SISSA
Scuola Internazionale Superiore di Studi Avanzati

Neuroscience Area – PhD course in

Molecular Biology

High-throughput screening to identify human microRNAs enhancing CRISPR/Cas9-based homologous recombination for cardiac gene correction

Candidate:

Edoardo Schneider

Advisor:

Mauro Giacca

Co-advisors:

Anna Cereseto

Thierry Pedrazzini

Academic Year 2018-19

Joint PhD Programme in Molecular Biology (JuMBO)

"High-throughput screening to identify human microRNAs enhancing CRISPR/Cas9-based homologous recombination for cardiac gene correction"

PhD student

Edoardo Schneider

International Centre for Genetic Engineering & Biotechnology

(ICGEB)

Trieste, Italy

Director of studies: Prof. Mauro Giacca

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The definitive treatment of genetic diseases through precise gene editing has been a long sought goal of gene therapy, yet hitherto unachieved at the clinical level. The CRISPR/Cas9-based technology ensures simple and efficient modification of the cellular genome and thus promises to change the perspective of treatment of hereditary disorders, including genetic cardiomyopathies. However, in mammals, gene correction through the homologous recombination (HR) machinery is largely less efficient than gene inactivation through the error prone, non-homologous end joining (NHEJ) route. This is particularly true in adult post-mitotic tissues, such as the heart, since HR commonly relies on S to G2/M-phase transition.

The main goal of this project is to identify genetic treatments able to enhance frequency of CRISPR/Cas9-induced HR *in vitro* and in the heart. By robotic high-throughput, high-content microscopy we systematically screened a library of 2,024 human microRNAs to search for regulators of HR-mediated gene correction. We identified 20 miRNAs that significantly increase CRISPR/Cas9-induced HR events compared to controls ($P<0.001$). Interestingly, 10 among the top identified miRNAs belong to only two associated miRNAs families sharing the same seed sequence. A common and distinctive feature of these miRNAs is to induce accumulation of key proteins of the HR

pathway, including MRE11, NBN, RAD50 and RAD51. Another highly effective miRNA, not belonging to either of these two families, regulates expression of p38 beta, suggesting involvement of this MAPK in the HR process.

To study the effect of these miRNAs in the cardiac context, we first designed a CRISPR-AAV-based HR detection tool allowing precise, in frame insertion of a promoterless GFP gene into the last exon of the Myosin regulating light chain 2 (Myl2) gene in Cas9-expressing cardiomyocytes. The selected miRNAs also markedly enhanced Cas9-induced HR in this model. The same approach was then tested *in vivo* by producing AAV9 particles coding for the HR reporter system and the miRNAs and systemically injecting them in newborn Cas9+ mice. Also in this setting, the selected miRNAs increased the HR frequency up to 4-folds over control.

Together, these results are encouraging in indicating that transient cardiac treatment with miRNAs enhancing HR together with the use of pro-recombinogenic AAVs might be sufficient to increase gene correction to a therapeutically sufficient level.

1. INTRODUCTION

1.1 Genome editing and engineering through gene therapy

1.1.1. How cells repair DNA damages

1.1.1.1 Non-homologous end joining (NHEJ)

The non-homologous end joining (NHEJ) pathway, as defined for the first time in 1993 by Moore and coworkers (Moore and Haber 1996), corresponds to the imprecise correction of a diverse range of DNA substrates at double-strand breaks (DSBs). This correction mechanism resembles the classical pathway by which the cells repair DNA damages using a three-step process, which involves a nuclease, a polymerase and a ligase. Upon DSBs formation, the Ku proteins recognize and bind the two DNA strands (Mimori and Hardin 1986). These proteins form a heterodimer (Ku70 and Ku80), which protects the DNA ends from degradation and promotes the recruitment of the DNA-dependent protein kinase catalytic subunit (DNA-PKcs), having serine-threonine kinase activity, and Artemis. This complex recognizes and cleaves different types of DNA damaged overhangs. Artemis:DNA.PKcs presents 5'-endonucleolitic activity that produces nicks inside the 5' overhang in order to create blunt DNA ends (Ma, Pannicke et al. 2002). Therefore, Artemis modifies the cleaved DNA overhangs to create substrates for the enzymatic ligation with the Ligase 4 complex (Yannone, Khan et al. 2008), which has the ability to bind a variety of DNA ends (Gu, Lu et al. 2007). Once this occurs, through their BRCT domains, the error-prone polymerases mu and lambda (pol x), can bind the Ku:DNA complex and synthesize the new DNA sequences by adding rNTPs or dNTPs to the resected strands (Ma, Lu et al. 2004). Finally, the Ligase 4 complex catalyzes the last step of the DSB resolution, which corresponds to the rejoining of the DNA ends. In summary, the NHEJ pathway consists of a very conserved DNA repair pathway, the components of which share an extensive tolerance in the configuration of the DNA recognized targets. Consequently, this DNA repair process is flexible and multifunctional.

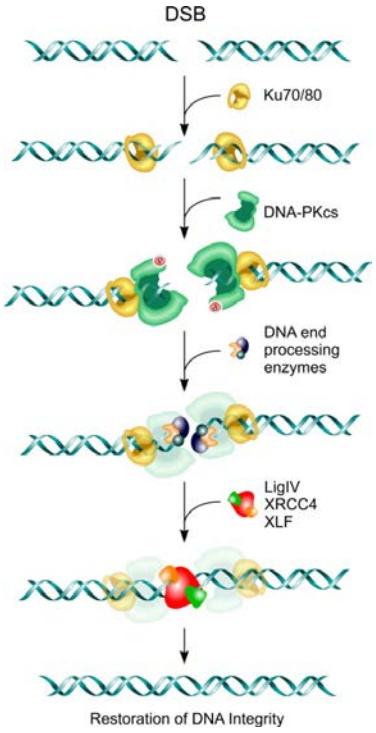


Figure 1: The NHEJ pathway is a very conserved DNA repair pathway, the components of whcih share an extensive tolerance in the configuration of the recognized DNA targets.

1.1.1.2 Single-strand DNA annealing (SSA)

Another mechanism by which cells repair DNA insults consists of the SSA pathway, which uses homologous repeats to solve the DSBs. The activation of this pathway involves the annealing of specific repeats flanking the DSBs, which produces deletions. One of the first evidences proposing SSA as a specific mode for DNA repair comes from the work of Sternberg and colleagues (Lin, Sperle et al. 1984). In their model, DSBs occur between homologous repeats and generate 3′ ssDNA, which are then annealed to create an intermediate. It has been shown that the generation of these 3′ ssDNA sequences at the DSB locus corresponds to a key step for SSA activation. According to this information, it was demonstrated that this pathway requires a DNA end resection factor called CtIP (Bennardo, Cheng et al. 2008). The DNA repair process then requires a ligation step and a subsequent endonucleolytic cleavage of the non-homologous 3′ ssDNA sequences, which are mediated by RAD52 (Symington 2002) and by the

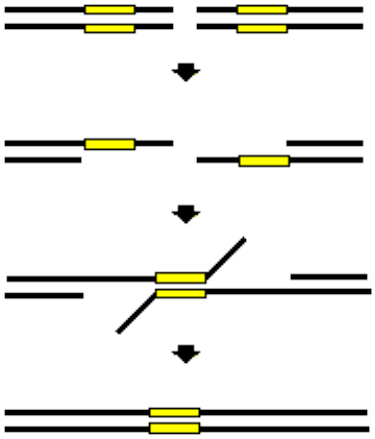


Figure 2: Model of Single Strand Anealling: The annealing of specific repeats The DNA repair process requires a ligation step and a subsequent endonucleolytic cleavage of the non-homologous 3' ssDNA sequences, which are mediated by RAD52 (Symington 2002) and by the ERCC1/XPF complexflanking the DSBs, which produces deletions.

ERCC1/XPF complex (Motycka, Bessho et al. 2004). Eventually, the activation of a polymerase results critical to fill the gaps.

1.1.1.3 Homologous recombination (HR)

Homologous recombination has always been the gold standard for gene therapy applications. Nevertheless, this is a very inefficient DNA repair mechanism in higher eukaryotes, which rarely occurs in dividing cells and is almost shutoff in differentiated and post-mitotic cells. Accordingly, it is known that many crucial components of the HR pathway are transcriptionally regulated depending on the different cell cycle phases (Cho, Campbell et al. 1998). During meiosis, HR drives many crucial biological processes such as prevention of the collapsing of damaged DNA replication forks (Michel, Grompone et al. 2004), chromosome segregation during the first meiotic division and telomere maintenance (McEachern and Haber 2006). Most relevant, the main role of HR during mitosis is to repair both spontaneous and induced DSBs with the substitution of damaged DNA with a similar or identical template sequence. The first evidence

of how this pathway works came from the study of Orr-Weaver and colleagues, who demonstrated how a non replicating plasmid, presenting a DSB and, within the same region, sequences with homology for the yeast genome, was able to efficiently recombine upon transformation (Orr-Weaver, Szostak et al. 1981). According to this study, the 3′ ssDNA tail of the DSB end invades the homologous duplex in the corresponding yeast genomic site and is then extended by DNA synthesis. The 3′ ssDNA invasion of the homologous duplex creates a displacement loop (D-loop), which induces binding of the tail with the sequence on the complementary side of the template DNA duplex by Watson-Crick interactions. On the other hand, the non-invading 3′ ssDNA strand is then completed by DNA synthesis. Consequently, this forms a double-Holliday-junction (dHJ) intermediate (Holliday 2007), which will generate, by random resolution, an identical number of crossover and noncrossover products. During meiosis, the production of crossover products corresponds to the asymmetric resolution of this intermediate, thus leading to the formation of new allelic combinations (Maloisel, Bhargava et al. 2004). This DNA correction model was defined as DSB repair (DSBR) (Szostak, Orr-Weaver et al. 1983) and corresponds to the canonical way by which cells repair DNA damages during meiosis. Indeed, it has been shown that DSBR pathway activation frequently occurs in hotspots for meiotic recombination, where DSBs can be generated (Kolodkin, Klar et al. 1986). Nevertheless, this system does not represent the only mechanism by which HR can repair DSBs. Another possible sub-pathway, defined as synthesis-dependent strand annealing (SDSA) (Ferguson and Holloman 1996), was proposed in 1996. Different from DSBR, SDSA corresponds to a conservative replication process, leading to the formation of non-crossover products, by which only one of the 3′ssDNA tail invades the homologous duplex, without the generation of the dHJ (Miura, Yamana et al. 2012). Upon strand invasion, a polymerase synthetizes the new DNA in the 5′ to 3′ direction. This step therefore creates a single HJ, which then dissociates from the DNA duplex allowing binding of the amplified displaced sequence to the other 3′ ssDNA on the other side of the DSB. Once the displaced DNA anneals with the opposite DNA strand

on the cleavage site, DNA synthesis fills the gaps and completes the repair process.

It is possible to classify the key components of HR pathways into two main groups: proteins associated with the DNA damage response (DDR) that work as a complex composed by MRE11, RAD50 and NBS1 (MRN complex) and proteins

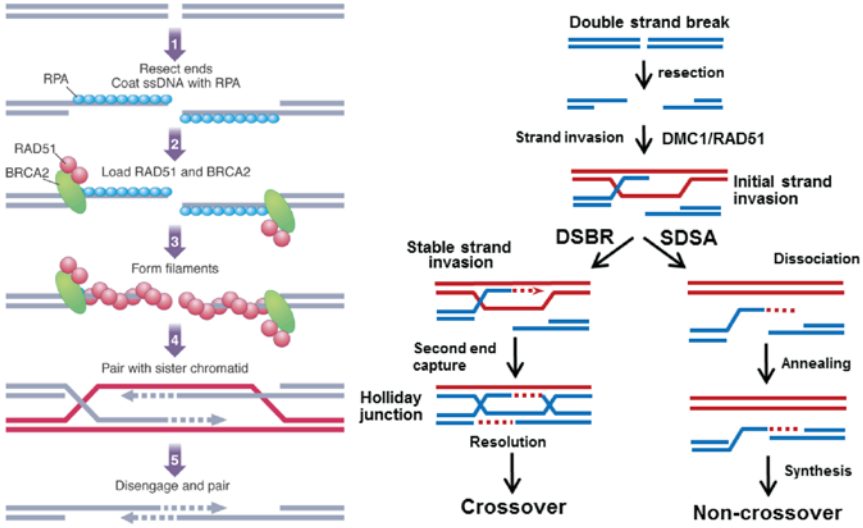


Figure 3: After the generation of a double strand break, cells can repair the damage through the HR pathway if a template DNA is present inside the cell. Two different sub-pathways (DSBR or SDSA) trigger the HR-mediated DSB resolution; they share the initial phases of cleavage recognition and stabilization, but differ in the consequent repair step. Depending on the specific process, HR-mediated DNA repair can generate crossover or non-crossover products.

specifically associated with HR (RAD51, RAD52, RAD54, RAD55, RAD57, RAD59, RDH54/TID1). As a first step, after DSBs formation, the MRN complex recognizes the damage and associates with the CtIP protein, enzyme that catalyzes the 5'-3' end resection in order to generate the 3' ssDNA tails (Davies, Forment et al. 2015).

Another key protein directly associated with the initialization of the HR process is replication protein A (Ward, Norat et al.). This protein is a heterotrimeric DNA-

binding protein that recognizes and binds ssDNA sequences at the DSBs loci, prevents ssDNA degradation by cellular nucleases (Chen and Wold 2014) and promotes generation of 3' ssDNA tails (Yan, Toczylowski et al. 2011). After DSBs formation, RPA is rapidly replaced, through the activity of RAD52, with the most important component of the pathway, the ATP-dependent recombinase RAD51. This protein exists as a ring composed by 7 monomers that binds the ssDNA and possesses motives able to hydrolyze ATP. After the first nucleation event, other RAD51 subunits bind at the end of the forming filament to build a long protein polymer defined as presynaptic filament (Bianco, Tracy et al. 1998).

Another key regulator of the HR pathway known to directly regulate RAD51 activity is BRCA2 (San Filippo, Chi et al. 2006). Indeed, multiple evidences underlines how this protein directly interacts with RAD51 and triggers ssDNA recognition and binding. Accordingly, Moynahan and collaborators demonstrated that, by introducing mutations in the BRCA2 gene, they could abolish the RAD51 foci formation, which directly reflects the synthesis of the presynaptic filament (Moynahan, Pierce et al. 2001).

Subsequently, RAD51 catalyzes binding of the DNA duplex (forming the synaptic complex), strand invasion and pairing of the ssDNA with the homologous sequences in the genome (Morrical 2015). The homologous template invasion is allowed by the intervention of accessory factors such as Rad54, Rad54B, Rdh54 and Mnd1 (Li and Heyer 2009), allowing the consequent formation of the D-loop. Then, upon Rad54-mediated RAD51 removal, polymerase δ induces the elongation of the invading DNA strand and, upon disassembling of the synaptic complex, DSBs can be resolved by either DSBR or SDSA sub-pathway activation.

1.1.2 Programmable nucleases as tool for precise gene modifications

Gene editing, by definition, indicates the possibility to insert, eliminate or modify specific gene sequences in the genomes of living organisms. Nevertheless, trying to correct inherited mutations, knockdown specific genes or inserting exogenous sequences, both *in vitro* and *in vivo*, still is a very

inefficient process, far away from therapeutic applications. Because of this, there is a strong need for new strategies aimed at improving the efficiency by which gene editing occurs. The discovery of programmable endonucleases, able to introduce site specific DSBs inside eukaryotic genomes, dramatically changed the prospective of gene editing. Zinc-finger endonucleases (ZFNs), transcriptional activator-like effector nucleases (TALENs) and, more recently, the clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 system, are some examples of this technology.

1.1.2.1 Zinc-finger endonucleases (ZFNs)

Zinc-finger endonucleases are the first example of the manipulation of DNA binding modules, associated with nucleases, employed to introduce precise DNA cleavages inside eukaryotic genomes. The Cys_2-His_2 zinc-finger (ZF) domain represents a well-characterized DNA binding domain, normally expressed in eukaryotic genomes. The ZF is composed by 30 amino acids organized in a conserved ββα configuration (Beerli and Barbas 2002), able to recognize and bind 3 bp sequences at the α-helix surface. Due to the modular characteristics of this DNA binding domain, it results able to specifically recognize all the 64 possible triplets combination. Moreover, a variety of ZF modules can be link together in order to recognize longer DNA sequences (Bhakta, Henry et al. 2013). Consequently, the possibility to create custom high specific DNA binding modules and link them with nucleases presenting a cleavage domain, such as FokI, allow researchers to deal with a technology virtually able to target and cleave any DNA sequence. These tools can be employed to create nicks, but, in order to induce DSBs inside the genome, ZFNs work as a dimers, recognizing and cutting the preselected DNA strands depending on the different zinc-finger arrays. These tools have been largely employed for genome editing approaches and many modifications have been developed during the years to improve the DNA cleavage specificity and efficiency. As an example, in the paper of Guo and collaborators, the authors developed a hyperactivated FokI cleavage domain, able to improve the

efficiency of DNA cleavage up to 15-fold if compared to the classical ZFNs (Guo, Gaj et al. 2010).

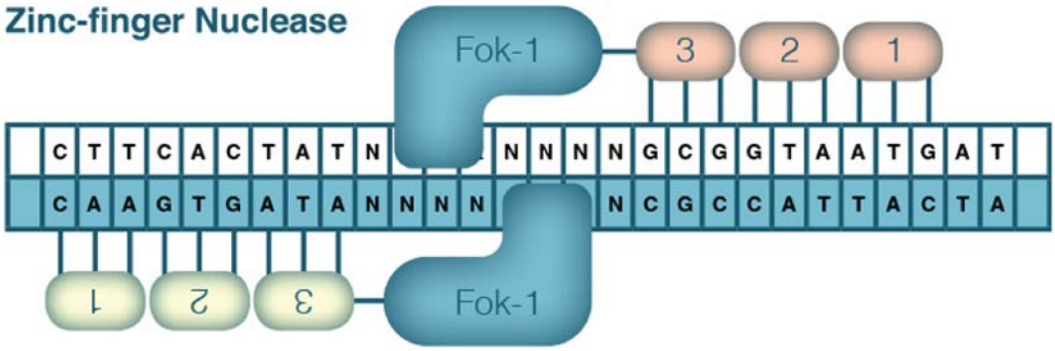


Figure 4: ZNFs are the first example of the manipulation of DNA binding modules, associated with nucleases. They are employed to introduce precise DNA cuts inside virtually any sequence in the genome. The Cys_2-His_2 zinc-finger (ZF) domain represents a well-characterized DNA binding domain, normally expressed in eukaryotic genomes, which confers cleavage specificity of ZNFs.

1.1.2.2 Transcriptional activator-like effector nuclease (TALEN) endonucleases

In nature, the TAL effectors (TALEs) are DNA binding proteins produced by specific bacteria. They are composed of 33-34 amino acids, which are highly variable in positions 13^{th} and 14^{th}. These two are defined as repeat variable di-residues (RVDs) and result crucial for high specific nucleotide recognition (Boch, Scholze et al. 2009, Moscou and Bogdanove 2009). Due to their DNA binding characteristics, different TAL effectors can be linked together and fused with endonucleases (TALENs) to generate high specific DNA cleavages inside genomes. Hockemeyer and co-workers provided one of the first evidence regarding the possibility to introduce precise DSBs inside the human genome using the TALEN technology (Hockemeyer, Wang et al. 2011). These authors modified human ES and iPS cells, by targeting the PPP1R12C, OCT_4 and $PITX_3$ genes using a TALEN approach and comparing the results with the ones obtained using the ZFNs technology (Hockemeyer, Soldner et al. 2009). As in

the case of ZFNs, TALENs work as a dimer to introduce two DSBs, each one recognizing one specific DNA strand.

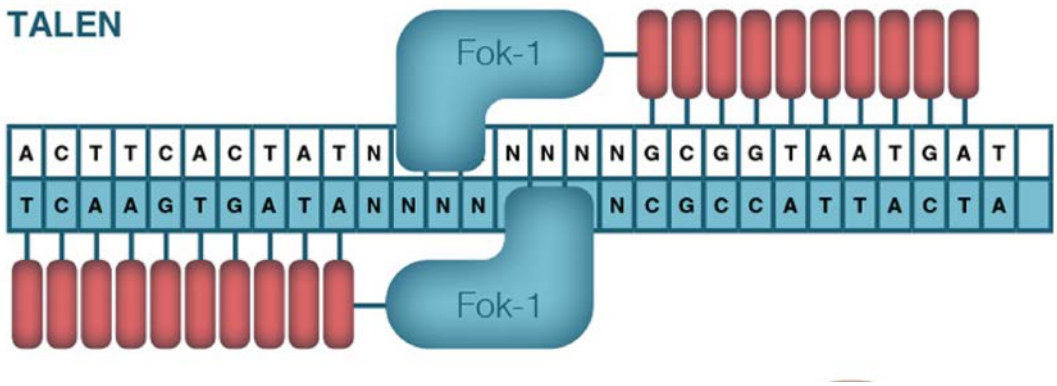


Figure 5: Diverse TAL effectors, known to recognize specific nucleotides inside the genome, can be linked together and fused with endonucleases (TALENs) to generate high specific DNA cleavages.

1.1.2.3 Clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 endonuclease

1.1.2.3.1 History of CRISPR/Cas9 discovery

The first annotation regarding CRISPR was published in the paper of Nakata and coworkers in 1987, when these authors recognized a particular set of 21-37 bp repeated sequences, interspaced by non-repetitive 32 bp spacers, in the *E. coli* iap gene, (Ishino, Shinagawa et al. 1987). Since then, the analysis of many other bacterial genomes has confirmed the presence of the same tandem repeats (40% of the sequenced genomes). For this reason, in the paper of Mojica and collaborators, these sequences were first classified as a unique family of clustered repeat elements (Mojica, Diez-Villasenor et al. 2000) and, in a second work, as CRISPR (Jansen, Embden et al. 2002). In the same publication, authors also showed the presence of conserved sequences adjacent to the tandem repeats, which were defined as CRISPR-associated (cas) genes. These publications lay the bases for the next classification of three different categories of the identified CRISPR system (type I-III). Furthermore, to understand the importance of the CRISPR system, an essential contribution was provided in 2005 from the work of Mojica and collaborators, in which these authors

demonstrated the extrachromosomal and viral-related origin of the spacer sequences (Mojica, Diez-Villasenor et al. 2005). Additionally, they also underlined how archeal cells expressing specific spacers were resistant to the infection of different phages. In the same year, the function of the protospacer adjacent motif (PAM), another CRISPR essential component, was also defined. Bolotin and collaborators discovered how the presence of a short conserved 3-nt sequence (PAM), was crucial to confer the type II Cas9 DNA cleavage specificity (Bolotin, Quinquis et al. 2005).

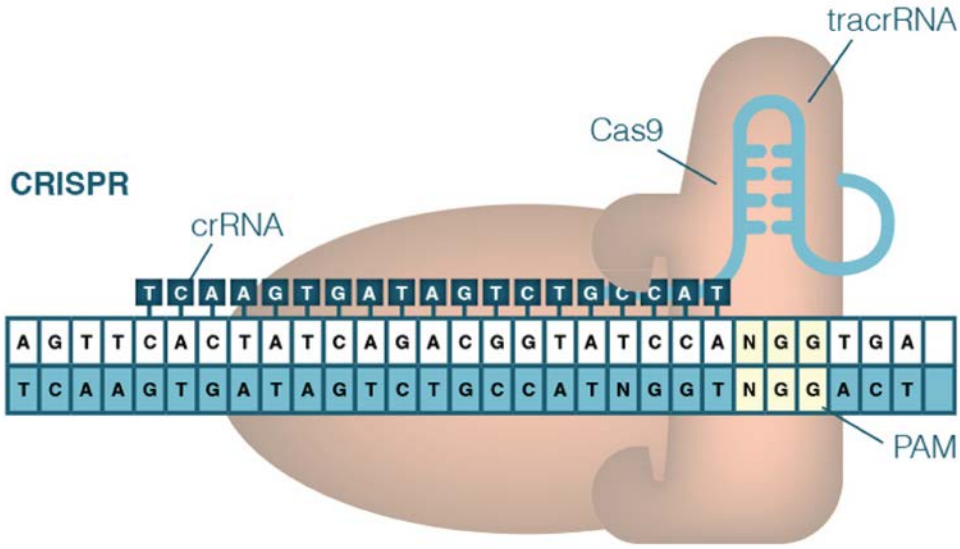


Figure 6: Thanks to the Waston-crick interactions between the gRNA, which corresponds to a crRNA fused with the tracrRNA and the target locus, the Cas9 can induce high specific DNA cleavage. Nevertheless it presents many limitations, it still results the most efficient tool to stimulate both precise knockdown and knock-in events in the eukaryotic genome.

Collectively, all these information allowed researchers to define the CRISPR system as a sort of bacterial immune response against viral infections (Pourcel, Salvignol et al. 2005). The first confirmation of this theory came from the work of Barrangou and coworkers, who, in 2007, confirmed the CRISPR system as a site-specific endonuclease, whose activity is controlled by the CRISPR spacers (crRNAs) (Barrangou, Fremaux et al. 2007). Three years later, the bacterial

mechanistic role of the CRISPR system was definitively clarified (Garneau, Dupuis et al. 2010). Then, in 2011, the role of another key component of CRISPR type II system biogenesis was elucidated. Delcheva and collaborators described the function of a trans-activating crRNA (tracrRNA) able to hybridize with the crRNAs and therefore allow the final maturation of the CRISPR array into a mature and active form (Deltcheva, Chylinski et al. 2011).

Despite the great enthusiasm that was elicited by this novel site-specific DNA cleavage system, genome-engineering applications were still far to be explored. One of the first evidence in this respect of this came in 2011, when Siksnys and coworkers, trying to control the endonucleolytic activity of the CRISPR system, demonstrated the possibility to transfer the *Streptococcus thermophilus* type II CRISPR system into *Escherichia coli (*Sapranauskas, *Gasiunas et al. 2011)*. A critical turning point came in 2013, when Cong and collaborators demonstrated the first successful adaptation of an engineered type II *Streptococcus thermophilus* CRISPR system to induce genome editing in mammalian cells (Cong, Ran et al. 2013). By using a single guide RNA (sgRNA), which corresponds to a crRNA fused with the tracrRNA, these authors were able to drive the Cas9 DNA cleavage activity and therefore stimulate eukaryotic cellular NHEJ or HR repair pathways.

CRISPR/Cas9 technology shows a series of advantages if compared with the previously established gene editing tools. It is very easy to customize, displays a very broad range of cleavage pattern and, most notably, the editing efficiency results much more improved, since target DNA recognition is mediated by nucleic acid hybridization rather than by protein-DNA binding, as it is the case for TALEN and ZNF. This also allows the simultaneous targeting of multiple genomic loci.

1.1.2.3.2 Natural mechanism of action

The first explanation of how the CRISPR system works inside bacteria cells came from the work of Mojica and collaborators, who, in 2003, showed how some spacers of the *E. Coli* CRISPR array perfectly matched with specific sequences

inside P1 phage genomes. This observation was confirmed looking at the same correspondence between many other spacers and known viral sequences. Taken together, these data allowed researchers to postulate a theory by which the CRISPR system is as a sort of bacterial adaptive immune response against phage infection. In other terms, prokaryotes evolved a nucleic acid-based immunity system, by which the specificity is determined by the CRISPR spacers and the enzymatic activity depends on the Cas enzymes. After cell invasion, phages release their genome and, as a consequence, Cascade and Cas3 enzymes start to acquire specific spacers that are then loaded iton the CRISPR locus inside the prokaryotic genome. This phase, known as "adaptation", represents the first step of the CRISPR-Cas mediated defense process and corresponds to the segregation of spacers between repeated sequences. There are two different forms of adaptation, depending on the origin of the spacer. If it derives from a not previously met phage the process is defined "naïve". In contrast, when a pre-existing evidence of viral infection is present, the adaptation is defined as "primed" (Fineran and Charpentier 2012). The Cas1 and Cas2 enzymes, both nucleases, control the spacer integration. In support to this notion, Yosef and coworkers demonstrated how the overexpression of Cas1 and 2 resulted in a significant increase in the spacer integration events. On the other hand, these authors also demonstrated that mutations in the active site totally abolish the adaptation process (Yosef, Goren et al. 2012). Cas1 and 2 work as a complex formed by a dimer of Cas2 that recognize two dimers of Cas1. Recent studies by Nunez and collaborators have defined how, in the *E. coli* Type I-E CRISPR system, the complex formed by Cas1 and 2 is able to integrate sequences into the CRISPR array using a retroviral integrase and transposase approach (Nunez, Lee et al. 2015). Moreover, Cas9 also participates in the integration process, possibly by driving the site-specific spacer integration. Finally, the adaptation also requires a spacer selection based on specific criteria. In support to this notion, it was demonstrated that the PAM sequence plays a key role also during this phase, allowing cells to discriminate between self and non-self (Mojica, Diez-Villasenor et al. 2009). The resulting CRISPR array is a noncoding RNA

transcript that has to pass through different maturation steps, depending on the different CRISPR system type (crRNA biogenesis and processing), before becoming active. During this step, the CRISPR RNA precursor (pre-crRNA) is first transcribed and then processed into the mature form by Cas proteins and other factors.

By comparing the three types of CRISPR systems, they all share a series of similarities in crRNA conformation and processing. Types I and III both use Cas6 as the key enzyme during crRNA maturation, excluding the Type I-C, which employs Cas5d (Nam, Haitjema et al. 2012). Nevertheless, the type II crRNA maturation is driven by a totally different mechanism that employs Cas9, cellular RNAse III and the base pairing with the tracrRNA (Deltcheva, Chylinski et al. 2011). The last step corresponds to the "interference", which represents the combined action of crRNA and Cas proteins that recognize and cleave the phage DNA (Brouns, Jore et al. 2008). Both Type I and II systems employ the PAM sequence to avoid self-CRISPR array DNA cleavage and therefore to induce the non-self activation (Semenova, Jore et al. 2011). The mature crRNAs are loaded on the Cascade complex, which allows the unspecific DNA binding and scanning, searching for possible PAM sequences throughout the viral genome. Once the crRNA binds DNA, this causes a conformational change of the Cascade complex, leading to Cas3 activation (Wiedenheft, Lander et al. 2011). This enzyme is able to nick the target DNA, which then undergoes degradation (Westra, van Erp et al. 2012). In contrast, the type II system only requires the Cas9 activity, driven by the crRNA and tracrRNA complex. The crRNA forms a heteroduplex with the target DNA sequence, which is then cleaved thanks to the activity of Cas9 HNH and RuvC nucleolytic domains on the complementary and non-complementary strands (Nishimasu, Ran et al. 2014).

1.1.2.3.3 Diversity between Cas9 orthologs

Since its discovery, a high level of conservation of the CRISPR system was evident among different bacterial species. In order to investigate CRISPR evolution, Fonfara and coworkers performed a bioinformatics research using a

BLAST approach to compare previously described Cas9 sequences with publicly available genome databases (Fonfara, Le Rhun et al. 2014). The results revealed the presence of 653 sequences corresponding to Cas9 orthologs, associated with 347 different bacterial species. Consequently, the comparison of the cas gene sequences, the direction of the CRISPR array and the sequence of different tracrRNAs allowed researchers to recognize three different main groups of Type II CRISPR-Cas: type II-A, corresponding to the *S. thermophilus* or *S. piogenes* Cas9 enzymes; type II-B, referring to the *F. novicida*; type II-C, typical of *P. multocida* and *N. meningitis*. These findings confirmed how the type II CRISPR loci result largely variable between species, but largely conserved through the different groups (Chylinski, Le Rhun et al. 2013).

Different Cas9 orthologs may also require different PAM recognition site. To characterize the different PAM motifs using the same BLAST approach, the group of Charpentier identified a number of conserved sequences downstream the target recognition site (Fonfara, Le Rhun et al. 2014). This information allowed researchers to elucidate the wide range of ortholog CRISPR systems characterized by specific Cas9 enzymes driven by specific dual-RNA structures and PAMs. The identification of novel Cas9 orthologs, showing different specificity, can therefore lay the basis for the generation of novel genome editing tools.

1.1.2.3.4 Adaptation of Cas9 system to perform genome editing in eukaryotic cells

Since the discovery of the CRISPR system as a programmable gene editing endonuclease (Marraffini and Sontheimer 2008), there has been growing interest in the possibility to modify Cas9 to transpose its activity in mammalian cells. Nevertheless, the huge biological differences between prokaryotic and eukaryotic cells may represent a great limit. The first evidences reporting a human codon-optimized *S. pyogenes* Cas9 (SpCas9), carrying nuclear localization signals, derived from the publications of Zhang and Doudna laboratories (Cong, Ran et al. 2013, Jinek, East et al. 2013). In the first paper, the

authors described how, by co-transfecting inside 293FT cells the tracrRNA and crRNA specific for the EMX1 locus together with the codon-optimized SpCas9 and the RNAse III (SpRNAse III) genes, they were able to efficiently target the EMX1 locus. The generation of DSBs in the target site was then evaluated using the SURVEYOR assay. Moreover, the same authors also observed that the presence of the SpRNAse III was not mandatory to achieve crRNA maturation and therefore DNA cleavage. Additionally, they also performed the same gene editing assays using an engineered chimeric single guide RNA (sgRNA) structure generated by fusing the tracrRNA and crRNA sequences. The second main contribute to the field came from the publication of Doudna and collaborators in 2013, when they described an efficient RNA-guided and site-specific DBS formation in CLTA locus using a type II Cas9 bacterial protein in human cells. In the same year, Church and collaborators demonstrated the possibility to target seven eukaryotic loci and underlined that both NHEJ and HR pathway are activated to repair the cleaved site (Mali, Yang et al. 2013).

