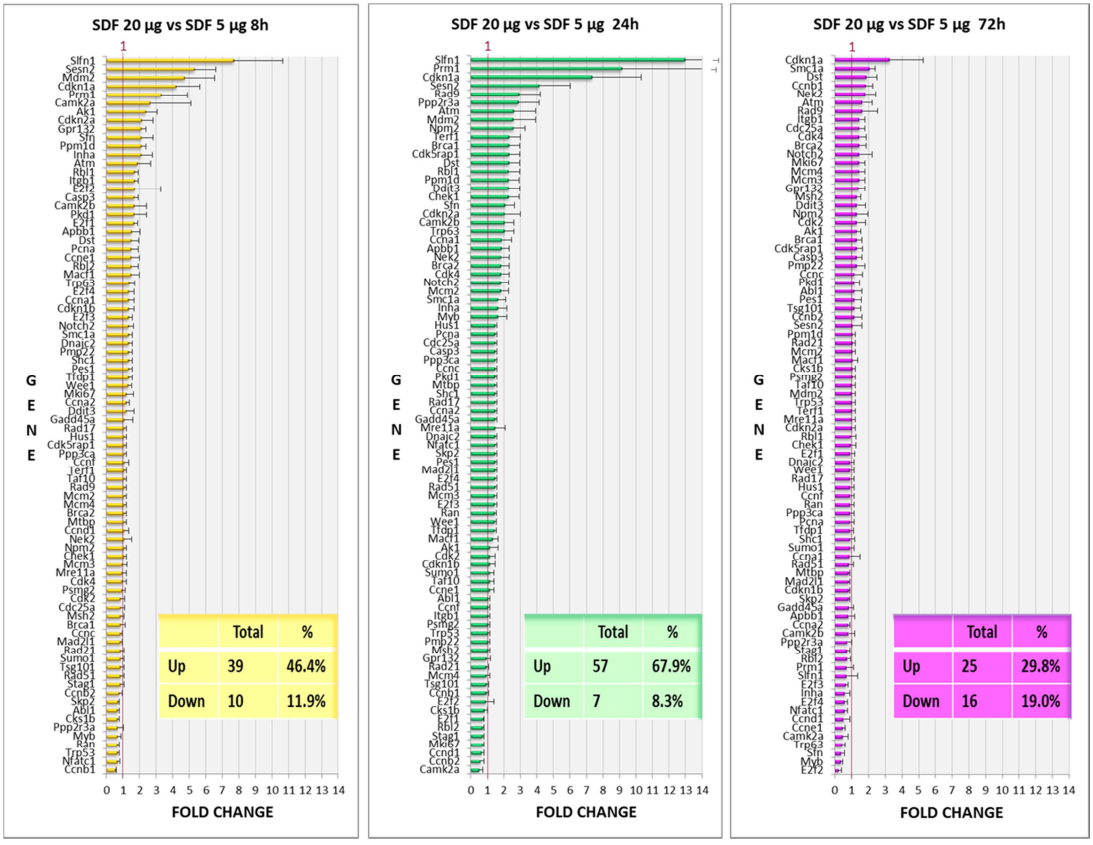


Figure 18 - C) cells treated with 20 µg of SDF as compared to those treated with 5 µg of SDF (set to 1).

As for DNA repair genes, to evaluate a possible effect of the nucleofection protocol on the expression of cell cycle genes, we analyzed cells that underwent the transfection protocol without SDF (Fig. 18 D). The temporal pattern after

nucleofection without SDF was similar to that observed for cells treated with SDF. From the number of upregulated genes and fold change point of view, this experimental condition treated with no SDF was very similar to that treated with 5 µg of SDF (Fig. 18 A), whereas quantitative differences could be evidence if compared to the experimental condition with high SDF dosage. Also in this case, the genes modulated by the transfection protocol are partially different from those modulated by the SDF.

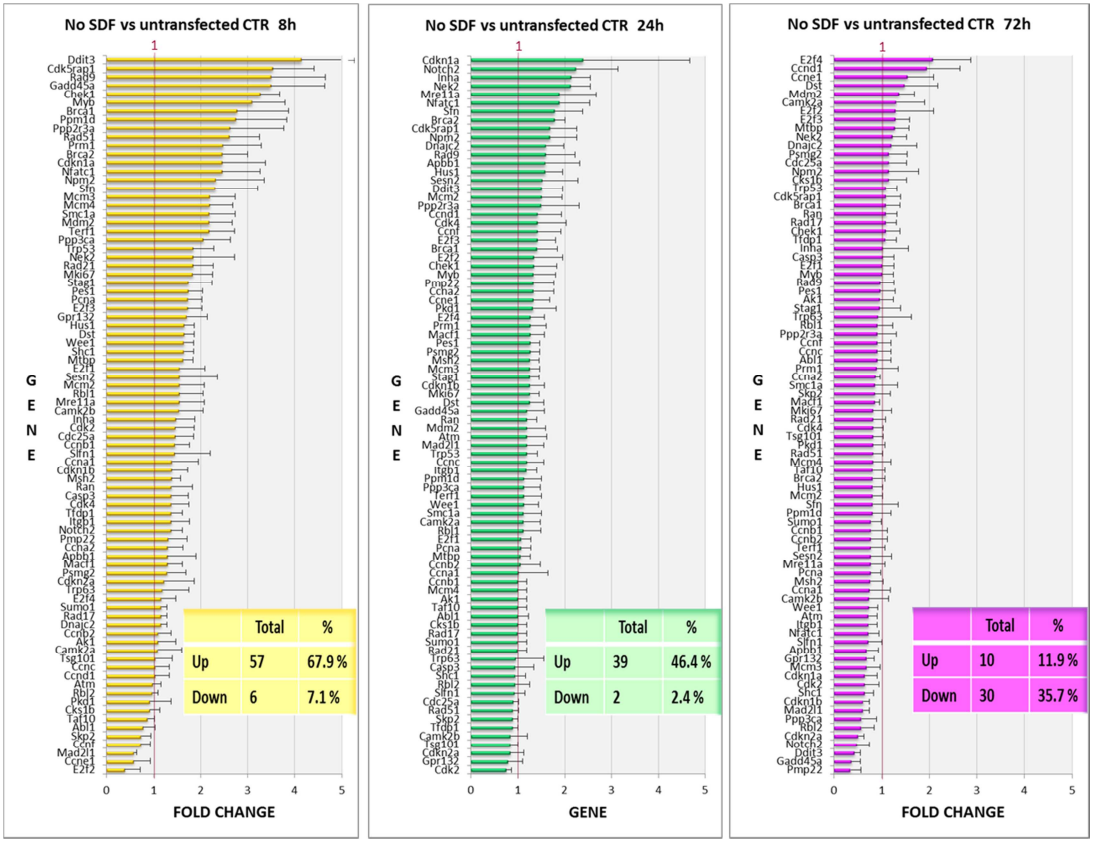


Figure 18 - D) cells treated (transfected) without SDF (indicated in the figure as No SDF) as compared to untransfected cells (CTR, set to 1).

To subtract the effect of the nucleofection protocol, which allows to distinguish the specific effect of SDF, we compared cells nucleofected with or without SDF, either at low dosage or at high dosage. After the subtraction of the effect of the

nucleofection protocol, the specific effect of 5 μg of SDF alone (Fig. 18 E) appeared quantitatively reduced, in terms of both the proportion of upregulated genes and fold change, in respect to the additive effects of 5 μg of SDF and nucleofection protocol (Fig. 18 A). Also in this case, the large number of upregulated genes at 8 h, and even at 24 h, evidenced after treatment with SDF and nucleofection (Fig. 18 A, left panel) seems to depend more from the nucleofection protocol (Fig. 18 D) than from a specific effect of the SDF itself (Fig. 18 E). A later persistence (up to 72 h) of a greater proportion of upregulated genes was evidenced. The temporal pattern specifically induced by the high dosage of SDF (20 μg), after subtraction of the effect of the nucleofection protocol (Fig. 18 F), is quite similar to that cumulative of SDF and nucleofection protocol (Fig. 18 B) at 24 h and 72 h, although partially reduced from quantitative point of view (in terms of fold change). Also in this case, a later persistence (up to 72 h) of a greater number of upregulated genes appeared to be specifically induced by high dosage of SDF (Fig. 18 F as compared with Fig. 18 D). However, also at high dosage, at early experimental time (8 h) the predominant effect seemed to be that of the nucleofection protocol than that of SDF itself.

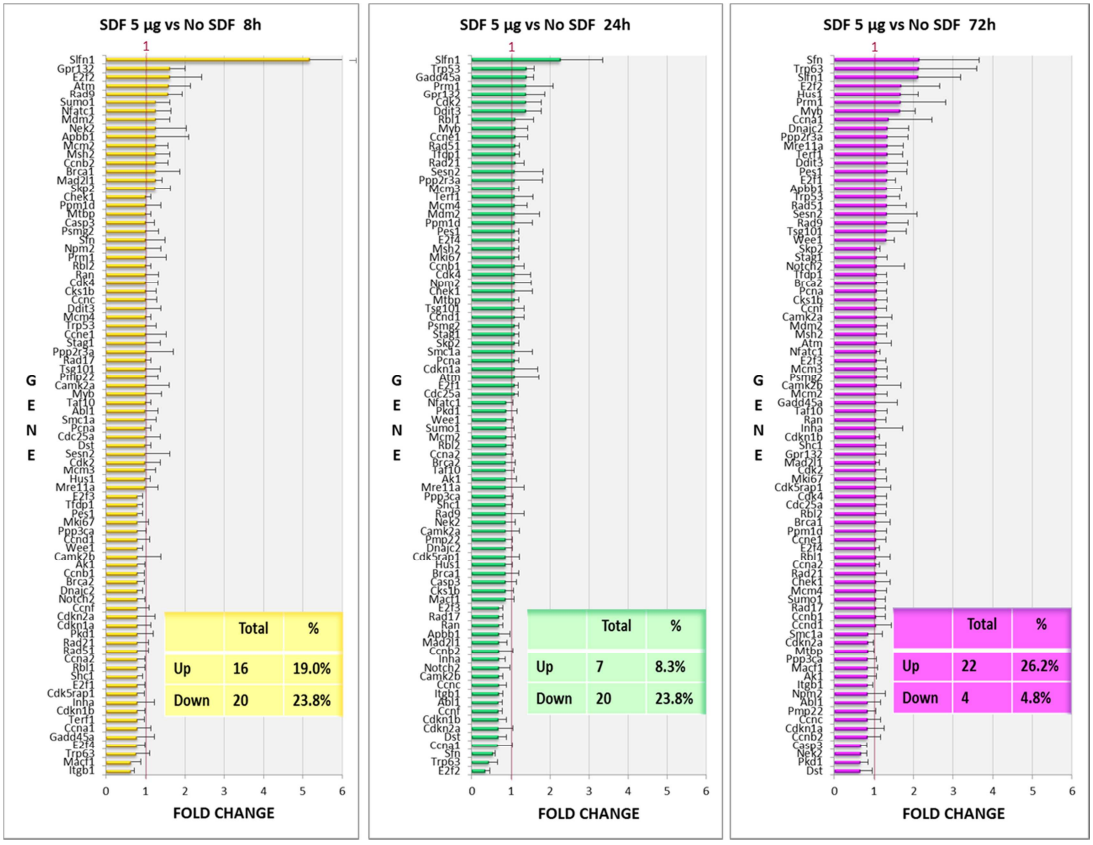


Figure 18 - E) cells treated with 5 µg of SDF as compared to cells treated (transfected) without SDF (indicated in the figure as No SDF, set to 1).

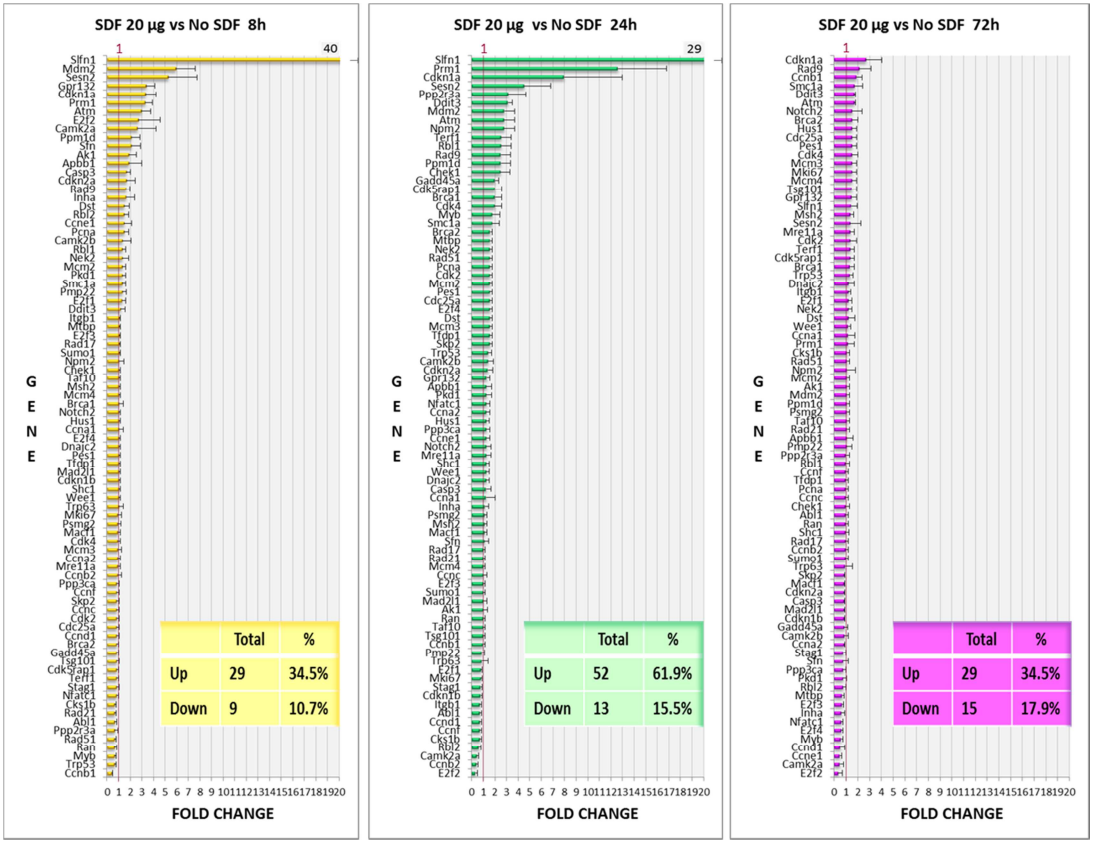


Figure 18 - F) cells treated with 20 µg of SDF as compared to cells treated (transfected) without SDF (indicated in the figure as No SDF, set to 1).

A synthetic view of these results about cell cycle gene expression is reported in Fig. 19.

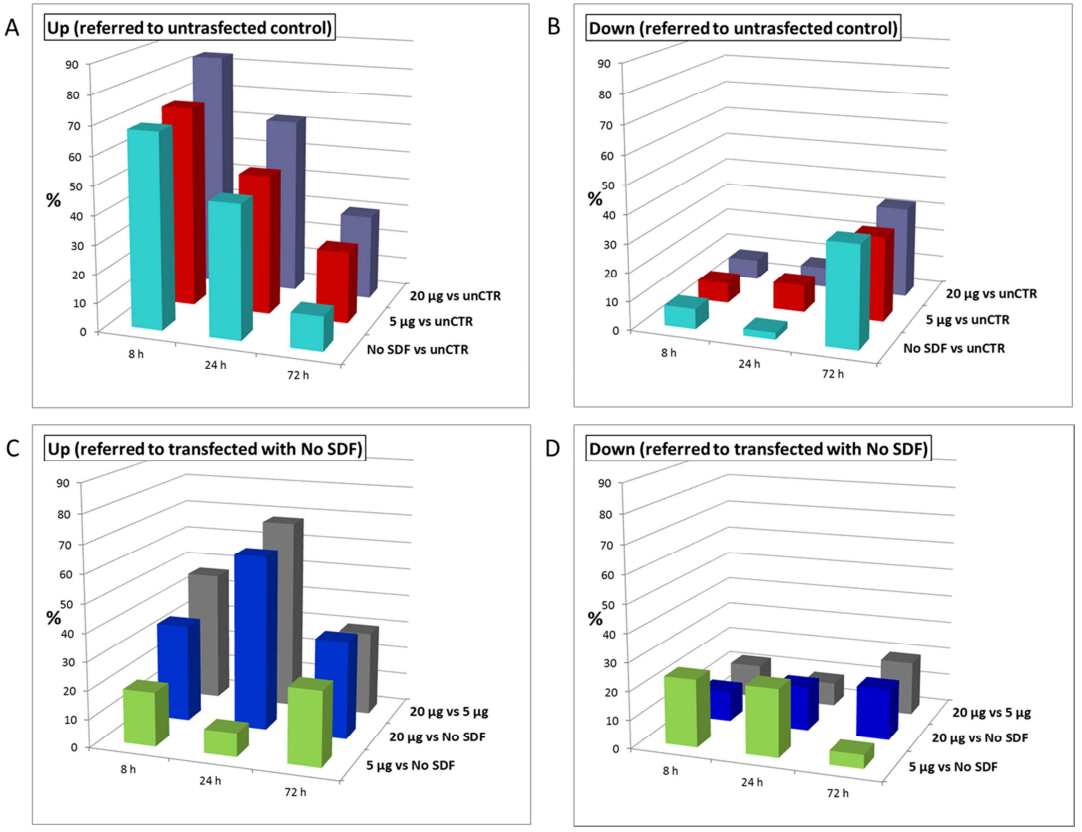
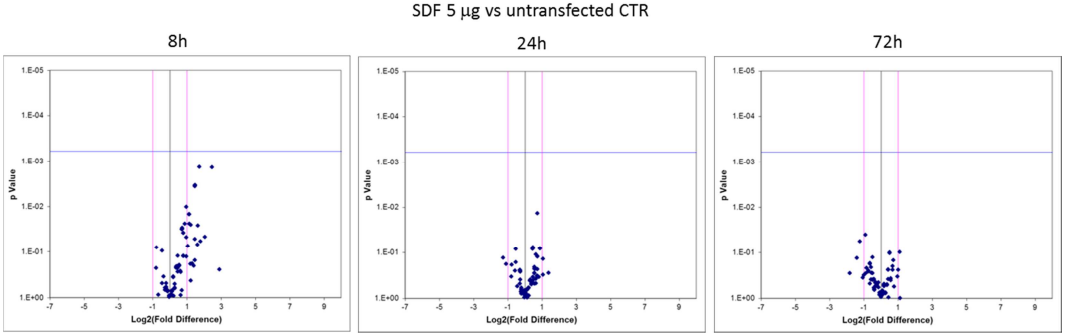


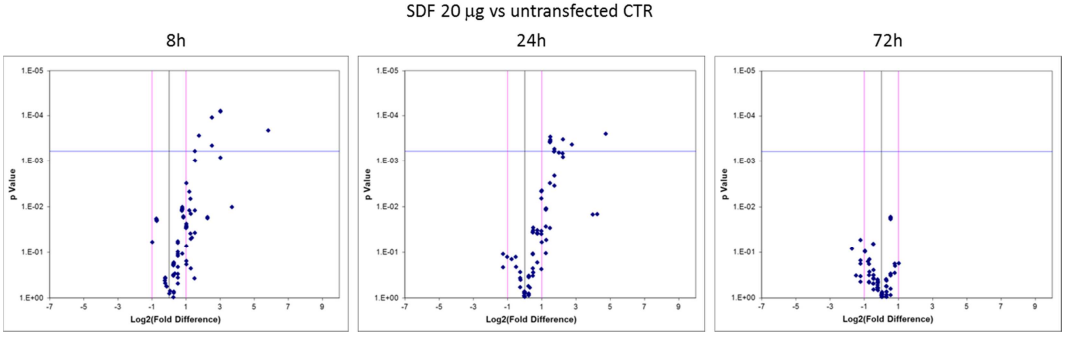
Figure 19 - Synthesis of results about the proportion of modulated cell cycle genes. A) Upregulated and downregulated (B) genes after transfection with no SDF (No SDF), low (5 µg SDF) or high (20 µg SDF) dosage of SDF, as compared with untransfected control (unCTR). C) Upregulated and D) downregulated genes after treatment with low (5 µg SDF) or high (20 µg SDF) dosage of SDF as compared with transfected control with no SDF (No SDF); a comparison for dosage effect evaluation (20 µg SDF vs 5 µg SDF) is also reported. Differences in the proportion of both upregulated and downregulated genes between the different experimental times (8 h, 24 h and 72 h) were statistically significant (every χ^2 or exact Fisher's tests, $p<0.001$). See text for the criteria used for the definition of upregulation and downregulation.