1.1.2.3.5 CRISPR/Cas9 system engineering

Since the first gene editing application of CRISPR/Cas9 in eukaryotic cells, many other laboratories started employing this tool. Despite the increasing enthusiasm due to the number of advantages of this technique, different limitations were identified. Off-target effects, the generation of mosaicism, or the creation of multiple alleles due to the DSB repair through NHEJ, represent some example of these limitations. To overcome these drawbacks, different newly engineered versions of the Cas9 system have been proposed. Since it is known that the Cas9 enzyme cleaves DNA on both strands using two catalytic domains (HNH and RuvC) and creating blunt-ended DSBs, by mutating the sequence of one of these two, the laboratory of Zhang and collaborators demonstrated the possibility to generate a site-specific nickase, the activity of which causes a nick on one strand of the target DNA (Ran, Hsu et al. 2013). This modified Cas9 can therefore create single-stranded breaks, which are then repaired preferentially through base excision repair (BER) or HR pathway.

Considering the great number of off-targets that characterize the classical Cas9 activity, by generating a two Cas9 nickase system targeting opposite strands, Zhang and coworkers were able to introduce a site-specific DSB, minimizing the off-target activity and stimulating the HR-mediated repair. Additionally, in the paper published in 2014 from Joung and collaborators, in the attempt to minimize the off-target effects of Cas9, the authors showed a significant improvement in specificity using a truncated version of the gRNA combined with pairs of Cas9 nickases (Fu, Sander et al. 2014).

Cas9 can be also used to alter the expression of specific genes, without inducing their knockout. An engineered version of Cas9, carrying mutations in HNH and RuvC-I sites and therefore defined as nuclease-deficient Cas9 (dCas9), was described in 2013 by Lim and collaborators (Qi, Larson et al. 2013). This group proved that the catalytically-dead Cas9 protein still maintained its DNA binding specificity leading to an alteration of the transcription of the targeted gene. This strategy, defined as CRISPR interference (CRISPRi), was then confirmed to repress targeted gene expression in *E. Coli* cells, without the identification of any off-target effects.

On the other hand, different laboratories also explored the possibility to use the Cas9 enzyme to boost the transcription of target genes. Many CRISPR/dCas9-mediated gene activation strategies have been developed to target specific promoters or induce the enhancing of regions of interest. One or many activation domains can be linked to the dCas9. Indeed, in the work of Gersbach and collaborators, they fused a SgRNA driven dead Cas9 with the VP64 acid transactivator domain, which corresponds to four repetition of the *Herpes Simplex* viral protein 16, and showed how its administration to human cells was able to specifically induce the expression of the target gene (Perez-Pinera, Kocak et al. 2013). Nevertheless, directly manipulate the epigenetic states at specific regions correlate with the transcription activation of target genes. Indeed, by fusing a dCas9 together with the transcription activator acetyltransferase p300, a protein driving the histone H3 lysine 27 acetylation, in the paper of Hilton and collaborators, they demonstrated the Il1RN, MYOD and

OCT-4 transcription enhancing (Hilton, D'Ippolito et al. 2015). Moreover, another technology initially developed for imaging purposes, is the fusion of a dCas9 together with a protein scaffold carrying peptide epitope array (SunTag). Specific Abs fused with transactivators can recognize this array and consequently activate the transcription of the dCas9-targeted genes. This was proved in the paper of Gilbert and co-workers where they screened a list of sgRNAs matching in proximity of the transcription start site (TSSs) of 49 genes that correlates with cellular susceptibility to ricin (Qi, Larson et al. 2013). Nevertheless, fusing different transactivator to the Cas9 is not the only way to achieve targeted gene expression modulation. Indeed, in the paper of Zalatan and co-workers they connected to the sgRNA specific viral RNA sequences, which are recognized by MCP, PCP and Com RNA-binding protein. By fusing the VP64 transactivator to these proteins, they demonstrated a significant activation of target genes transcription. Moreover, by fusing to the same proteins repression domains (such as the KRAB domain), they underlined a corresponding transcription inhibition (Zalatan, Lee et al. 2015).

Another important addition to the Cas9 technology system refers to the characterization of Cas9 from different bacterial species. In particular, a substantial progress came from the discovery of Cas9 from *S. aureus* (SaCas9), which displays similar efficiency as Cas9 from *S. pyogenes* but is significantly smaller in size (1053 amino acids vs. 1368 of SpCas9) (Friedland, Baral et al. 2015). Indeed, many viral based approaches to deliver transgenes to cells, such as those taking advantage of the Adeno-Associated Vectors (AAV), have a precise space limitation for the inserted transgene they can accommodate. Friedland and collaborators showed that an optimized *S. aureus* Cas9 nickase sequence, together with a double expression cassette for sgRNA, efficiently fits within the AAV transgene packaging requirements (Friedland, Baral et al. 2015).

Since the continuous expression of Cas9 inside the target cell in the presence of guide RNA can be detrimental in the long term, increasing the probability of unwanted off-target genome modifications, many laboratories are designing strategies to induce self-inactivation of the transduced Cas9. For this purpose,

Bleris and collaborators proposed a self-inactivating transgene expression Cas9 system (Moore, Spinhirne et al. 2015). These authors used an all-in-one plasmid coding for the Cas9 sequence, for the target DNA (mKate2) and for two gRNAs (T1 and T2). The U6 promoter drives expression of the two gRNAs, while Cas9 and mKate2, separated by a 2A peptide, are transcribed starting from a CMV promoter. Upon HEK 293 cells transfection, both the mKate2 and the CRISPR system are transcribed. However, since the two gRNAs were designed to recognize sequences inside the mKate2 ORF, Cas9 activity consequently induced the vector degradation. Thus, this becomes a self-inactivating vehicle, able to control both the expression of desired transgenes and limit the Cas9 activity inside the transfected cells.

Again concerning the strategies aimed at limiting the off-target effects and improving the specificity of the CRISPR system, the Zhang laboratory recently identified a novel *Francisella novicida* U112 CRISPR system characterized by several interesting features (Zetsche, Gootenberg et al. 2015). It contains a putative class II CRISPR system, named Cpf1, as a single RNA-guided endonuclease able to functionally defend bacterial cells to plasmid interference. Cpf1 differs from the Cas9 protein both in structure and function, with potential significant advantages. First, different form Cas9, it presents only one RuvC-like endonucleolytic domain. Second, Cpf1 cleavage specificity is driven by a single stem loop in the CRISPR array (42 nt) and a short T-rich PAM domain, without the presence of the tracrRNA. This single RNA oligo therefore results much easier to synthesize if compared to the longer crRNA-tracrRNA duplex (110 nt). Furthermore, Cpf1 cleaves the double stranded DNA molecule producing staggered end termini. These represent a possible facilitation to mediate gene insertion through the NHEJ-repair, a mechanism much more efficient than HR in mammalian cells and, in particular, in post mitotic tissues. More recently, in 2018, the Cereseto laboratory, to improve the enzyme specificity, without acting on the cleavage activity, identified, using a high-throughput yeast-based approach, new Cas9 variants (Casini, Olivieri et al. 2018). The most promising variant, called evoCas9, resulted particularly efficient at improving the rationally

designed on-target activity (4-fold enriched), with a editing efficiency comparable with the WT enzyme.

1.1.2.4 Genome editing using CRISPR/Cas9

1.1.2.4.1 Ex-vivo Biomedical application of CRISPR/Cas9 system

1.1.2.4.1.1 Genetic modified animal models

The simplicity of the CRISPR/Cas9 system has immediately attracted attention for its potential use towards the generation of genetic modified animal models, able to mimic disease-causing mutations in patients. Due to its flexibility to target virtually any gene, the CRIRPR system is very suitable for these purposes. Indeed, the direct co-administration of the Cas9 protein complexed to the specific gRNA into the zygotes or early stage embryos of mice or rats can easily and efficiently generate specific knockout animals. A very clear example of the feasibility of this process came from the work of Wang and coworkers, in which these authors showed the generation of mice carrying a single mutation in multiple genes, using the CRISPR/Cas9 system with no evident toxicity (Wang, Yang et al. 2013). By direct injection of Cas9 mRNA and sgRNAs targeting Tet1 and Tet2 genes into mice zygotes, they were able to obtain 95% of the newborn mice carrying a biallelic mutation in both genes. This process therefore resulted easier and faster if compared to the cumbersome and time-consuming intercrossing between single-mutant mice.

Conditional gene modification in mice and rats can also be achieved by providing embryos DNA templates for HR-mediated gene knock-in. Accordingly, Chu and collaborators reported in 2016 a 20% success rate for an 11 kb transgene insertion inside the *Rosa26* locus obtained by direct injection in the C57BL/6 mice zygotes of a modified Cas9 mRNA, together with the gRNA and the homologous template (Chu, Weber et al. 2016).

1.1.2.4.1.2 Ex-vivo modification of cells

Another key application of this technology is the possibility to use Cas9 to specifically modify ex-vivo, for therapeutic purposes, human clinical cell samples, which are then re-administered to patients. T-lymphocytes and CD34+ hematopoietic stem cells are particularly suitable for his application. A first example of this use came from the work by Schumann and collaborators, in which these authors targeted the CXCR4 gene, which codes for a co-receptor essential for HIV cell entry (Schumann, Lin et al. 2015). Electroporation of the Cas9 enzyme and a gRNA targeting the CXCR4 into human CD4+ T-cells led to the knockdown of the receptor and the acquisition of resistance to HIV infection (Schumann, Lin et al. 2015).

The same authors also demonstrated that the direct cell transfection of the ribonucleoprotein complex together with a single stranded oligo DNA template complementary to specific sequences of both CXCR4 and PD-1, a cell surface receptor known to be associated with inhibition of cancer cell killing, resulted in efficient knock-out genome modification (Schumann, Lin et al. 2015).

This work laid the basis for the approval of an ongoing clinical trial that received green light in 2016, the purpose of which is to engineer CD4+ T-cells, harvested from patients, with a Cas9-mediated ex vivo gene editing approach to induce expression of a T cell receptor (TCR) recognizing the NY-ESO-1 tumor.

More recently, the novel Cpf1 endonuclease has also been employed for correcting genetic mutations. The most frequent alterations leading to the development of Duchenne muscular dystrophy (DMD) consists of X-linked mutations that induce the skipping of three exons preceding exon 51, which interrupt the open reading frame (Bozkurt, Aguilar et al.) by juxtaposing out-of-frame exons. In the attempt to correct this disease-causing mutation, Olson and collaborators administered Cpf1 endonuclease and the corresponding sgRNAs targeting the T-rich splice acceptor site of exon 51 into DMD iPS-derived cardiomyocytes (CMs) (Zhang, Long et al. 2017). In this way, the NHEJ-mediated resolution of the DSB and sequence disruption

by InDel introduction, laid to skipping of exon 51. This restored the reading frame allowed the production of an incomplete, but partially functional dystrophin protein.

1.1.2.4.1.3 Genomic high-throughput screenings

If compared with RNA interference mediated by the use of short harpin RNAs (shRNAs) or small interfering RNAs (siRNAs), the CRISPR/Cas9 system presents several advantages. Different from the transient mRNA degradation due to shRNAs/siRNAs interference, Cas9 mediated gene knockout produces the biallelic alteration of the genomic target sequence, and, consequently, its permanent silencing. Moreover, the CRISPR technology id easily applicable and has low cost, which is attractive for high-throughput screening studies. In 2014, the group of Koike-Yusa published the results of a genome-wide sgRNA screening in mouse ES cells, aimed at identifying new silenced genes directly associated with the resistance against the *Clostridium falciparum* alpha-toxin (Koike-Yusa, Li et al. 2014). Since then, different set of sgRNA libraries have been generated in order to study different biological events. For example, Doench and collaborators generated a lentiviral-based sgRNAs library to study the knockdown of human or mouse cell-surface proteins (Doench, Hartenian et al. 2014). Furthermore, by using a catalytically dead Cas9 fused with a Krüppel associated box (KRAB) (a well described transcriptional repressor of RNA polymerase I, II and III) (Witzgall, O'Leary et al. 1994) and controlled by a doxocyclin-inducible promoter, Gilbert and coworkers expressed in K562 cells a customized sgRNA library to identify targets regulating the resistance to the cholera toxin (Gilbert, Larson et al. 2013). The results obtained showed the enrichment and the loss of many sgRNAs, corresponding to the downregulation of specific genes possibly related with cholera toxin resistance and confirming the feasibility of the approach.

1.1.2.4.2 In-vivo biomedical applications

Despite the CRISPR/Cas9 technology has been dramatically improved over the last years, still the gene editing technology needs further improvements. Nevertheless, there are cases in which the efficiency of the treatment may be sufficient to show an evident clinical effect. For example, monogenic recessive disorders caused by loss-of-function mutations, such as cystic fibrosis or DMD, may represent good candidates to be treated with genome editing approaches. Hereditary tyrosinemia type I (HTI) is another example of a genetic disease that can potentially be treated using the CRISPR/Cas9 technology. This condition is characterized by mutations in the fumarylacetoacetate hydrolase (Fah) gene, one of the key components of the tyrosine degradation process, which produces the toxic accumulation of this amino acid in the blood. To correct these alterations, Grompe and coworkers demonstrated how, in a mouse model of HTI, it is possible to restore the WT sequence of the Fah gene using Cas9 induced HR-mediated knock-in in the liver of adult mice (Yin, Xue et al. 2014). Upon treatment, only 0.4% of the hepatocytes was shown to express the corrected Fah sequence; nevertheless, this was sufficient to significantly ameliorate the pathological phenotype.

Another example of a potential good candidate for a gene editing treatment was provided by Regan and coworkers, who demonstrated how the Cas9-mediated permanent knock-out of the Proprotein convertase subtilisin/kexin type 9 (PCSK9) in mice led to 40% decrease in plasma cholesterol (Ding, Strong et al. 2014). These findings further confirmed the protective effect of PCSK9 depletion against the development of cardiovascular diseases. Different from the previous approach, in this work the authors used Cas9 to inactivate the target gene, a process that results much more efficient if compared to precise HR-mediated gene insertion. Nevertheless, also in Autosomal dominant mutations can be considered good targets for possible CRISPR-based gene correction therapies. Indeed, in the publication of Bakondi and collaborators, they proposed a Cas9-based approach to successfully treat S334ter-3 rats, a well-described model for Retinitis Pigmentosa (Bakondi, Lv et al. 2016).

Despite the great progresses in improving the efficiency and precision of the CRISPR technology, a major limitation still remains *in vivo* delivery. To develop a plasmid-free transfection method, the group of Kim and coworkers developed a gRNA and Cas9 system fused with cell penetrating peptides (CPPs) (Ramakrishna, Kwaku Dad et al. 2014). The treatment of different human cell lines with this technology led to the efficient knockdown of the target genes, with reduced number of off-target events.

Another optimized CRISPR delivery approach was reported by Beisel and collaborators, who described a nanoparticle-based delivery system (Sun, Ji et al. 2015). The complex of the sgRNA and the Cas9 enzyme was loaded inside DNA nanoclews, which are DNA-based nanoparticles, coated with polyethylenimine (McGill, Ismail et al.), which allowed the endosomal escape thanks to its positive charge. Other approaches to vectorize the CRISPR system derived from the work of Mangeot and collaborators (Mangeot, Risson et al. 2019). In this study they proposed a lentiviral-based approach to package the CRISPR system and to treat primary cells, embryos and animal models. This technology, called Nanoblades, resulted much more efficient at inducing dsDNA breaks if compared with lipids-based or elettroporation approaches, both in tumor cells (U2OS) or iPS cells.

Despite these examples represent an important progress in this field still the most promising CRISPR delivery method *in vivo* remains the use of viral vectors, such as AAVs. Indeed, due to their favorable characteristics, such as high transduction efficiency of many tissues and low immunogenicity, these vectors can be exploited to efficiently transfer the Cas9 gene and different gRNAs into many cell types. Of interest, AAVs have also been employed before the discovery of CRISPR for different gene editing approaches. As an example, an endonuclease free approach was proposed by Kay and collaborators in 2012 (Wang, Lisowski et al. 2012). The purpose of this study was to inset the human Fah transgene into the safe harbor 28S rDNA locus of $Fah^{-/-}$ HTI mice by HR. (Paulk, Wursthorn et al. 2010). These investigators packaged the rDNA-hFah AAV sequence inside an AAV8 vector, which is known to efficiently transduce

hepatocytes. Upon treatment with NTBC, a drug known to induce liver damage in the HTI mouse model and consequent fast hepatocytes proliferation, AAV8 were administered to mice. The episomal non-integrating viruses were diluted and therefore lost, while the recombined sequences were selected as resistant cells and the site-specific integration was confirmed by PCR. By repeating the AAV8 administration, the investigators were therefore able to improve the number of $hFah^{+}$ hepatocytes until the functional restoration of the liver physiological parameters.

More recently, several groups started to employ AAVs to transfer the Cas9 system inside the desired tissues. Rivera and collaborator reviewed the possibility to treat mdx mice, the mouse model for DMD, with the Cas9 technology to restore the production of the protein (Nelson, Hakim et al. 2016). These authors proposed a strategy based on the NHEJ-mediated removal of exon 23 of the dmd gene, which carries a nonsense mutation that induces the premature block of protein production. The possibility to remove exons from the coding sequence can therefore restore the correct frame and allow the production of an incomplete micro-dystrophin. This strategy generates a less severe form of the pathology, defined as Becker muscular dystrophy (BMD). By using an AAV8 vector coding for the SaCRISPR/Cas9 system and another AAV8 vector carrying two sgRNAs against intron 22 and 23, they were able to induce the skipping of the exon 23. This gene editing approach can restore the 8% of the dystrophin protein levels, with a 67% of DMD-positive myotubes. Accordingly, the muscle histological parameters, such as morphology, fibrosis, number of infiltrating macrophages and the consequent inflammation, were also significantly improved after the CRISPR treatment. Finally, the strength parameters, such as twitch and the tetanic force were also improved, confirming the beneficial effect of the AAV-mediated Cas9 gene editing in this particular pathological context. This NHEJ-mediated approach can therefore be applied to many mutations occurring in the dystrophin gene just by generating precise DSBs and consequently forcing the skipping of specific exons.

Since the administration of the Cas9 using AAV determines its prolonged expression, to limit the generation of possible off-target mutations other groups have developed a combined delivery method using both viral and non-viral vectors. An example of this approach can be found in the paper of Wu and coworkers, where these investigators described the possibility to treat a hereditary form of tyrosinemia (HTI) by a nanoparticle-mediated Cas9 delivery combined with the AAV-mediated expression of the sgRNAs and the template for correction (Yin, Song et al. 2016). The feasibility of this approach was tested in a mouse model of the human pathology, characterized by a mutation in the Fah gene, which leads to splicing alterations. The results from one single administration of the combined treatment in adult mice confirmed the restoration of the correct splicing of Fah gene in 6% of the hepatocytes.

Based on these first successful applications *in vivo*, a number of monogenic conditions can be considered for CRISPR/Cas9-mediated repair, especially when a low efficiency of gene editing may still be clinically relevant. Of note, correction through NHEJ can be considered in virtually all cell types provided that a suitable gene delivery system is available. On the contrary, gene editing by HR and thus precise gene editing is conceptually limited to cycling cells.

1.2 Adeno-Associated virus (AAV)

Adeno-Associated virus (AAV) is a single-stranded DNA Parvovirus that became very popular in the last few decades with the advent of gene therapy.

The first evidence of the existence of AAV came from the experiments of Atchison and collaborators, who, in 1965, identified a new viral contaminant of Adenoviral preparation (Atchison, Casto et al. 1966). Despite AAV was shown to grow together with Adenovirus, there were no identified biological similarities between them. AAV was classified as a dependovirus, due to the requirement of the presence of a helper virus, such as Herpesvirus (Georg-Fries, Biederlack et al. 1984) or Vaccinia virus (Schlehofer, Ehrbar et al. 1986), in addition to Adenovirus, to achieve a functional infection. Indeed, it is known that without the presence of a helper virus, AAV is able to integrate into a specific region of

chromosome 19 (19q13.4), defined AAVS1, where it persists indefinitely in a latent form (Berns and Linden 1995). Of note, AAVS1 represents only 10% of the all the possible integration sites in the human genome; indeed, many other random integrations have been described (Huser, Gogol-Doring et al. 2010, Huser, Gogol-Doring et al. 2014). In contrast, during a productive infection, the AAV genome persist in an episomally concatamerized form and activates the viral gene expression (Cheung, Hoggan et al. 1980).

In nature, many different AAVs have been identified over the years and classified depending on their different infection specificity. Several serotypes were shown to efficiently infect specific tissues depending on their peculiar capsid protein composition. The first evidence reporting AAV as a possible gene transfer tool came from the publication of Tratschin and coworkers, in which these investigators demonstrated how this virus was able to allow the expression of the prokaryotic chloramphenicol acetyltransferase (Huser, Gogol-Doring et al.) gene in HEK 293 cells (Tratschin, West et al. 1984). Many of the information that we know about the AAV structure, genome and intracellular trafficking came from the study of AAV serotype 2 (AAV2). Indeed, this is also the first AAV that was engineered to produce non-replicative viral particles (Samulski, Berns et al. 1982). Due to their favorable characteristics, such as the absence of a recognized pathology induced in humans, the low immunogenicity, the ability to infect wide range of tissues, especially post-mitotic tissues, and the long lasting transgene expression, AAV appeared perfectly suitable to be used for the development of a gene transfer vector. Over the years, recombinant AAVs have been engineered to transduce with high efficiency both cells *in vitro* or tissues and organs *in vivo*. According to the wide range of possible applications, many groups started to employ AAV to induce virtually life-long-lasting expression of the desired transgenes, for both basic research purposes and in clinical trials.

The first evidence reporting the possible clinical impact of AAV-based treatment came from the paper of Buch and collaborators, who described a safe AAV-mediated treatment of Leber's congenital amaurosis disease (LCA). This

condition is caused by mutations in the RPE65 gene, coding for a key component of phototransduction process (Buch, Bainbridge et al. 2008). These investigators confirmed in several animal models of inherited retinal disorders the efficient AAV-mediated RPE65 transgene transport inside the retina, with a consequent improvement of the vision, without the identification of side effects. These proof-of-concept experiments laid the basis for the first clinical trial, where young LCA patients were treated with AAV2 coding for the WT form of the RPE65 cDNA. One year after treatment, patients did not develop any adverse reactions against the virus, showing indeed a strong improvement in the visual sensitivity (Cideciyan, Hauswirth et al. 2009). These publications laid the basis for the further Food and Drug Administration (FDA) approval in 2018 of the AAV2-hRPE65v2 based therapy, to cure biallelic-mutation associated retinal dystrophy, with the name of Luxturna.

During the years many other clinical trials started to use AAV for the treatment of several pathologies, such as haemophilia B (Nathwani, Reiss et al. 2014), Alzheimer's disease (Mandel 2010), Parkinson's disease (Christine, Starr et al. 2009), Duchenne muscular dystrophy (Rodino-Klapac, Chicoine et al. 2007) or familial lipoprotein lipase deficiency (LPLD).

The first AAV-based medical product that received the approval from the European commission to be commercialized in the Western world, under the name of Glybera (Salmon, Grosios et al. 2014), was an AAV1-based drug, coding for lipoprotein lipase (LPL), which was developed for the treatment of LPLD, a rare disease leading to an accumulation of triglycerides in the plasma and eventually fatal pancreatitis (Santamarina-Fojo, Haudenschild et al. 1998). Nevertheless, after few years, it was recalled from the European market due to its high cost. More recently, in 2019, another AAV-based therapy received the FDA approval for the treatment of Spinal Muscle Atrophy (SMA), under the name of Zolgensma. SMA is caused by the loss or the mutation of SMN1 gene, coding for SMN protein, a component of SMN complex, which plays a critical role in the biosynthesis of small nuclear ribonucleoproteins (snRNPs) and in the

pre-mRNA splicing process. Zolgensma is an AAV9-based therapy targeting the motor neurons and codifying for the WT sequence of SMN1 gene.
Currently, over 230 clinical trials using AAV are listed in the Wiley Gene Therapy Clinical Trial Worldwide Database (http://www.abedia.com/wiley/).

1.2.1 Biology of the wild type Adeno-Associated Virus

AAV is characterized by a single-stranded DNA genome of approximately 5 kb of size. It was classified in the Parvovirus family and Dependovirus genus due to its requirement of the presence of a helper virus to achieve productive infection. Therefore, this characteristic distinguishes AAV from other parvoviruses that are able to replicate independently (Galev, Afanas'ev et al. 1989).
The first characterization of the AAV genome structure was provided by Srivastava and Lusby, who defined it as composed by two open reading frames (ORFs) flanked by two 145-nt regulatory elements (inverted terminal repeats, ITRs) (Lusby, Bohenzky et al. 1981, Srivastava, Lusby et al. 1983).
By mutating the AAV genome, Tratschin and collaborators first demonstrated the specific roles of the two ORFs in viral infection (Tratschin, Miller et al. 1984). Alternative splicing of the first ORF (Rep) generate 4 mRNAs, coding for essential non-structural proteins (Muller, Kaul et al.)ANNO (Trempe and Carter 1988). These are critical for viral replication, viral DNA integration into the host genome and gene expression (Senapathy, Tratschin et al. 1984). Moreover, these investigators also described an essential role of the Rep proteins in enhancing AAV promoters in the presence of a helper virus infection or exerting an inhibitory effect in its absence (Trempe and Carter 1988). More precisely, the Rep genes codes for Rep78, Rep68, Rep52 and Rep40, whose transcription requires two different promoters, P5 and P19. Rep78 and 68 mRNAs starts from the P5 promoter, while Rep52 and 40 require the P19 promoter (Marcus, Laughlin et al. 1981). The former two proteins were identified to play a central role in all the life cycle of AAV by also regulating its DNA integration into the chromosome 19 of the host genome (Young, McCarty et al. 2000). In contrast, Rep52 and 40 were shown to regulate the viral packaging (King, Dubielzig et al.

2001). The second ORF, defined as Cap, instead codes for 3 structural protein-coding genes (VP1, VP2 and VP3) and one additional protein essential for capsid assembly (Ramakrishna, Kwaku Dad et al.). At difference from the first ORF, the transcription of the three VP genes is regulated by one single promoter (P40) and the 3 capsid proteins are generated primarily by alternative splicing a non-canonical ACG translation start codon for VP2 (Becerra, 1988; Trempe and Carter, 1988). The three distinct protein products share the C-terminus sequence that corresponds to VP3. The relative abundance of the individual proteins determined the stoichiometry of 1:1:10 of VP1, VP2 and VP3 in assembled capsids, respectively (Kronenberg, Kleinschmidt et al. 2001).

Additionally, the Assembly-Activating protein (AAP), required for the correct capsid assembly (Naumer, Sonntag et al. 2012), becomes stable upon VP3 protein expression and represents a scaffolding protein able to interact with Cap proteins to trigger the correct organization of the viral particle formation (Sonntag, Schmidt et al. 2010). The AAV genome also contains two 145-bp long T-shaped harpin sequences, flanking the already described ORFs. These are essential for all the functional steps of viral live circle (Ashktorab and Srivastava 1989). As already mentioned before, Rep78/68 play a critical role the absence of helper virus infection. Indeed, it was demonstrated that they negatively regulate the AAV promoters and induce latency. Despite this process does not alter the host-cell viability, it induces the activation of specific cellular genes that make cells more sensitive to DNA damage agents, probably due to the low levels of Rep protein production (Marcello, Massimi et al. 2000).

1.2.2 Generation of recombinant AAV (rAAV)

Due to their favorable characteristics, over the last three decades there has been a progressive interest in the possibility to modify the AAV genome and capsid structure for gene therapy applications. In fact, it was demonstrated that the viral genome, cloned into a plasmid, was capable of packaging and formation of an infectious viral particle and that, in principle, any sequence could be placed between the two ITRs (Gao, Qu et al. 1998). Consequently, to

obtain a recombinant vector able to transduce specific cell lines and tissues the AAV genome was modified. Different from the genome of a WT AAV, recombinant vectors typically contain, between the flanking ITRs, a promoter followed by a transgene expression cassette and the regulatory sequences. Therefore, the only parts of the AAV genome that remains unmodified are the two ITRs, since these are crucial components of viral genome packaging and intracellular processing. Noteworthy, since the Rep ORF is essential to achieve viral DNA integration inside the host genome, its elimination generates vectors able to persist only in an episomal and concatamerized form. Nevertheless, in order to achieve the functional transgene expression, the ssDNA sequence of recombinant vectors has to be duplicated thanks to the activity of an endogenous DNA polymerase. This crucial phase was described as a rate-limiting step for the successful rAAV transduction (Ferrari, Samulski et al. 1996). Trying to overcome this limitation, self-complementary AAV (scAAV) was engineered from the natural occurring AAVs. This is characterized by a self-hybridizing genome, able to create a double-stranded (ds) DNA directly through Watson-Crick interaction. Therefore, this class of vectors does not require the intervention of the cellular DNA polymerase for dsDNA conversion. Therefore, scAAVs was reported to be more efficient at transducing cells, generating higher levels of gene transfer in several tissues, such as brain, liver or retina (McCarty, Monahan et al. 2001, Petersen-Jones, Bartoe et al. 2009). The two flanking WT ITRs and a mutated ITR in the center compose the genome of scAAV vectors. Removing the terminal resolution site (Petrs-Silva, Dinculescu et al.) to a WT ITR generates this peculiar DNA structure. DNA replication in these vectors starts from a binding site present on the 5′ WT ITR, then, the absence of the TRS, allows the synthesis of DNA through the modified ITR and again on the opposite strand until the enzyme reaches the 3′ WT ITR. Despite scAAV is more efficient at transducing several cells and tissues, the limited packaging capacity, corresponding to half of the conventional rAAVs capacity (2.1 kb compared with 4.7 kb of conventional ssAAV), represents a main weakness of this vector.

1.2.3 Natural and synthetic serotypes

The first isolated serotype was AAV serotype 2, which was recognized as a contaminant of Adenoviral preparations. From its discovery, over the years, many other serotypes were identified. These share many genome structure feature however differ by peculiar serological profiles they elicit in the host due to different capsid protein structure (Hoggan, Blacklow et al. 1966, Rutledge, Halbert et al. 1998, Gao, Alvira et al. 2002, Gao, Vandenberghe et al. 2004). The first step in the infectious cycle of AAV is the recognition of specific membrane structures, which is essentially mediated by the Cap proteins. AAV2, AAV3 and AAV13 recognize heparan sulfate proteoglycans; AAV1, AAV4, AAV5 and AAV6 bind N-linked or O-linked sialic acids, while AAV9 was shown to bind galactose residues (Wu, Asokan et al. 2006, Shen, Bryant et al. 2011). Consequently, the different identified serotypes show specific infection profiles. AAV8 and AAV9 were shown to be more efficient vector for hepatocytes transduction (Ohashi, Nakai et al. 2005), while AAV9 and AAV6 efficiently transduced the airway epithelium (Halbert, Allen et al. 2001). The brain can be efficiently transduced by several AAV serotypes (AAV1, 4, 5, 7 and 8). In the study of Inagaki and coworkers, these authors demonstrated efficient pancreas transduction using AAV1, AAV8 and AAV9 (Inagaki, Fuess et al. 2006). Noteworthy, the heart is one of the most permissive tissues, since several AAVs (AAV1, 6, 8 and 9) can efficiently transfer genes into cardiomyocytes. To date, 13 primate AAV serotypes have been identified and many others were artificially generated in the laboratory by fusing different capsid components. Furthermore, trying to improve transduction efficiency in different tissues, many chimeric viral particles have been engineered by combining the same genome with different AAV capsids. To create new recombinant capsids, Petrs-Silva and collaborators isolated new variants from capsid libraries (Petrs-Silva, Dinculescu et al. 2009), obtained with different methods. For example, to generate the first capsid library, the laboratory of Muller and coworkers used viral particles loaded with random peptides on their surface (Muller, Kaul et al. 2003). Libraries can also be generated using other strategies, such as random mutation of the Cap gene, error-prone PCR or by DNA-shuffling techniques (Perabo, Endell et al. 2006). An

example of an engineered vector comes from the paper of Zhong and collaborators, who described the generation of a more efficient AAV thanks to the substitution of a specific tyrosine in the capsid proteins with a phenylalanine. With this approach, these investigators generated a 30 time-more efficient AAV2 compared to the canonical serotype 2 vector (Zhong, Li et al. 2008).