Also for cell cycle genes, after the overall analysis, the more statistically stringent analysis by the Student's t test after the Bonferroni's correction for multiple comparisons was performed (using $p<0.0006$ as significance threshold). The cell cycle genes with statistically significant expression differences, at least at one experimental time resulted to be 15. The statistically significant increase of expression of these genes was evidenced (using the $p<0.0006$ level) only after nucleofection with 20 μg of SDF (Fig. 20 B, C and F); the increase occurred for 2 genes at both 8 h and 24 h, for 5 genes only at 8 h and for 8 genes only at 24 h. No statistically significant effects (at $p<0.0006$ level) could be evidenced after nucleofection with 5 μg of SDF (Fig. 20 A and E) or after transfection with no SDF (Fig. 20 D). Each of these genes showed a quantitative expression value greater than 2. Also these results suggested a dose effect and a specific temporal pattern.

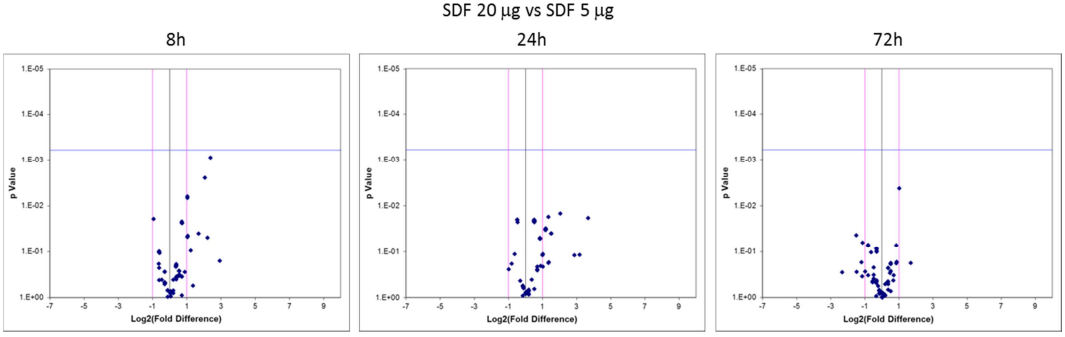
Figure 20 - Volcano plots of studied cell cycle genes in MEF-mutEGFP at 8 h, 24 h and 72 h after treatment, according to different experimental conditions (from A to F). The spots above the blue line mark those genes with statistically significant expression difference in respect to control (Student's t-test, after Bonferroni's correction for multiple comparisons, $p<0.0006$); the spots on the left and on the right of the pink lines mark those genes with expression level 2-fold, respectively, lower or greater than control.



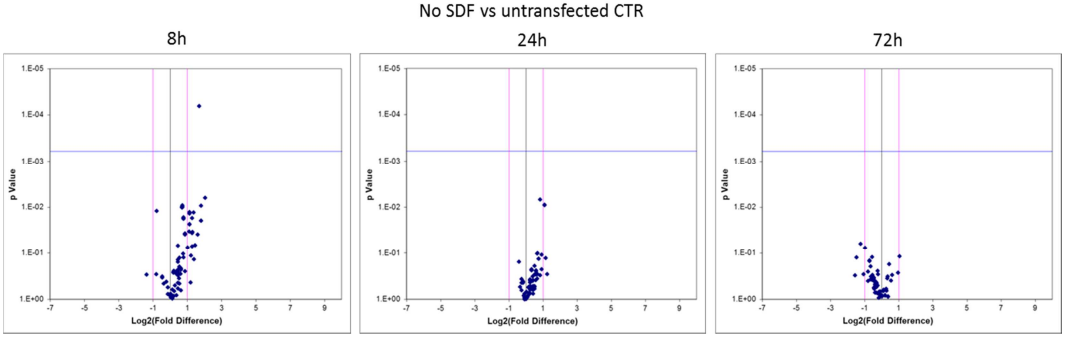
A) cells treated with 5 µg of SDF as compared to untransfected cells (CTR).



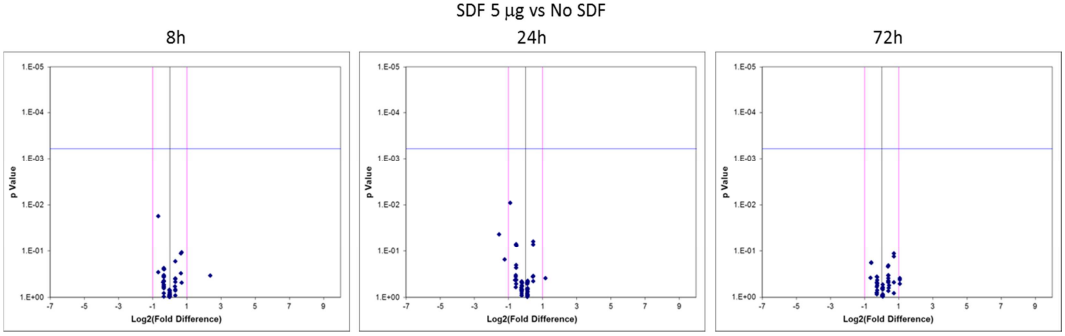
B) cells treated with 20 µg of SDF as compared to untransfected cells (CTR).



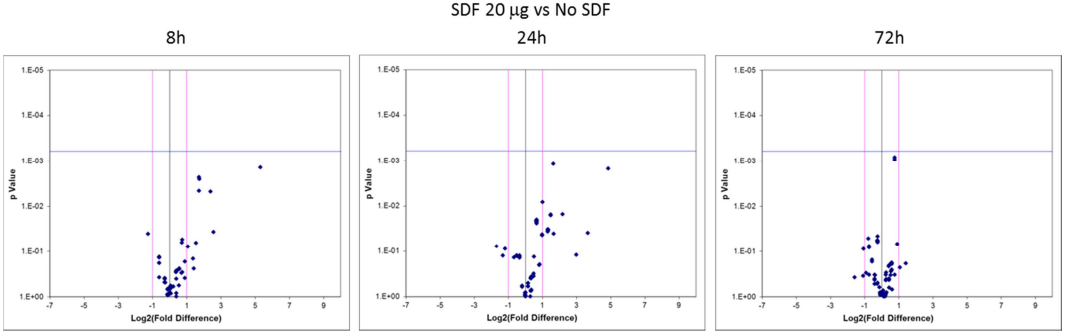
C) cells treated with 20 µg of SDF as compared to those treated with 5 µg of SDF.



D) cells treated (transfected) without SDF (indicated in the figure as No SDF) as compared to untransfected cells (CTR).



E) cells treated with 5 µg of SDF as compared to cells treated (transfected) without SDF (indicated in the figure as No SDF).



F) cells treated with 20 µg of SDF as compared to cells treated (transfected) without SDF (indicated in the figure as No SDF).

4.7 Selection of specific targets affected by SFHR within the DNA repair and cell cycle genes.

To select the best specific gene targets within both the DNA repair and cell cycle genes with an expression modulation induced by SFHR, a combination of statistical and biological criteria were used. Because a gene was selected, the following 4 conditions had to be simultaneously verified:

- a statistical significant expression difference with control at least at one experimental time (Student's t test after Bonferroni's correction, $p<0.0006$);
- a statistical significant dose effect with clear differences between the experimental conditions transfected with no SDF, 5 μg and 20 μg of SDF (Student's t test, $p<0.05$);
- a fold change in respect to control greater than 3, at least at one experimental condition;
- a temporal pattern reflecting the overall temporal pattern of all genes taken together with an early (8 h) and / or intermediate (24 h) upregulation followed from a progressive return to baseline and / or late (72 h) downregulation.

According to these criteria, the following 18 genes were selected. Five genes more specific of the DNA repair pathway (Figure 21): Trex1, Neil2, Parp3, Pms2, Rad51l3. Seven genes more specific of the cell cycle pathway (Figure 22):

Slfn1, Cdkn1a, Prm1, Mdm2, Gpr132, Ddit3, Chek1. Six genes directly involved in both pathways (Figure 23): Sesn2, Rad9, Ppm1d, Atm, Brca1, Brca2.

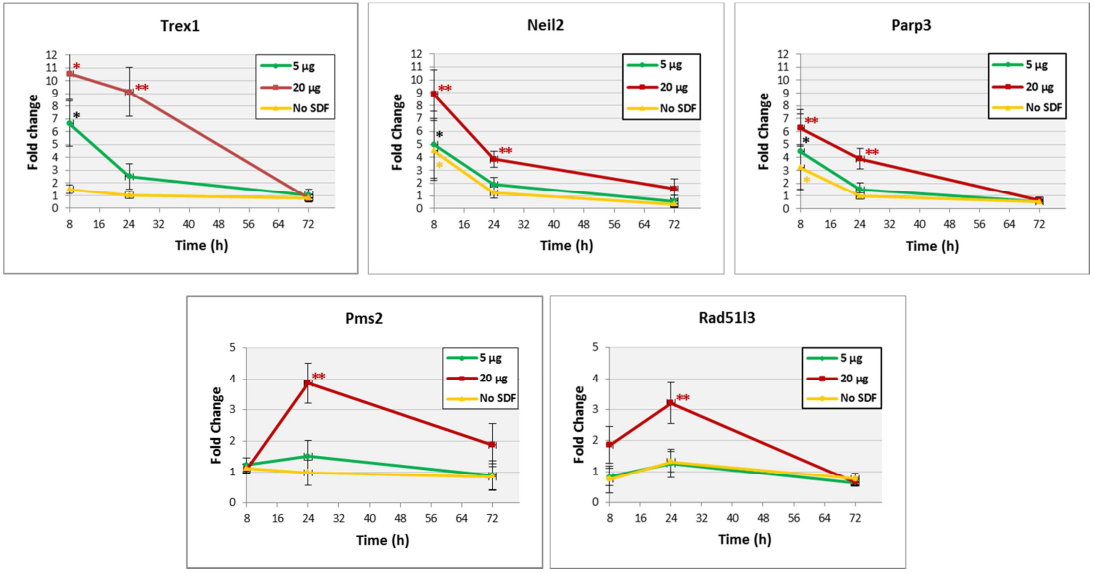


Figure 21 - Temporal pattern after SFHR of selected genes within DNA repair pathway (see text for selection criteria). **Student t-test with Bonferroni's correction, p<0.0006; *Student t-test, p<0.05.

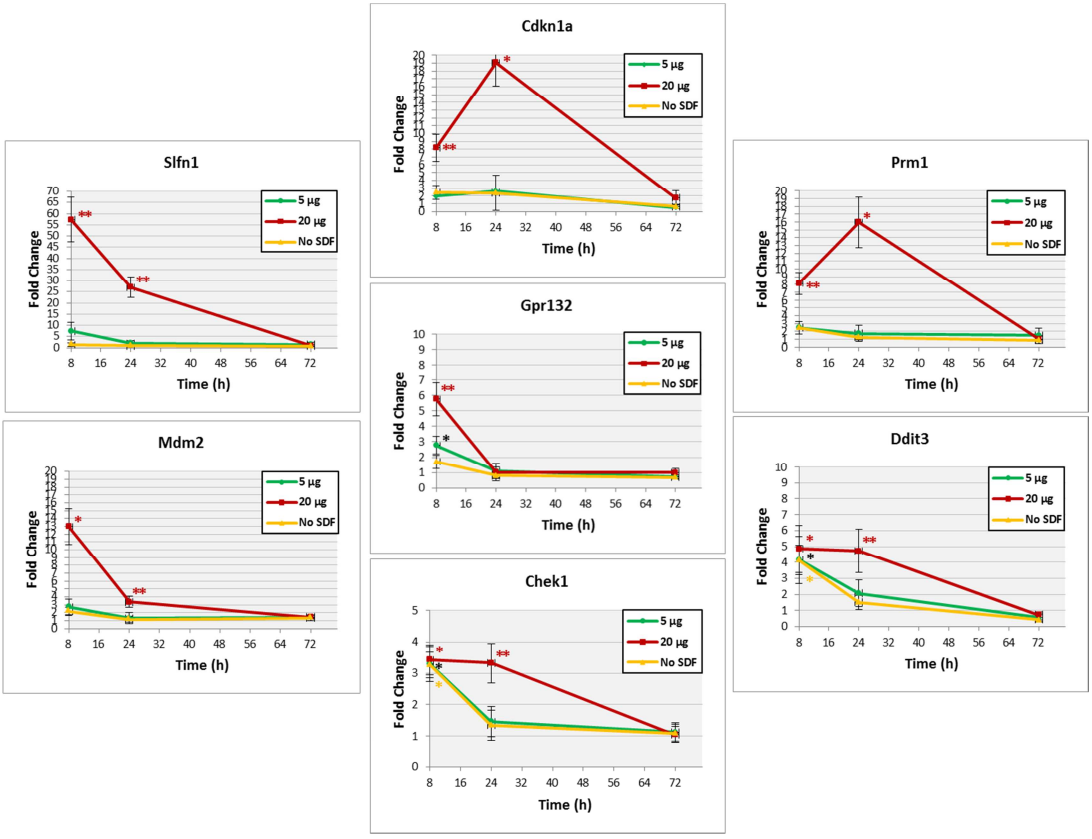


Figure 22 - Temporal pattern after SFHR of selected genes within cell cycle pathway (see text for selection criteria). **Student t-test with Bonferroni's correction, p<0.0006; *Student t-test, p<0.05.

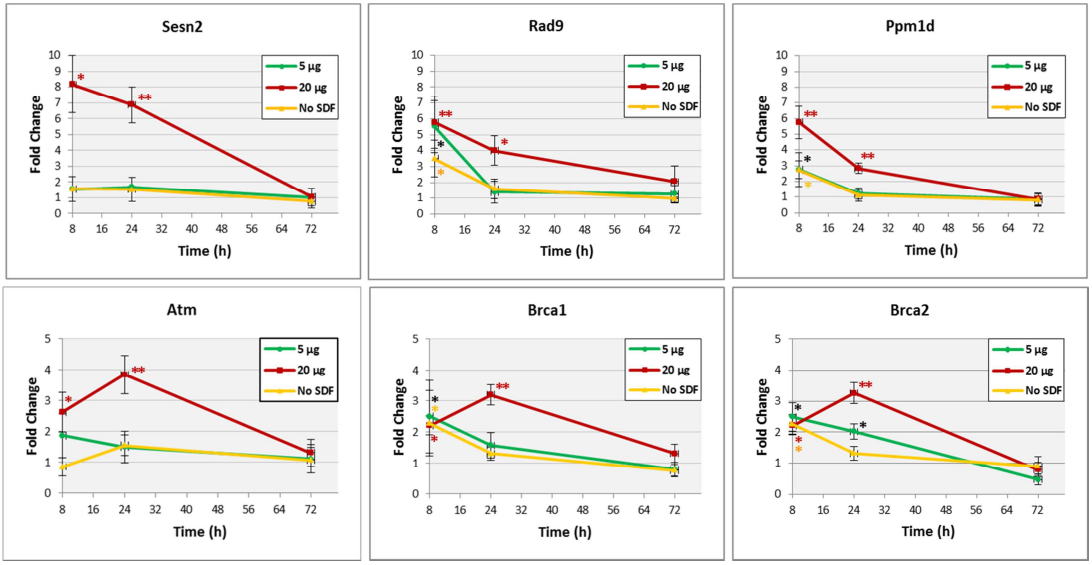


Figure 23 - Temporal pattern after SFHR of selected genes with a role within both the DNA repair and the cell cycle pathways (see text for selection criteria). **Student t-test with Bonferroni's correction, p<0.0006; *Student t-test, p<0.05.

A very synthetic description of the selected genes, along with some links to databases where full information may be found, are reported below.

The **Trex1** (three prime repair exonuclease 1) gene encodes for a protein with both a DNA damage checkpoint activity (upstream ORF) and an exonuclease activity (downsteam ORF). It has a critical role in the MMR pathway and a proofreading function for DNA polymerase. It is the major 3'->5' DNA

exonuclease in mammalian cells, with a direct role in the degradation of genomic double-stranded DNA to minimize potential autoimmune activation by persistent self DNA during cell death.
http://www.genenames.org/data/hgnc_data.php?hgnc_id=12269
http://www.informatics.jax.org/marker/MGI:1328317
http://genatlas.medecine.univ-paris5.fr/fiche.php?symbol=TREX1

The **Neil2** (nei endonuclease VIII-like 2 (E. coli)) gene encodes for a DNA glycosylase which initiate the first step in BER. It recognizes apurinic / apyrimidinic sites. It is involved in repair of lesions in DNA generated during transcription and / or replication, mainly by removing oxidative products of cytosine.
http://www.genenames.org/data/hgnc_data.php?hgnc_id=18956
http://www.informatics.jax.org/marker/MGI:2686058
http://genatlas.medecine.univ-paris5.fr/fiche.php?symbol=NEIL2

The **Parp3** (poly (ADP-ribose) polymerase family, member 3) gene encodes for a protein that belongs to a superfamily of ADP-rybosyl transferases (ARTs) that catalyze the poly(ADP-rybosyl)ation of a limited number of acceptor proteins involved in chromatin architecture and DNA metabolism. It is involved in

detection / signaling pathway leading to the repair of both DNA single-strand breaks (SSBs) by the BER mechanism and DSBs by the NHEJ. It links the DNA damage surveillance network to the mitotic fidelity checkpoint. It is also involved in transcriptional silencing. It is preferentially localized to the daughter centriole throughout the cell cycle.
http://www.genenames.org/data/hgnc_data.php?hgnc_id=273
http://www.informatics.jax.org/marker/MGI:1891258
http://genatlas.medecine.univ-paris5.fr/fiche.php?symbol=PARP3

The **Pms2** (postmeiotic segregation increased 2 (S. cerevisiae)) gene encodes for an endonuclease involved in the DNA MMR. It allows the correction of base-base mismatches and insertion-deletion loops resulting from DNA replication and recombination events.
http://www.genenames.org/data/hgnc_data.php?hgnc_id=9122
http://www.informatics.jax.org/marker/MGI:104288
http://genatlas.medecine.univ-paris5.fr/fiche.php?symbol=PMS2

The **Rad51l3** (RAD51 homolog D (S. cerevisiae)) gene encodes for a protein involved in the DSBs repair by HR. In complexes formed with other members of the RAD51 family, it catalyzes homologous pairing between single- and double-

stranded DNA playing a role in the early stage of recombination-based repair of DNA. It is involved in the disruption of Holliday junctions.
http://www.genenames.org/data/hgnc_data.php?hgnc_id=9823
http://www.informatics.jax.org/marker/MGI:1261809
http://genatlas.medecine.univ-paris5.fr/fiche.php?symbol=RAD51D

The **Slfn1** (schlafen family member 1) gene encodes for a protein that belongs to the Schlafen family of proteins. It is involved in the checkpoint and cell cycle arrest by negative regulation of G1 / S transition.
http://www.genenames.org/data/hgnc_data.php?hgnc_id=25500
http://www.informatics.jax.org/marker/MGI:1313259
http://genatlas.medecine.univ-paris5.fr/fiche.php?symbol=SLFN12

The **Cdkn1a** (cyclin-dependent kinase inhibitor 1A, also known as p21) gene encodes for a protein involved in the checkpoint and cell cycle arrest by negative regulation of G1 / S transition. It binds and inhibits the activity of cyclin-CDK2 or cyclin-CDK1 complexes. Its expression is tightly controlled by p53. It is a negative upstream regulator of DNA methyl transferase 1 (Dnmt1).
http://www.genenames.org/data/hgnc_data.php?hgnc_id=1784
http://www.informatics.jax.org/marker/MGI:104556

http://genatlas.medecine.univ-paris5.fr/fiche.php?symbol=CDKN1A

The **Prm1** (protamine 1) gene encodes for a basic chromosomal protein that substitute for histones in the chromatin of sperm during haploid phase of spermatogenesis. It is also involved in mitotic chromosome condensation (M phase), replacing histones during chromatin remodeling.
http://www.genenames.org/data/hgnc_data.php?hgnc_id=9447
http://www.informatics.jax.org/marker/MGI:97765
http://genatlas.medecine.univ-paris5.fr/fiche.php?symbol=PRM1

The **Mdm2** (p53 E3 ubiquitin protein ligase homolog (mouse), also known as murine double minute 2) gene encodes for a protein involved in the bypassing of G1 checkpoint and in the suppression of the cell cycle arrest. It acts as a negative regulator of p53, achieving p53 repression by binding and blocking the N-terminal trans-activation domain of p53, by acting as an E3 ubiquitin ligase targeting both itself and p53 for degradation, and inhibiting p53 transcriptional activation. Although the major function of Mdm2 is to suppress p53 activities, emerging evidences has identified p53-independenty roles of Mdm2. The overall effect of Mdm2 overexpression is to induce genomic instability through inhibiting DSBs repair and suppressing cell cycle arrest.

http://www.genenames.org/data/hgnc_data.php?hgnc_id=6973
http://www.informatics.jax.org/marker/MGI:96952
http://genatlas.medecine.univ-paris5.fr/fiche.php?symbol=MDM2

The **Gpr132** (G protein-coupled receptor 132) gene encodes a member of the G protein-coupled receptor superfamily. It is able to induce cell cycle arrest in G2 / M transition, to delay mitosis.
http://www.genenames.org/data/hgnc_data.php?hgnc_id=17482
http://www.informatics.jax.org/marker/MGI:1890220
http://genatlas.medecine.univ-paris5.fr/fiche.php?symbol=GPR132