1.2.4 Intracellular processing

To understand all the possible limitations directly associated with a successful transduction, it is very important to underline all the critical steps associated with AAV infection (Figure 7). As outlined above, specific serotypes can infect specific tissues depending on the peculiar proteins recognized on the cellular surface. For example, starting from binding to membrane heparan sulfate proteoglycans (HSPGs), AAV2 is rapidly internalized by the cells through clathrin-mediated endocytosis (Xiao and Samulski 2012). The clathrin vesicles then fuse with the endosomes, and AAV2 remains there until the acidification of the compartment. This event consequently induces the exposure, by the viral particles, of the N-terminus of the VP1 capsid protein, which contains phospholipase activity and facilitates endosome escape. Once viral particles reach the cytosol, they need transport into the nucleus. For this purpose, AAV contains two nuclear localization signals (NLS) on both the VP1 and VP2 capsid proteins, which allow full-packaged capsids nuclear internalization. Moreover, Nicolsoln and Samulski demonstrated how the nuclear pore complex (NPC) and importin-β are also essential for AAV nuclear entry (Nicolson and Samulski 2014). Despite this available information, this phase of the viral life cycle still remains one of the key limiting step for viral transduction, as it can also be concluded from the paper of Bartlett and coworkers, who demonstrated how viral particles accumulate in perinuclear regions (Bartlett, Wilcher et al. 2000). The first report on the visualization of the viral internalization steps in real time was provided the work of Lux and collaborators, who substituted the AAV2 WT VP2 protein with a GFP-VP2 (Lux, Goerlitz et al. 2005). There are two different theories trying to explain how nuclear import occurs. In 2002, Xiao and

coworkers showed that AAV particles nuclear entry was not dependent on the active transport through the NPC (Xiao, Warrington et al. 2002). In contrast, in 2014 Nicolson and Samulski demonstrated how the internalization was triggered by an active importin-β mediated mechanism (Nicolson and Samulski 2014). Additionally, it is not clear whether particles uncoating happens outside the nucleus, in the perinuclear area, or inside the nucleus (Johnson and Samulski 2009), however it is known that all the crucial components for dsDNA conversion are located inside the nucleus (Hansen, Qing et al. 2001).

Once AAV reaches the nucleus its genome still needs to be converted from ssDNA to dsDNA and this corresponds to another rate-limiting step to achieve an efficient viral transduction (Ferrari, Samulski et al. 1996). Indeed, AAV genomes can be recognized from the DNA damage response (DDR) proteins. In fact, as first shown by our laboratory, the viral genomes physically interact with the cellular DDR MRN complex (Cervelli, Palacios et al. 2008). To demonstrate this, by infecting cells with a LacR-GFP and an AAV coding for LacO target sites Cervelli and coworkers were able to observe the viral genome conversion in specific foci inside the nucleus, which overlapped with those of DDR proteins. Since it is known that post-mitotic cells present very low levels of DDR protein expression, this observation can explain why AAV preferentially infects post-mitotic tissues if compared with replicating cells. Indeed, Lovric and collaborators demonstrated how AAV transduction efficiency of CMs progressively increased after birth during CM maturation and correlated with the reduction of DDR protein expression (Lovric, Mano et al. 2012).

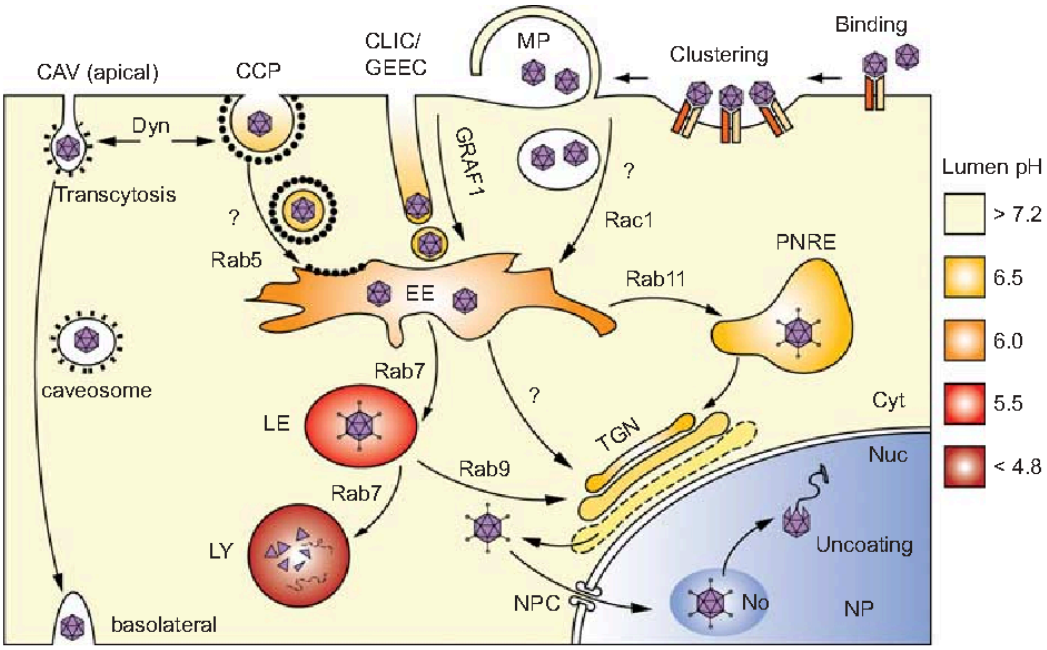


Figure 7: Graphic representation of AAV intracellular trafficking. First, depending on the serotype, vectors bind their receptor and co-receptor complex and, after endocytosis, enter into the target cells. Different pathways are responsible of AAV internalization, for example CLIC/GEEC endocytosis, clathrin-mediated endocytosis (CCP) or caveolar endocytosis CAV). After this step, AAV vectors accumulate in the trans-Golgi (TGN), probably through a retrograde transport pathway that involves the endosome and its maturation from early to late endosome. During this step, the capsid conformation changes and vectors are released form the Golgi to the cytoplasm thanks to the exposure on the surface of the phospholipase A2 (PLA2) domain. After this step, AAVs enter into the nucleus through the nuclear pore complex (NPC) and importin-β, where they release their genome. (Nonnenmacher and Weber 2012).

1.2.5 AAV production

Several techniques have been developed over the years to generate recombinant AAV vectors. One possibility is to employ mammalian cells, using a double/triple plasmid transfection procedure using a protocol that has been optimized in order to minimize the generation of WT particles.

Typically, HEK 293 cells are transfected with two plasmids coding for the Rep and Cap genes and for the desired transgene, flanked by the two ITRs. Since the Rep and Cap plasmid does not contain the ITRs, no WT particles will be produced; on the other hand, only the particles bearing the desired transgene are synthesized. This procedure results very efficient for the production of any serotype just by modifying the cap genes of choice, carrying any desired transgene flanked by the two IRTs (Laughlin, Tratschin et al. 1983). Nevertheless, to achieve a massive rAAV production, a combined infection with a helper virus, such as Adenovirus, is needed. This corresponds to a limitation of the technique, since it also leads to the production of Adenoviral particles. In order to avoid the production of helper-virus particles, another plasmid, coding for all the key Adenoviral genes (E4, VA and E2a), can be transfected. Since this third plasmid does not contain any virulent gene, it results completely safe and non-infectious, coding only for the viral genes essential for AAV synthesis. Due to its simplicity, plasmid transfection still remains the most used method to synthesize AAV particles. Moreover, it is also a very fast procedure, since AAV particles can be harvested at 48-72h post-transfection. The subsequent purification steps are aimed at both eliminating contaminating cell debris and recovering only capsids containing the viral DNA, thus eliminating the considerable proportion of empty capsids generated by the cells. This can be typically achieved by cesium chloride (CsCl) density gradient centrifugation. To increase the titer and purity of the preparation, this step can be repeated several times. Other techniques that can be employed are based ligand affinity chromatography (Zolotukhin, Byrne et al. 1999) however not all the AAV serotypes can be purified with this method.

1.2.6 Limitation in the use of AAV

Despite the wide range of possible applications, the use of AAV to transfer genes *in vivo* has some important limitations.

1.2.6.1 immune response against AAV treatment

Over the last years, several of the investigators who have used AAV vectors in clinical trials have started reporting the possible activation of an immunological response after AAV transduction. Zhu and collaborators underlined the correlation between AAV treatment and the innate immune system activation (Zhu, Huang et al. 2009). Other studies have underlined the possible production of neutralizing antibodies (NAb) from the adaptive immune system. This corresponds to a strong limitation of gene transfer efficiency, since NAb can completely inhibit AAV transduction (Manno, Pierce et al. 2006). In addition, these antibodies completely prevent the possibility of re-administering AAV vectors with the same serotype as the initial administration. Finally, a gene trial for hemophilia B, entailing liver gene transfer of the Factor IX cDNA, reported the long-term development of CD8 T cells against the AAV capsid protein (Hui, Edmonson et al. 2015) suggesting that incomplete processing and degradation of the capsid inside the transduced hepatocytes might lead to the generation of reactive cytotoxic lymphocytes that eventually clear the transduced cells.

As a consequence, it appears very important to develop strategies trying to avoid the immune response system. Work from different investigators has indeed demonstrated the possibility to use rationally designed capsids, which generate less immunogenic serotypes during the first rAAV treatment (Li, Diprimio et al. 2012, Hui, Basner-Tschakarjan et al. 2013). In addition, to minimize the immune response against AAV, patients can be transiently immunosuppressed pharmacologically (Jiang, Couto et al. 2006).

Despite this evidence of immune reactivity, it worth underlining that no severe reactions against the AAV treatment have been so far described.

1.2.6.2 Limitation in AAV tropism and transduction

As previously described, AAV capsid proteins are responsible for cell target specificity of the different serotypes. Zincarelli and collaborators, in 2008, described how infection of different systemically delivered AAV serotypes (AAV1-9) in mice resulted in specific distribution and consequent transgene expression of each vector serotype inside the tissues. Furthermore, it was also demonstrated that different serotypes sharing the same target molecules infect the same cell types, however they can differ in their internalization rate (Pasquinelli, Reinhart et al. 2000). Despite many natural serotypes have been described infecting several tissues, not all the cell types can be transduced. For example, stem cells or endothelial cells, despite expressing receptor molecules that are common targets for different AAV serotypes are resistant to AAV transduction. This clearly underlines that the main determinants of AAV tropism lie beyond the capacity of the vector to be internalized, and involved subsequent steps such as endosomal escape, nuclear transport, de-capsidation and conversion from ssDNA to dsDNA, as also discussed above.

1.2.6.3 Limited transgene packaging capacity

Since the natural viral transgene packaging capacity is 4.7 kb, a large number of cDNAs can be delivered to cells and tissues using the AAV technology – the average size of human cDNAs is approximately 2.7 kb -. Grieger and Samulski, explored the possibility to synthesize capsids composed of only 2 Cap proteins, in order to permit a larger transgene loading capacity (Grieger and Samulski 2005). The results from this study demonstrated that, by engineering different serotypes (AAV1-5), a maximum size of 6 Kb could be reached. Furthermore, Allocca and collaborators demonstrated the possibility to synthesize AAV particles carrying genomes of more than 8 Kb (Dong, Nakai et al. 2010). Nevertheless, a substantial lower infectivity was detected by using particles containing a >5.2 Kb genome. The same observation was evident in the study by

Wu and collaborators, who in 2009 demonstrated lower transduction efficiency of rAAV containing genomes larger than 5 kb (Wu, Yang et al. 2010).

1.3 MicroRNA

MicroRNAs (miRNAs) are small non-coding RNAs, coded by the genome of plants, animals and viruses, which have been discovered to post-transcriptionally regulate gene expression. In humans, 1,984 microRNA precursors have been annotated, which correspond to 2,693 mature and active sequences (miRBase release 22.1, October 2018, http://www.mirbase.org). MiRNAs drive the process of RNA-interference, since, by direct base-pairing recognition of target mRNAs, they can silence the expression of their target genes through different mechanisms. In particular, they recognize specific sequences, often located in the 3'-UTRs of transcripts, both through perfect of imperfect base pairing with their "seed sequence", located between nucleotides 2 and 7 of 5'-UTR of mature miRNAs. This specific sequence is essential to trigger genes knockdown (Doench and Sharp 2004). It has been demonstrated that 60% of the human genome is directly regulated by the activity of miRNAs (Lewis, Burge et al. 2005); indeed, hundreds mRNAs can be regulated by a single miRNA (Friedman, Farh et al. 2009) and many miRNAs can repress a single gene. According to this information, miRNA are now considered as strong regulators of many physiological and pathological processes.

1.3.1 MicroRNAs discovery

The first discovered miRNA was described by Victor Ambros and collaborators, when these investigator were working on the identification of possible mutations in the genome of *C. elegans*, which were able to modify the growth of this nematode (Lee, Feinbaum et al. 1993). They discovered that the lin-4 gene encoded for two RNAs, of 22 and 61 nt. Instead of producing a protein-coding mRNA, the shorter RNA was found to give rise to a non-coding transcript able to inhibit lin-4 gene expression by directly targeting the 3'-UTR of the gene. Subsequently, in 2000, another gene associated with *C. elegans* larval timing

development, Lin-7, was seen to produce a similar 21-nt non-coding transcript (Reinhart, Slack et al. 2000). Moreover, it was demonstrated that its sequence was complementary to the 3'-UTR of many other genes related with *C. elegans* development. Some years later, other 33 new miRNAs were identified both in invertebrate and vertebrate models as conserved regulatory sequences (Lagos-Quintana, Rauhut et al. 2001).

1.3.2 MicroRNAs biogenesis

The Biogenesis of miRNAs is a multi-step process that begins inside the nucleus and completes in the cytoplasm. MiRNAs precursor sequences are located in intronic and intragenic regions and their transcription is carried out by RNA polymerase II (Lee, Kim et al. 2004). This enzyme generates the so-called pri-miRNAs, which correspond to long transcripts, presenting a harpin conformation, a cap-structure and a poly-adenylation site (Cai, Hagedorn et al. 2004). Pri-miRNAs are then recognized by Drosha, one of the key components the miRNA biogenesis pathway. This enzyme is a RNAse III that, together with a

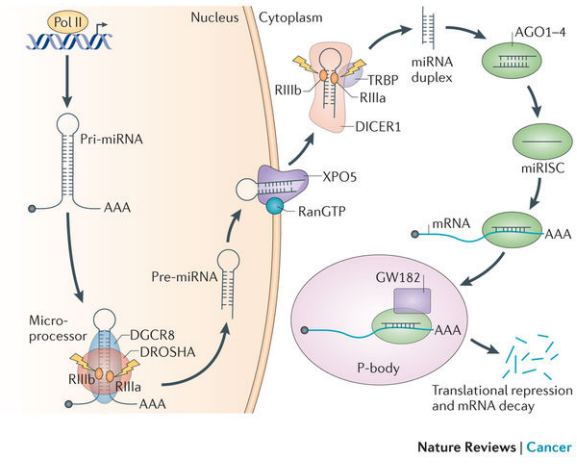


Figure 8: miRNAs biogenesis corresponds to multi step process that begins inside the nucleus and completes in the cytoplasm. MiRNAs precursor sequences are located in intronic and intragenic regions and their transcription is triggered by the RNA polymerase II (Lee, Kim et al. 2004). This enzyme generates the so-called pri-miRNAs, which correspond to long transcripts, presenting a harpin confirmation, a cap-structure and a poly-adenylation site (Cai, Hagedorn et al. 2004).

cofactor, called Pasha, modifies the pri-miRNA structure by converting it into a shorter RNA intermediate, defined as pre-miRNA (Lee, Ahn et al. 2003). These molecules of 70-130 nt of length maintain the harpin conformation, but present, at their 3' ends, short overhang sequences (Lund, Guttinger et al. 2004). Subsequently, Exportin 5 recognizes the overhang sequences and allows the transport of pre-miRNAs into the cytosol. Here, they are recognized by Dicer, another RNAse III protein, that cleaves the intermediates into a double stranded RNA of 22 nt of length (Ketting, Fischer et al. 2001). Only one of the two strands will become the mature miRNA form, able to modulate gene expression, while, the other strand will be degraded.

1.3.3 Mechanism of action of microRNAs

After the conclusion of the maturation process, miRNAs are recognized by Argonaute proteins and consequently internalized inside the RNA induced silencing complex (Vatta, Mohapatra et al.) (Hammond, Bernstein et al. 2000). This protein complex therefore allows miRNAs to recognize and bind by Watson-Crick interaction, the target mRNAs and consequently induce their degradation and/or translation inhibition, depending on the perfect or imperfect base pairing with the seed-sequence. Consequently, looking at the different seed-sequences it is possible to clusterize the miRNAs into different families, targeting similar sets of mRNAs. In contrast, the other nucleotides are required for stabilization of the binding. An example of a miRNA family came from the paper of Pasquinelli and collaborators, who described the characteristics of the let-7 miRNA family (Pasquinelli, Reinhart et al. 2000).

The perfect pairing with the seed-sequence then leads to the degradation of the target mRNAs, which is triggered by the Argonaute-mediated cleavage of the sequence (Llave, Xie et al. 2002). In contrast, the imperfect base pairing preferentially leads to the inhibition of mRNA translation, a mechanism that more frequently occurs in vertebrates. Eichhorn and collaborators demonstrated that mRNA destabilization is to the most common process by which miRNAs modulate gene expression (Eichhorn, Guo et al. 2014). To achieve the

destabilization, the poly-A tail of the target mRNAs is removed thanks to the RISC-associated protein GW182, the activity of which activity is crucial for target gene silencing (Braun, Huntzinger et al. 2011).

1.3.4 Modulation of endogenous microRNAs levels

To better understand the role of miRNAs in both physiological and pathological processes, or to develop possible miRNA-based treatments, a possible approach consists in the artificial up or down regulation of their expression. With this goal, it is possible to use the so-called miRNA mimics technology, which uses double-stranded RNA sequences, identical to the endogenous miRNAs, but chemically modified to force RISC to use only the desired miRNA strand. In contrast, to inhibit the expression of specific miRNAs it is possible to use miRNA inhibitors. These are usually small antisense sequences, able to efficiently bind the endogenous miRNAs by Watson-Crick interaction (Velu and Grimes 2012). As opposed to miRNA mimics, which are incorporated into the RISC complex and thus tolerate minimal structural variations, antisense miR can be modified by a portfolio of chemical modifications, including inclusion of locked nucleic acids (Michailidou, Beesley et al.), modifications of the backbone with phosphorothioate linkages at the 5′ or 3′ ends, and 2′-O-methyl or 2′-fluoro modifications of the bases (Davis, Lollo et al. 2006, Davis, Propp et al. 2009, Lennox and Behlke 2010). Among these, the LNA results the most efficient modification showing the highest affinity in complementary RNAs targeting. It corresponds to a class of bicyclic RNA analogues showing a N-type (C3′-endo) conformation chemically maintained by the addition of 2′-O, 4′-C methylene bridge (Stenvang and Kauppinen 2008).

Several viral and non-viral based techniques have been proposed over the last years for the delivery of miRNA mimics into the cells both *in vitro* and *in vivo* (Pereira, Rodrigues et al. 2013). In our laboratory, Lesizza and collaborators, tried to identify the best non-viral *in vivo* delivery method to administer miRNA mimics (in the specific case, hsa-miR-590-3p and hsa-miR-199-3p) to the infarcted heart tissue in order to boost CM proliferation. These investigators

compared the efficiency of miRNA mimic transport of different lipids compositions after a single intracardiac injection, leading to the identification of the best protocol allowing the persistence of the mimics for 12 days after treatment using a lipofectamine-based commercial nanoformulation (Lesizza, Prosdocimo et al. 2017).

Alternatively, the miRNA-encoded gene corresponding to ether the pri-RNA or the pre-miRNA can be embedded into a viral vector, most notably an AAV vector. This design permits the choice of a suitable promoter (constitutive or tissue-specific) and the prolonged expression of the miRNA. Using this approach however, both miRNA strands can be used by RISC from the precursor miRNA to generate the final effector molecule. Thus, in contrast with the use of synthetic miRNA mimics, the unwanted effect of the strand opposite to the one of choice is unavoidable.

1.3.5 MicroRNAs target identification

Since it is known that miRNAs can modulate the expression of several mRNA, trying to identify the miRNA targets is commonly a very demanding and complex process. A first approach takes advantage of bioinformatics prediction. Over the last years, many tools have actually been developed to bioinformatically predict miRNA targeting (Targetscan, miRanda, Diana, miRDB, RNA22), based on the sequence matching between the seed sequence and the 3'-UTR of genes. Still, even when considered collectively, these tools continue to remain largely imprecise in predicting the actual miRNA targets.

Experimental techniques are clearly needed to identify the actual miRNA target genes. A first approach is commonly the analysis of the transcriptome in cells transfected with miRNA mimics or anyhow expressing the miRNA of interest. Since miRNAs act by downregulating their targets, the downregulated genes in the transcriptome include the miRNA direct targets. Unfortunately, a disadvantage of this approach is the fact that many down-regulated genes can be the result of an indirect effect, rather than being the real targets. Thus, the definition of a real miRNA target eventually requires direct validation. This is

commonly achieved first by studying the putative target mRNA levels upon miRNA mimic transfection and then by validating the specific seed interaction of the miRNA with the target, by using segments of the mRNA target cDNA (commonly, the 3'-UTR) downstream a reporter gene (commonly, luciferase) to verify its effective downregulation upon miRNA transfection or expression (Jin, Chen et al. 2013).

As previously mentioned, the activity of miRNAs requires the correct assembly of several molecules inside the RISC complex. Thus, miRNA target identification can also be attempted through biochemical techniques based on immunoprecipitation of this complex, followed by transcriptome analysis. Easow and collaborators proposed a cross-linking immunoprecipitation (CLIP) approach combined with the microarray analysis of the selected transcripts (Easow, Teleman et al. 2007). The CLIP technique permits crosslinking and stabilization of the RNA/proteins interactions and consequently facilitates the recovery of the mRNAs linked inside RISC. Moreover, to improve the specificity of the RNA/protein crosslinking, a method named PAR-CLIP was developed (Thomson, Bracken et al. 2011). This requires the random or site-specific integration of photoactivable nucleosides analogs inside the nascent RNAs of living cells in order to probe the RNA/protein interaction. Upon UV irradiation and protein immunoprecipitation, a 100- to 1000- fold increase in the RNA recovery was observed, compared to the classic CLIP techniques (Baigude, Ahsanullah et al. 2012). Finally, Orom and Lund proposed an alternative technique to determine the miRNA:mRNA interactions based on the administration of biotinylated miRNAs to the cells, followed by the isolation of the deriving complexes using streptavidin-linked beads (Orom and Lund 2007).

1.4 Human genetic cardiovascular disorders, the clinical demand

Genetic abnormalities have been largely recognized as one of the major etiological causes of cardiomyopathies. Indeed, in the last decade, high-throughput sequencing analysis has revealed many pathological mutations present in both sporadic and familial cardiomyopathies (Herman, Lam et al. 2012, Ware, Seidman et al. 2016). Historically, one of the first evidence of genetic cardiovascular diseases (CVD) derives from the studies of Lehrman and co-workers, who sequenced the low-density lipoprotein receptor (LDLR) gene in a patient affected by familial hypercholesterolemia and characterized by a homozygous mutation. Their results confirmed a 5 Kb depletion in the sequence of the gene (Lehrman, Goldstein et al. 1985). A few years later, in 1989, other mutations in β cardiac myosin heavy chain were recognized as the main cause of hypertrophic cardiomyopathy (Jarcho, McKenna et al. 1989). A number of subsequent studies provided a comprehensive picture of the contribution of genetic mutation to cardiac disorders. Three main classes of human genetic cardiomyopathies can be considered as the main causes of inherited cardiac disease. These are hypertrophic, dilated and arrhythmogenic right ventricular cardiomyopathies.

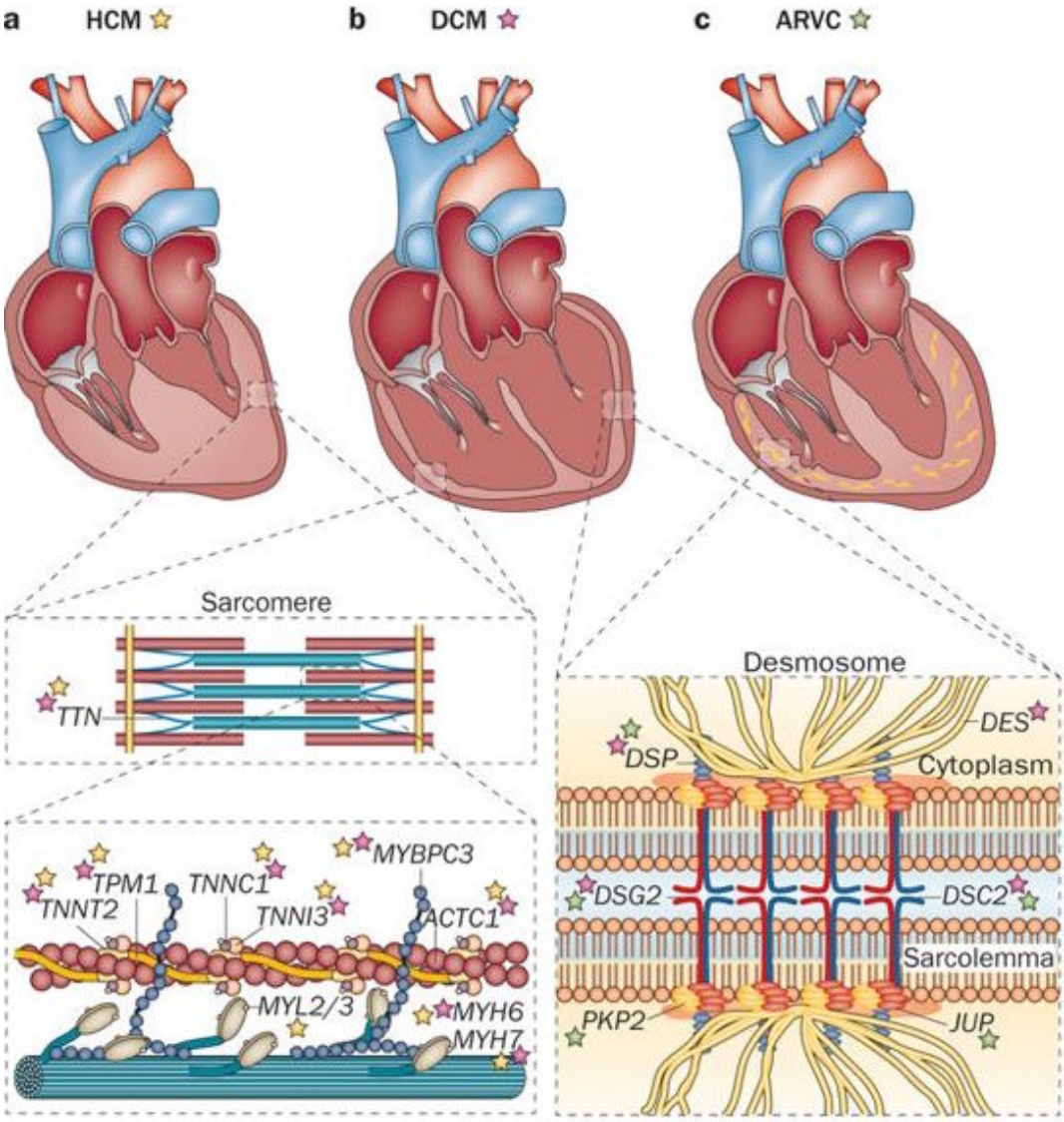


Figure 9: Schematic representation of the most frequent disease causing mutations. There are three main classes of genetic determined cardiomyopathies: a) Hypertrophic cardiomyopathies (HCM); b) Dilated cardiomyopathies (DCM); c) Arrhythmogenic right ventricular cardiomyopahies (ARVC). HCMs and DCMs are caused by mutations in genes coding for sarcomeric proteins, while ARVCs correlate with desmosome genes alterations. Adapted from (Yacoub 2014).

1.4.1 Hypertrophic cardiomyopathy

Hypertrophic cardiomyopathies (HCMs) are one of the most prevalent classes of cardiac dysfunctions worldwide. They are characterized by an unexplained ventricular hypertrophy that can be clinically ranked from asymptomatic, until causing sudden cardiac death (SCD) (Maron, Towbin et al. 2006). It is typically considered as a single gene disorder and usually mutations concentrate on the genes encoding for sarcomeric proteins. The most frequent variants are present in two genes coding for the β-myosin heavy chain (MYL7) (Geisterfer-Lowrance, Kass et al. 1990) and the myosin biding protein C (MYBPC3) (Ho 2012). In particular, in MYBPC3, more then 180 mutations have been annotated, the majority of which correspond to frameshift or nonsense mutations. In particular, one frequent alteration is a G to A transition identified in the last nucleotide of exon 6 (Richard, Charron et al. 2003); this mutation, occurring in 14% of all HCM patients in Tuscany, is associated with a severe phenotype and a poor prognosis in humans. To study the molecular determinant of the disease and to develop new possible therapeutic approaches, the group of L. Carrier created a knock-in transgenic mouse carrying this particular human point mutation in the corresponding mouse exon 6 of cMyBP-C sequence. Using this model, the Carrier group reported that the G to A transition in the exon 6 induced the formation of 3 different mutant mRNAs with missense, nonsense and deletion/insertions mutations (Vignier, Schlossarek et al. 2009). In another paper (Gedicke-Hornung, et al. 2013) the same authors demonstrated that, by inducing the exclusion of exons 5 and mutated exon 6 through targeted 2′-O-methyl phosphorothioate-(2OMePS)-antisense oligoribo-nucleotides (AONs), the alternative spliced mRNA gave rise to a functional protein. The systemic administration of an AAV9 vector coding for the AON to newborn Mybpc3-knocked-in mice prevented systolic dysfunction and the consequent ventricular hypertrophy (Gedicke-Hornung, Behrens-Gawlik et al. 2013). A number of other mutations have been reported as responsible for HCM development, including alterations in cardiac alpha-actin (ACTC1) (Morita, Rehm et al. 2008) and myosin regulatory light chain (MYL2) (Macera, Szabo et al. 1992).

Despite the large range of HCM related mutations, the metabolic and mechanical dysfunction in all HCM cases relies on an increased energetic cost and an overall reduction of the sarcomeric power of contraction (Burke, Cook et al. 2016).

1.4.2 Dilated cardiomyopathy

Dilated cardiomyopathy (DCM) commonly presents with ventricular dilatation and impaired myocardial contractility. The disease is typically reported in combination with heart failure, but, in the case of genetic alterations, it is commonly associated with arrhythmia (Bozkurt, Colvin et al. 2016). The genetic background that correlates with the development of the disease is usually very heterogeneous. Many DCM-causing mutations in key protein coding genes may generate the disruption or dysfunctions of the sarcomere, alterations in the cytoskeleton, in mitochondrial components, in desmosomes or in the nuclear membrane.

One of the first annotations in this direction comes from the paper of Gerull and collaborators, where they described specific alteration in the TTN gene, encoding for the giant muscle filament titin (Gerull, Gramlich et al. 2002). Moreover, in 2012 Herman and co-workers demonstrated that between 15 and 25% of DCM cases are caused by truncations in TTN sequence (Herman, Lam et al. 2012). Alterations in the same gene were also underlined in a subset of peripartum cardiomyopathies, usually associated with the LV function recovery after pregnancy (van Spaendonck-Zwarts, Posafalvi et al. 2014). Another class of DCM causing mutations includes non-synonymous variants of the LMNA gene, coding for lamin A and C. Approximately 6% of DCM cases show variants in the LMNA gene, with a presumed mechanism involving atypical and pathological mechanosignalling to the nucleus (Worman 2018). Autosomal dominant DCMs are commonly caused by missense or frameshift mutations that may occur throughout the coding sequence. This class of DCM shows a peculiar phenotype with sinus node dysfunction, atrioventricular block, atria and ventricular fibrillation, atrial flutter and ventricular tachycardia (Malhotra and Mason 2009).

Another class of DCM-causing mutation was recognized in the PLN gene. It encodes for a 52 amino acids residue transmembrane protein (phospholambdan), that, if unphosphorilated, inhibits the activity of the sarcoplasmic reticulum Ca^{2+}-ATPase (SERCA). Alterations in this gene are associated with lethal ventricular arrhythmias (Haghighi, Kolokathis et al. 2006).
In the paper of Brauch and co-workers they identified RBM20 as another key mutated gene associated with DCM development (Brauch, Karst et al. 2009). Indeed, RBM20 regulates the splicing of the TTN gene (Guo, Schafer et al. 2012). Consequently, its mutation could represent a downstream molecular effect that clinically associates with the alterations occurring in the TTN variants.
Moreover, alterations in ion channels were also described as associated with DCM development. Indeed, in the paper of McNair and collaborators they recognized missense mutations of SCN5A gene in familial cases of DMCs with high risk of arrhythmias (McNair, Sinagra et al. 2011).

1.4.3 Arrhythmogenic right ventricular cardiomyopathy (ARVC)

This class of diseases determines a progressive loss of the right ventricular tissue with its consequent substitution with fibrous and adipose tissue (Fontaine, Frank et al. 1984). The disease commonly correlates with the detachment of cardiac myocytes due to cardiac desmosome proteins aberrations. To date, mutations in 5 genes have been described as causative of with ARVC development: plakophili2 (PKP2) (Gerull, Heuser et al. 2004), desmoplakin (Reilly, McKoy et al.) (Rampazzo, Nava et al. 2002), desmoglein-2 (DSG2) (Pilichou, Nava et al. 2006), desmoscollin-2 (DSC2) (Syrris, Ward et al. 2006) and junctional-plakophilin (JUP) (McKoy, Protonotarios et al. 2000). Despite not all the patients present one or more of these mutations, together they represent the 50% of the cases with an autosomal dominant transmission.

1.4.4 Perspectives of gene correction for the treatment of inherited cardiomyopathies

In the last decade many gene correction strategies were optimized to implement possible clinical applications. The optimization of delivery systems (AAV), the designing of RNA-based approaches and the improvement of specific nuclease activity (CRISPR/Cas9), drastically paved the way towards the development of successful therapies in cardiac field.

Moreover, an increased understanding of the genetic alterations leading to the development of several cardiomyopathies represents a grate starting point in the preclinical setting. Indeed, many of the heterogeneous disease causing mutations described in the precedent paragraphs represent monogenic and well-characterized alterations in structural and metabolic genes. Most of them are autosomal dominant or recessive nucleotides variants or little indels, which create in frame or out of frame sequence modifications. Moreover, the mutation in the stop codon or the generation of a premature stop codon can create truncated proteins. Usually, this category of alterations represents a perfect target for HDR-mediated gene correction. Nevertheless, this repair pathway is largely inefficient if compared with NHEJ in post-mitotic tissues, such as the heart. Indeed, NHEJ was recently exploited in many preclinical studies to induce a partial correction of disease causing mutations in the Duchenne Muscular Dystrophy (DMD). This pathology affects both heart and muscles and 13 % of patients present mutations in exons 48-50, with the consequent loss of the protein (Aartsma-Rus, Fokkema et al. 2009).

Therefore, by the use of CRISPR/Cas9 and NHEJ-mediated genomic inactivation of the splicing inclusion sequence of mutated exon 51, it was possible to efficiently and permanently rescue the reading frame and expression of dystrophin in postnatal mdx mice and converting the DMD into a milder Becker Muscular Dystrophy (BDM) phenotype (Long et al 2017). A similar approach was also reviewed in the paper of Aartsma-Rus and collaborators, where they proposed an antisense oligonucleotide (Macera, Szabo et al.) to promote the exclusion of exon 51 (Aartsma-Rus, Straub et al. 2017). This molecule received the FDA approval for the treatment of DMD patients in 2016 under the name of eteplirsen. The same technology was successfully tested on LMNA mutations

known to alter its splicing and promoting the production of prelamin A (Lee, Nobumori et al. 2016). Furthermore, AONs were also used to induce an exon skipping in the genome of a mouse model carrying the Mybpc mutation to produce a different transcript with the consequent improvement in cardiac function (Gedicke-Hornung, Behrens-Gawlik et al. 2013).

Nevertheless, HDR still remains the gold standard for precise gene correction. Despite it is clear that this pathway infrequently occurs in post-mitotic tissues, CRISPR/Cas9-based gene editing approaches demonstrated a successful HDR-mediated dystrophin correction in a fraction of post-mitotic myofibers of a DMD mouse model (Bengtsson, Hall et al. 2017). Moreover, other papers reported the possibility to use HDR to correct specific sequences also in the heart of postnatal mice (Xie, Zhang et al. 2016, Ishizu, Higo et al. 2017).

2. MATERIALS AND METHODS

2.1 Cells

HeLa (cervical epithelial) and U2OS (osteosarcoma) cells were seeded in T75 flasks (Falcon) and grown in Dulbecco's modified Eagle medium (DMEM) + [GlutaMAX-I] + [Pyruvate] 1 g/L D-Glucose (Life Technologies), supplemented with 10% Fetal bovine serum (FBS, Life Technologies), 100 U/ml penicillin and 100 µg/ml streptomycin (Sigma), at 37°C in a humidified 5% CO_2 incubator. Mouse cardiomyocytes plated in primaria™ (Corining) plates, were kept in culture during the experiments in Dulbecco's modified Eagle medium (DMEM) + [GlutaMAX-I] + [Pyruvate] 4.5 g/L D-Glucose (Life Technologies) a higher Glucose concentration, 20 mg/ml vitamin B12 (Sigma) and 5% FBS. Differently, HL-1 cells, an immortalized cardiomyocyte mouse cell line, were grown in Claycomb medium (Sigma), supplied with 10% FBS, 100 U/ml penicillin and 100 µg/ml streptomycin (Sigma) and 2 mM L-Glutamine and kept in culture in primaria™ cell culture dish (Corning). Subconfluent cultures of HeLa, U2OS and HL-1 were split every 2-3 days using Trypsin. Briefly, the medium was aspirated, cells were washed with phosphate-buffered saline (PBS) (KH_2PO_4 1.8 mmol/L; NaCl 137 mmol/L; Na_2HPO_4 10 mmol/L; KCl 2.7 mmol/L) and 1 ml of 1X Trypsin

(diluted in PBS) were added to detach cells. After cell detachment, cells were resuspended in culture medium and re-seeded for propagation of the cultures (typically 1:5 dilution) or for experiments.

2.2 MicroRNA mimics and siRNAs

MiRNA mimics are small double-stranded RNA sequences, corresponding to the same sequences of endogenous human miRNAs. All the miRNA mimics used in this project (miRIDIAN miRNA mimics) and siRNAs (siGENOME SMARTpools) were purchased from Dharmacon (GE Healthcare).

2.3 MicroRNAs/siRNAs transfection

Transfection of human miRNA mimics or siRNAs was performed using a standard reverse transfection protocol, at a final miRNA/siRNA concentration of 50 nM. For experiments performed in 96-well plates, 15 µl of 500 nM miRNAs/siRNAs were transferred to 96 well plates. Transfection reagent, Lipofectamine RNAiMAX Transfection Reagent (0.2 µl per 96-well, Life Technologies) diluted in 35 µl of OPTI-MEM (Life Technologies) was then added to the miRNAs/siRNAs arrayed on 96-well plates and incubated at RT for 30 min. HeLa, U2OS cells, and mouse CMs, suspended in 100 µl antibiotic-free culture medium, were then seeded onto the 96-well plates at different densities depending on the cell line (HeLa: $8x10^3$ cells per well; U2OS: $6x10^3$ cells per well; CMs: $1x10^4$ cells per well) and incubated at 37°C, 5% CO_2. For experiments in 384-well plates, 5 µl of 500 nM miRNA/siRNA were transferred to each well and 0.1 µl of Lipofectamine RNAiMAX Transfection Reagent (Life Technologies) diluted in 15 µl OPTIMEM was added. After 30 min, cells were seeded onto the plates at a density of $1.5x10^3$ cells per well and incubated at 37°C, 5% CO_2. Transfections performed in 6-well plate format were done using the same procedure, but transferring 7 ul from the 20 uM miRNAs/siRNAs stock to each well. 5.7 ul of Lipofectamine RNAiMAX were diluted in 600 ul of OPTI-MEM and then added to miRNAs/siRNAs. 30 min. after, $3x10^5$ cells per well, resuspended in 2.5 ml of culture medium, were seeded on the transfection reagents in the plate.

2.4 Production, purification and characterization of rAAV vectors

The AAV Vector Unit (AVU) at ICGEB Trieste (http://www.icgeb.org/avu-core-facility.html), generated as described previously (see paragraph 1.2.5) all the AAV vectors used in this project. Briefly, infectious AAV6 and AAV9 vector particles were generated in HEK293T cells by co-transfecting each vector plasmid together with the packaging plasmid expressing AAV capsid proteins and Adenovirus helper proteins, pDF6 for AAV6 (King, Dubielzig et al. 2001) and p5E18-VD2/9 for AAV9 (Inagaki, Fuess et al. 2006). Plasmid p5E18-VD2/9 (providing AAV2 Rep and AAV9 Cap genes) was provided by Dr James M. Wilson. Viral stocks were obtained by CsCl gradient centrifugation; rAAV titers, determined by measuring the copy number of viral genomes in pooled, dialyzed gradient fractions, were in the range of $1x10^{12}$ to $1x10^{13}$ genome copies per ml. A list of the AAVs used in this project is summarized in the table 1.

Tab. 1	
Recombinant Vector serotype	Coding cassette
AAV serotype 6	SgRNA_ES#1
AAV serotype 6	Myl2_T2A_eGFP
AAV serotype 9	SgRNA_ES#1
AAV serotype 9	Myl2_T2A_eGFP
AAV serotype 9	hsa-miR-302d
AAV serotype 9	hsa-miR-520c

2.5 The high-throughput screening

The high content screening was performed in U2OS cells transfected with a human microRNA mimics library, composed by 2,042 mature double-stranded RNA sequences (miRBASE, version 21.1, October 2018) arrayed in 384 well plates. A standard reverse transfection protocol was used to transfect cells with the library and the day after miRNAs transfection, U2OS were treated with the

triple-plasmid HR detection assay to evaluate HR-mediated GFP correction and expression. A final concentration of 36 ng of DNA/well was transfected to cells (5 ng of pCUB-eGFP-Y66S; 11.5 of px330_SpCas9; 19.5 ng of pGEM T easy_delta20GFP*), diluted in 20 ul of OPTiMEM and using FuGENE HD (Promega) (with a ratio 3.5 ul FuGENE HD each ug of DNA) as transfection reagent. The GFP accumulation was monitored 72 hours after plasmids transfection. Cells were then washed twice with PBS, fixed with PFA 4% for 15 min. at RT, permeabilized with 0.5% Tryton-X diluted in PBS for 10 min., washed again with PBS, blocked with 2% BSA diluted in PBS and stained with an Ab (ab6556, Abcam) recognizing the total amount of GFP to normalize the results over transfection efficiency. The screening was run in duplicate. With this unbiased method we identified a list of human miRNAs strongly increasing CRISPR/Cas9-based HR, up to 3-fold over control.

2.6 AAV transduction in HL-1 and cardiomyocytes

Twenty-four hours after miRNA/siRNA transfection, recombinant AAV6 vectors coding for the sgRNA and for the template for correction were added to Cas9+ mouse cardiomyocytes and HL-1 cells with different ratio between the two viruses. Upon transduction condition optimization, the vectors were added to cells at specific multiplicities of infection (m.o.i.), corresponding to 10,000 vgp/cell in total. Ninety-six hours after transduction, the medium was removed, cells were washed with PBS and processed for fluorescence imaging.

2.7 Fluorescence Microscopy, image acquisition and analysis

To prepare the samples for fluorescence imaging, U2OS cells were fixed with 4% paraformaldehyde (Hummel, Li et al.) for 15 min. and washed with PBS. Cells were then permeabilyzed using 0.5% Tryton X-100 (Sigma), washed once more with PBS and blocked with 2% bovine serum albumin (BSA), diluted in PBS. After 1 hour, we immunostained cells with an Ab recognizing both folded and misfolded eGFP, produced in rabbit (ab6556, Abcam), diluted in blocking solution (1:2,000) keeping cells overnight in darkness at 4°C. By doing this, we

were able to normalize the data over the transfection efficiency. Cells were then washed twice with PBS and a secondary fluorescent (Alexa 594) anti-Rabbit Ab (diluted 1:500 in blocking solution) was added to cells and incubated for 2 hours at room temperature (RT). After this time, cells were washed again with PBS and Hoechst 33342 (diluted 1:5,000, Life Technologies) was used to counterstain nuclei. Differently, mCMs, after fixation (PFA 4%, 10 min.), permeabilization (0.5% TrionX-100 for 10 min.) and blocking (2% BSA for 1 hour), were stained with a combination of two Abtibodies: the first corresponds to the same used to stain U2OS cells (ab6556, Abcam), while the second recognizes the sacromeric alpha-actinin (produced in Mouse), diluted in blocking solution 1:250 (EA-53, Abcam). Plates were kept at 4 °C overnight in darkness. The day after, cells were washed twice with PBS and anti-mouse 594 (Invitrogen) and anti-rabbit 488 (Invitrogen) secondary Ab (diluted 1:500 in blocking solution) were added together and incubated for two hours at RT. After this time, cells were washed again with PBS and Hoescht (Invitrogen, #33342) was added. The screening high content image acquisition was performed using ImageXpress Micro automated high-content screening fluorescence microscope (Molecular Devices). Images were acquired at 10x magnification to reach approximately 1,500 cells analyzed per condition. Image analysis was performed using the "Multi-Wavelength Cell Scoring" application module implemented in MetaXpress software (Molecular Devices). Whereas, all the experiments not related to the screening plates were acquired at 20x magnification with a total of 15 images per well using the Operetta high-content imaging system (PerkinElmer) and analyzed with Columbus image data storage and analysis platform (PerkinElmer).

2.8 Total RNA extraction from 6-well plates

Total RNA was extracted from U2OS and cardiomyocytes seeded onto 6-well plates by using TRIzol Reagent (Life Technologies) taking care at recovering the aqueous phase, without interphase contamination. Cells were lysed in 1 ml of TRIzol Reagent per well and homogenate was recovered, put in Pre-spun Phase Lock Gel-Heavy tubes and incubated for 5 min at RT. 0.2 ml of chloroform per ml

of TRIzol was added and the tubes were shaken vigorously for 15 s. Samples were then centrifuged at 12,000 x g for 10 min at 4°C. After the centrifugation, the upper aqueous phase, was transferred to a new tube and 1.5 volumes of 100% ethanol was added. 700 ul of the samples, including any precipitate, were transferred into RNAeasy mini spin columns (Qiagen) and centrifuged at 8000 x g for 30 s. After this step, the flow-through was discarded, 700 µl of RWT Buffer was added to the columns and the tubes were centrifuged for 15 s. at 8000 x g to wash the columns. After the washing step, 500 µl of RPE Buffer was added and the samples were centrifuged for 15 s. at 8,000 x g. This step was repeated a second time by adding 500 µl of RPE Buffer and centrifuging tubes for 2 min. at 8000 x g. The RNeasy Mini spin columns were placed into a new 2 ml collection tube and centrifuged at full speed for 1 min. This additional step was done to eliminate any possible carryover of Buffer RPE or other residual. Tubes were then left opened under chemical hood to allow the complete evaporation of the EtOH and further dry the membrane. Finally RNeasy Mini spin columns were transferred to new 1.5 ml collection tubes and 50 µl of RNase-free water were added directly onto the RNeasy Mini spin columns and incubated for 2 min.. Columns were centrifuged at 8000 x g for 1 min. to elute the RNA. This step was repeated twice to improve the RNA concentration.

2.9 Reverse Transcription PCR

Total RNA concentration was measured using Nanodrop 1000 (Thermo Scientific). 1 µg of total RNA was used for cDNA synthesis. The RNA was diluted into 11 µl of RNAse-free water and treated with 2 µl of RNAse-free DNAse I (Roche), 1.5 µl of Buffer 10x (Roche) and 0.5 µl of PRI (Roche) for 15 min, at RT. DNAse I was then inactivated by rising the temperature to 75°C for 5 min. After DNAse digestion, RNA was incubated with 2 µl of 10 mM dNTPs (Promega), 2 µl of random primers (Invitrogen, 1:20 dilution) and 5 µl of RNAse-free water for 5 min at 65°C. Reverse transcription was performed by using 8 µl of 5x First-Strand Buffer (Invitrogen), 4 µl of DTT 0,1 M (Invitrogen) and 2 µl of PRI (Protector Rnase Inhibitor). The samples were incubated for 2 min at 37°C and

then 2 µl of Maloney murine leukemia virus reverse transcriptase (M-MLV RT, Invitrogen) were added. The reverse transcription reaction was performed as follows: 10 min at 25°C (RT activation), 50 min at 37°C (RT reaction), 15 min at 72°C (RT inactivation).

2.10 Real-Time PCR

To measure how different miRNAs affect expression of particular genes, we performed Real-Time PCR using the TaqMan assay (Applied Biosystems). GAPDH was used as housekeeping gene for normalization. cDNA (1 µl) was treated with 10 µl of TaqMan mix (Qiagen – not correct), 8 µl of RNAse-free water, 1 µl of the TaqMan probes and analized by using a standard TaqMan protocol using a CFX96^{TM} Real-Time System (Bio-Rad) (95°C for 10 min, followed by 40 cycles of 95°C for 15 sec and 60°C for 1 min), according to the manufacturer's recommendations. TaqMan probes consist of a fluorophore (R) covalently bind to the 5'-terminus of an oligonucleotide probe and a quencher (Q) on the 3'-terminus.
In the beginning, quencher and reporter are very close and Q quenches the fluorescence emitted by R, when exited by a specific wavelength. During the annealing, TaqMan probes bind to the target sequence. Then, during the elongation phase, the Probe is partially displaced, the reporter is cleaved and becomes able to produce fluorescence. The resulting fluorescence is proportional to the quantity of amplified product in the samples. A list of the used probes is summarized in the tab. 2 and 3.

Tab. 2	
TARGET	Probe ID
hsa-MRE11	Hs00967437_m1
hsa -NBN	Hs01039834_m1
hsa -RAD50	Hs00990023_m1
hsa -RAD51	Hs00947967_m1

hsa -MAPK11	Hs00177101_m1
Tab. 3	
hsa-GAPDH	1505059 (Applied Biosystems)

2.11 Total RNA extraction from Heart sections

Heart sections of treated animals were lysate with TRYzol (Invitrogen) and dissociated using MagNA Lyser green beads (Roche) at 6500 rpm for 60 seconds. After this, the homogenate was put in 1.5 ml tubes and incubated for 5 min at RT. 0.2 ml of chloroform per ml of TRIzol were added and the tubes were shaken vigorously for 15 s. Samples were then centrifuged at 12,000 x g for 10 min at 4°C. Total RNA was extracted using NucleoSpin miRNA kit (Machery-Nagel), according with manufacture's instruction.

2.12 miRCURY miRNAs quantification

To determine the miRNAs expression in the extracted hearts we used the miRCURY LNA™ Universal RT microRNA PCR (EXIQON). Total RNA was extracted (as described 9n the section 2.11) and then diluted to a final concentration of 5 ng/ul using nuclease free water. Reverse transcription reaction was prepared by using 2 ul of diluted RNA, according with manufacture's instructions and following the subsequent PCR conditions: 60 min. at 42°C; 5 min. at 95°C; immediately at 4°C. The deriving cDNA was then diluted 1:80 in nuclease-free water. Real-Time reaction was prepared according with manufacture's instructions and using the primers listed in the Tab. 4 to detect hsa-miR-520c-3p and hsa-miR-302d-3p. As internal control we used the U6 small nuclear RNA, stably expressed in the tissue.

Tab. 4	
Probes	Product num.

hsa-miR-520c-3p	YP00204497
hsa-miR-302d-3p	YP00204311
RNU6-1 RNA, U6 small nuclear 1	203907

2.13 Cell cycle profile evaluation

U2OS cells were seeded in 6-well plate and reverse transfected with hsa-miR-520c-3p, hsa-miR-302d-3p, hsa-miR-4469 and mimic#4. 72 hours after transfection, 1X Trypsin was used to detach the cells and centrifuged at 1300 rpm for 1 min. The pellet was then resuspended and washed twice with PBS and the surnatant was removed. Upon the washing step, cells were resuspended in 1 ml of 50 mg/ml propidium iodide (PI) (Sigma #P4864) solution diluted in water and supplied with 0.1% $Na_3C_6H_5O_7$, 0.1% NP40 and 12.5 ul of 0.5mg/ml RNAse, DNAse free (Roche, #11119915001). Cells were incubated for 20 min. at RT in the darkness. Cytofluorimetric analysis was performed by considering a population of 10,000 cells for each sample.

2.14 Western-Blot quantification

Protein expression was evaluated by western-blot quantification: treated cells were washed in ice-cold PBS and Proteins were harvested in RIPA lysis buffer (20 mM TrisHCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 0.5 % sodium deoxycholate, 0.5 % NP40, 0.1 % phosphatase inhibitors (Sigma P0044), 1 % PMSF, EDTA free proteinase inhibitor tablet (Sigma 04693159001) (1 tablet/10 ml buffer)). Cell lysates were then sonicated with Bioruptor (Diagenode) for 30 minutes using 3 pulses of 10 minutes each. After this, samples were then centrifuged at 13,000 rpm for 20 min. at 4°C, the surnatants were collected and protein concentrations were quantified by Bradford colorimetric assay (BioRad, CA, USA). Samples were resuspended with Laemmli sample buffer (30% glycerol, 3% SDS, 0.19 M Tris HCl, pH 6.8, 0.015% bromophenol blue, 3% 2-mercaptoethanol), denatured by boiling for 5 min at 90 °C and resolved by SDS-PAGE of the relevant percentage according to the protein of interest (8-12% gels). Proteins were then

transferred to a nitrocellulose membrane (GE Healthcare, 10600015), at 350 mA for 1 hour at 4°C. After this step, membranes were blocked for 1 hour at RT using 5% milk or 5% BSA, diluted in TBS-T (0.1% Tween20). Primary Abs were diluted in blocking buffer, added to membranes and incubated overnight at 4°C. Blots were then incubated with HRP conjugated secondary Abs (see tab. 5) for 1 hour at RT. GAPDH was used as standard control through all the experiments. Band intensity was analyzed using ImageJ software. In the tab. 5 are listed the Ab used during the project.

Tab. 5		
PRIMARY Antibodies		
TARGET	COMPANY CODE	SPECIES
CHK-2	Cell Signaling, BK2662S	Rabbit
P-CHK-2	Novus Biologicals, NB-100 92502	Rabbit
H2AX(ser139)	Millipore, 05-636	Mouse
MRE11	Santa Cruz, sc-5859	Rabbit
NBS1	Santa Cruz, sc-11431	Rabbit
P-NBS1 (ser343)	Novus Biologicals, NB100-92610	Rabbit
RAD50	Novus Biologicals, NB110-60438	Rabbit
RAD51	Santa Cruz, sc-8349	Rabbit
P38α/β	Cell Signaling, 9212	Rabbit
P-P38	Cell Signaling, 9211	Rabbit
GAPDH	Santa Cruz, sc-25778	Rabbit
HSC-70	Stressgen, SPA-815F	Rat
SECONDARY Antibodies		
HRP-Anti-Mouse	DAKO, P044701	Goat
HRP-Anti-Rabbit	Thermo Scientific, 31460	Goat
HRP-Anti-Rat	DAKO, P0450	Rabbit

2.15 Deep-sequencing analysis

The RNA-seq analysis was performed in collaboration with Macrogen Inc. (Seoul, Korea). After total RNA extraction from un-transfected samples or samples treated with the negative control (mimic#4) or with the selected miRNAs, the company provided us the subsequent specification for RNA processing: Illumina TruSeq Stranded mRNA Library construction (library prep,

stranded, polyA enrichment); Illumina NovaSeq 2x100bp sequencing (library run, 30 million reads per sample). The transcriptomic analysis was run in triplicate and data analysis was performed in collaboration with Danilo Licastro from "Consorzio per il centro di Biomedicine Molecolare" (CBM), AREA science park, Basovizza, Trieste.

2.16 T7 endonuclease Assay

The T7E1 assay was performed following the manufacturer's instructions (M0302, New England Biolabs, MA, United States). 100 ng of genomic DNA (extracted from cells treated with plasmids expressing Cas9 and the different sgRNAs or transfected only with Cas9 coding plasmid) were used to amplify the target region. The PCR reactions were conducted, using the oligonucleotides (Sigma-Aldrich) listed in Tab. 6, according to the protocol of the Taq polymerase used (Roche). The PCR program was optimized with the following cycles: 1 cycle of 2 minutes at 94°C; 35 cycles of 30 seconds at 94°C, 45 seconds at 58°C and 1 minute at 72°C; 1 cycle of 7 minutes at 72°C. The PCR amplicons were purified using Wizard SV Gel and PCR Clean-Up System (Promega) following the manufacturer's instructions. 500 ng of purified amplicons were heat denatured and slowly re-annealed to facilitate heteroduplex formation between the WT and the cleaved sequence by using optimized PCR conditions: 5 min. at 95°C; 3 min. at 90°C; 3 min at 65°C; 3 min at 63°C; 3 min at 37°C and 10 min at 25°C. The eteroduplex, diluted in 1X NEB2 Buffer, were then digested with the T7E1 enzyme (New England Biolabs, NEB) at the mismatch locus for 60 min. at 37°C. After this step, samples were loaded in a polyacrylamide gel to separate the resulting digested bands, using 100 bp DNA ladder (Promega) as referring marker. To evaluate the separation the gel was soaked in diluted Ethidium Bromide and then irradiated with UV at the transilluminator.

Tab. 6	
T7 endonuclease Primers	Sequences
My2_Exon7_Fw Primer	CTCTTGCTTGCAGATCGACC
My2_Exon7_Rev Primer	GGCAGGTGGATCTCTGAGTT

2.17 3'UTR-Luciferase Assay

In order to determine possible targeting of miRNA-4469 to the 3'UTR of MAPK11, a portion of the 3'UTR sequence was purchased from GenScript and cloned into the reporter vector psiCheck2 (Promega) using the restriction enzyme PmeI and NotI (New England Biolabs). HeLa cells were reverse transfected on day 0 in a 96 well comparing UNTR_LIPO with the selected miRNAs. After 24h cells were transfected with empty psiCheck2 vector of the vector containing the MAPK11 3'UTR sequence. After additional 48 hours cells were washed with PBS, harvested in 60 ul GloLysis Buffer (Promega) and assayed using the DualGlo Luciferase kit (Promega). Renilla luminescence was normalized over firefly luminescence. In addition, to account for the effect of the miRNAs on plasmid transfection, an additional normalization step was carried out by dividing the psiCheck2-3'UTR measurements by the corresponding empty vector measurements for each miRNA treatment.

2.18 Mouse cardiomyocyte preparation

Ventricular CMs were isolated from both male and female neonatal mice. Briefly, ventricles from neonatal (P0) mice were cut into pieces inside a 2.6 cm petri dish filled with and ice-cold calcium and bicarbonate-free Hanks with HEPES (CBFHH) buffer. Upon cutting the hearts in pieces, the fragments were transferred in a falcon, resuspended in an appropriate volume of Trypsin and dissociated using mechanical and physical stirring. Digestion was then performed at 37°C in 4-min steps for 6-8 times with collection of the supernatant to ice-cold fetal bovine serum (FBS, Life Technologies) after each step. The collected supernatant was passed through a cell strainer (40 mm, BD Falcon) and centrifuged to pellet the cells, which were then resuspended in Dulbecco's

modified Eagle medium 4.5 g/l glucose (DMEM, Life Technologies) supplemented with 5% FBS, 20 mg/ml vitamin B12 (Sigma) and 100 U/ml penicillin and 100 mg/ml streptomycin (Sigma). After centrifugation, cells were then seeded onto 100-mm plastic dishes for 3 hours at 37°C in 5% CO_2 and humidified atmosphere to allow fibroblast attachment to the dish surface. After this time, the supernatant, composed mostly of CMs, was then collected, cells were counted and plated at the appropriate density into collagen-coated ULTRA Primaria 96-well plates (PerkinElmer).

2.19 Mice injection, hearts extraction and slices preparation

All the animals were housed with optimal environmental conditions. All animal experiments followed the European guidelines and international laws and policies (ECC Council Directive 87/609, OJL 348, 12 December 1987) with the approval of the ICGEB Animal Welfare Board, the Italian Minister of Health and the Ethical Committee.

Neonatal P1 Cas9+ mice were intraperitoneally injected twice in a day using Micro-Fine™ 0.3 ml syringes (30G) (Jansen, Embden et al.). Animals were therefore treated firstly in the morning with 30 ul of high titers of AAV9 coding for hsa-miR-520c-3p and hsa-miR-302d-3p ($4x10^{11}$ vgp/mouse) and secondly, in the evening, they were injected with 30 ul of the combination of AAV9 coding for the guide and the template for eGFP HR-mediated insertion ($2.5x10^{11}$/mouse for each virus). After one month, animals were sacrificed, the hearts were perfused using PBS to allow a complete washout of the remaining blood, extracted and immediately cut on the longitudinal direction. Half of the heart was fixed overnight at 4°C using PFA 4% and processed for fluorescence microscopy, while the other half was use for RNA extraction in order to quantify the miRNAs overexpression.

3. RESULTS

3.1 High content screening to identify human microRNAs affecting CRISPR-Cas9-based homologous recombination

Genome editing based on homologous recombination (HR) (Thomas and Capecchi 1987), is the gold standard treatment for gene therapy of inherited mutations. However, this process is based on a very rare event, in particular in post-mitotic cells and tissues, such as neurons or cardiomyocytes. The advent of the CRISPR-Cas9 technology has dramatically changed the perspectives in this field, allowing researchers to introduce specific double stranded cuts in the genome to boost HR. Nevertheless, this tool is not yet optimized for clinical applications and there is a strong need to find strategies to improve the efficiency of HR-based gene correction.

To identify possible factors able to improve the efficiency of CRISPR-Cas9-based HR, we optimized a robotized high content screening approach to study these events. Since HR is a very sporadic and complex biological process, we decided to investigate the possibility to improve its efficiency using endogenous, multi-target molecules, such as human microRNAs, given the property of these molecules of targeting multiple genes at the same time and their involvement in numerous biological processes, as outlined in the Introduction.

To optimize the possibility to identify a detectable number of events in a high-content setup, we designed a transient HR recognition assay. For this purpose, we transfected the human U2OS osteosarcoma cell line with a combination of three plasmids: an HR reporter, a DNA template for correction and a CRISPR-Cas9 coding plasmid also containing information for a single-guide RNA sequence (figure 10 panel A). As an HR reporter we took advantage from an already published plasmid, pCUB-eGFP-Y66S (Tran, Liu et al. 2003). This codes for an eGFP sequence carrying a single point mutation (A to C) at position 201 of the EGFP gene, resulting in Tyr66 to Ser66 amino acid conversion (pUC_eGFP_Y66S). This mutation allows the production of an improperly folded eGFP and completely abrogates its fluorescence. The homologous template for correction is provided by the pGEM T easy_delta20GFP* plasmid and

corresponds to a truncated eGFP sequence lacking the ATG and the last 20 nucleotides and carrying a silent point mutation inside the PAM region, in order to avoid Cas9 recognition after gene correction. A third plasmid codes for the humanized sequence of the *Streptococcus pyogenes* Cas9 and for the single guide RNA (gRNA-eGFP5) targeting the eGFP sequence and driving the nuclease specificity. Cas9 and sgRNA expression are driven by the constitutive CMV and the human U6 RNA promoters, respectively.

By co-transfecting these plasmids into U2OS cells, we could visualize the appearance of green fluorescent cell clusters exclusively in the samples treated with the complete combination of the three plasmids (figure 10, panel B). After fixation, we stained the cells with an antibody recognizing both the folded and the unfolded eGFP proteins, in order to normalize the data for transfection efficiency. Representative images in figure 10 panel C show the green GFP+ cells corresponding to successful gene targeting events, and the total transfected cells immunostained in red (Alexa 594); nuclei are stained I blue with Hoescht.