The **Ddit3** (DNA-damage inducible transcript 3, also known as C / EBP homologous protein (CHOP)) gene encodes for a protein which is a basic leucine zipper transcription factor of the dimer forming C / EBP family. It acts mainly as dominant negative regulator of several other transcription factors, but can also induce transcription of downstream target genes. Promoters of target genes showed no common sequence motifs, reflecting that Ddit3 forms heterodimers with several alternative transcription factors that bind to different motifs. It is a stress-inducible nuclear protein involved in G1 / S

checkpoint and cell cycle arrest, as well as in the possible induction of apopotosis.

http://www.genenames.org/data/hgnc_data.php?hgnc_id=2726

http://www.informatics.jax.org/marker/MGI:109247

http://genatlas.medecine.univ-paris5.fr/fiche.php?symbol=DDIT3

The **Chek1** (checkpoint kinase 1) gene encodes for a Serine / Threonine protein kinase which modulate a signaling that, finally, prevents cells proliferation. It plays a role in the G1 / S and intra-S phase checkpoints leading to cell cycle arrest. It phosphorylates the histone H3 inducing a chromatin-mediated transcriptional repression. It acts as an integrator for Atm and Atr signaling and is activated by BRCA1.

http://www.genenames.org/data/hgnc_data.php?hgnc_id=1925

http://www.informatics.jax.org/marker/MGI:1202065

http://genatlas.medecine.univ-paris5.fr/fiche.php?symbol=CHEK1

The **Sesn2** (sestrin 2) gene encodes for a protein implicated in the link between DNA damage and growth. It protects cells against oxidative, genotoxic and energetic stresses. It exerts its cytoprotective function by regenerating overoxidized peroxiredoxins, with a major role in the antioxidant defense of the

cell. The functional interconnection between the response to DNA damage and cell cycle originates from the fact that Sesn2 protein is a target of the p53 and a negative regulator of mTOR (mammalian target of rapamycin), a positive regulator of cell growth that belongs to the phosphatidylinositol kinase-related kinase (PIKK) family.

http://www.genenames.org/data/hgnc_data.php?hgnc_id=20746

http://www.informatics.jax.org/marker/MGI:2651874

http://genatlas.medecine.univ-paris5.fr/fiche.php?symbol=SESN2

The **Rad9** (Rad 9 homolog A (S. pombe)) gene encodes for a protein that plays a role in both DNA repair (HR and MMR by base-pair excision repair) and cell cycle checkpoint and arrest. It possesses a 3'->5' exonuclease activity required for cell cycle arrest at the G2 checkpoint in response to incompletely replicated or damaged DNA. It plays a role in telomere stability and homologous recombinational repair as a mechanism for promoting cell survival after ionizing radiation exposure. It plays a role in locating Claspin to sites of DNA damage, facilitating its role during the Chek1-mediated checkpoint response.

http://www.genenames.org/data/hgnc_data.php?hgnc_id=9827

http://www.informatics.jax.org/marker/MGI:1328356

http://genatlas.medecine.univ-paris5.fr/fiche.php?symbol=RAD9A

The **Ppm1d** (protein phosphatase, Mg2+ / Mn2+ dependent, 1D) gene is a member of the PP2C family of Serine / Threonine protein phosphatases. It negatively contributes to both DNA repair and growth inhibitory pathways activated in response to DNA damage in a p53-dependent manner. It appears to suppress the NER, turn off DNA damage checkpoint responses, restore chromatin structure, as well as inhibit senescence and apoptosis.
http://www.genenames.org/data/hgnc_data.php?hgnc_id=9277
http://www.informatics.jax.org/marker/MGI:1858214
http://genatlas.medecine.univ-paris5.fr/fiche.php?symbol=PPM1D

The **Atm** (ataxia telangiectasia mutated) gene encodes for a serine / threonine protein kinase that belongs to the superfamily of phosphatidylinositol 3-kinase-related kinases (PIKKs). It is recruited and activated by the DSBs. The ATM-mediated DNA damage response consists of both a rapid and a delayed response. The rapid response, including the maintenance of a phosphorylated state of the CDK2-cyclin, results in cell cycle arrest at G1 / S checkpoint, that possibly allows the DSB repair. If DSB cannot be repaired, subsequent phosphorylation events lead to stabilization and activation of p53 and

transcription of several p53 downstream genes, causing long-term cell cycle arrest or even apoptosis.

http://www.genenames.org/data/hgnc_data.php?hgnc_id=795

http://www.informatics.jax.org/marker/MGI:107202

http://genatlas.medecine.univ-paris5.fr/fiche.php?symbol=ATM

The **Brca1** (breast cancer 1, early onset) gene encodes for a multifunctional protein, involved in several pathways, which undergo functional modulation by ATM / ATR kinases. It activates the DNA repair of DSBs by HR in cooperation with, among others, Brca2 and Rad51. It is a negative regulator of cell growth acting at G2 / M checkpoint by activating Chek1 upon DNA damage. It may induce apoptosis through induction of Ddit1 (Gadd45). It interacts with RNA polymerase II and histone deacetylase complexes, playing a role in chromatin remodeling and transcriptional regulation (for example of Cdkn1a (p21)). It also mediates ubiquination.

http://www.genenames.org/data/hgnc_data.php?hgnc_id=1100

http://www.informatics.jax.org/marker/MGI:104537

http://genatlas.medecine.univ-paris5.fr/fiche.php?symbol=BRCA1

The **Brca2** (breast cancer 2, early onset) gene, despite to a very different structure in respect to Brca1, encodes for a multifunctional protein with at least some functions similar and / or interrelated to those of Brca1 protein. It is involved in the DSBs repair by HR pathway; in particular, the localization of Rad51 to the DSBs requires the formation of a Brca1-Palb2-Brca2 complex. It is involved in the cell cycle regulation at G2 / M checkpoint. It has an intrinsic histone acetyltransferase activity and it is involved in chromatin remodeling and transcriptional control; peculiarly, it is involved in transcription-associated recombination.

http://www.genenames.org/data/hgnc_data.php?hgnc_id=1101

http://www.informatics.jax.org/marker/MGI:109337

http://genatlas.medecine.univ-paris5.fr/fiche.php?symbol=BRCA2

5. DISCUSSION

Gene repair strategies attempt to directly correct genetic mutations *in situ*, maintaining gene regulation under the endogenous promoter control. During last years, gene repair approaches received increasing attention as compared to traditional gene therapy strategies, by which additional copies of therapeutic genes are delivered and expressed in transduced cells under the control of strong exogenous promoters (Vasquez et al., 2001; Sorell et al., 2005; Fischer et al., 2008). Furthermore, it is intriguing the possibility of gene targeting in stem cells, a primary aim of gene therapy (Tenzen et al., 2010). Main advantages of the SFHR approach, and gene targeting in general, are the easy production and use of SDFs without the need for complex vectors, the permanent homologous modification of the genomic sequence of interest, the inheritable and physiologically expressed modification and the potential to correct any genetic disease (also dominant negative). However, issues associated with this approach are still numerous and many steps involved in gene correction process are still unknown (Sargent et al., 2011), although it is mandatory that it becomes an efficient, safe and reproducible strategies before its application to clinical medicine. Common drawbacks are the low and variable correction efficiency, the possible random integration of SDF and the

often ineffective transfer of the SDF to the nucleus. In this thesis, an EGFP-based mammalian reporter assay system, in MEF (Luchetti et al., 2012), was used to study the main mechanisms involved in the cellular response to exogenous DNA invasion and influencing the SFHR procedure. To this purpose, epigenetics, DNA repair and cell cycle control mechanisms, as well as their interconnection, were studied.

The topic of the relationship between DNA methylation and SFHR may be dissected at least at 3 levels: the effect of SDF methylation pattern on the efficiency of correction, the effect of correction event on the methylation of corrected genomic locus (*i.e.* after correction) and the effect of the pattern of methylation of the genomic locus to be corrected (*i.e.* before correction) on the efficiency of correction. In effect, a low correction efficiency using methylated SDFs was evidenced in this thesis. This could arise from a still unknown mechanism, possibly involving methyl-binding protein that could inhibit SDF integration within genomic DNA. The use of corrector SDFs with no methylation pattern, among the other characteristics influencing its correction efficiency (Luchetti et al. 2012), is strongly advised. Moreover, during the experiments we noted that both C1 and D1 sorted positive (so, respectively, wt and fluorescent corrected) cells progressively lost fluorescence with cell passages, although a retromutation was excluded by molecular analysis. It

should be noted that D1 corrected cells have been sorted twice: soon after transfection (isolating corrected fluorescent from non-corrected non-fluorescent cells) and after *in vitro* fluorescence decrease (isolating in this case two cell populations both corrected: fluorescent and silenced non-fluorescent). DNA methylation analysis of the integrated EGFP locus evidenced a good correlation between EGFP silencing and *Hpa II* / *Aci I* DNA methylation patterns. The DNA methylation and consequent locus silencing, both reversible by a hypomethylating intervention, appeared to arise with cell replication and culture senescence in both parental C1 and in D1 corrected clone. The methylation patterns resulted site-specific (as *Hpa II* and *Aci I* methylation patterns resulted different) and also qualitatively and quantitatively different in C1 and in D1 re-sorted clones. The first appeared to be more prone to methylate the 3' end of the SFHR target region (*d* amplicon), the second preferentially underwent to methylation of regions upstream and downstream the SFHR target region (*c* and *e* amplicon). It should be taken into account that this effect may depend on the fact that the correction of D1 clone was achieved by using for the correction a non methylated PCR product that can somehow "mark" the zone, which consequently could undergo a slower methylation dynamics during cell replication. However a change in the methylation patterns may specifically arise as consequence of the recombination event possibly

recognized by cellular machinery of defense from invading DNA. Even the hypothesis that the differences in methylation patterns between C1 parental cell line and D1 corrected clone are due to the epigenetic background of the genomic zone where the EGFP construct inserted cannot be ruled out. The DNA methylation that arises after correction leads to corrected locus silencing and, consequently, to an underestimation of correction events (Cuozzo et al., 2007). These results indicate DNA methylation as an experimental variable to be considered, because partially masking the real efficiency of SDF-mediated correction.