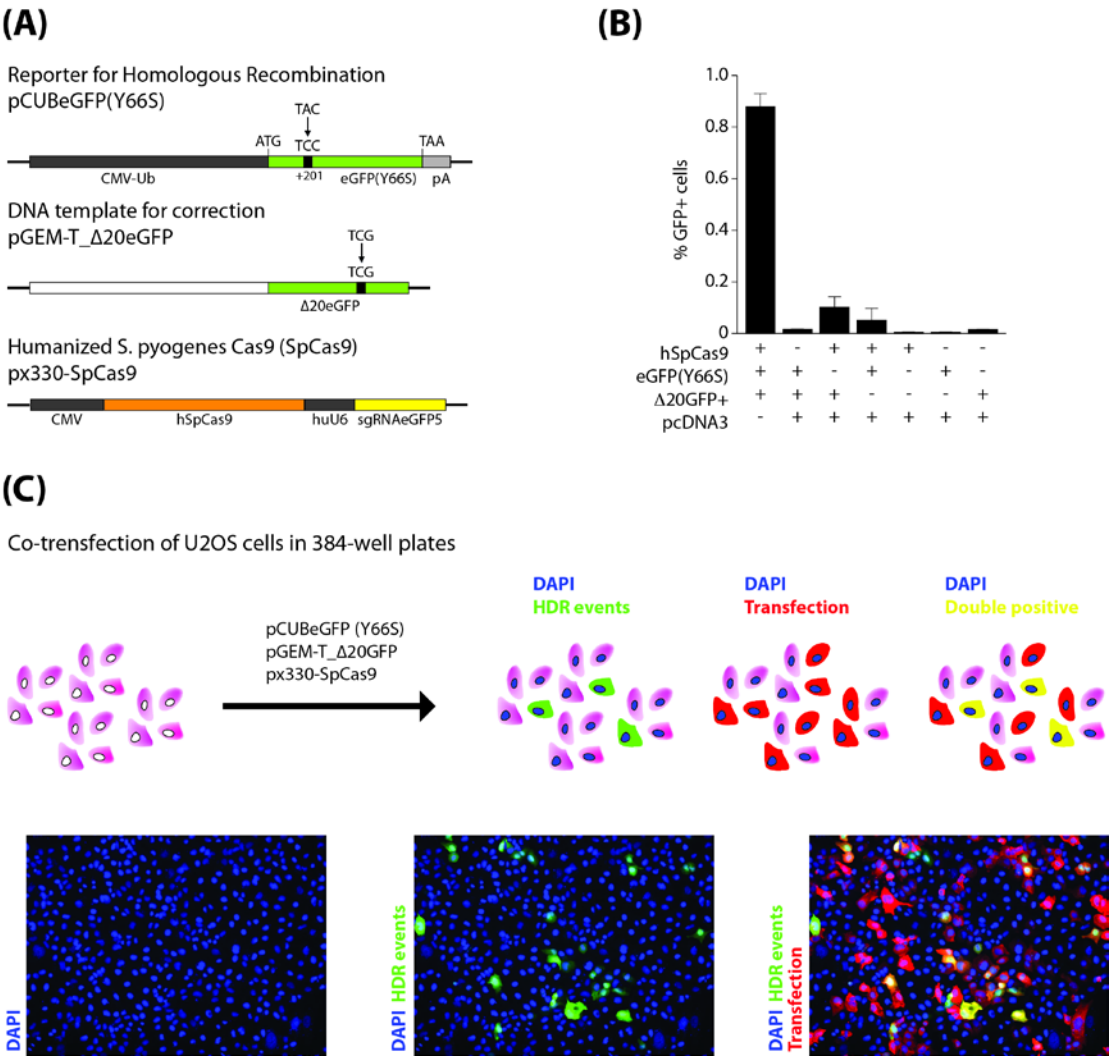
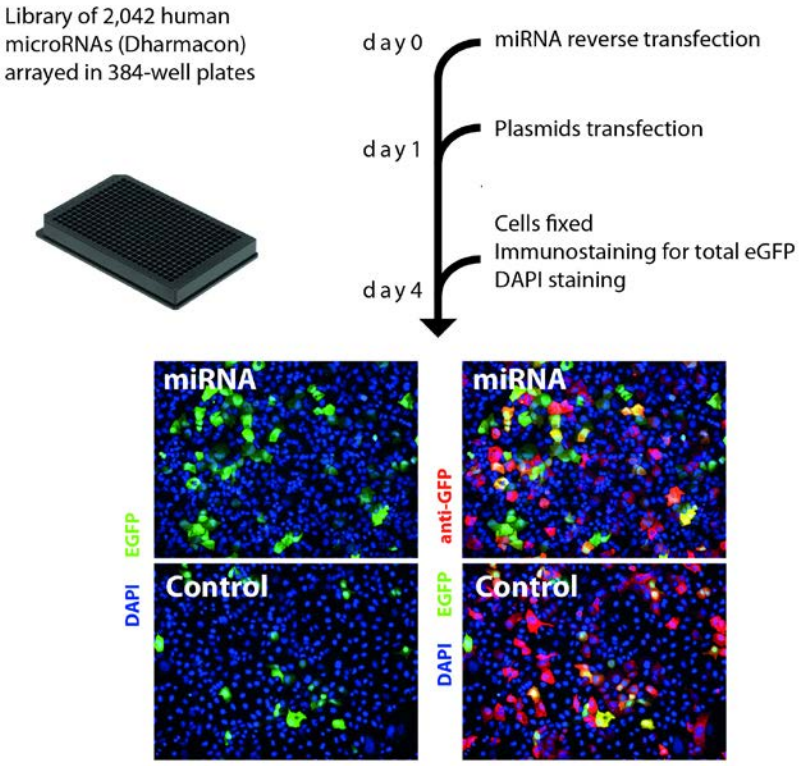


Figure 10: (A) The HR detection assay consists of a triple plasmid transfection protocol in U2OS cells: the first plasmid codes for the HR reporter, a mutated eGFP sequence, point-mutated in lysine 201, which creates a misfolded non-fluorescent protein; the second plasmid is the template for correction (a truncated eGFP sequence carrying a silent point-mutation in the PAM recognition site); the third plasmid codes for the humanized version of the *Streptococcus Pyogenes* Cas9, together with a guide RNA targeting the eGFP DNA sequence (SgEGFP5). (B) Basal level of HR detection in samples co-transfected with the complete combination of the three plasmids. (C) Schematic summary of images acquisition and representative images of cells in which recombination had occurred. GFP+ cells correspond to HR events; the red cells (stained with an Alexa594-labeled antibody against both folded and unfolded GFP) permits visualization of transfection efficiency. Nuclei are stained with Hoescht.

3.2 High-throughput screening protocol and image processing

Once we set up the experimental conditions for the HR assay, we proceeded with the high-throughput screening. For this purpose, we plated and reverse transfected U2OS cells with a genome wide library of 2,042 human miRNAs (human miRNA mimic library, miRBASe version 19) arrayed in a 384-well plate format. After 24 hrs, we transfected the three plasmids of our assay and, after additional 72 hours, we fixed the cells and immunostained them for total eGFP expression in order to quantify transfection efficiency. We acquired a total of 4 images for each well and the screening was performed in duplicate (Figure 11).



As negative controls in our screening, we used two miRNA mimics (mimic#1 – cel-miR-67 and mimic#4 – cel-miR-231) derived from from *C. elegans*, selected for the absence of any homology with human miRNAs and known to not induce any toxicity. As an additional control, we also included treatment with a siRNA against Ubiquitin C (UBC), a component of the proteasome-ubiquitin system that is essential for cell survival, the silencing of which leads to cell death. The use of this siRNA allows an easy evaluation of transfection efficiency, by simply counting the number of surviving cells.

3.3 The identification of human microRNAs able to positively modulate homologous recombination pathway

Using this unbiased approach, we identified a number of human microRNAs that are able to positively modulate HR frequency. The graph in the figure 12, panel A shows the distribution of the miRNAs exerting positive and negative effect on HR. The screening was performed in duplicate and the comparison of the results

Figure 11: We arrayed the 2,042 human microRNA mimics in 384-well plate format and plated U2OS cells in each well to achieve reverse-transfection. At 24 hours after transfection we added the 3 plasmids of the assyas and, 72 hours later, we fixed the cells and stained them with an Ab recognizing transfected eGFP (Alexa 594). The screening was performed in duplicate.

showed good reproducibility (P value = 0.10) (Figure 12, panel A). We identified 50 miRNAs positively affecting efficiency of HR by more than 2 fold (figure 12, panel B) and 20 miRNAs that significantly improved HR events up to 3-fold over mock treated control (UNTR_LIPO), without affecting plasmid transfection or cell viability.

Representative images in figure 12, panel C show the effect of six selected miRNAs, miR-302a-3p, miR-302d-3p, miR-520b, miR-520c-3p, miR-520e and miR-520f, which all increased the number of GFP-positive cell clusters, and indication of successful gene targeting events.

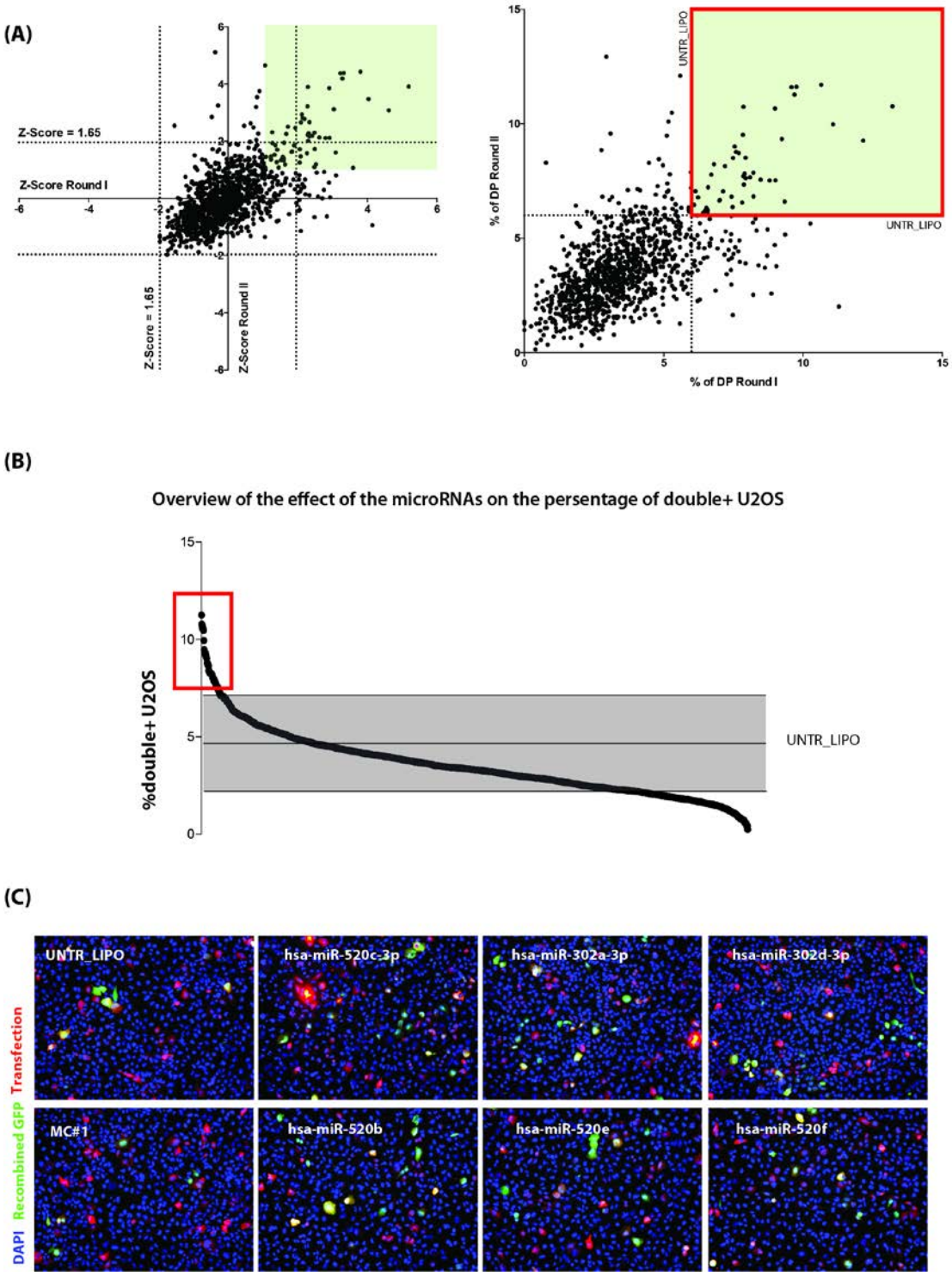


Figure 12: (A) The comparison between the two replicates of the screening resulted statistically significant (pvalue < 0.10) and reported a list of 50 miRNAs as positive modulators of HR up to 3-fold over UNTR_LIPO. (B) An overview of the results underlined how miRNAs both positively and negatively influence Cas9-mediated HR. (C) Representative images from the top miRNA hits.

We focused on the top 20 miRNA hits of our screening for further analysis.

First, we re-tested their ability to increase HR efficiency in U2OS cells using the recombination assay described above. As shown in figure 13, panel A, for all these miRNAs we could confirm the induction of GFP-positive, recombinant cells with an increment of two folds over control.

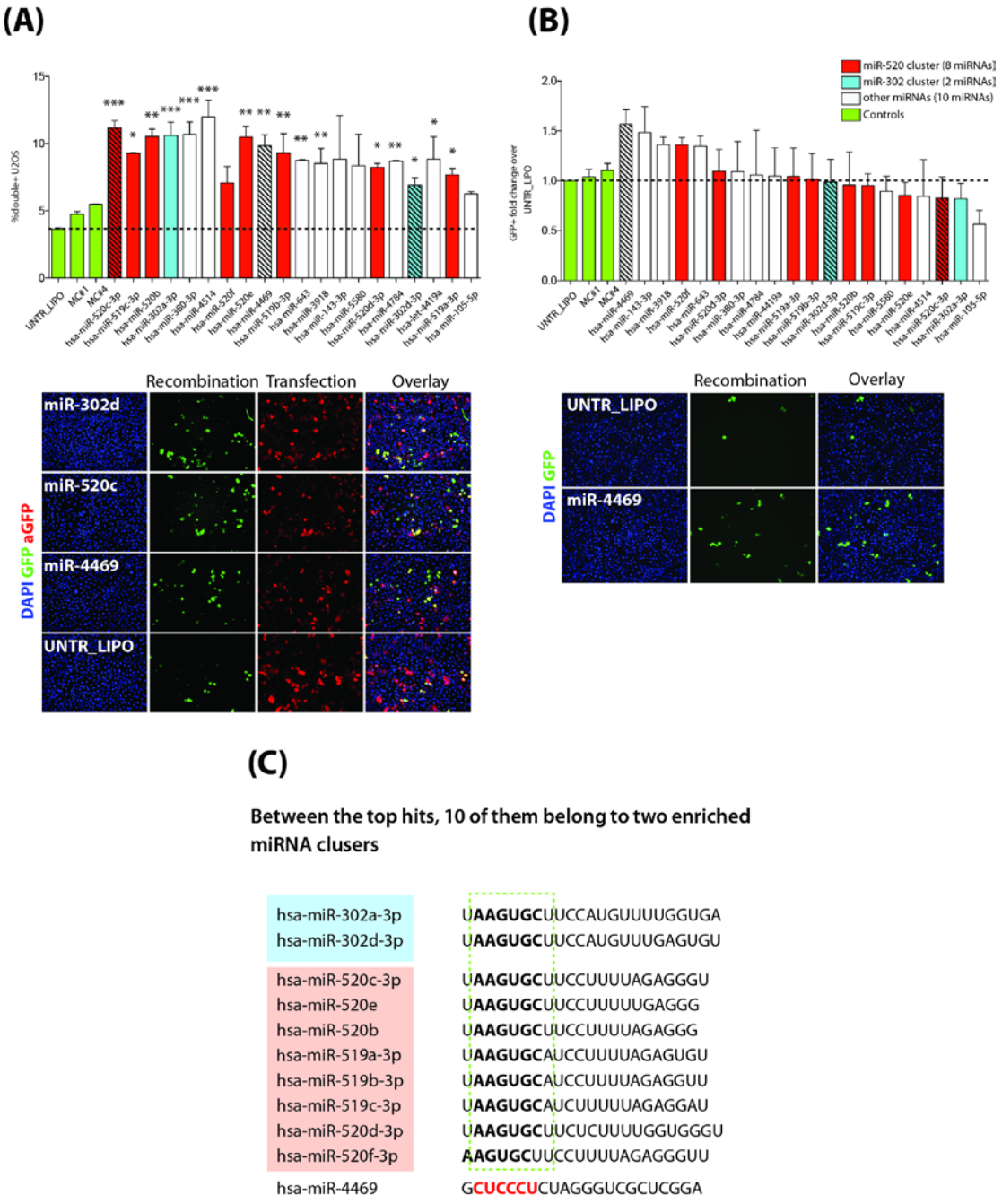


Figure 13: The graph in panel A corresponds to the validation of the miRNAs selected from the screening while panel B shows the effect of HR improvement in genome-integrated targets. Data are mean ± SEM (N = 3 individual experiments); *p < 0.05, **p < 0.01, ***p < 0.001; one-way ANOVA. The 8 miRNAs that belong to cluster 520 are marked in red, while the 2 from cluster 302, are in blue. The miRNAs in white are not included in the two clusters, while in green are the controls. In both graphs, the selected miRNAs used for further experiments are marked with striped columns. Panel C shows the conserved seed sequence shared by the hits of the screening that belong to the enriched clusters. The seed sequence of miR-4469 is

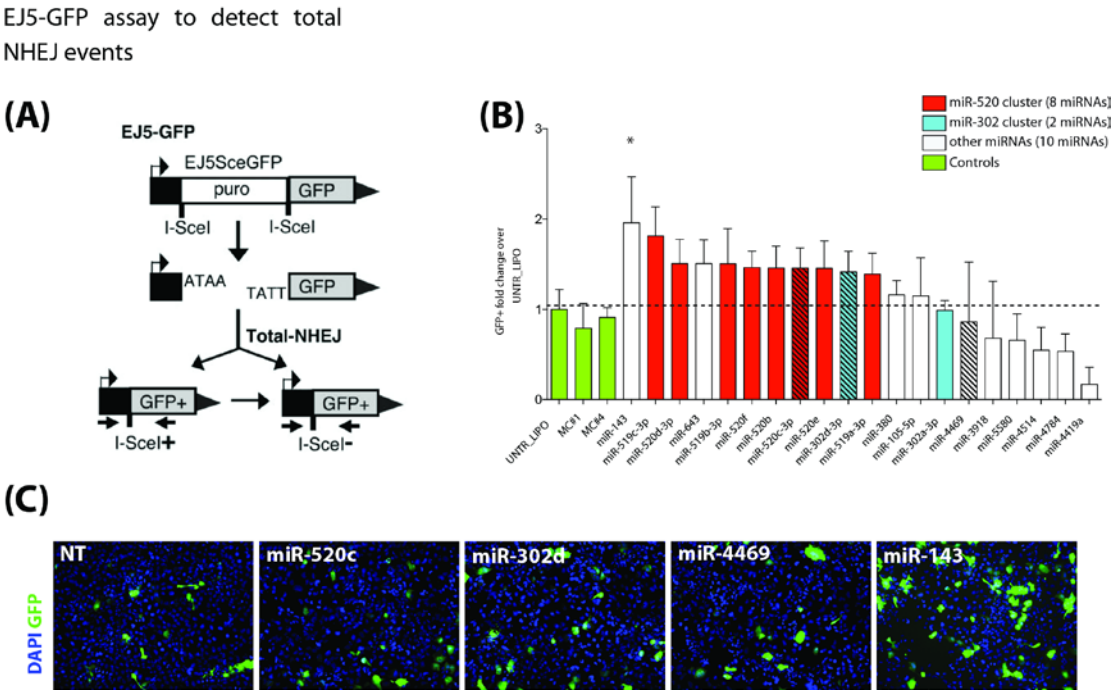
Interestingly, an accurate analysis of these miRNAs revealed that 10 out of the top 20 molecules belong to two enriched miRNA clusters (miR-520 and miR-302, figure 13, panel C) known to be overexpressed in embryonic and pluripotent stem cells (Ren, Jin et al. 2009) and almost absent in post-mitotic cells. The miRNAs belonging to the same cluster share the same seed sequence (AAGUGC), which most likely indicates a similar gene regulation profile and a common mechanism of action. Therefore, these miRNAs possibly lead to the improvement of HR frequency by modulating conserved pathways and gene networks.

In order to evaluate the effect of the top selected miRNAs hits on a genome-integrated target, we generated a stable cell line, carrying the eGFP-Y66S reporter sequence stably integrated into the AAVS1 locus of U2OS cells. Using these cells, we screened the top 20 miRNAs for their ability to induce recombination events and the appearance of GFP+ cells. Twenty-four hours after miRNAs treatment, we co-transfected cells with the template and the SpCas9 expressing plasmids and incubated them at 37°C for further 72 h. In this experimental setting, we obtained a different ranking of efficacy of the 20 different miRNAs, with miR-4469 resulting the most efficient at inducing HR events (figure 13 panel B). This miRNA has a different seed sequence (CUCCCU) compared to the miR-302 and miR-520 clusters (figure 13, panel B), meaning that its effect on HR is probably mediated by a different gene modification profile. Taken together the results of the validation experiments, we decided to concentrate our study on three molecules, hsa-miR-4469 plus two representatives of the miRNA clusters (hsa-miR-520c-3p and hsa-miR-302d-3p) figure 13 panels A and B (striped marked columns).

3.4 Evaluation of the effect of the selected miRNAs on the modulation of the non homologous end joining DNA repair pathway

Since HR is one of the two possible DNA repair mechanisms and mammalian cells preferentially use NHEJ to resolve DNA double strand breaks, we wanted to investigate whether the effect of the selected miRNA is restricted to the HR

pathway or also extends to NHEJ. We evaluated the possible effect of the miRNAs on activation of the NHEJ pathway by taking advantage of a reporter plasmid designed to identify all the possible classes of NHEJ events (Bennardo, Cheng et al. 2008). This plasmid contains a promoter and a GFP sequence that are separated by the presence of a puromycin-resistance cassette , flanked by



two I-Sce-I recognition sites that are in the same orientation (figure 14, panel A). When the reporter plasmid is co-trasfected in U2OS cells together with an I-Sce-I expression plasmid, the I-Sce-I enzyme cuts at the corresponding recognition sites eliminating the Puro sequence. Then, the two staggered end generated by the cut are repaired by NHEJ reconstituting the GFP expression cassette.

In cotransfection experiments, among the 20 different miRNAs tested, only miR-143 showed a significant improvement in the number of GFP+ cells (up to 2-fold over UNTR_LIPO), while all the other miRNAs were ineffective at stimulating the NHEJ-mediated DNA repair, underling their specific effect on HR.

3.5 The selected microRNAs modify the cell cycle profile

To start understanding the mechanisms of HR enhancement and since it is known from literature that the HR events occur in the S and G2/M phases of the cell cycle (Aylon, Liefshitz et al. 2004, Ira, Pellicioli et al. 2004), we decided to evaluate the cell cycle profile of U2OS cells after treatment with the selected miRNAs. In this experimental setup we used as negative control mimic#4 (previously used in the screening) and compared the possible cell cycle modifications induced by the 3 selected miRNAs (hsa-miR-302d-3p, hsa-miR-

Figure 14: (A) Schematic representation of EJ5-GFP reporter plasmid. The reporter corresponds to a WT eGFP sequence separated from its promoter by a puromycin resistance cassette flanked by two I-SceI recognition sites. By co-transfecting this construct together with the I-SceI expression plasmid, the puromycin cassette is excised and the GFP gene is expressed due to the cut followed by NHEJ-mediated DNA repair. Data are mean ± SEM (N = 3 individual experiments); *p < 0.05; one-way ANOVA. (B) Effect of the selected miRNAs on NHEJ. (C) Representative images of the selected miRNAs.

520c-3p and hsa-miR-4469) to the effect of UNTR_LIPO.

At 72 h after miRNA transfection, we collected the cells for propidium iodide (PI) staining and FACS analysis to evaluate the different cell cycle profiles (figure 15 panel A). The results of these experiments confirmed a significant reduction of cells in G1 (from 59% to 45-47%) together with a consequent accumulation in S

and G2/M (from 20% to 28-29%) (figure 15, panel B) upon treatment wih the pro-recombinogenic miRNAs. Interestingly, while this cell cycle alteration resulted evident in all the miRNA-treated samples, the S-G2/M shift was more evident in the samples treated with hsa-miR-520c-3p and hsa-miR-302d-3p compared to hsa-miR-4469, underlining their conserved pathway regulation.

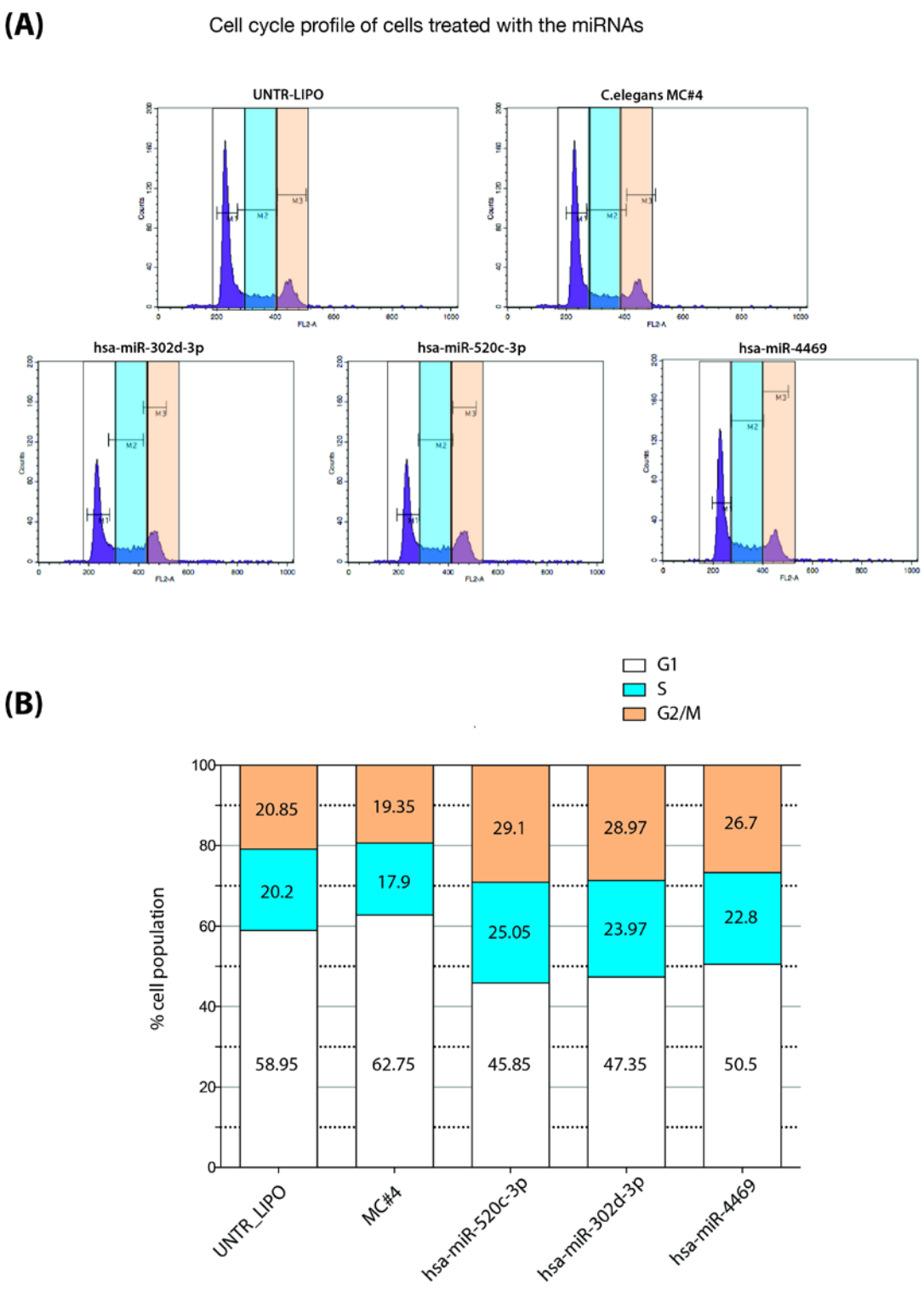


Figure 15: (A) Results of the cell cycle profile analysis of U2OS cells un-transfected and transfected with the negative control (mimic#4) or with the selected miRNAs (hsa-miR-520c-3p; hsa-miR-302d-3p; hsa-miR-4469). (B) The graph shows the percentage of the cells population in G1, S and G2/M phases.

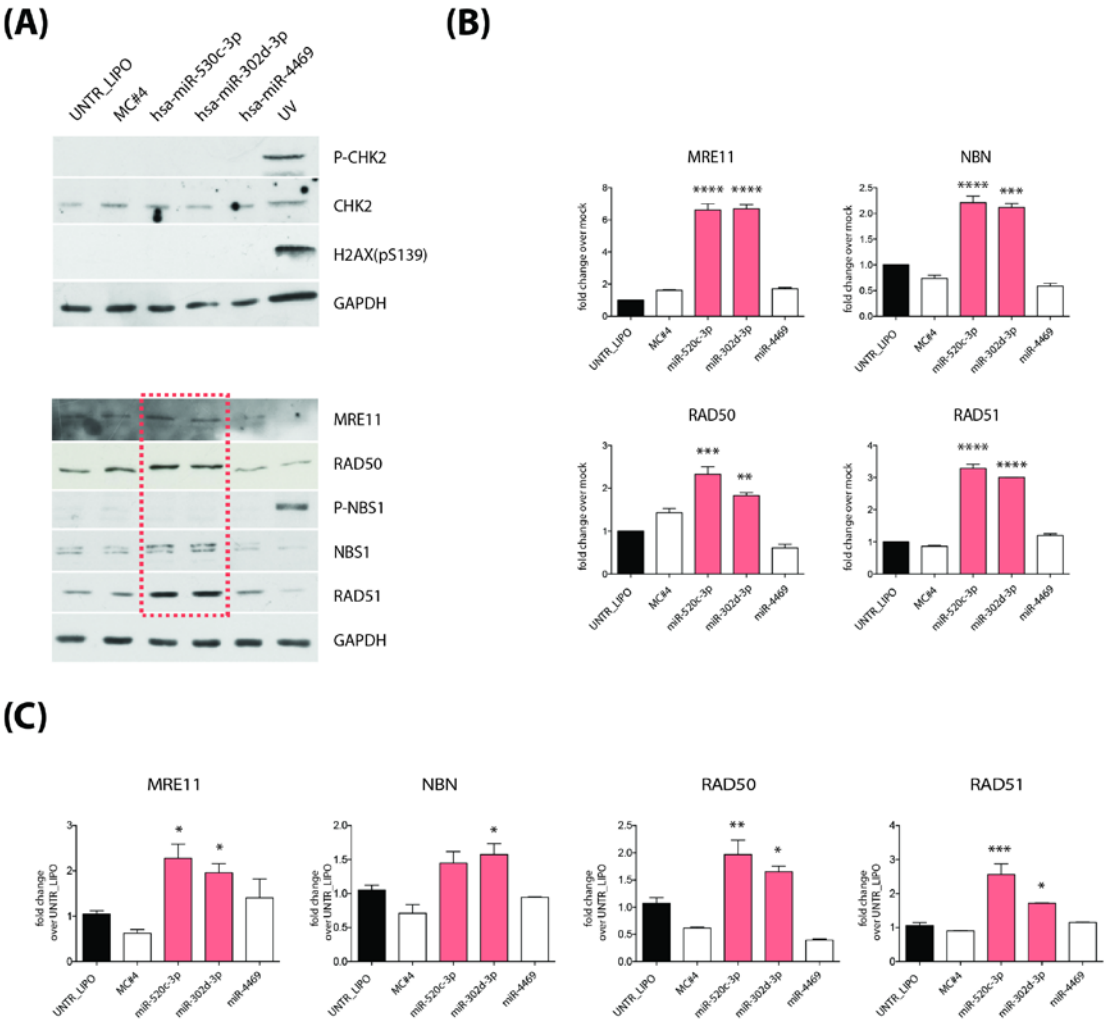
3.6 The selected molecules do not induce DNA damage and upregulate key components of homologous recombination machinery

Cellular DNA damage response (DDR) and repair mechanisms are triggered by stimuli that threaten DNA integrity. For this reason, to evaluate whether the miRNA treatment might induced by itself DNA damage, we performed a comprehensive survey of activation/expression of DNA damage markers in cells treated with the selected miRNAs. By western-blot analysis, using protein extracts from treated U2OS cells, we evaluated the activation and accumulation of Chk2, P-Chk2 and γ-H2AX proteins. The carboxy-terminal phosphorylation of H2AX is one of the earliest DNA damage responses of DSB formation (Fillingham, Keogh et al. 2006). As a positive control for DNA damage activation, we used UV treated U2OS cells (at a fluence of 20 J/m^2) (Hanasoge and Ljungman 2007).

The results of these western blots did not show any DNA damage marker activation in the miRNA-treated samples, in comparison with the UV positive control, in which the levels of both P-Chk2 and γ-H2AX increased (figure 16, panel A).

We then decided to evaluate the effect of the selected miRNAs on well-described components of the HR pathway. Upon DNA DSB formation, one of the first DNA repair steps is the recognition of the broken DNA ends by the MRN complex (Zha, Boboila et al. 2009, Taylor, Cecillon et al. 2010), composed by three key proteins: MRE11, NBS and RAD50. Once this occurs, another HR-associated protein, replication protein A (Ward, Norat et al.), accumulates in the region (Chen and Wold 2014). Then, this is rapidly replaced, through the activity of RAD52, by the most important component of the HR pathway, the ATP-dependent recombinase RAD51 (Bianco, Tracy et al. 1998). For these reasons, we decided to evaluate possible variations in gene expression and protein accumulation of the three MRN components and the ATP-dependent recombinase RAD51. Western blot analysis showed an accumulation of the three MRN proteins in both hsa-miR-520c-3p- and hsa-miR-302d-3p-treated samples (figure 16, panel B). Accordingly, in the same samples, we also detected a

significant RAD51 protein enrichment. A similar scenario was also confirmed by checking the mRNA levels of these proteins by Real Time PCR of cells in which we transfected the selected miRNAs, after normalization of the results over the same controls (figure 16, panel C). Therefore, despite the absence of DNA damage induction, these miRNAs contribute to upregulate the transcription and activation of components of the HR machinery. In contrast, and interestingly, we did not observe an equivalent phenotype in samples treated with miR-4469,



underlying that this miRNA affects HR through different pathways.

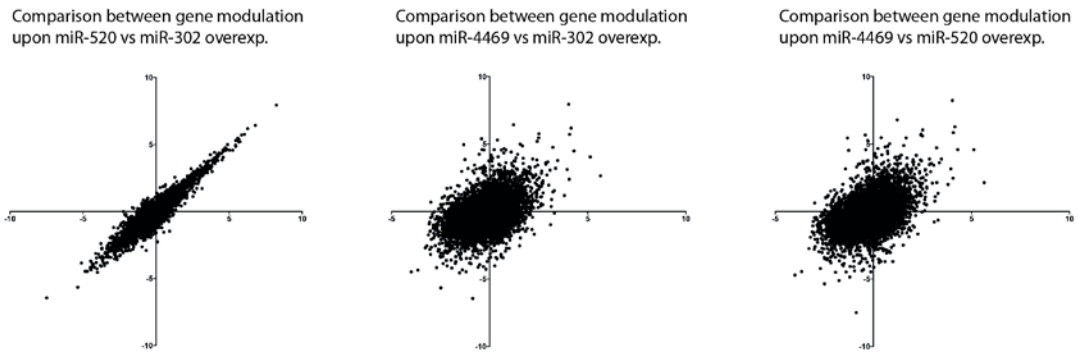
3.7 Analysis of the transcriptome in cells transfected with hsa-miR-520c-3p, hsa-miR-302d-3p and has-miR-4469

Trying to better understand the mechanism of action of the three selected miRNAs, we transfected U2OS cells with hsa-miR-520c-3p, hsa-miR-302d-3p and hsa-miR-4469 and, 72 hours after the transfection, we extracted total RNA and performed deep sequencing analysis (the sequencing was performed in collaboration with Macrogen, while data analysis was completed by Danilo Licastro from CBM, AREA science Park, Trieste). As expected, the results highlighted a high correlation between hsa-miR-520c-3p and hsa-miR-302d-3p transcriptomic profiles (figure 17 panel A), which was not detected by comparing the results obtained from hsa-miR-4469-transfected cells. On these data, bioinformatics analysis were performed by focusing on the most enriched and downregulated genes. The total gene modulation induced by each single treatment is summarized in the Heatmap in figure 17, panel B.

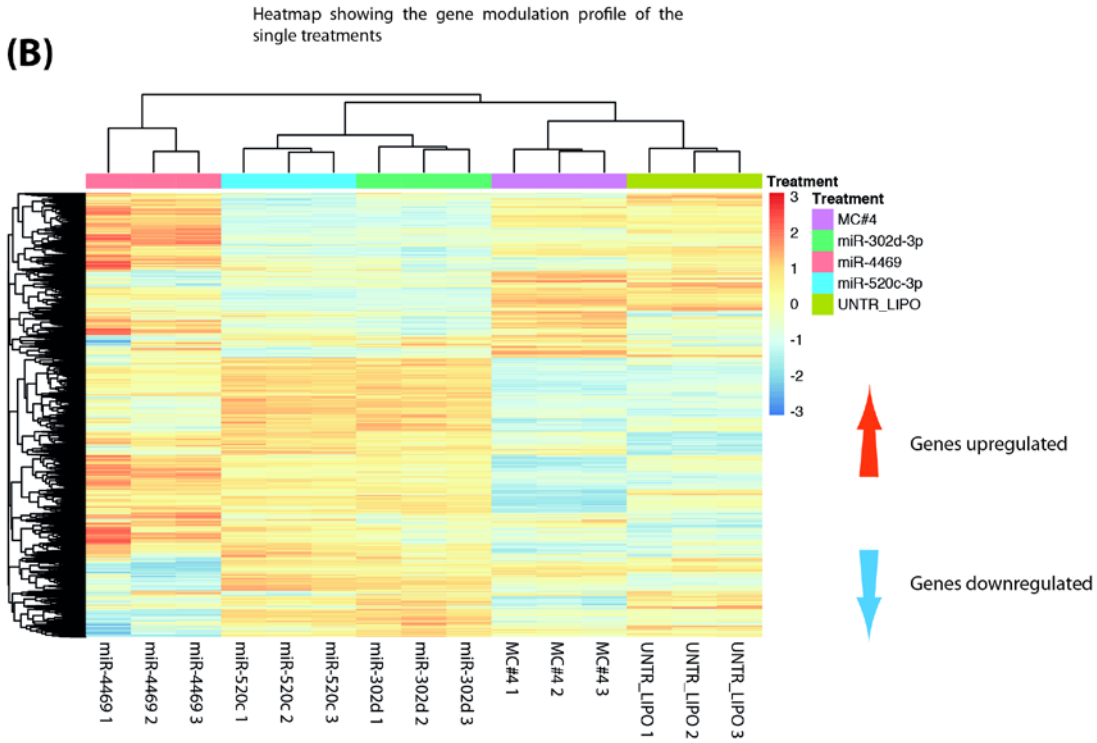
The results of this unbiased analysis highlighted the enrichment of key gene networks correlated with DNA repair and recombination. Moreover, accordingly with our previous experiments, we also confirmed how the miRNAs positively affected many HR-related genes, among which, RAD51 and the MRN complex (figure 17, panel C). Interestingly, we confirmed G1-related cyclins (CCND2) and CDKs (CDKN1A) as down regulated in this context. Furthermore, between the genes down regulated by hsa-miR-4469, we recognized MAPK11 (p38β), whose pathway alteration, together with that of the other MAPKs (ERK and JNK), is known to play a role in the HR-mediated DNA repair process (Golding, Rosenberg et al. 2007).

Figure 16: (A) Wester-Blot analysis and quantification (B) of un-transfected U2OS compared with U2OS transfected with mimic#4 or the miRNA hits to detect the activation of the DNA damage response and the accumulation of the HR related proteins (MRE11, NBN, RAD50 and RAD51). (C) Results deriving from real time quantification of MRE11, NBN, RAD50 and RAD51 mRNA expression in un-transfected and miRNAs treated cells. Data are mean ± SEM (N = 3 individual experiments); *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001; one-way

(A)


Comparison between gene modulation upon miR-520 vs miR-302 overexp.
Comparison between gene modulation upon miR-4469 vs miR-302 overexp.
Comparison between gene modulation upon miR-4469 vs miR-520 overexp.

(B)


Heatmap showing the gene modulation profile of the single treatments
Treatment
Treatment
MC#4
miR-302d-3p
miR-4469
miR-520c-3p
UNTR_LIPO
Genes upregulated
Genes downregulated
miR-4469 1
miR-4469 2
miR-4469 3
miR-520c 1
miR-520c 2
miR-520c 3
miR-302d 1
miR-302d 2
miR-302d 3
MC#4 1
MC#4 2
MC#4 3
UNTR_LIPO 1
UNTR_LIPO 2
UNTR_LIPO 3

(C)

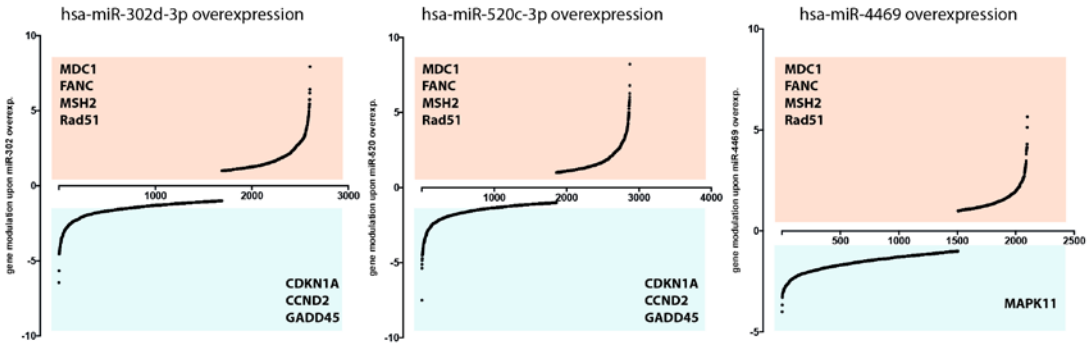

hsa-miR-302d-3p overexpression
hsa-miR-520c-3p overexpression
hsa-miR-4469 overexpression
MDC1
FANC
MSH2
Rad51
CDKN1A
CCND2
GADD45
MAPK11

Figure 17: (A) Comparison between the gene modulation profile of U2OS cells treated with hsa-miR-520c-3p, hsa-miR-302d-3p and hsa-miR-4469. (B) Heatmap of the single gene modulation for each treatment. (C) The graphs show the up-regulation and down-regulation of some conserved genes between the treatments.

3.8 MAPK11 (p38β), a possible hsa-miR-4469 target

Trying to understand the mechanism of action of miR-4469, we concentrated on the analysis of its transcriptomic profile. As previously mentioned, in the list of the down regulated genes we detected p38β (MAPK11) as a possible direct miRNA target. A previously published work (Golding, Rosenberg et al. 2007) recognized the signaling from the mitogen-activated protein kinases ERK, JNK and p38, as a possible regulator of HR. In particular, these investigators reported that, among the other MAPKs, p38β inhibition enhanced HR up to 2-folds over control. Therefore, we speculated on the possibility to correlate p38β inhibition with the effects of miR-4469 treatment described above. For this purpose, we first confirmed the specificity of miR-4469 in targeting MAPK11; both Real Time (figure 18, panel A) and western blot (figure 18, panel B) quantifications highlighted a significant reduction of MAPK11, at both the RNA and protein levels, only in miR-4469-treated samples. Furthermore, the predicted seed sequence of miR-4469 (according to TargetScan 7.2) exactly matches a corresponding sequence in two positions of the 3'UTR of the human gene (figure 18, panel C). Finally, we obtained functional evidence that MAPK11 is a direct target of miR-4469 using a 3' UTR-luciferase assay performed in HeLa cells. More specifically, transfection of a luciferase reporter plasmid (psiCheck2, from Promega) containing the human 3'UTR sequence of MAPK11 gene cloned at 3' end of the luciferase coding sequence, induced a large and statistically significant downregulation of luciferase activity in samples co-transfected with miR-4469 (figure 18, panel D). All together these data suggest that miR-4469 directly regulates p38β.

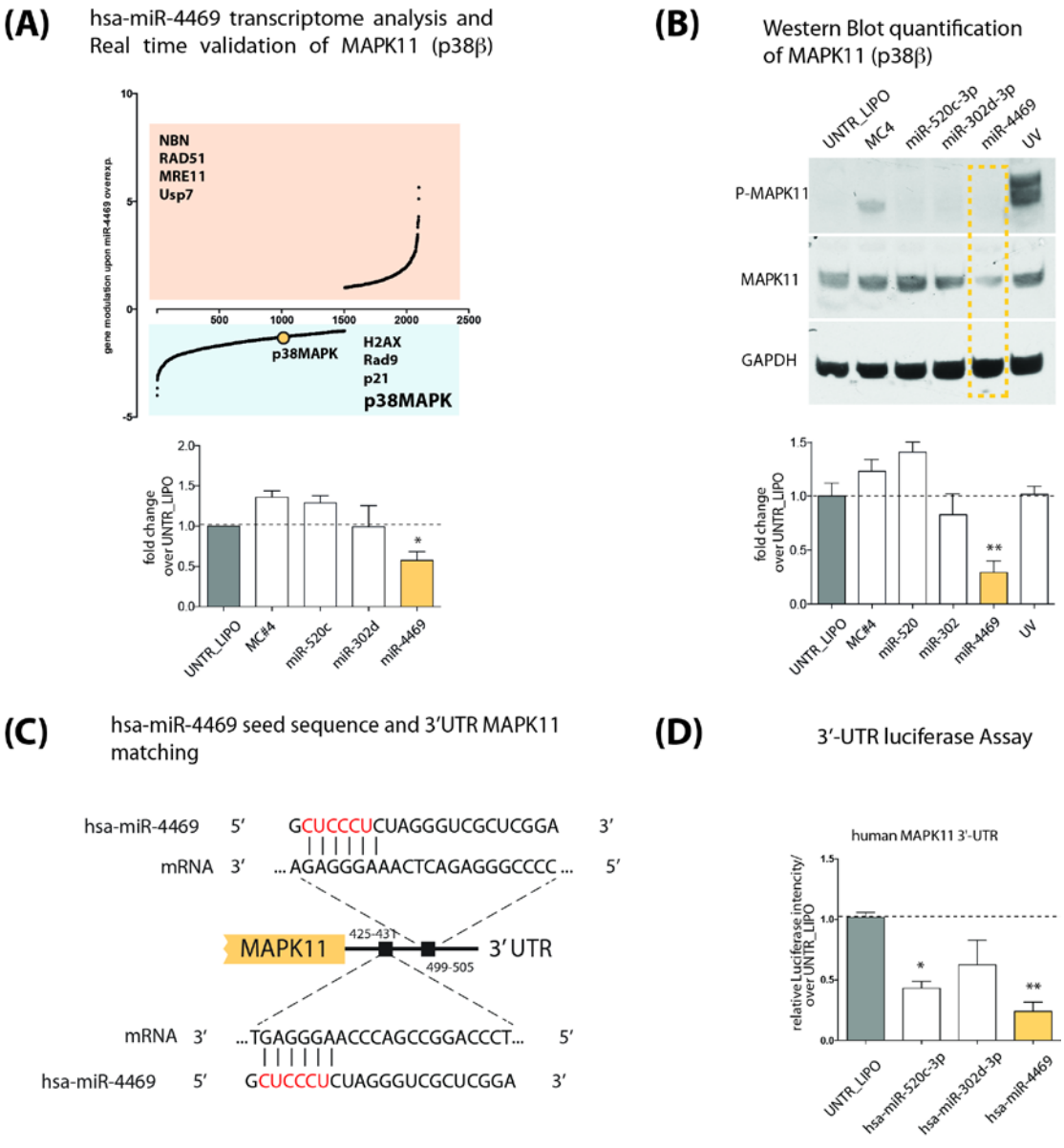


Figure 18: (A) Gene expression profile of miR-4469 treated U2OS cells. (B) Western Blot quantification of MAPK11 and P-MAPK11 total protein samples deriving from U2OS cells treated with the selected miRNAs. (C) MiR-4469 seed sequence matching with the human MAPK11 3'-UTR. (D) 3'-UTR luciferase assay performed in HeLa cells with or without the selected miRNAs overexpression. Data are mean ± SEM (N = 3 individual experiments); *p < 0.05, **p < 0.01; one-way ANOVA.

3.9 Evaluation of homologous recombination modulation by specific MAPK11 inhibition

Starting from these observations, we decided to evaluate the effect of p38β inhibition on HR-mediated DNA repair by using the same assay used for the screening. First, we tested different concentrations of two well-described highly specific p38 chemical inhibitors (SB203580 and ARRY757) to detect a possible improvement on the frequency of HR events. The results obtained (data not shown) indicated that the inhibitor SB203580, at the concentration of 10 µM, was the most effective at improving the HR frequency in these experimental settings and therefore it was used in all the subsequent tests.

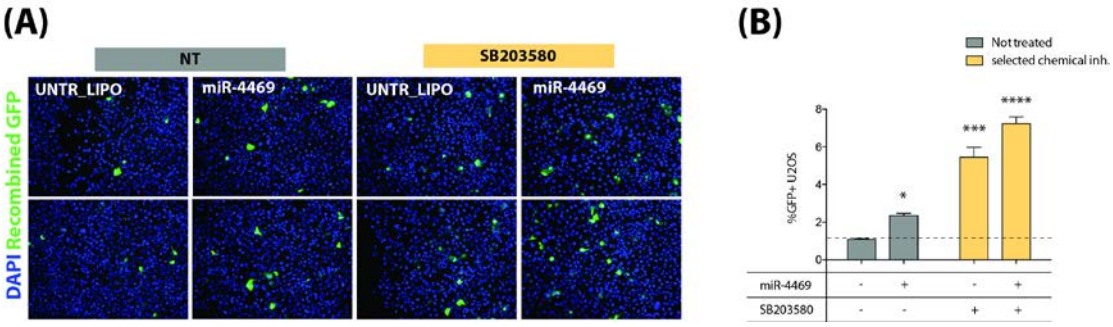


Figure 19: (A) Representative images of U2OS cells transfected with the HR detection system with or without both the miR-4469 and the p38 beta chemical inhibitor (SB203580). (B) The graph reports the percentage of recombination reached by co-treating cells with the chemical alone or in combination with miR-4469 (up to 5-fold over UNTR_LIPO). Data are mean ± SEM (N = 3 individual experiments); *p < 0.05, ***p < 0.001, ****p < 0.0001; one-way ANOVA.

To evaluate possible synergistic effects of the MAPK inhibitor and the miRNA, we reverse transfected U2OS with miRNA-4469 and, 24 hours later, we added SB203580 (10 µM) and transfected cells with our triple plasmids for the HR assay (figure 19, panel B). The results of this experiment confirmed the effect of miR-4469 and also showed that MAPK11 inhibition *per se* were very efficient at improving HR frequency (up to 5-fold over control). The combined treatment with SB203580 and miR-4469 further improved the number of HR events, up to 7-fold over control (figure 19, panel B).

These observations confirmed the important role of p38β inhibition in enhancing recombination efficiency, but also emphasized the additive role of miR-4469, possibly due to the modulation of other associated pathways.

These observations confirmed the critical role of p38β inhibition at enhancing recombination efficiency, but emphasized also the additive role of miR-4469, probably due to the modulation of other associated pathways.

3.10 A strategy to investigate homologous recombination and the effect of the selected microRNAs in mouse CMs *in vitro*

Since in all the previous experiments we used highly proliferating cells, where the HR machinery is still active, we decided to test the action of our miRNA hits in primary cells, specifically in neonatal cardiomyocytes (CMs) characterized by limited *in vitro* replication potential.

First, we decided to find a possible detection system to evaluate the number of recombination events also in these cells. With this aim, we took inspiration from a previously published paper (Ishizu, Higo et al. 2017) and designed an AAV-based HR assay that allows the in-frame insertion of a promoterless GFP cassette into a cardiac-specific expressing gene. As a target locus we used the Myosin light chain 2 (Myl2) sequence.

Using Benchling (https://www.benchling.com/crispr/), an access-free CRISPR guide design software, we designed two sgRNAs targeting the terminal part of the Myl2 coding sequence. Then, we generated a template DNA, which corresponds to a promoterless WT eGFP sequence, in frame with a self cleaving peptide sequence (T2A), flanked by two homologous arms (HA) and matching with part of the Myl2 locus (figure 20, panel A). The presence of the T2A peptide inside the template ensures the production of two separated proteins that fold independently. We evaluated the efficiency of the two sgRNAs driving Cas9 activity using the T7 endonuclease assay after transfection into an immortalized mouse muscle-like cell line (HL-1) with or without Cas9 (figure 20, panel B). Both guides led to a significant Cas9-induced Myl2 sequence cleavage. Eventually, we

selected the sgRNA_ES#1 (TGGGATGGAGAGTGGGCTGT), which scored the best in Benchling, to be used in further experiments.

Next we subcloned the guide and the template in an AAV backbone (pAAV adapted). These vectors were used to produce AAV6 particles, a serotype known to efficiently transduce CMs *in vitro* (figure 20, panel C). By transducing mouse CMs with the complete AAV6 vector combination, we predicted that, upon HR, the eGFP sequence would be inserted in frame into the end of the last exon of Myl2 locus. As a consequence, the eGFP expression cassette becomes under the transcriptional control of Myl2 promoter and is expressed only in recombined CMs. We decided to use Cas9-heart specific expressing mice to optimize the probability to detect efficient HR in these conditions (Carroll, Makarewich et al. 2016).

We started by reverse transfecting neonatal Cas9 mouse CMs with the selected miRNAs and, the day after, we transduced them with the AAV6 HR-cocktail with a total m.o.i. of $1x10^4$ vgp/CM. At 72 hours after infection, we fixed the cells and performed immunostaining against GFP and alpha-actinin, the latter antibody to selectively identify CMs. The results of these experiments confirmed that all the three selected miRNAs uperegulated the HR event, as shown by an increased number of recombined cells up to 3-fold over UNTR_LIPO ($P<0.05$; Figure 20, panel D).

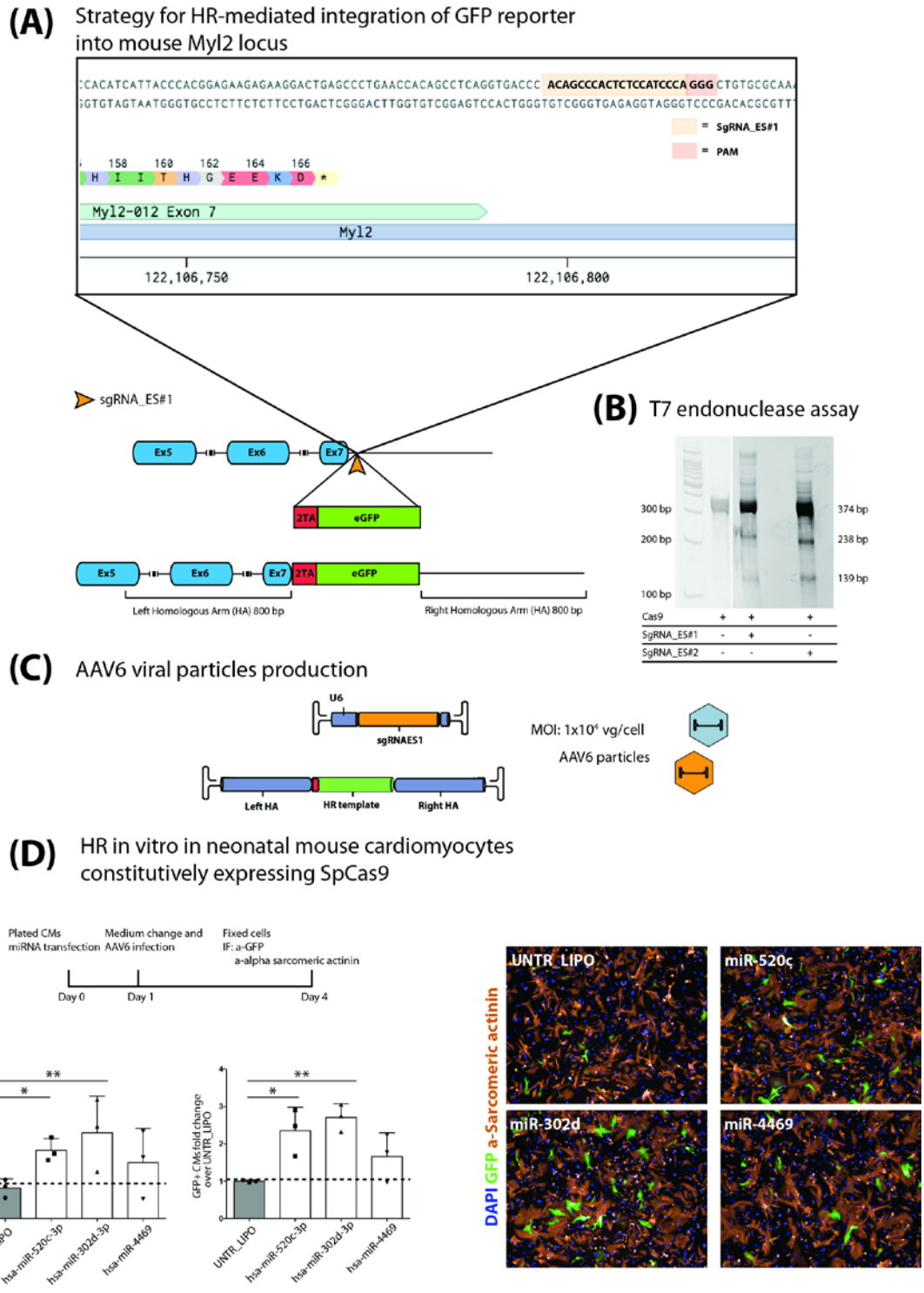


Figure 20: (A) Scheme of sgRNA and template DNA designing. The guide RNA (Sg_ES#1) targets the terminal part of the Myl2 locus, a cardiac specific gene. The template corresponds to a promoterless WT eGFP sequence in frame with the self-cleaving peptide (T2A), flanked by two homologous arms of 800 bp each. (B) T7-assay performed in HL-1 cells to evaluate the driving activity of the guideRNAs at inducing DNA cleavage in Myl2 locus. (C) Scheme of high-titer AAV6 particles production. (D) Evaluation of the effect of the selected miRNAs at improving HR-mediated eGFP insertion in mouse neonatal Cas9+ CMs. Data are mean ± SEM (N = 3 individual experiments); *p < 0.05, **p < 0.01; one-way ANOVA.

3.11 *In vivo* evaluation of the effect of the selected microRNAs on Cas9-based homologous recombination

Moving from *in vitro* to *in vivo* experiments, we used the same constructs to produce AAV serotype 9 particles (known to effectively transduce the cardiac tissue when injected, either systemically or intracardiacally, in mice) coding for the guide and the template. Moreover, we also produced AAV9 vectors expressing hsa-miR-520c or hsa-miR-302d (figure 21, panel A). With these tools, we decided to evaluate the effect of the AAV-mediated miRNAs overexpression on the Cas9-based insertion of the eGFP cassette into the Myl2 locus *in vivo*.

We injected intraperitoneally twice in a day the combination of high titer ($3x10^{11}$ vgp each virus) AAV9 vectors coding for the HR-cocktail in P1 Cas9+ mice with or without the miRNAs coding vectors (n=3). We performed the first injection in the morning, giving to Cas9+ mice a high dose ($1x10^{12}$ vgp/mouse) of AAV9 coding for hsa-miR-520c, hsa-miR-302d or an empty MCS vecto (AAV-Control). The second injection was performed in the evening of the same day. After one month, we sacrificed the animals, extracted the hearts, fixed them in PFA 4% and processed for microscopic analysis. In order to visualize the GFP fluorescent signal and to overcome possible fixation quenching artifacts, we immunostained histological criosections using a specific anti-GFP antibody in conjunction with a second antibody recognizing mouse sarcomeric alpha actinin.

According to the results obtained *in vitro*, we detected the appearance of GFP+ myocytes in all the treated animals. Notably, also in these experimental setting, the miRNAs significantly improved the number of recombination events, up to 4-fold over untreated AAV-Control vector ($P<0.01$) (figure 21, panel C). Moreover, we confirmed the miRNAs overexpression by Real Time quantification using RNA samples deriving from a portion of unfixed treated hearts. Figure 21, panel D shows some representative fluorescence microscope images of myocardium from control and microRNAs treated mice; the increased number of GFP+ fibers in the hsa-miR-520c and hsa-miR-302d treated samples in comparison to control is evident. Quantification of the homologous recombination events is reported Figure 21, panel C.

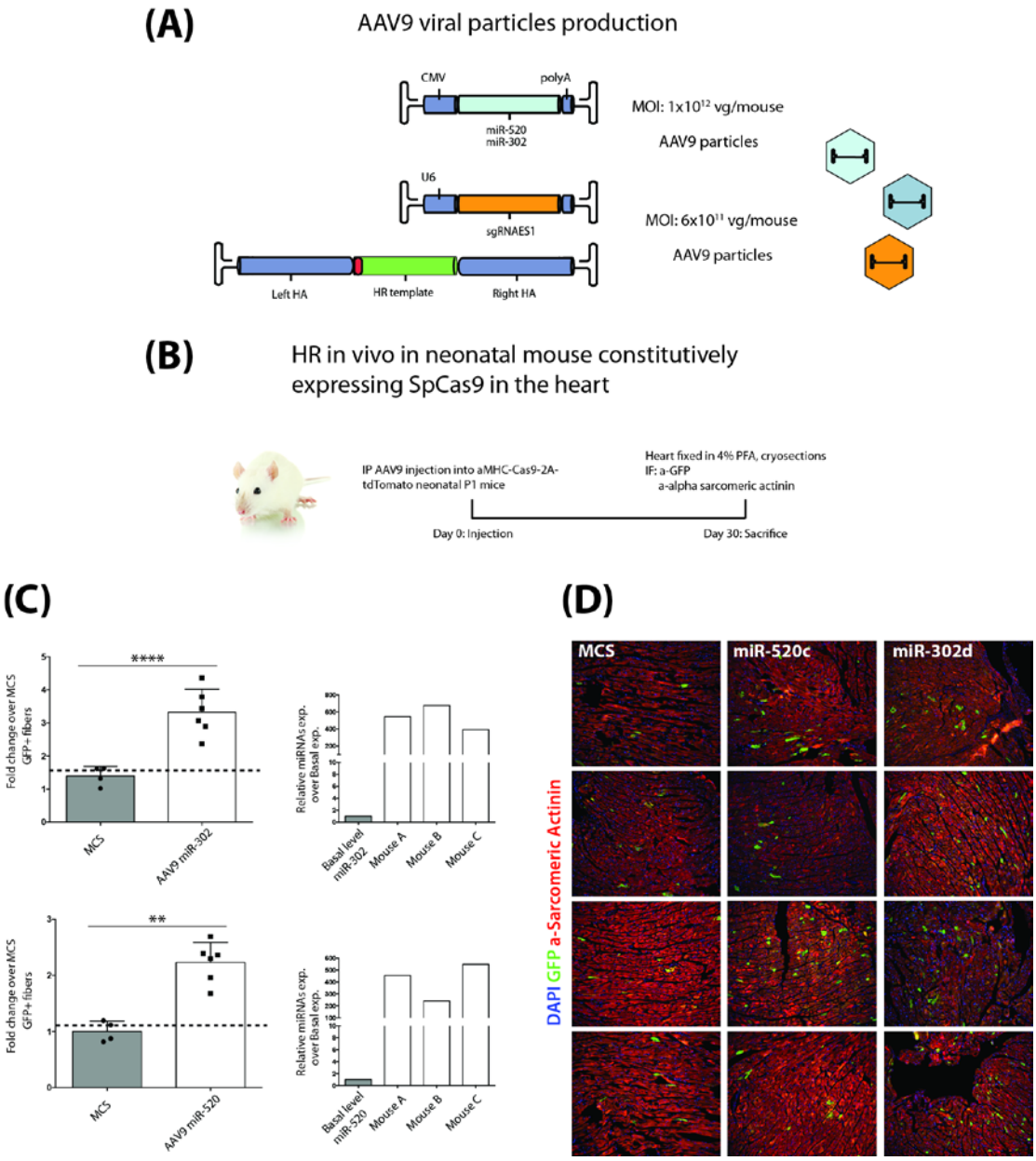


Figure 21: (A) Representative images of hi-titer AAV9 particles production. (B) Experimental workflow. (C) Results deriving from the eGFP+ fiber contents in longitudinal heart sections of three miR-520c-3p and miR-302d-3p treated Cas9+ mice compared to the basal level of HR (Menke-van der Houven van Oordt, McGeoch et al.). (D) Representative images of heart sections of control mice and miRNAs treated animals. Data are mean ± SEM (N = 3 individual mice, 2 slides each); **p < 0.01, ****p < 0.0001; t-test.

4. DISCUSSION

In the last decade, the discovery and optimization of the CRISPR/Cas9 technology, firstly described as a sort of bacterial immune system (Ishino, Shinagawa et al. 1987) and then modified to become an eukaryote genome editing tool (Cong, Ran et al. 2013), has dramatically changed the perspectives of gene therapy. Among the plethora of possible applications, the chance to target and cleave with high precision specific sequences in the human genome has encouraged researchers to look at this technique as a possible way to achieve the precise correction of hereditary, disease-causing mutations.

Thanks to the RNA-guided activity of Cas9, a double-stranded DNA break is created at the target site of the genomic DNA. Then, mammalian cells repair it, however, they preferentially employ the NHEJ DNA repair mechanism. Unfortunately, this is an error-prone mechanism active throughout the cell cycle, which might lead to either insertion or deletion of random nucleotides into the cleaved sequence. On the other hand, in mammalian cells, double-stranded DNA breaks might be fixed with the HR-dependent mechanism as well. Homologous directed repair, in the presence of a suitable identical or similar template DNA, mediates the precise restoration of the damaged sequence; hence, it has always been considered as the gold standard for gene therapy of inherited mutations (Thomas and Capecchi 1987). Nevertheless, in mammals, HR is most active only in the G2/S phase of the cell cycle and almost turned off in post mitotic cells and tissues, such as cardiomyocytes and neurons. Despite Cas9 activity has strongly improved the possibility to insert, eliminate or modify peculiar sequences in the genome, this technology is still far from being used in therapeutic applications. Consequently, there is a strong need to find new strategies that could improve HR efficiency upon Cas9 activity.

4.1 A high-content robotized screening identified novel factors to improve CRISPR/Cas9-mediated homologous recombination

This project has addressed the question how HR efficiency might be improved. We focused on the identification of innovative factors to specifically increase the Cas9-based HR activity and adapt this technology for cardiac gene therapy applications. Since HR relies on a very complex process, we decided to study the effect of endogenous, multi-target molecules, such as microRNAs using a robotized, unbiased high-throughput approach. For this purpose, we designed a triple-plasmids transfection assay that allowed us to recognize single HR events in a high-content format and to evaluate the single miRNAs effect. The HR reporter (pCUB-eGFP-Y66S) corresponds to an eGFP expression cassette bearing a single point mutation in lysine 201 (resulting in Tyr66 to Ser66 amino acid conversion), which leads to the formation of a misfolded and not fluorescent protein (Tran, Liu et al. 2003). The DNA template for correction is a promoterless and truncated WT eGFP sequence, also carrying a silent, single point mutation in the PAM recognition site, in order to avoid DNA cleavage once HR occurs (pGEM T easy_delta20GFP*). Lastly, the CRISPR/Cas9 system is provided by the third plasmid, which corresponds to the humanized sequence of the *Streptococcus pyogenes* Cas9 and the sgRNA, driving the cleavage specificity of the nuclease. We are well aware that this system is based on transient plasmid transfection and thus it monitors HR occurring on episomal DNA, therefore not necessarily reflecting the complexity of epigenetic modification of chromosomal targets. As an alternative, we considered the creation of a stable U2OS cell clone containing an integrated version of the template. However, we eventually preferred the transient approach for essentially two reasons. First, it avoids the problem of target site selection. Multiple evidences indicated that not all targets in the genome are equally accessible to Cas9 and, even less, to HR events. Operating from an episomal target avoids a bias due to the integration locus while it also maximizes the probability of recombination events, being both the target and the homologous template present in multiple copies after transfection. Second, a transient system is simpler and more versatile, especially

in light of adapting it to screenings in primary cells, such as cardiomyocytes, for which stable clones cannot be obtained while transfection using AAV6 vectors is feasible and efficient.