Due to evidences, both of our research group and from literature, about the involvement of DNA repair and cell cycle control mechanisms in SFHR, we targeted different proteins involved in these pathways using an ATM kinase inhibitor, a PARP-1 inhibitor and a RNA polymerase II inhibitor. Each treatment was also performed adding a hypomethylating agent. ATM and PARP-1 are well known effectors of the response to DNA damage and of cellular checkpoint activation (Abraham, 2001; Shiloh et al., 2003; Aguilar-Quesada et al., 2007; Krishnamukar et al., 2010). Intriguingly, PARP-1 has recently been linked to the regulation of chromatin structure and transcription, DNA methylation and imprinting, insulator activity and chromosome organization, playing key roles in a number of nuclear processes (Semionov et al., 1999; Althaus, 2005; Bryant

et al., 2006). About the inhibition on RNA polymerase II, that prevent transcription initiation and elongation, it should be taken into account that transcribed parts of the genome are more efficiently repaired and DNA damage is removed faster from transcribed strands (TS) than non- transcribed strands (NTS) (Selby et al., 1993). Only cells simultaneously treated with the hypomethylating agent and the PARP-1 inhibitor, showed a synergistic effect of increase in fluorescence. The influence that PARP-1 inhibition has on SFHR-modification efficacy could be explained by recent findings reporting how PARP-1 inhibition leads to stalled replication fork with consequent formation of DNA double strand breaks that are resolved by HR through ATM activation (Aguilar-Quesada et al., 2007). On the other hand it was also reported that the inhibition of PARP-1 may increase the methylation of genomic DNA (Althaus et al., 2005; Caiafa et al., 2009). This may be the reason for the synergistic effect: only when methylation is prevented, PARP-1 inhibition effect becomes evident. Anyhow a more accessible hypomethylated chromatin may enhance events otherwise not clearly visible. However, on the basis of experimental data, also an opposite and more direct involvement of the PARP-1 repair pathway possibly limiting the efficiency of SFHR may be proposed. The inhibition of PARP-1 might favor the SDF integration (with a mechanism to be discovered), particularly when the chromatin switches to a hypomethylated open

conformation allowing the integration of those residual intracellular SDFs not yet integrated. Additional experiments are needed to distinguish between these different hypothesis.

The studies about the interconnection between SFHR, DNA repair and cell cycle control were then expanded by analyzing a greater number of involved genes. The general results obtained by the quantitative analysis of 84 DNA repair gene expression, point to a peculiar temporal and quantitative pattern depending on additional effects of the nucleofection protocol and of the cellular invasion by exogenous DNA. The nucleofection protocol appears to induce an early (8 h) upregulation of these genes, which progressively return to the basal expression and, finally (72 h), mostly become downregulated. The specific effect of the SDF appeared to be quantitatively additive, in a dose-dependent manner, in respect to that of the nucleofection. At low dosage of SDF, most of the genes appear still upregulated at intermediate experimental times (24 h) but appear to return to basal levels or become downregulated later (72 h). At high dosage of SDF the effect is enhanced from both the quantitative point of view and the persistence of the effect, as still at intermediate time (24 h) there is a further increase of upregulated genes and at late time (72 h) most of the genes still appear upregulated. Using gene-targeting and gene-transfer approaches based on electroporation methods, the relevant effect exerted from

the technique used, additive in respect to that specifically exerted from the SDF, has to be taken into account. On the contrary to current opinions, the genes significantly affected from the SFHR belong to different DNA repair pathways, namely MMR, BER and HR, and not only to HR pathway. Their early upregulation may constitute the molecular basis for the correction, whereas their late downregulation may be responsible for the reduced correction efficiency, possibly as an adaptive response to cell invasion by exogenous DNA. In previous work (Luchetti et al., 2012) Southern blot experiments evidenced the absence of detectable levels of random integration. However the possibility that it can occur below the level of detection of the analytical system could not be dismissed. In this thesis, the quantitative expression analysis of 7 NHEJ genes did not show any statistically significant increase of their expression. This seems to be a confirmation that with this approach there is, at least, no increment of the basal activity of the error-prone NHEJ sub-pathway.

From our previous results (Luchetti et al., 2012) emerged that SDF shows a greater probability to access the target locus during G2 / M phase possibly because some cellular characteristics in this phase such as the absence of nuclear envelope, the tetraploid status of the cells, the loosely packing of chromosome and the higher activity of HR DNA repair pathway. The quantitative expression study of 84 cell cycle genes reported in this thesis,

revealed a wide impact of SFHR on the cell cycle regulation. A dramatic effect of cell cycle arrest, delayed up to the latest experimental time (72 h) was evidenced to be induced from both the nucleofection protocol and the SDF. As for DNA repair genes, the quantitative effect in terms of both gene number and level of expression resulted to be higher and more prolonged if SDF is involved, in a dose dependent manner, in respect to the effect induced by the nucleofection protocol alone. The impact seems to be exerted on each of the 3 main cellular checkpoints, as the cell cycle genes with the expression modulated by the SFHR resulted to be functionally linked to both G1 / S and G2 / M transitions, as well as to the intra-S-phase checkpoint. The effect of cell cycle arrest well correlates with the observation that most of the modulated genes are upregulated and that, among them, almost all are well recognized negative regulators of checkpoints and cell cycle. These results are in excellent agreement with those obtained for DNA repair genes. In fact cellular checkpoint activation and cell cycle arrest are mandatory for DNA damage repair. From this point of view, this expression enhancement of negative regulators of cell cycle may constitute part of the molecular basis for the correction. In the meantime, the induced modulation of cell cycle genes may be insufficient, from quantitative and / or qualitative point of view, to warrant higher levels of correction. The marked SDF dose effect on the expression of both DNA repair

and cell cycle genes, well correlates with the enhanced correction efficiency reached at high SDF dosage.

Among the studied genes, 18 were selected as the best target / effectors of SFHR. Analyzing their functional role, what emerged is that appear very difficult to assign most of them either to the DNA repair or to the cell cycle regulation pathway. Often these genes have a dual (sometimes multiple) role. This further stresses the interconnection between SFHR, DNA repair effectors and cellular checkpoint regulators and clearly indicates that only an integrated study approach may highlight the network. Interestingly, 2 of the selected genes, Trex1 and Rad9, have a 5'->3' exonuclease activity on double strand DNA. Their hypothesized favorable role in enhancing the SFHR correction efficiency may be based on the fact that Trex1 have a role in MMR and in a proofreading function for DNA polymerase and Rad9 have a role in MMR and HR, as well as a in the cell cycle arrest acting at the G2 / M checkpoint. However, a contrasting hypothesis may be posed that their exonucleasic activity may degrade the SDF with consequent unfavorable effects on SFHR efficiency. Additional experiments are needed to distinguish between these two working hypothesis.

6. CONCLUSIONS

These results contribute to the comprehension of the molecular mechanisms underlying cell invasion by exogenous DNA and its genomic integration.

The three main pathways of cell cycle, DNA methylation and DNA repair appear to mediate the cellular response to this invasion.

The interplay of these pathways and their specific temporal patterns appear to influence the correction efficiency of SFHR.

The selection of specific molecular targets to manipulate, provides suggestions for increasing gene repair efficiency to achieve an higher correction for a practical SFHR application to *ex vivo* therapeutic approaches.

REFERENCES

1. Abraham RT (2001). Cell cycle checkpoint signaling through the ATM and ATR kinases. Genes Dev 15: 2177-2196.
2. Aguila-Quesada R, Muñoz-Gamez JA, Martin-Oliva D, Peralta A, Valenzuela MT et al. (2007). Interaction between ATM and PARP-1 in response to DNA damage and sensitization of ATM deficient cells through PARP inhibition. BMC Mol Biol 8:29.
3. Althaus FR (2005). Poly(ADP-ribose): a co-regulator of DNA methylation?. Oncogene 24:11-2.
4. Auriche C, Di Domenico EG, Pierandrei S, Lucarelli M, Castellani S, Conese M, Melani R, Zegarra-Moran O, Ascenzioni F (2010). CFTR expression and activity from the human CFTR locus in BAC vectors, with regulatory regions, isolated by a single-step procedure. Gene Ther 17(11):1341-54.
5. Bedayat B, Abdolmohamadi A, Ye L, Maurisse R, Parsi Het et al. (2009). Sequence-specific correction of genomic hypoxanthine-guanine phosphoribosyl transferase mutations in lymphoblasts by small fragment homologous replacement. Oligonucleotides 20:7-16.
6. Boris-Lawrie K, Temin HM (1994). The retroviral vector. Replication cycle and safety considerations for retrovirus-mediated gene therapy. Annals of the New York Academy of Science 716:59–70.
7. Bryant HE, Helleday T (2006). Inhibition of poly (ADP-ribose) polymerase activates ATM which is required for subsequent homologous recombination repair. Nucleic Acid Research 27:4526-4531.

8. Brinster RL, Chen HY, Trumbauer ME, Senear R, Warren R, Palmiter RD (1985). Factors affecting the efficiency of introducing foreign DNA into mice by microinjecting eggs. Proc Natl Acad Sci USA 82:4438-4442.
9. Brummelkamp, TR, Bernards R, Agami R (2002). A system for stable expression of short interfering RNAs in mammalian cells. Science 296:550–553.
10. Bryant HE, Helleday T (2006). Inhibition of poly (ADP-ribose) polymerase activates ATM which is required for subsequent homologous recombination repair. Nucleic Acids Research 34: 1685–1691.
11. Bukrinsky MI, Haffar OK (1999). HIV-1 nuclear import: in search of a leader. Front Biosci 4:772-781.
12. Caiafa P, Zampieri M (2009). Epigenetics: poly(ADP-ribosyl)ation of PARP-1 regulates genomic methylation patterns. FASEB J 23:672-8.
13. Capecchi MR (1989). Altering the genome by homologous recombination. Science 244:1288-92.
14. Capecchi MR (1994). Targeted gene replacement. Scientific American March 52-59.
15. Cavazzana-Calvo M, Hacein-Bey S, de Saint Basile G et al. (2000). Gene therapy of human severe combined immunodeficiency (SCID). X1 disease Science 288:669-672.
16. Cheung P, Lau P (2005). Epigenetic regulation by histone methylation and histone variants. Mol Endocrinol 19:563-573.
17. Colosimo A, Goncz KK, Holmes AR, Kunzelmann K, Novelli G, Malone RW, Bennett MJ, Gruenert DC (2000). Transfer and expression of foreign genes in mammalian cells. Biotechniques 29:314-331.