For the screening, we arrayed in 384 well plates a series of 2,042 human microRNA mimics (miRBASe version 21.1) and we reverse transfected them in U2OS cells. The results of the screening confirmed our assumption that microRNAs, being pleiotropic modulators of virtually all biological functions, may also act as strong positive and negative regulators of HR. We ended up with a list of 20 molecules that significantly boosted HR frequency (up to 3-fold over control). Interestingly, 10 out of the top 20 miRNA hits belong to two enriched miRNA clusters: miR-520 and miR-302. These miRNAs share the same seed sequence (AAGUGC) and probably modulate the same gene targets and pathways. Moreover, many of them are known to be commonly expressed as clusters in embryonic and pluripotent stem cells, while turned off in fully differentiated tissues. Indeed, to identify the miRNAs profile of hESC lines and comparing it with primary human mesenchimal stem cells (hMSCs), Barroso-del Jesus and collaborators performed real time PCR quantification of 170 human miRNAs. The results of these experiments confirmed a significant expression of the miR302-367 cluster in hESC, which was barely detectable in hMSCs and normal adult tissues. According to the previous observation, in the publication of Ren and co-workers the miR-520 cluster was enriched in hESC in comparison to adult cells (Barroso-delJesus, Romero-Lopez et al. 2008, Ren, Jin et al. 2009).

Importantly, the selected 20 miRNAs were effective at improving the efficiency of HR on a gene target present on an episomal plasmid, as in the case of our screening, but also on genes integrated into the host cell genome. Of note, when the relative efficiency of the selected miRNAs was compared after transfection into U2OS cells bearing the reporter eGFP_Y66S stably integrated into the AAVS-1 locus, the most effective molecule was hsa-miR-4469.

Barely anything is known about this microRNA. The expression of miR-4469 was reported for the first time in malignant human B cells (Jima, Zhang et al. 2010) and then identified in tumor breast tissue (Persson, Kvist et al. 2011). More

details were provided by a study from Cao and collaborators, in which these authors described the role of miR-4469 in increasing tumor metastasis motility by the direct inhibition of CDK3 (Cao, Xiao et al. 2017). The specific seed sequence of miR-4469 (CUCCCU) is completely different from the one characterizing the miR-520 and miR-302 clusters, implying that it probably regulates a different and distinct set of target genes.

4.2 The effect of the selected microRNAs was HR-specific, without affecting NHEJ

Considering the tendency of mammalian cells to choose NHEJ as the preferred repair mechanism to resolve double-stranded DNA breaks, we investigated whether the miRNAs identified from the screening only act specifically on HR. More specifically, we evaluated the effect of the top 20 selected molecules on NHEJ using a plasmid reporter designed for this purpose (Bennardo, Cheng et al. 2008), again using an assay based on transient transfection. The results of this test confirmed that none of the hits belonging to the miR-520 or miR-302 clusters were effective with this reporter, underlying their peculiar action on HR. Interestingly, in this experimental setting, only hsa-miR-143 was capable to also improve NHEJ. This microRNA was recently described as a potential tumor suppressor in several human cancers (Borralho, Kren et al. 2009, Noguchi, Yasui et al. 2013, Yan, Chen et al. 2014). Indeed, its expression was reported to increase p53 levels in the presence of DNA damage and to mediate MDM2 downregulation, therefore playing a role in the DNA damage-induced cell death program (Zhang, Sun et al. 2013). Whether, and to what extent, these findings might be related to the effect of miR-143 on HR and NHEJ that we discovered in our study remains a matter for future investigation

4.3 Effect of the selected microRNAs on cell proliferation

To search for possible molecular mechanisms at the basis of the observed miRNAs effect, we first focused on the observation that HR occurs only in dividing cells and therefore it should be precisely regulated throughout the

different phases of cell cycle (Cho, Campbell et al. 1998). Hence, we speculated on the possibility that our selected miRNAs might induce cell cycle alterations that could push cells in S or G2/M, when HR is maximally active (Aylon, Liefshitz et al. 2004). By an accurate analysis of the cell cycle profile of treated and untreated control U2OS cells, we indeed demonstrated that all the three miRNAs were able to increase the percentage of cells in S and G2/M phases (from 20% to almost 30%). This effect was more evident for hsa-miR-520c-3p and hsa-miR-302d-3p. It is of interest to remark that these two microRNA are known to be highly expressed in ES cells, as also discussed above. In these cells, HR is known to occur at very high frequency compared to other cell types (Tichy, Pillai et al. 2010). In addition, it is well established that these cells spend a significant proportion of their cell cycle time in G2/M (White and Dalton 2005). It is tempting to speculate that one or both of these occurrences might be due to the expression of these microRNAs. As a matter of fact, downregulation of members of the miR-302 family in ES cells stimulates differentiation of these cells (Parchem, Moore et al. 2015) and thus, presumably, also the reduction of the proportion of cells in G2/M. It will be interesting to analyze, therefore, whether blockage of miR-302 or miR-520 family members in hES cells also results in diminished rates of HR.

Other reports in the literature link these two miRNAs with the control of cell proliferation also in other cell types. Li and collaborators have shown that hsa-miR-520c was able to boost the proliferation and progression of gastric cancer by directly targeting IRF2 (Li, Wen et al. 2016). Xu and co-workers revealed that miR-302d activity promotes proliferation of pluripotent stem cell-derived cardiomyocytes by suppressing LATS2 (Xu, Yang et al. 2019). Work from our own laboratory on microRNAs stimulating the capacity of neonatal cardiomyocyte to proliferate revealed that the miR-302 and miR-520 families contain microRNAs that are powerful inducer of proliferation of these cells (Eulalio, Mano et al. 2012). In this respect, however, it is important to observe that there is no apparent correlation between the capacity of microRNA to stimulate cardiomyocyte proliferation and that to promote HR, as concluded by

the comparison of the screening reported in this Thesis and the results by Eulalio and collaborators. Hsa-miR-4469 itself (which was not present in the original screening for cardiomyocyte proliferation) is very effective at stimulating HR but ineffective in activating cardiomyocyte proliferation (data not shown).

4.4 Effect of the selected microRNAs and the DNA damage response

An issue that stemmed from the observation that the selected microRNAs perturb the normal cell cycle profile was to understand whether they might directly induce DNA damage and the activation of cell cycle checkpoints.

To answer this question, we evaluated the activation and accumulation of CHK2 and H2AX proteins. The first is a factor belonging to the DNA damage-spreading cascade, which is rapidly activated and phosphorylated by ATM in response to ionizing radiations mediating G2 phase arrest (Matsuoka, Rotman et al. 2000). The histone variant H2AX regulates one of the earliest steps in DSB recognition. Upon occurrence of a DSB, ATM phosphorylates the carboxyl-terminal of H2A present on the chromatin surrounding the damage site, (γ-H2AX); this corresponds to a mark for triggering the recruitment of many chromatin modifiers and DSB repair factors, allowing efficient DNA repair. (Downs, Allard et al. 2004, Morrison, Highland et al. 2004).

Interestingly, in contrast to UV-treated control cells that served as a positive control, neither CHK2 nor H2AX was activated following the sole transfection of the selected microRNAs into the cells.

On the other hand, we observed a significant accumulation of RAD51 and of the three MRN complex components (MRE11, NBN and RAD50) at the mRNA and protein levels in both hsa-miR-520c-3p and hsa-miR-302d-3p-treated samples. The MRN complex is known to mediate the early recognition of the broken ends at the DSB site (Zha, Boboila et al. 2009, Taylor, Cecillon et al. 2010). RAD51 instead is a HR specific ATP-dependent recombinase, which forms a nucleoprotein filament on ssDNA in a gap or at the 3' end of a DSB to prevent its degradation. It also creates a catalytic center that allows the invasion and annealing of the homologous template DNA (Son and Hasty 2018). Therefore,

these data clearly point to a direct modulation of factors belonging to the HR machinery by the overexpression of hsa-miR-302 and hsa-miR-520 microRNA clusters. However, this property is not shared by hsa-miR-4469.

4.5 Mechanism of action of hsa-miR-4469

Since miRNAs are multi-target, endogenous molecules that regulate hundreds of genes, we decided to analyze the transcriptome of miRNA-treated U2OS cells, to identify both direct and indirect targets. There was a strict correlation between the modifications induced by hsa-miR-502c-3p and hsa-miR-302d-3p, both in the upregulated and downregulated genes; on the contrary, hsa-miR-4469-treated samples exhibited a different expression pattern. Among the upregulated genes, we confirmed the presence of the MRN complex components and RAD51 protein; in addition, we also pointed out new HR-associated genes, such as MSH2 (a component of the BRCA-1-associated genome surveillance complex) (Wang, Cortez et al. 2000) and MDC1 (a nuclear protein known to interact with RAD51 during HR) (Zhang, Ma et al. 2005). Moreover, we confirmed the downregulation of G1-related cyclins and kinases. Interestingly, among the hsa-miR-4469 possible targets, there was MAPK11 (p38β), which, together with the other MAPKs (ERK and JNK), is known to play a role in the HR-mediated DNA repair process (Golding, Rosenberg et al. 2007). Golding and coworkers described p38β inhibition as a positive stimulus for HR. This observation encouraged us to further study the correlation between MAPK11 inhibition and the effect of miR-4469 on HR enhancement.

We indeed discovered that MAPK11 is a direct target of hsa-miR-4469. By Real Time PCR and western blot quantification, we observed a significant MAPK11 inhibition at both the mRNA and protein levels in hsa-miR-4469-transfected cells. Moreover, TargetScan analysis revealed that a perfect match exists between the hsa-miR-4469 seed sequence and the 3'-UTR of the human MAPK11 mRNA. When we transfected HeLa cells with hsa-miR-4469 and a reporter plasmid carrying the luciferase gene with appended the MAPK11 3'UTR

we observed a significant downregulation of luciferase activity, indicative of direct targeting of this sequence by the microRNA.

As a further prove of the role of MAPK11 inhibition on HR modulation, the highly specific p38 chemical inhibitor SB203580 (Jarnicki, Conroy et al. 2008) was able to induce alone a 5-fold enrichment of GFP-positive cells in U2OS cells transfected with the triple plasmid HR reporter assay. The effect was additive in the presence of hsa-miR-4469 (up to 7-fold of HR over control). Interestingly, SB203580 is a well studied and described molecule already used as a potent inhibitor of the inflammatory response and cytokine production *in vivo* (Badger, Bradbeer et al. 1996, Herman, Krawczynska et al. 2014). A phase II study is now evaluating the effect of another MAPK11 chemical inhibitor, ARRY 797, to prevent the ventricular dilatation in patients affected by a Laminin A/C-related cardiomyopathy (https://clinicaltrials.gov/ct2/show/NCT03439514).

Future experiments may focus on testing the effect of these chemicals on HR *in vivo*, by administering these drugs together with the heart-specific HR reporter system described in paragraph 3.10 and figure 20.

4.6 Gene editing in cardiomyocytes

Developing a precise, HR-mediated gene therapy method is a promising approach for the treatment of many monogenic and inherited disease-causing mutations, especially in terminally differentiated tissues, including the heart. Multiple indeed evidence indicates that precise gene targeting and repair through NHEJ is possible in cardiomyocytes ex vivo and in vivo in muscles, retina or heart using the CRISPR/Cas9 components delivered using AAV9 vectors (Bengtsson, Hall et al. 2017, Johansen, Molenaar et al. 2017, Yu, Mookherjee et al. 2017). However, it is known that HR-based DSB repair is a molecular process largely restricted to proliferating cells, which occurs mainly in the S and G2 phases of the cell cycle. This process is believed to be almost turned off in post mitotic cells and tissues, such as adult cardiomyocytes.

Given the properties of our microRNAs, we decided to investigate the possibility of obtaining HR-mediated gene targeting in cardiomyocytes both *in vitro* and *in*

vivo upon the delivery of hsa-miR-520c, hsa-miR-302d and hsa-miR-4469. We used the same system described in the manuscript by Ishizu and collaborators, entailing the precise genomic insertion of a promoterless GFP cassette with a self-cleaving T2A peptide at its 5' end, in frame with the cardio-specific myosin-light chain 2 (Myl2) gene (Ishizu, Higo et al. 2017).

Primary neonatal CMs deriving from mice expressing Cas9 exclusively in the heart (Carroll, Makarewich et al. 2016) were infected with a combination of two AAV6 vectors. One delivered the single guide RNA targeting an intron sequence at the end of the last exon of the Myl2 locus, while the other one introduced the template DNA for HR (the T2A-GFP cassette bracketed by two 800 bp homologous arms). In accordance with the Ishino et al. findings, also in our experimental settings we recognized the appearance of recombined cells, reaching, at the $1x10^4$ m.o.i., 1% of HDR-mediated gene targeting in neonatal cardiomyocytes. Notably, by transfecting CMs with the selected synthetic miRNAs, we reached a 3-fold increase in HR positive events in hsa-miR-520c-3p- and hsa-miR-302d-3p-treated samples. Hsa-miR-4469 treatment was less effective at improving HR, inducing up to 2-fold more HR-mediated GFP insertions over control. This finding is important because it suggests that differentiated, although neonatal, cardiomyocytes are able to respond to the stimulus imposed by miRNAs treatment and can repair the Cas9-generated DSBs by HR. Moreover, our results highlight a conserved spectrum of regulation of the three human miRNAs in mouse cardiomyocytes. It is known that hsa-miR-302d-3p is important for cardiomyocyte proliferation during development, when it is expressed at high levels, and is sufficient to induce cardiomyocyte proliferation and promote cardiac regeneration if re-expressed in the adult (Tian, Liu et al. 2015); moreover, its reactivation in the adult induced cardiomyocyte dedifferentiation. This may suggest that this also promotes the re-expression of the HR machinery protein complex.

4.7 Gene editing in the heart

Considering our *in vitro* results showing the feasibility to stimulate homologous directed repair in differentiated, largely post-mitotic neonatal mice cardiomyocytes, we decided to investigate whether this process also might occur in neonatal mice *in vivo*.

For these experiments, we produced AAV9 particles (known to efficiently and uniformly transduce the heart tissue *in vivo*) coding for the sgRNA, for hsa-miR-302d and hsa-miR-520c, and containing the sequence of the template for HR. Unfortunately, although we obtained also a plasmid construct designed to express miR-4469, we failed to detect any RNA expression after transfecting it into U2OS cells. Thus, we did not proceed with hsa-miR-4469 AAV9 particle production and optimization experiments are still ongoing.

We pre-treated Cas9+ mice at 1-day post birth by intraperitoneally injecting the AAV9 coding for the different miRNAs ($1x10^{12}$ vgp/mouse) and using as control an empty AAV9-MCS vector. After 8 hours, we injected a mixture of equal amounts of vector genomes for each of the different AAV9 HR system components ($6x10^{11}$ vgp/mouse).

In these conditions, fluorescence microscope analysis of histological preparations from heart samples one month after treatment allowed the detection of a significant number of HR events (200 GFP+ fibers/section, 2 sections for each heart were analyzed), in the form of GFP+ myocardial fibers. Furthermore, in these experiments we were also able to detect a significant increase in the number of HR events in hsa-miR-520c-3p- and hsa-miR-302d-3p-treated animals (3-4 fold change over control mice infected with AAV9-MCS). From the same samples we also evaluated the levels of miRNAs overexpression, which resulted statistically significant in both the hsa-miR-520c-3p- and hsa-miR-302d-3p-treated animals.

4.8 Conclusions and future perspectives

To the best of our knowledge, the experimental work described in this Thesis is the first one that systematically analyzed the effect of all human miRNAs on HR. The identified, pro-recombinogenic miRNAs (hsa-miR-520c-3p, hsa-miR-302d-3p and hsa-miR-4469) have never been previously associated with HR. Additional work is clearly needed to bring these results further.

First, our understanding of the mechanism of action of the selected miRNAs is still scanty. The finding that all three miRNAs increase the time cells spend in G2/M is consistent with their pro-recombinogenic effect but the mechanism by which this occurs is still obscure. Consistent with this conclusion, it also remains unclear whether the finding that hsa-miR-520c-3p and hsa-miR-302d-3p increase the levels of pro-recombinogenic proteins is secondary to the G2/M increase or is a separate, direct effect of the microRNAs. In the latter case, the mechanism by which this is exerted needs to be understood.

Second, the results obtained by delivering the miRNAs in neonatal hearts are very encouraging towards the development of a precise gene correction approach in vivo. However, we still miss data on the efficacy of the system in adult hearts. Neonatal cardiomyocytes are largely different from adult cardiomyocytes in terms of proliferative capacity, metabolism, cytoskeletal and centrosome organization, connections at the intercalated disk, and, finally, sarcomeric structure. To what extent these differences, in particular, their withdrawal from the cell cycle, can be overcome by any of the three selected microRNAs needs to be experimentally tested and understood.

Third, the HDR process in mammalian cells unfavorably competes with the most frequent and error-prone NHEJ repair mechanism. We still do not understand the determinants that regulate the choice of a repair pathway. Whether the

inhibition of NHEJ, for example by transient RNA interference against the Ku proteins or DNA-PKcs, might boost the efficacy of the selected miRNAs is an interesting experiment that we plan to perform in the near future.

Fourth, our experiments clearly indicate that the presence of the homologous template for recombination is not a limiting factor for the HR process to occur. This is most likely due to the very efficient AAV vector system, which could have stimulated HR repair. Moreover, AAV genomes by themselves are known to preferentially interact with the DNA damage response and HR protein machinery (Cervelli, Palacios et al. 2008). To what extent the selected miRNAs also directly affect AAV biology and might boost AAV transduction efficiency is another very interesting spin-off project consequent to our results.

Fifth and final, it will also be important to understand what is the role of the chromatin status and architecture in proximity of the Cas9-induced DSB. From the paper of Mano and collaborators in our laboratory we already know that chromatin remodeling factors strongly influence AAV transduction (Mano, Ippodrino et al. 2015). Furthermore, other investigator have reported that a modification of histone dosage can modulate HR frequency (Liang, Burkhart et al. 2012). Starting by taking advantage of the transcriptome data we have generated, we wish to explore the possible alterations in chromatin modifiers and remodelers that could also partially explain effect of the miRNAs we have selected.

5. Bibliography

(1990). "[Analysis of a program for atypical familial microcytosis. Molecular basis for alpha-thalassemia. GEHBTA]." Sangre (Barc) **35**(2): 102-113.

Aartsma-Rus, A., I. Fokkema, J. Verschuuren, I. Ginjaar, J. van Deutekom, G. J. van Ommen and J. T. den Dunnen (2009). "Theoretic applicability of antisense-mediated exon skipping for Duchenne muscular dystrophy mutations." Hum Mutat **30**(3): 293-299.

Aartsma-Rus, A., V. Straub, R. Hemmings, M. Haas, G. Schlosser-Weber, V. Stoyanova-Beninska, E. Mercuri, F. Muntoni, B. Sepodes, E. Vroom and P. Balabanov (2017). "Development of Exon Skipping Therapies for Duchenne Muscular Dystrophy: A Critical Review and a Perspective on the Outstanding Issues." Nucleic Acid Ther **27**(5): 251-259.

Ashktorab, H. and A. Srivastava (1989). "Identification of nuclear proteins that specifically interact with adeno-associated virus type 2 inverted terminal repeat hairpin DNA." J Virol **63**(7): 3034-3039.

Atchison, R. W., B. C. Casto and W. M. Hammon (1966). "Electron microscopy of adenovirus-associated virus (AAV) in cell cultures." Virology **29**(2): 353-357.

Aylon, Y., B. Liefshitz and M. Kupiec (2004). "The CDK regulates repair of double-strand breaks by homologous recombination during the cell cycle." EMBO J **23**(24): 4868-4875.

Badger, A. M., J. N. Bradbeer, B. Votta, J. C. Lee, J. L. Adams and D. E. Griswold (1996). "Pharmacological profile of SB 203580, a selective inhibitor of cytokine suppressive binding protein/p38 kinase, in animal models of arthritis, bone resorption, endotoxin shock and immune function." J Pharmacol Exp Ther **279**(3): 1453-1461.

Baigude, H., Ahsanullah, Z. Li, Y. Zhou and T. M. Rana (2012). "miR-TRAP: a benchtop chemical biology strategy to identify microRNA targets." Angew Chem Int Ed Engl **51**(24): 5880-5883.

Bakondi, B., W. Lv, B. Lu, M. K. Jones, Y. Tsai, K. J. Kim, R. Levy, A. A. Akhtar, J. J. Breunig, C. N. Svendsen and S. Wang (2016). "In Vivo CRISPR/Cas9 Gene Editing Corrects Retinal Dystrophy in the S334ter-3 Rat Model of Autosomal Dominant Retinitis Pigmentosa." Mol Ther **24**(3): 556-563.

Barrangou, R., C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D. A. Romero and P. Horvath (2007). "CRISPR provides acquired resistance against viruses in prokaryotes." Science **315**(5819): 1709-1712.

Barroso-delJesus, A., C. Romero-Lopez, G. Lucena-Aguilar, G. J. Melen, L. Sanchez, G. Ligero, A. Berzal-Herranz and P. Menendez (2008). "Embryonic stem cell-

specific miR302-367 cluster: human gene structure and functional characterization of its core promoter." Mol Cell Biol **28**(21): 6609-6619.

Bartlett, J. S., R. Wilcher and R. J. Samulski (2000). "Infectious entry pathway of adeno-associated virus and adeno-associated virus vectors." J Virol **74**(6): 2777-2785.

Beerli, R. R. and C. F. Barbas, 3rd (2002). "Engineering polydactyl zinc-finger transcription factors." Nat Biotechnol **20**(2): 135-141.

Bengtsson, N. E., J. K. Hall, G. L. Odom, M. P. Phelps, C. R. Andrus, R. D. Hawkins, S. D. Hauschka, J. R. Chamberlain and J. S. Chamberlain (2017). "Corrigendum: Muscle-specific CRISPR/Cas9 dystrophin gene editing ameliorates pathophysiology in a mouse model for Duchenne muscular dystrophy." Nat Commun **8**: 16007.

Bengtsson, N. E., J. K. Hall, G. L. Odom, M. P. Phelps, C. R. Andrus, R. D. Hawkins, S. D. Hauschka, J. R. Chamberlain and J. S. Chamberlain (2017). "Muscle-specific CRISPR/Cas9 dystrophin gene editing ameliorates pathophysiology in a mouse model for Duchenne muscular dystrophy." Nat Commun **8**: 14454.

Bennardo, N., A. Cheng, N. Huang and J. M. Stark (2008). "Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair." PLoS Genet **4**(6): e1000110.

Berns, K. I. and R. M. Linden (1995). "The cryptic life style of adeno-associated virus." Bioessays **17**(3): 237-245.

Bhakta, M. S., I. M. Henry, D. G. Ousterout, K. T. Das, S. H. Lockwood, J. F. Meckler, M. C. Wallen, A. Zykovich, Y. Yu, H. Leo, L. Xu, C. A. Gersbach and D. J. Segal (2013). "Highly active zinc-finger nucleases by extended modular assembly." Genome Res **23**(3): 530-538.

Bianco, P. R., R. B. Tracy and S. C. Kowalczykowski (1998). "DNA strand exchange proteins: a biochemical and physical comparison." Front Biosci **3**: D570-603.

Boch, J., H. Scholze, S. Schornack, A. Landgraf, S. Hahn, S. Kay, T. Lahaye, A. Nickstadt and U. Bonas (2009). "Breaking the code of DNA binding specificity of TAL-type III effectors." Science **326**(5959): 1509-1512.

Bolotin, A., B. Quinquis, A. Sorokin and S. D. Ehrlich (2005). "Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin." Microbiology **151**(Pt 8): 2551-2561.

Borralho, P. M., B. T. Kren, R. E. Castro, I. B. da Silva, C. J. Steer and C. M. Rodrigues (2009). "MicroRNA-143 reduces viability and increases sensitivity to 5-fluorouracil in HCT116 human colorectal cancer cells." FEBS J **276**(22): 6689-6700.

Bozkurt, B., D. Aguilar, A. Deswal, S. B. Dunbar, G. S. Francis, T. Horwich, M. Jessup, M. Kosiborod, A. M. Pritchett, K. Ramasubbu, C. Rosendorff, C. Yancy, F. American Heart Association Heart, C. Transplantation Committee of the Council on Clinical, S. Council on Cardiovascular, Anesthesia, C. Council on, N. Stroke, H. Council on, Q. Council on and R. Outcomes (2016). "Contributory Risk and Management of Comorbidities of Hypertension, Obesity, Diabetes Mellitus, Hyperlipidemia, and Metabolic Syndrome in Chronic Heart Failure: A Scientific Statement From the American Heart Association." Circulation **134**(23): e535-e578.

Bozkurt, B., M. Colvin, J. Cook, L. T. Cooper, A. Deswal, G. C. Fonarow, G. S. Francis, D. Lenihan, E. F. Lewis, D. M. McNamara, E. Pahl, R. S. Vasan, K. Ramasubbu, K. Rasmusson, J. A. Towbin, C. Yancy, F. American Heart Association Committee on Heart, C. Transplantation of the Council on Clinical, Y. Council on Cardiovascular Disease in the, C. Council on, N. Stroke, E. Council on, Prevention, C. Council on Quality of and R. Outcomes (2016). "Current Diagnostic and Treatment Strategies for Specific Dilated Cardiomyopathies: A Scientific Statement From the American Heart Association." Circulation **134**(23): e579-e646.

Brauch, K. M., M. L. Karst, K. J. Herron, M. de Andrade, P. A. Pellikka, R. J. Rodeheffer, V. V. Michels and T. M. Olson (2009). "Mutations in ribonucleic acid binding protein gene cause familial dilated cardiomyopathy." J Am Coll Cardiol **54**(10): 930-941.

Braun, J. E., E. Huntzinger, M. Fauser and E. Izaurralde (2011). "GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets." Mol Cell **44**(1): 120-133.

Brouns, S. J., M. M. Jore, M. Lundgren, E. R. Westra, R. J. Slijkhuis, A. P. Snijders, M. J. Dickman, K. S. Makarova, E. V. Koonin and J. van der Oost (2008). "Small CRISPR RNAs guide antiviral defense in prokaryotes." Science **321**(5891): 960-964.

Buch, P. K., J. W. Bainbridge and R. R. Ali (2008). "AAV-mediated gene therapy for retinal disorders: from mouse to man." Gene Ther **15**(11): 849-857.

Burke, M. A., S. A. Cook, J. G. Seidman and C. E. Seidman (2016). "Clinical and Mechanistic Insights Into the Genetics of Cardiomyopathy." J Am Coll Cardiol **68**(25): 2871-2886.

Cai, X., C. H. Hagedorn and B. R. Cullen (2004). "Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs." RNA **10**(12): 1957-1966.

Cao, T., T. Xiao, G. Huang, Y. Xu, J. J. Zhu, K. Wang, W. Ye, H. Guan, J. He and D. Zheng (2017). "CDK3, target of miR-4469, suppresses breast cancer metastasis via inhibiting Wnt/beta-catenin pathway." Oncotarget **8**(49): 84917-84927.

Carroll, K. J., C. A. Makarewich, J. McAnally, D. M. Anderson, L. Zentilin, N. Liu, M. Giacca, R. Bassel-Duby and E. N. Olson (2016). "A mouse model for adult cardiac-

specific gene deletion with CRISPR/Cas9." Proc Natl Acad Sci U S A **113**(2): 338-343.

Casini, A., M. Olivieri, G. Petris, C. Montagna, G. Reginato, G. Maule, F. Lorenzin, D. Prandi, A. Romanel, F. Demichelis, A. Inga and A. Cereseto (2018). "A highly specific SpCas9 variant is identified by in vivo screening in yeast." Nat Biotechnol **36**(3): 265-271.

Cervelli, T., J. A. Palacios, L. Zentilin, M. Mano, R. A. Schwartz, M. D. Weitzman and M. Giacca (2008). "Processing of recombinant AAV genomes occurs in specific nuclear structures that overlap with foci of DNA-damage-response proteins." J Cell Sci **121**(Pt 3): 349-357.

Chen, R. and M. S. Wold (2014). "Replication protein A: single-stranded DNA's first responder: dynamic DNA-interactions allow replication protein A to direct single-strand DNA intermediates into different pathways for synthesis or repair." Bioessays **36**(12): 1156-1161.

Cheung, A. K., M. D. Hoggan, W. W. Hauswirth and K. I. Berns (1980). "Integration of the adeno-associated virus genome into cellular DNA in latently infected human Detroit 6 cells." J Virol **33**(2): 739-748.

Cho, R. J., M. J. Campbell, E. A. Winzeler, L. Steinmetz, A. Conway, L. Wodicka, T. G. Wolfsberg, A. E. Gabrielian, D. Landsman, D. J. Lockhart and R. W. Davis (1998). "A genome-wide transcriptional analysis of the mitotic cell cycle." Mol Cell **2**(1): 65-73.

Christine, C. W., P. A. Starr, P. S. Larson, J. L. Eberling, W. J. Jagust, R. A. Hawkins, H. F. VanBrocklin, J. F. Wright, K. S. Bankiewicz and M. J. Aminoff (2009). "Safety and tolerability of putaminal AADC gene therapy for Parkinson disease." Neurology **73**(20): 1662-1669.

Chu, V. T., T. Weber, R. Graf, T. Sommermann, K. Petsch, U. Sack, P. Volchkov, K. Rajewsky and R. Kuhn (2016). "Efficient generation of Rosa26 knock-in mice using CRISPR/Cas9 in C57BL/6 zygotes." BMC Biotechnol **16**: 4.

Chylinski, K., A. Le Rhun and E. Charpentier (2013). "The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems." RNA Biol **10**(5): 726-737.

Cideciyan, A. V., W. W. Hauswirth, T. S. Aleman, S. Kaushal, S. B. Schwartz, S. L. Boye, E. A. Windsor, T. J. Conlon, A. Sumaroka, J. J. Pang, A. J. Roman, B. J. Byrne and S. G. Jacobson (2009). "Human RPE65 gene therapy for Leber congenital amaurosis: persistence of early visual improvements and safety at 1 year." Hum Gene Ther **20**(9): 999-1004.

Cong, L., F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, L. A. Marraffini and F. Zhang (2013). "Multiplex genome engineering using CRISPR/Cas systems." Science **339**(6121): 819-823.

Davies, O. R., J. V. Forment, M. Sun, R. Belotserkovskaya, J. Coates, Y. Galanty, M. Demir, C. R. Morton, N. J. Rzechorzek, S. P. Jackson and L. Pellegrini (2015). "CtIP tetramer assembly is required for DNA-end resection and repair." Nat Struct Mol Biol **22**(2): 150-157.

Davis, S., B. Lollo, S. Freier and C. Esau (2006). "Improved targeting of miRNA with antisense oligonucleotides." Nucleic Acids Res **34**(8): 2294-2304.

Davis, S., S. Propp, S. M. Freier, L. E. Jones, M. J. Serra, G. Kinberger, B. Bhat, E. E. Swayze, C. F. Bennett and C. Esau (2009). "Potent inhibition of microRNA in vivo without degradation." Nucleic Acids Res **37**(1): 70-77.

Deltcheva, E., K. Chylinski, C. M. Sharma, K. Gonzales, Y. Chao, Z. A. Pirzada, M. R. Eckert, J. Vogel and E. Charpentier (2011). "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III." Nature **471**(7340): 602-607.

Ding, Q., A. Strong, K. M. Patel, S. L. Ng, B. S. Gosis, S. N. Regan, C. A. Cowan, D. J. Rader and K. Musunuru (2014). "Permanent alteration of PCSK9 with in vivo CRISPR-Cas9 genome editing." Circ Res **115**(5): 488-492.

Doench, J. G., E. Hartenian, D. B. Graham, Z. Tothova, M. Hegde, I. Smith, M. Sullender, B. L. Ebert, R. J. Xavier and D. E. Root (2014). "Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation." Nat Biotechnol **32**(12): 1262-1267.

Doench, J. G. and P. A. Sharp (2004). "Specificity of microRNA target selection in translational repression." Genes Dev **18**(5): 504-511.

Dong, B., H. Nakai and W. Xiao (2010). "Characterization of genome integrity for oversized recombinant AAV vector." Mol Ther **18**(1): 87-92.

Downs, J. A., S. Allard, O. Jobin-Robitaille, A. Javaheri, A. Auger, N. Bouchard, S. J. Kron, S. P. Jackson and J. Cote (2004). "Binding of chromatin-modifying activities to phosphorylated histone H2A at DNA damage sites." Mol Cell **16**(6): 979-990.

Easow, G., A. A. Teleman and S. M. Cohen (2007). "Isolation of microRNA targets by miRNP immunopurification." RNA **13**(8): 1198-1204.

Eichhorn, S. W., H. Guo, S. E. McGeary, R. A. Rodriguez-Mias, C. Shin, D. Baek, S. H. Hsu, K. Ghoshal, J. Villen and D. P. Bartel (2014). "mRNA destabilization is the dominant effect of mammalian microRNAs by the time substantial repression ensues." Mol Cell **56**(1): 104-115.

Eulalio, A., M. Mano, M. Dal Ferro, L. Zentilin, G. Sinagra, S. Zacchigna and M. Giacca (2012). "Functional screening identifies miRNAs inducing cardiac regeneration." Nature **492**(7429): 376-381.

Ferguson, D. O. and W. K. Holloman (1996). "Recombinational repair of gaps in DNA is asymmetric in Ustilago maydis and can be explained by a migrating D-loop model." Proc Natl Acad Sci U S A **93**(11): 5419-5424.