18. Colosimo A, Goncz KK, Novelli G, Dallapiccola B, Gruenert DC (2001). Targeted correction of a defective selectable marker gene in human epithelial cells by small DNA fragments. Molec Therapies 3:178-185.
19. Colosimo A, Guida V, Antonucci I, Bonfini T, Stuppia L, Dallapiccola B (2007). Sequence-specific modification of a beta-thalassemia locus by small DNA fragments in human erythroid progenitor cells. Haematologica 92:29-30.
20. Colot V, Rossignol JL (1999). Eukaryotic DNA methylation as an evolutionary device. Bioessays 21:402-411.
21. Conese M, Boyd AC, Di Gioia S, Auriche C, Ascenzioni F (2007). Genomic context vectors and artificial chromosomes for cystic fibrosis gene therapy. Curr Gene Ther 7: 175–187.
22. Constantinou A, Davies AA, West SC (2001). Branch migration and Holliday junction resolution catalyzed by activities from mammalian cells. Cell 104, 259–268.
23. Cromie GA, Connelly JC, Leach DR (2001). Recombination at double-strand breaks and DNA ends: conserved mechanisms from phage to humans. Mol Cell 8:1163–1174.
24. Crystal RG, McElvaney NG, Rosenfeld MA, et al. (1994). Administration of an adenovirus containing the human CFTR cDNA to the respiratory tract of individuals with cystic fibrosis. Nat Genet 8:42-51.
25. Cuozzo C, Porcellini A, Angrisano T, Morano A, Lee B, et al. (2007). DNA damage, homology-directed repair, and DNA methylation. Plos Genet 3: e110.
26. D'Adda di Fagagna F (2008). Living on a break: cellular senescence as a DNA-damage response. Nat Rev Cancer 7:512–522.

27. Daicumakos EG (1973). Methods for micromanipulation of human somatic cells in culture. Methods in Cell Biology, Academic Press New York p287-311.
28. Davis BR, Yannariello-Brown J, Prokopishyn NL, Lou Z, Smith MR, Wang J, Carsrud NDV, Brown DB (2000). Glass needle-mediated microinjection of macromolecules and transgenes into primary blood stem/progenitor cells. Blood 95:437-444.
29. De Pamphilis ML, Herman SA, Martinez-Salas E, Chalifour LE, Wirak DO, Cupo DY, Miranda M (1988). Microinjecting DNA into mouse ova to study DNA replication and gene expression and to produce transgenic animals. BioTechniques 6:662-679.
30. Dikomey E, Dahm Daphi J, Brammer I, Martensen R, Kaina B (1998). Correlation between cellular radiosensitivity and non-repaired double-strand breaks studied in nine mammalian cell lines. Int J Radiat Biol 73, 269–278.
31. Eggler AL, Inman RB, Cox MM (2002). The Rad51-dependent pairing of long DNA substrates is stabilized by replication protein A. J Biol Chem 277, 39280–39288.
32. Einfeld DA, Roelvink PW (2002). Advances towards targetable adenovirus vectors for gene therapy. Curr Opin Mol Ther 4: 444–51.
33. Fang WH, Modrich P (1993). Human strand-specific mismatch repair occurs by a bidirectional mechanism similar to that of the bacterial reaction. J Biol Chem 268, 11838–11844.
34. Fischer A, Cavazzana-Calvo M (2008). Gene therapy of inherited diseases. Lancet 371:2044-7.

35. Flotte TR, Carter B, Conrad C, et al. (1996). A phase I study of an adeno-associated virus CFTR gene vector in adult CF patients with mild disease. Human Gene Ther 7:1145-49.
36. Flotte TR, Laube B (2001). Gene therapy in cystic fibrosis. Chest 120: 124S-131S .
37. Friend DS, Papahadjopoulus D, Debs RJ (1996). Endocytosis and intracellular processing accompanying transfection mediated by cationic liposomes. Biochim Biophys Acta 1278:41-50.
38. Gao X, Huang L (1995). Cationic liposome-mediated gene transfer. Gene Ther 21: 710-22.
39. Goncz KK (2000). Site-direct alteration of genomic DNA by small fragment homologous replacement. Methods Mol Biol 133:85-99.
40. Goncz KK, Colosimo A, Dallapiccola B, Gagne L, Hong K,Novelli G, Papahadjopoulos D, Sawa T, Schreier H, Wiener-Kronish J, Xu Z, Gruenert DC (2001). Expression of DeltaF508 CFTR in normal mouse lung after site-specific modification of CFTR sequences by SFHR. Gene Ther 8: 961-965.
41. Goncz KK, Gruenert DC (1999). Site-directed alteration of DNA by small fragment homologous replacement (SFHR). Gene Targeting Vector Protocols, Humana Press Totowa NJ p85-99.
42. Goncz KK, Gruenert DC (2000). Site-directed alteration of genomic DNA by small-fragment homologous replacement. Methods Mol Biol 133, 85–99.
43. Goncz KK, Kunzelmann K, Xu Z, Gruenert DC (1998). Targeted replacement of normal and mutant CFTR sequences in human airway epithelial cells using DNA fragments. Hum Mol Genet 7, 1913–1919.
44. Goncz KK, Prokopishyn NL, Abdolmohammadi A, Bedayat B, Maurisse R, Davis BR, Gruenert DC (2006). Small fragment homologous replacement-

mediated modification of genomic beta-globin sequences in human hematopoietic stem/progenitor cells. Oligonucleotides 16: 213-224.

45. Goncz KK, Prokopishyn NL, Chow BL, Davis BR, Gruenert DC (2002). Application of SFHR to gene therapy of monogenic disorders. Gene Ther 9: 691-694.
46. Goto T, Monk M (1998). Regulation of X-chromosome inactivation in development in mice and humans. Microbiol Mol Biol Rev 62, 362-378.
47. Gruenert DC (1998). Gene Correction with small DNA fragments. Curr Res Molec Therapeutics: Gene Therapy and Oligonucleotides 1: 607-613.
48. Gruenert DC (1999). Opportunities and challenges in targeting genes for therapy. Gene Ther 6: 1347–1348.
49. Haber JE (2000). Partners and pathways repairing a double-strand break. Trends Genet 16, 259–264.
50. Hammarsten O, De Fazio LG, Chu G (2000). Activation of DNA-dependent protein kinase by single-stranded DNA ends. J Biol Chem 275, 1541–1550.
51. Hickson I, Zhao Y, Richardson CJ, Green SJ, Martin NM, Orr AI, Reaper PM, Jackson SP, Curtin NJ, Smith GC (2004). Identification and characterization of a novel and specific inhibitor of the ataxia-telangiectasia mutated kinase ATM. Cancer Res 64:9152-9159.
52. Holliday R (1964). A mechanism for gene conversion in fungi. Genet Res 5:282–304.
53. Jaenisch R, Bird A (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 33 Suppl 245-254.
54. Jeggo PA, Hafezparast M, Thompson AF, Broughton BC, Kaur, GP, Zdzienicka MZ, Athwal RS (1992). Localization of a DNA repair gene

(XRCC5) involved in double-strand-break rejoining to human chromosome 2. Proc Nat Acad Science USA 89, 6423–6427.

55. Johnson RD, Jasin M (2000). Sister chromatid gene conversion is a prominent double-strand break repair pathway in mammalian cells. EMBO J 19:3398–3407.

56. Kapsa RM, Quigley A, Lynch GS et al. (2001). In vivo and in vitro correction of the mdx dystrophin gene nonsense mutation by short-fragment homologous replacement. Hum Gen Ther 12:629-624.

57. Kapsa RM, Quigley A, Vadolas J, Steeper K, Ioannou PA, Byrne E, Kornberg AJ (2002). Targeted gene correction in the mdx mouse using short DNA fragments: towards application with bone marrow-derived cells for autologous remodeling of dystrophic muscle. Gene Ther 9:695-699.

58. Kass SU, Pruss D, Wolffe AP (1997). How does DNA methylation repress transcription?. Trends Genet 13: 444-449.

59. Kawasaki H, Taira K (2004). Induction of DNA methylation and gene silencing by short interfering RNAs in human cells. Nature 431:211-217.

60. Klein TM, Fitzpatrick-McElligott S (1993). Particle bombardment: a universal approach for gene transfer to cells and tissues. Curr Opin Biotechnol 4:583-590.

61. Krishnakumar R, Kraus WL (2010). The PARP side of the ucleus: molecular actions, physiological outcomes and clinical targets. Mol Cell 39:8-24.

62. Kunzelmann, K, Legendre JY, Knoell DL, Escobar LC, Xu Z, Gruenert DC (1996). Gene targeting of CFTR DNA in CF epithelial cells. Gene Ther 3:859–867.

63. Lakshmipathy U, Pelacho B, Sudo K, Linehan JL, Coucouvanis E, Kaufman DS, Verfaillie CM (2004). Efficient transfection of embryonic and adult stem cells. Stem Cells 22:531–543.

64. Lenz P, Bacot SM, Frazier-Jessen MR, Feldman GM (2003). Nucleoporation of dendritic cells: efficient gene transfer by electroporation into human monocyte-derived dendritic cells. FEBS Lett 538:149–154.
65. Li E, Beard C, Jaenisch R (1993). Role for DNA methylation in genomic imprinting. Nature 366:362-365.
66. Lindahl T, Wood RD (1999). Science 286, 1897.
67. Lips J and Kaina B (2001). DNA double-strand breaks trigger apoptosis in p53-deficient fibroblasts. Carcinogenesis 22:579– 585.
68. Luchetti A, Filareto A, Sanchez M, Ferraguti G, Lucarelli M, Novelli G, Sangiuolo F, Malgieri A (2012). Small Fragment Homologous Replacement: Evaluation of Factors Influencing Modification Efficiency in an Eukaryotic Assay System. PLoS One 7(2):e30851.
69. Maasho K, Marusina A, Reynolds NM, Coligan JE, Borrego F (2004). Efficient gene transfer into the human natural killer cell line, NKL, using the Amaxa nucleofection system. J Immunol Methods 284:133-40.
70. Maloisel L, Rossignol JL (1998). Suppression of crossing-over by DNA methylation in Ascobolus. Genes Dev 12:1381-1389.
71. Martensson S and Hammarsten O (2002). DNA-dependent protein kinase catalytic subunit. Structural requirements for kinase activation by DNA ends. J Biol Chem 277:3020–3029.
72. Matsui H, Johnson LG, Randell SH, Boucher RC (1997). Loss of binding and entry of liposome-DNA complexes decreases transfection efficiency in differentiated airway epithelial cells. J Biol Chem 272:1117-1126.
73. Maurisse R, Cheung J, Widdicombe J, Gruenert DC (2006). Modification of the pig CFTR gene mediated by small fragment homologous replacement. Annals of the New York Academy of Science 1082: 120-123.

74. McNally MA, Lebkowski JS, Okarma TB, Lerch LB (1988). Optimizing electroporation parameters for a variety of human hematopoietic cell lines. BioTechniques 6:882-885.
75. Modrich P, Lahue R (1996). Mismatch repair in replication fidelity, genetic recombination, and cancer biology. Ann Rev Biochem 65:101–133.
76. Muller K, Heller H, Doerfler W (2001). Forein DNA integration. Genome-wide perturbations of methylation and transcription in the recipient genomes. J Biol Chem 276: 14271-14278.
77. Nagasawa H, Brogan JR, Peng Y, Little JB, Bedford JS (2010). Some unsolved problems and unresolved issues in radiation cytogenetics: a review and new data on roles of homologous recombination and non-homologous end joining. Mutat Res 701:12–22.
78. Nemunaitis J, Cunningham C, Buchanan A, Blackburn A, Edelman G, Maples P, Netto G, Tong A, Randlev B, Olson S, Kirn D et al. (2001). Intravenous infusion of a replication-selective adenovirus (ONYX-015) in cancer patients: safety, feasibility and biological activity. Gene Ther 8: 746-759.
79. New JH, Sugiyama T, Zaitseva E, Kowalczykowski SC (1998). Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A. Nature 391:407–410.
80. Nishikawa M, Huang L (2001). Nonviral vectors in the new millennium: delivery barriers in gene transfer. Hum Gene Ther 12: 861-870.
81. Ogris M (2003). Non-viral cancer gene therapy: what is best?. Drug Discov Today 8: 63.
82. Okano M, Bell DW, Haber DA, Li E (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99:247-257.

83. Orkin SH (1986). Molecular genetics and potential gene therapy. Clin Immunol Immunopathol 40: 151–6.
84. Pegg AE, Dolan ME, Moschel RC (1995). Structure, function, and inhibition of O6 -alkylguanine–DNA alkyltransferase. Prog. Nucleic Acid Res. Mol Biol 51:167–223.
85. Pfeiffer P, Goedecke W, Obe, G (2000). Mechanisms of DNA double-strand break repair and their potential to induce chromosomal aberration. Mutagenesis 15:289–302.
86. Plass C, Soloway PD (2002). DNA methylation, imprinting and cancer. Eur J Hum Genet 10:6-16.
87. Quinonez R, Sutton RE (2002). Lentivirus vectors for gene delivery into cells. DNA Cell Biol 21:937-951.
88. Reeves WH, Sthoeger ZM (1989). Molecular cloning of cDNA encoding the p70 (Ku) lupus autoantigen. J Biol Chem 264:5047–5052.
89. Remus R, Kammer C, Heller H, Schmitz B, Schell G, Doerfler W (1999). Insertion of foreing DNA into an established mammalian genome can alter the methylation of cellular DNA sequences. J Virology 73:1010-1022.
90. Rich T, Allen RL, Wyllie AH (2000). Defying death after DNA damage. Nature 407:777–783.
91. Robertson KD, Wolffe AP (2000). DNA methylation in health and disease. Nat Rev Genet 1(1):11-9.
92. Rotman G, Shiloh Y (1999). ATM from gene to function. Hum Mol Genet 7: 1555-63.
93. Sanford JC, Smith FD, Russell JA (1993). Optimizing the biolistic process for different biological applications. Methods in Enzymology Academic Press San Diego CA:p217.

94. Sangiuolo F, Bruscia E, Serafino A, Nardone AM, Bonifazi E, Lais M, Gruenert DC, Novelli G (2002). In vitro correction of cystic fibrosis epithelial cell lines by small fragment homologous replacement (SFHR) technique. BMC Med Genet 3:42.
95. Sangiuolo F, Filareto A, Spitalieri P, Scaldaferri ML, Mango R, et al. (2005). In vitro restoration of functional SMN protein in human trophoblast cells affected by spinal muscular atrophy by small fragment homologous replacement. Hum Gene Ther 16:869-880.
96. Sangiuolo F, Scaldaferri ML, Filareto A, Spitalieri P, Guerra L et al. (2008). CFTR gene targeting in mouse embryonic stem cells mediated by Small Fragment Homologous Replacement (SFHR). Frontiers in Bioscience 1: 2989-2899.
97. Sargent RG, Kim S, Gruenert DC (2011). Oligo/Polynucleotide-based gene modifications: strategies and therapeutic potential. Oligonucleotides 21:55-75.
98. Scott-Taylor TH, Gansbacher B, Sadelain M. (1998). Efficient retroviral infection of human cells utilising an adenoviral vector expressing the ecotropic receptor. Adv Exp Med Biol.;451:423-30.
99. Selby CP, Sancar A (1993). Molecular mechanism of transcription-repair coupling. Science 260: 53-58.
100. Semionov A, Cournoyer D, Chow TYK (1999). Inhibition of poly(ADP-ribose) polymerase stimulates extrachromosal homologous recombination in mouse Ltk-fibroblasts. Nucleic Acid Research 27:4526-4531.
101. Semionov A, Cournoyer D, Chow TYK (2003). 1,5-isoquinolinediol increases the frequency of gene targeting by homologous recombination in mouse fibroblasts. Biochemistry and Cell Biology 81: 17-24.

102. Shen YM, Hirschhorn RR, Mercer WE, Surmacz E, Tsutsui Y, Soprano KJ, Baserga R (1982). Gene transfer:DNA microinjection compared with DNA transfection with a very high efficiency. Mol Cell Biol 2:1145-1154.
103. Shiloh Y (2003). ATM and related protein kinases: safeguarding genome integrity. Nat Cancer Rev 3: 155-68.
104. Silman NJ, Fooks AR (2000). Biophysical targeting of adenovirus vectors for gene therapy. Curr Opin Mol Ther 2: 524–31.
105. Smithies O, Gregg RG, Boggs SS, Koralewski MA, Kucherlapati RS (1985). Insertion of DNA sequences into the human chromosomal B-globin locus by homologous recombination. Nature 317: 230-234.
106. Sonoda E, Takata M, Yamashita YM, Morrison C, Takeda S (2001). Homologous DNA recombination in vertebrate cells. Proc Nat Acad Science USA 98:8388–8394.
107. Sorrell DA, Kolb AF (2005). Targeted modification of mammalian genomes. Biotechnol Adv 23:431-69.
108. Spitalieri P, Cortese G, Pietropolli A, Filareto A, Dolci S, et al. (2009). Identification of multipotent cytotrophoblast cells from human first trimester chorionic villi. Cloning and Stem Cell 11:535-546.
109. Stasiak AZ, Larquet E, Stasiak A, Muller S, Engel A, Van Dyck E, West SC, Egelman EH (2000). The human Rad52 protein exists as a heptameric ring. Curr Biol 10:337–340.
110. Sullenger BA, (2003). Targeted genetic repair: an emerging approach to genetic therapy. The Journal of clinical investigation 112(3):310-1.
111. Suzuki MM, Bird A (2008). DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet 9:465-476.
112. Takata M, Sasaki MS, Sonoda E, Morrison C, Hashimoto M, Utsumi H, Yamaguchi-Iwai Y, Shinohara A, Takeda S (1998). Homologous

recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. EMBO J 17:5497–5508.
113. Tenzen T, Zembowicz F, Cowan CA (2010). Genome modification in human embryonic stem cells. J Cell Physiol 222:278-81.
114. Thomas KR, Capecchi MR (1986). Targeting of genes to specific sites in the mammalian genome. Cold Spring Harb Symp Quant Bio:3472755.
115. Thomas KR, Folger KR, Capecchi MR (1986). High frequency targeting of genes to specific sites in the mammalian genome. Cell 44:419-428.
116. Thorpe P, Stevenson BJ, Porteous DJ (2002). Optimising gene repair strategies in cell culture.Gene Ther 9:700-2.
117. Todaro M, Quigley A, Kita M, Chin J, Lowes K, Kornberg AJ, Cook MJ, Kapsa R (2007). Effective detection of corrected dystrophin loci in mdx mouse myogenic precursors. Human Mutation 28: 816-823.
118. Trompeter HI, Weinhold S, Thiel C, Wernet P, Uhrberg M (2003). Rapid and highly efficient gene transfer into natural killer cells by nucleofection. J Immunol Methods 1: 245-56.
119. Tsai TF, Bressler J, Jiang YH, Beaudet AL (2003). Disruption of the genomic imprint in trans with homologous recombination at Snrpn in ES cells. Genesis 37:151-161.
120. Umar A, Kunkel TA (1996). DNA-replication fidelity, mismatch repair and genome instability in cancer cells. Eur J Biochem 238,297–307.
121. Van Dyck E, Stasiak AZ, Stasiak A, West SC (1999). Binding of double-strand breaks in DNA by human Rad52 protein. Nature 398, 728–731.
122. Vasquez KM, Marburger K, Intody Z, Wilson JH (2001). Manipulating the mammalian genome by homologous recombination. Proc Natl Acad Sci USA 98:8403-10.

123. Venkitaraman AR (2009). Linking the cellular functions of BRCA genes to cancer pathogenesis and treatment. Ann Rev Pathol 4:461–487.
124. Vermeulen W, de Boer J, Citterio E, van Gool AJ, van der Horst GT, Jaspers NG, de Laat WL, Sijbers AM, van der Spek PJ, Sugasawa K et al. (1997). Mammalian nucleotide excision repair and syndromes. Biochem Soc Trans 25, 309– 315.
125. Woodcock CL (2006). Chromatin architecture. Curr Opin Struct Biol 16:213-220.
126. Yanez RJ, Porter AC (1998). Therapeutic gene targeting. Gene Ther 2.
127. Yang NS, Sun WH (1995). Gene gun and other non-viral approaches for cancer gene therapy. Nat Med 1:481-483.
128. Yang Y, Haecker SE, Su Q., et al. (1996). Immunology of gene therapy with adenoviral vectors in mouse skeletal muscle. Hum Mol Genet 5:1703–12.
129. Yoder JA, Walsh CP, Bestor TH (1997). Cytosine methylation and the ecology of intragenomic parasites. Trends Genet 13:335-340.
130. Zabner J, Coulture LA, Gregory RJ, Graham SM, Smith AE, Welsh MJ (1993). Adenovirus-mediated gene transfer transiently corrects the chloride transport defect in nasal epithelia of patients with cystic fibrosis. Cell 75:207:216.
131. Zayed H, McIvor RS, Wiest DL, Blazar BR (2006). In vitro functional correction of the mutation responsible for murine severe combined immune deficiency by small fragment homologous replacement. Hum Gene Ther 17:158-166.
132. Zelein AV, Alimov AA, Zelenina IA, Semenova ML, Rodova MA, Chernov BK, Kolesnikov VA (1993). Transfer of foreign DNA into the cells of developing mouse embryos by microprojectile bombardment. FEBS Lett 315:29-32.