Ferrari, F. K., T. Samulski, T. Shenk and R. J. Samulski (1996). "Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors." J Virol **70**(5): 3227-3234.

Fillingham, J., M. C. Keogh and N. J. Krogan (2006). "GammaH2AX and its role in DNA double-strand break repair." Biochem Cell Biol **84**(4): 568-577.

Fineran, P. C. and E. Charpentier (2012). "Memory of viral infections by CRISPR-Cas adaptive immune systems: acquisition of new information." Virology **434**(2): 202-209.

Fonfara, I., A. Le Rhun, K. Chylinski, K. S. Makarova, A. L. Lecrivain, J. Bzdrenga, E. V. Koonin and E. Charpentier (2014). "Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems." Nucleic Acids Res **42**(4): 2577-2590.

Fontaine, G., R. Frank, J. L. Tonet, G. Guiraudon, C. Cabrol, G. Chomette and Y. Grosgogeat (1984). "Arrhythmogenic right ventricular dysplasia: a clinical model for the study of chronic ventricular tachycardia." Jpn Circ J **48**(6): 515-538.

Friedland, A. E., R. Baral, P. Singhal, K. Loveluck, S. Shen, M. Sanchez, E. Marco, G. M. Gotta, M. L. Maeder, E. M. Kennedy, A. V. Kornepati, A. Sousa, M. A. Collins, H. Jayaram, B. R. Cullen and D. Bumcrot (2015). "Characterization of Staphylococcus aureus Cas9: a smaller Cas9 for all-in-one adeno-associated virus delivery and paired nickase applications." Genome Biol **16**: 257.

Friedman, R. C., K. K. Farh, C. B. Burge and D. P. Bartel (2009). "Most mammalian mRNAs are conserved targets of microRNAs." Genome Res **19**(1): 92-105.

Fu, Y., J. D. Sander, D. Reyon, V. M. Cascio and J. K. Joung (2014). "Improving CRISPR-Cas nuclease specificity using truncated guide RNAs." Nat Biotechnol **32**(3): 279-284.

Galev, E. E., B. N. Afanas'ev, L. P. Buchatskii, V. Kozlov Iu and A. A. Baev (1989). "[Features of the organization of the densovirus genome]." Dokl Akad Nauk SSSR **307**(4): 996-1000.

Gao, G., L. H. Vandenberghe, M. R. Alvira, Y. Lu, R. Calcedo, X. Zhou and J. M. Wilson (2004). "Clades of Adeno-associated viruses are widely disseminated in human tissues." J Virol **78**(12): 6381-6388.

Gao, G. P., M. R. Alvira, L. Wang, R. Calcedo, J. Johnston and J. M. Wilson (2002). "Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy." Proc Natl Acad Sci U S A **99**(18): 11854-11859.

Gao, G. P., G. Qu, L. Z. Faust, R. K. Engdahl, W. Xiao, J. V. Hughes, P. W. Zoltick and J. M. Wilson (1998). "High-titer adeno-associated viral vectors from a Rep/Cap cell line and hybrid shuttle virus." Hum Gene Ther **9**(16): 2353-2362.

Garneau, J. E., M. E. Dupuis, M. Villion, D. A. Romero, R. Barrangou, P. Boyaval, C. Fremaux, P. Horvath, A. H. Magadan and S. Moineau (2010). "The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA." Nature **468**(7320): 67-71.

Gedicke-Hornung, C., V. Behrens-Gawlik, S. Reischmann, B. Geertz, D. Stimpel, F. Weinberger, S. Schlossarek, G. Precigout, I. Braren, T. Eschenhagen, G. Mearini, S. Lorain, T. Voit, P. A. Dreyfus, L. Garcia and L. Carrier (2013). "Rescue of cardiomyopathy through U7snRNA-mediated exon skipping in Mybpc3-targeted knock-in mice." EMBO Mol Med **5**(7): 1128-1145.

Geisterfer-Lowrance, A. A., S. Kass, G. Tanigawa, H. P. Vosberg, W. McKenna, C. E. Seidman and J. G. Seidman (1990). "A molecular basis for familial hypertrophic cardiomyopathy: a beta cardiac myosin heavy chain gene missense mutation." Cell **62**(5): 999-1006.

Georg-Fries, B., S. Biederlack, J. Wolf and H. zur Hausen (1984). "Analysis of proteins, helper dependence, and seroepidemiology of a new human parvovirus." Virology **134**(1): 64-71.

Gerull, B., M. Gramlich, J. Atherton, M. McNabb, K. Trombitas, S. Sasse-Klaassen, J. G. Seidman, C. Seidman, H. Granzier, S. Labeit, M. Frenneaux and L. Thierfelder (2002). "Mutations of TTN, encoding the giant muscle filament titin, cause familial dilated cardiomyopathy." Nat Genet **30**(2): 201-204.

Gerull, B., A. Heuser, T. Wichter, M. Paul, C. T. Basson, D. A. McDermott, B. B. Lerman, S. M. Markowitz, P. T. Ellinor, C. A. MacRae, S. Peters, K. S. Grossmann, J. Drenckhahn, B. Michely, S. Sasse-Klaassen, W. Birchmeier, R. Dietz, G. Breithardt, E. Schulze-Bahr and L. Thierfelder (2004). "Mutations in the desmosomal protein plakophilin-2 are common in arrhythmogenic right ventricular cardiomyopathy." Nat Genet **36**(11): 1162-1164.

Gilbert, L. A., M. H. Larson, L. Morsut, Z. Liu, G. A. Brar, S. E. Torres, N. Stern-Ginossar, O. Brandman, E. H. Whitehead, J. A. Doudna, W. A. Lim, J. S. Weissman and L. S. Qi (2013). "CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes." Cell **154**(2): 442-451.

Golding, S. E., E. Rosenberg, S. Neill, P. Dent, L. F. Povirk and K. Valerie (2007). "Extracellular signal-related kinase positively regulates ataxia telangiectasia mutated, homologous recombination repair, and the DNA damage response." Cancer Res **67**(3): 1046-1053.

Grieger, J. C. and R. J. Samulski (2005). "Packaging capacity of adeno-associated virus serotypes: impact of larger genomes on infectivity and postentry steps." J Virol **79**(15): 9933-9944.

Gu, J., H. Lu, B. Tippin, N. Shimazaki, M. F. Goodman and M. R. Lieber (2007). "XRCC4:DNA ligase IV can ligate incompatible DNA ends and can ligate across gaps." EMBO J **26**(4): 1010-1023.

Guo, J., T. Gaj and C. F. Barbas, 3rd (2010). "Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nucleases." J Mol Biol **400**(1): 96-107.

Guo, W., S. Schafer, M. L. Greaser, M. H. Radke, M. Liss, T. Govindarajan, H. Maatz, H. Schulz, S. Li, A. M. Parrish, V. Dauksaite, P. Vakeel, S. Klaassen, B. Gerull, L. Thierfelder, V. Regitz-Zagrosek, T. A. Hacker, K. W. Saupe, G. W. Dec, P. T. Ellinor, C. A. MacRae, B. Spallek, R. Fischer, A. Perrot, C. Ozcelik, K. Saar, N. Hubner and M. Gotthardt (2012). "RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing." Nat Med **18**(5): 766-773.

Haghighi, K., F. Kolokathis, A. O. Gramolini, J. R. Waggoner, L. Pater, R. A. Lynch, G. C. Fan, D. Tsiapras, R. R. Parekh, G. W. Dorn, 2nd, D. H. MacLennan, D. T. Kremastinos and E. G. Kranias (2006). "A mutation in the human phospholamban gene, deleting arginine 14, results in lethal, hereditary cardiomyopathy." Proc Natl Acad Sci U S A **103**(5): 1388-1393.

Halbert, C. L., J. M. Allen and A. D. Miller (2001). "Adeno-associated virus type 6 (AAV6) vectors mediate efficient transduction of airway epithelial cells in mouse lungs compared to that of AAV2 vectors." J Virol **75**(14): 6615-6624.

Hammond, S. M., E. Bernstein, D. Beach and G. J. Hannon (2000). "An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells." Nature **404**(6775): 293-296.

Hanasoge, S. and M. Ljungman (2007). "H2AX phosphorylation after UV irradiation is triggered by DNA repair intermediates and is mediated by the ATR kinase." Carcinogenesis **28**(11): 2298-2304.

Hansen, J., K. Qing and A. Srivastava (2001). "Infection of purified nuclei by adeno-associated virus 2." Mol Ther **4**(4): 289-296.

Herman, A. P., A. Krawczynska, J. Bochenek, H. Antushevich, A. Herman and D. Tomaszewska-Zaremba (2014). "Peripheral injection of SB203580 inhibits the inflammatory-dependent synthesis of proinflammatory cytokines in the hypothalamus." Biomed Res Int **2014**: 475152.

Herman, D. S., L. Lam, M. R. Taylor, L. Wang, P. Teekakirikul, D. Christodoulou, L. Conner, S. R. DePalma, B. McDonough, E. Sparks, D. L. Teodorescu, A. L. Cirino, N. R. Banner, D. J. Pennell, S. Graw, M. Merlo, A. Di Lenarda, G. Sinagra, J. M. Bos, M. J. Ackerman, R. N. Mitchell, C. E. Murry, N. K. Lakdawala, C. Y. Ho, P. J. Barton, S. A. Cook, L. Mestroni, J. G. Seidman and C. E. Seidman (2012). "Truncations of titin causing dilated cardiomyopathy." N Engl J Med **366**(7): 619-628.

Hilton, I. B., A. M. D'Ippolito, C. M. Vockley, P. I. Thakore, G. E. Crawford, T. E. Reddy and C. A. Gersbach (2015). "Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers." Nat Biotechnol **33**(5): 510-517.

Ho, C. Y. (2012). "Genetic considerations in hypertrophic cardiomyopathy." Prog Cardiovasc Dis **54**(6): 456-460.

Hockemeyer, D., F. Soldner, C. Beard, Q. Gao, M. Mitalipova, R. C. DeKelver, G. E. Katibah, R. Amora, E. A. Boydston, B. Zeitler, X. Meng, J. C. Miller, L. Zhang, E. J. Rebar, P. D. Gregory, F. D. Urnov and R. Jaenisch (2009). "Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases." Nat Biotechnol **27**(9): 851-857.

Hockemeyer, D., H. Wang, S. Kiani, C. S. Lai, Q. Gao, J. P. Cassady, G. J. Cost, L. Zhang, Y. Santiago, J. C. Miller, B. Zeitler, J. M. Cherone, X. Meng, S. J. Hinkley, E. J. Rebar, P. D. Gregory, F. D. Urnov and R. Jaenisch (2011). "Genetic engineering of human pluripotent cells using TALE nucleases." Nat Biotechnol **29**(8): 731-734.

Hoggan, M. D., N. R. Blacklow and W. P. Rowe (1966). "Studies of small DNA viruses found in various adenovirus preparations: physical, biological, and immunological characteristics." Proc Natl Acad Sci U S A **55**(6): 1467-1474.

Holliday, R. (2007). "A mechanism for gene conversion in fungi." Genet Res **89**(5-6): 285-307.

Hui, D. J., E. Basner-Tschakarjan, Y. Chen, R. J. Davidson, G. Buchlis, M. Yazicioglu, G. C. Pien, J. D. Finn, V. Haurigot, A. Tai, D. W. Scott, L. P. Cousens, S. Zhou, A. S. De Groot and F. Mingozzi (2013). "Modulation of CD8+ T cell responses to AAV vectors with IgG-derived MHC class II epitopes." Mol Ther **21**(9): 1727-1737.

Hui, D. J., S. C. Edmonson, G. M. Podsakoff, G. C. Pien, L. Ivanciu, R. M. Camire, H. Ertl, F. Mingozzi, K. A. High and E. Basner-Tschakarjan (2015). "AAV capsid CD8+ T-cell epitopes are highly conserved across AAV serotypes." Mol Ther Methods Clin Dev **2**: 15029.

Hummel, M., Z. G. Li, D. Pfaffinger, L. Neven and A. M. Scanu (1990). "Familial hypercholesterolemia in a rhesus monkey pedigree: molecular basis of low density lipoprotein receptor deficiency." Proc Natl Acad Sci U S A **87**(8): 3122-3126.

Huser, D., A. Gogol-Doring, W. Chen and R. Heilbronn (2014). "Adeno-associated virus type 2 wild-type and vector-mediated genomic integration profiles of human diploid fibroblasts analyzed by third-generation PacBio DNA sequencing." J Virol **88**(19): 11253-11263.

Huser, D., A. Gogol-Doring, T. Lutter, S. Weger, K. Winter, E. M. Hammer, T. Cathomen, K. Reinert and R. Heilbronn (2010). "Integration preferences of wildtype AAV-2 for consensus rep-binding sites at numerous loci in the human genome." PLoS Pathog **6**(7): e1000985.

Inagaki, K., S. Fuess, T. A. Storm, G. A. Gibson, C. F. McTiernan, M. A. Kay and H. Nakai (2006). "Robust systemic transduction with AAV9 vectors in mice: efficient global cardiac gene transfer superior to that of AAV8." Mol Ther **14**(1): 45-53.

Ira, G., A. Pellicioli, A. Balijja, X. Wang, S. Fiorani, W. Carotenuto, G. Liberi, D. Bressan, L. Wan, N. M. Hollingsworth, J. E. Haber and M. Foiani (2004). "DNA end resection, homologous recombination and DNA damage checkpoint activation require CDK1." Nature **431**(7011): 1011-1017.

Ishino, Y., H. Shinagawa, K. Makino, M. Amemura and A. Nakata (1987). "Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product." J Bacteriol **169**(12): 5429-5433.

Ishizu, T., S. Higo, Y. Masumura, Y. Kohama, M. Shiba, T. Higo, M. Shibamoto, A. Nakagawa, S. Morimoto, S. Takashima, S. Hikoso and Y. Sakata (2017). "Targeted Genome Replacement via Homology-directed Repair in Non-dividing Cardiomyocytes." Sci Rep **7**(1): 9363.

Jansen, R., J. D. Embden, W. Gaastra and L. M. Schouls (2002). "Identification of genes that are associated with DNA repeats in prokaryotes." Mol Microbiol **43**(6): 1565-1575.

Jarcho, J. A., W. McKenna, J. A. Pare, S. D. Solomon, R. F. Holcombe, S. Dickie, T. Levi, H. Donis-Keller, J. G. Seidman and C. E. Seidman (1989). "Mapping a gene for familial hypertrophic cardiomyopathy to chromosome 14q1." N Engl J Med **321**(20): 1372-1378.

Jarnicki, A. G., H. Conroy, C. Brereton, G. Donnelly, D. Toomey, K. Walsh, C. Sweeney, O. Leavy, J. Fletcher, E. C. Lavelle, P. Dunne and K. H. Mills (2008). "Attenuating regulatory T cell induction by TLR agonists through inhibition of p38 MAPK signaling in dendritic cells enhances their efficacy as vaccine adjuvants and cancer immunotherapeutics." J Immunol **180**(6): 3797-3806.

Jiang, H., L. B. Couto, S. Patarroyo-White, T. Liu, D. Nagy, J. A. Vargas, S. Zhou, C. D. Scallan, J. Sommer, S. Vijay, F. Mingozzi, K. A. High and G. F. Pierce (2006). "Effects of transient immunosuppression on adenoassociated, virus-mediated, liver-directed gene transfer in rhesus macaques and implications for human gene therapy." Blood **108**(10): 3321-3328.

Jima, D. D., J. Zhang, C. Jacobs, K. L. Richards, C. H. Dunphy, W. W. Choi, W. Y. Au, G. Srivastava, M. B. Czader, D. A. Rizzieri, A. S. Lagoo, P. L. Lugar, K. P. Mann, C. R. Flowers, L. Bernal-Mizrachi, K. N. Naresh, A. M. Evens, L. I. Gordon, M. Luftig, D. R. Friedman, J. B. Weinberg, M. A. Thompson, J. I. Gill, Q. Liu, T. How, V. Grubor, Y. Gao, A. Patel, H. Wu, J. Zhu, G. C. Blobe, P. E. Lipsky, A. Chadburn, S. S. Dave and C. Hematologic Malignancies Research (2010). "Deep sequencing of the small RNA transcriptome of normal and malignant human B cells identifies hundreds of novel microRNAs." Blood **116**(23): e118-127.

Jin, Y., Z. Chen, X. Liu and X. Zhou (2013). "Evaluating the microRNA targeting sites by luciferase reporter gene assay." Methods Mol Biol **936**: 117-127.

Jinek, M., A. East, A. Cheng, S. Lin, E. Ma and J. Doudna (2013). "RNA-programmed genome editing in human cells." Elife **2**: e00471.

Johansen, A. K., B. Molenaar, D. Versteeg, A. R. Leitoguinho, C. Demkes, B. Spanjaard, H. de Ruiter, F. Akbari Moqadam, L. Kooijman, L. Zentilin, M. Giacca and E. van Rooij (2017). "Postnatal Cardiac Gene Editing Using CRISPR/Cas9 With AAV9-Mediated Delivery of Short Guide RNAs Results in Mosaic Gene Disruption." Circ Res **121**(10): 1168-1181.

Johnson, J. S. and R. J. Samulski (2009). "Enhancement of adeno-associated virus infection by mobilizing capsids into and out of the nucleolus." J Virol **83**(6): 2632-2644.

Ketting, R. F., S. E. Fischer, E. Bernstein, T. Sijen, G. J. Hannon and R. H. Plasterk (2001). "Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans." Genes Dev **15**(20): 2654-2659.

King, J. A., R. Dubielzig, D. Grimm and J. A. Kleinschmidt (2001). "DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids." EMBO J **20**(12): 3282-3291.

Koike-Yusa, H., Y. Li, E. P. Tan, C. Velasco-Herrera Mdel and K. Yusa (2014). "Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library." Nat Biotechnol **32**(3): 267-273.

Kolodkin, A. L., A. J. Klar and F. W. Stahl (1986). "Double-strand breaks can initiate meiotic recombination in S. cerevisiae." Cell **46**(5): 733-740.

Kronenberg, S., J. A. Kleinschmidt and B. Bottcher (2001). "Electron cryo-microscopy and image reconstruction of adeno-associated virus type 2 empty capsids." EMBO Rep **2**(11): 997-1002.

Lagos-Quintana, M., R. Rauhut, W. Lendeckel and T. Tuschl (2001). "Identification of novel genes coding for small expressed RNAs." Science **294**(5543): 853-858.

Laughlin, C. A., J. D. Tratschin, H. Coon and B. J. Carter (1983). "Cloning of infectious adeno-associated virus genomes in bacterial plasmids." Gene **23**(1): 65-73.

Lee, J. M., C. Nobumori, Y. Tu, C. Choi, S. H. Yang, H. J. Jung, T. A. Vickers, F. Rigo, C. F. Bennett, S. G. Young and L. G. Fong (2016). "Modulation of LMNA splicing as a strategy to treat prelamin A diseases." J Clin Invest **126**(4): 1592-1602.

Lee, R. C., R. L. Feinbaum and V. Ambros (1993). "The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14." Cell **75**(5): 843-854.

Lee, Y., C. Ahn, J. Han, H. Choi, J. Kim, J. Yim, J. Lee, P. Provost, O. Radmark, S. Kim and V. N. Kim (2003). "The nuclear RNase III Drosha initiates microRNA processing." Nature **425**(6956): 415-419.

Lee, Y., M. Kim, J. Han, K. H. Yeom, S. Lee, S. H. Baek and V. N. Kim (2004). "MicroRNA genes are transcribed by RNA polymerase II." EMBO J **23**(20): 4051-4060.

Lehrman, M. A., J. L. Goldstein, M. S. Brown, D. W. Russell and W. J. Schneider (1985). "Internalization-defective LDL receptors produced by genes with nonsense and frameshift mutations that truncate the cytoplasmic domain." Cell **41**(3): 735-743.

Lennox, K. A. and M. A. Behlke (2010). "A direct comparison of anti-microRNA oligonucleotide potency." Pharm Res **27**(9): 1788-1799.

Lesizza, P., G. Prosdocimo, V. Martinelli, G. Sinagra, S. Zacchigna and M. Giacca (2017). "Single-Dose Intracardiac Injection of Pro-Regenerative MicroRNAs Improves Cardiac Function After Myocardial Infarction." Circ Res **120**(8): 1298-1304.

Lewis, B. P., C. B. Burge and D. P. Bartel (2005). "Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets." Cell **120**(1): 15-20.

Li, C., N. Diprimio, D. E. Bowles, M. L. Hirsch, P. E. Monahan, A. Asokan, J. Rabinowitz, M. Agbandje-McKenna and R. J. Samulski (2012). "Single amino acid modification of adeno-associated virus capsid changes transduction and humoral immune profiles." J Virol **86**(15): 7752-7759.

Li, X. and W. D. Heyer (2009). "RAD54 controls access to the invading 3'-OH end after RAD51-mediated DNA strand invasion in homologous recombination in Saccharomyces cerevisiae." Nucleic Acids Res **37**(2): 638-646.

Li, Y. R., L. Q. Wen, Y. Wang, T. C. Zhou, N. Ma, Z. H. Hou and Z. P. Jiang (2016). "MicroRNA-520c enhances cell proliferation, migration, and invasion by suppressing IRF2 in gastric cancer." FEBS Open Bio **6**(12): 1257-1266.

Liang, D., S. L. Burkhart, R. K. Singh, M. H. Kabbaj and A. Gunjan (2012). "Histone dosage regulates DNA damage sensitivity in a checkpoint-independent manner by the homologous recombination pathway." Nucleic Acids Res **40**(19): 9604-9620.

Lin, F. L., K. Sperle and N. Sternberg (1984). "Model for homologous recombination during transfer of DNA into mouse L cells: role for DNA ends in the recombination process." Mol Cell Biol **4**(6): 1020-1034.

Llave, C., Z. Xie, K. D. Kasschau and J. C. Carrington (2002). "Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA." Science **297**(5589): 2053-2056.

Lovric, J., M. Mano, L. Zentilin, A. Eulalio, S. Zacchigna and M. Giacca (2012). "Terminal differentiation of cardiac and skeletal myocytes induces permissivity

to AAV transduction by relieving inhibition imposed by DNA damage response proteins." Mol Ther **20**(11): 2087-2097.

Lund, E., S. Guttinger, A. Calado, J. E. Dahlberg and U. Kutay (2004). "Nuclear export of microRNA precursors." Science **303**(5654): 95-98.

Lusby, E., R. Bohenzky and K. I. Berns (1981). "Inverted terminal repetition in adeno-associated virus DNA: independence of the orientation at either end of the genome." J Virol **37**(3): 1083-1086.

Lux, K., N. Goerlitz, S. Schlemminger, L. Perabo, D. Goldnau, J. Endell, K. Leike, D. M. Kofler, S. Finke, M. Hallek and H. Buning (2005). "Green fluorescent protein-tagged adeno-associated virus particles allow the study of cytosolic and nuclear trafficking." J Virol **79**(18): 11776-11787.

Ma, Y., H. Lu, B. Tippin, M. F. Goodman, N. Shimazaki, O. Koiwai, C. L. Hsieh, K. Schwarz and M. R. Lieber (2004). "A biochemically defined system for mammalian nonhomologous DNA end joining." Mol Cell **16**(5): 701-713.

Ma, Y., U. Pannicke, K. Schwarz and M. R. Lieber (2002). "Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination." Cell **108**(6): 781-794.

Macera, M. J., P. Szabo, R. Wadgaonkar, M. A. Siddiqui and R. S. Verma (1992). "Localization of the gene coding for ventricular myosin regulatory light chain (MYL2) to human chromosome 12q23-q24.3." Genomics **13**(3): 829-831.

Malhotra, R. and P. K. Mason (2009). "Lamin A/C deficiency as a cause of familial dilated cardiomyopathy." Curr Opin Cardiol **24**(3): 203-208.

Mali, P., L. Yang, K. M. Esvelt, J. Aach, M. Guell, J. E. DiCarlo, J. E. Norville and G. M. Church (2013). "RNA-guided human genome engineering via Cas9." Science **339**(6121): 823-826.

Maloisel, L., J. Bhargava and G. S. Roeder (2004). "A role for DNA polymerase delta in gene conversion and crossing over during meiosis in Saccharomyces cerevisiae." Genetics **167**(3): 1133-1142.

Mandel, R. J. (2010). "CERE-110, an adeno-associated virus-based gene delivery vector expressing human nerve growth factor for the treatment of Alzheimer's disease." Curr Opin Mol Ther **12**(2): 240-247.

Mangeot, P. E., V. Risson, F. Fusil, A. Marnef, E. Laurent, J. Blin, V. Mournetas, E. Massourides, T. J. M. Sohier, A. Corbin, F. Aube, M. Teixeira, C. Pinset, L. Schaeffer, G. Legube, F. L. Cosset, E. Verhoeyen, T. Ohlmann and E. P. Ricci (2019). "Genome editing in primary cells and in vivo using viral-derived Nanoblades loaded with Cas9-sgRNA ribonucleoproteins." Nat Commun **10**(1): 45.

Manno, C. S., G. F. Pierce, V. R. Arruda, B. Glader, M. Ragni, J. J. Rasko, M. C. Ozelo, K. Hoots, P. Blatt, B. Konkle, M. Dake, R. Kaye, M. Razavi, A. Zajko, J. Zehnder, P. K.

Rustagi, H. Nakai, A. Chew, D. Leonard, J. F. Wright, R. R. Lessard, J. M. Sommer, M. Tigges, D. Sabatino, A. Luk, H. Jiang, F. Mingozzi, L. Couto, H. C. Ertl, K. A. High and M. A. Kay (2006). "Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response." Nat Med **12**(3): 342-347.

Mano, M., R. Ippodrino, L. Zentilin, S. Zacchigna and M. Giacca (2015). "Genome-wide RNAi screening identifies host restriction factors critical for in vivo AAV transduction." Proc Natl Acad Sci U S A **112**(36): 11276-11281.

Marcello, A., P. Massimi, L. Banks and M. Giacca (2000). "Adeno-associated virus type 2 rep protein inhibits human papillomavirus type 16 E2 recruitment of the transcriptional coactivator p300." J Virol **74**(19): 9090-9098.

Marcus, C. J., C. A. Laughlin and B. J. Carter (1981). "Adeno-associated virus RNA transcription in vivo." Eur J Biochem **121**(1): 147-154.

Maron, B. J., J. A. Towbin, G. Thiene, C. Antzelevitch, D. Corrado, D. Arnett, A. J. Moss, C. E. Seidman, J. B. Young, A. American Heart, H. F. Council on Clinical Cardiology, C. Transplantation, C. Quality of, R. Outcomes, G. Functional, G. Translational Biology Interdisciplinary Working, E. Council on and Prevention (2006). "Contemporary definitions and classification of the cardiomyopathies: an American Heart Association Scientific Statement from the Council on Clinical Cardiology, Heart Failure and Transplantation Committee; Quality of Care and Outcomes Research and Functional Genomics and Translational Biology Interdisciplinary Working Groups; and Council on Epidemiology and Prevention." Circulation **113**(14): 1807-1816.

Marraffini, L. A. and E. J. Sontheimer (2008). "CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA." Science **322**(5909): 1843-1845.

Matsuoka, S., G. Rotman, A. Ogawa, Y. Shiloh, K. Tamai and S. J. Elledge (2000). "Ataxia telangiectasia-mutated phosphorylates Chk2 in vivo and in vitro." Proc Natl Acad Sci U S A **97**(19): 10389-10394.

McCarty, D. M., P. E. Monahan and R. J. Samulski (2001). "Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis." Gene Ther **8**(16): 1248-1254.

McEachern, M. J. and J. E. Haber (2006). "Break-induced replication and recombinational telomere elongation in yeast." Annu Rev Biochem **75**: 111-135.

McGill, L. A., T. F. Ismail, S. Nielles-Vallespin, P. Ferreira, A. D. Scott, M. Roughton, P. J. Kilner, S. Y. Ho, K. P. McCarthy, P. D. Gatehouse, R. de Silva, P. Speier, T. Feiweier, C. Mekkaoui, D. E. Sosnovik, S. K. Prasad, D. N. Firmin and D. J. Pennell (2012). "Reproducibility of in-vivo diffusion tensor cardiovascular magnetic resonance in hypertrophic cardiomyopathy." J Cardiovasc Magn Reson **14**: 86.

McKoy, G., N. Protonotarios, A. Crosby, A. Tsatsopoulou, A. Anastasakis, A. Coonar, M. Norman, C. Baboonian, S. Jeffery and W. J. McKenna (2000). "Identification of a deletion in plakoglobin in arrhythmogenic right ventricular cardiomyopathy with palmoplantar keratoderma and woolly hair (Naxos disease)." Lancet **355**(9221): 2119-2124.

McNair, W. P., G. Sinagra, M. R. Taylor, A. Di Lenarda, D. A. Ferguson, E. E. Salcedo, D. Slavov, X. Zhu, J. H. Caldwell, L. Mestroni and G. Familial Cardiomyopathy Registry Research (2011). "SCN5A mutations associate with arrhythmic dilated cardiomyopathy and commonly localize to the voltage-sensing mechanism." J Am Coll Cardiol **57**(21): 2160-2168.

Menke-van der Houven van Oordt, C. W., A. McGeoch, M. Bergstrom, I. McSherry, D. A. Smith, M. Cleveland, W. Al-Azzam, L. Chen, H. Verheul, O. S. Hoekstra, D. J. Vugts, I. Freedman, M. Huisman, C. Matheny, G. van Dongen and S. Zhang (2019). "Immuno-PET Imaging to Assess Target Engagement: Experience from (89)Zr-Anti-HER3 mAb (GSK2849330) in Patients with Solid Tumors." J Nucl Med **60**(7): 902-909.

Michailidou, K., J. Beesley, S. Lindstrom, S. Canisius, J. Dennis, M. J. Lush, M. J. Maranian, M. K. Bolla, Q. Wang, M. Shah, B. J. Perkins, K. Czene, M. Eriksson, H. Darabi, J. S. Brand, S. E. Bojesen, B. G. Nordestgaard, H. Flyger, S. F. Nielsen, N. Rahman, C. Turnbull, Bocs, O. Fletcher, J. Peto, L. Gibson, I. dos-Santos-Silva, J. Chang-Claude, D. Flesch-Janys, A. Rudolph, U. Eilber, S. Behrens, H. Nevanlinna, T. A. Muranen, K. Aittomaki, C. Blomqvist, S. Khan, K. Aaltonen, H. Ahsan, M. G. Kibriya, A. S. Whittemore, E. M. John, K. E. Malone, M. D. Gammon, R. M. Santella, G. Ursin, E. Makalic, D. F. Schmidt, G. Casey, D. J. Hunter, S. M. Gapstur, M. M. Gaudet, W. R. Diver, C. A. Haiman, F. Schumacher, B. E. Henderson, L. Le Marchand, C. D. Berg, S. J. Chanock, J. Figueroa, R. N. Hoover, D. Lambrechts, P. Neven, H. Wildiers, E. van Limbergen, M. K. Schmidt, A. Broeks, S. Verhoef, S. Cornelissen, F. J. Couch, J. E. Olson, E. Hallberg, C. Vachon, Q. Waisfisz, H. Meijers-Heijboer, M. A. Adank, R. B. van der Luijt, J. Li, J. Liu, K. Humphreys, D. Kang, J. Y. Choi, S. K. Park, K. Y. Yoo, K. Matsuo, H. Ito, H. Iwata, K. Tajima, P. Guenel, T. Truong, C. Mulot, M. Sanchez, B. Burwinkel, F. Marme, H. Surowy, C. Sohn, A. H. Wu, C. C. Tseng, D. Van Den Berg, D. O. Stram, A. Gonzalez-Neira, J. Benitez, M. P. Zamora, J. I. Perez, X. O. Shu, W. Lu, Y. T. Gao, H. Cai, A. Cox, S. S. Cross, M. W. Reed, I. L. Andrulis, J. A. Knight, G. Glendon, A. M. Mulligan, E. J. Sawyer, I. Tomlinson, M. J. Kerin, N. Miller, I. kConFab, A. Group, A. Lindblom, S. Margolin, S. H. Teo, C. H. Yip, N. A. Taib, G. H. Tan, M. J. Hooning, A. Hollestelle, J. W. Martens, J. M. Collee, W. Blot, L. B. Signorello, Q. Cai, J. L. Hopper, M. C. Southey, H. Tsimiklis, C. Apicella, C. Y. Shen, C. N. Hsiung, P. E. Wu, M. F. Hou, V. N. Kristensen, S. Nord, G. I. Alnaes, Nbcs, G. G. Giles, R. L. Milne, C. McLean, F. Canzian, D. Trichopoulos, P. Peeters, E. Lund, M. Sund, K. T. Khaw, M. J. Gunter, D. Palli, L. M. Mortensen, L. Dossus, J. M. Huerta, A. Meindl, R. K. Schmutzler, C. Sutter, R. Yang, K. Muir, A. Lophatananon, S. Stewart-Brown, P. Siriwanarangsan, M. Hartman, H. Miao, K. S. Chia, C. W. Chan, P. A. Fasching, A. Hein, M. W. Beckmann, L. Haeberle, H. Brenner, A. K. Dieffenbach, V. Arndt, C. Stegmaier, A. Ashworth, N. Orr, M. J. Schoemaker, A. J. Swerdlow, L. Brinton, M. Garcia-Closas, W. Zheng, S. L. Halverson, M. Shrubsole, J. Long, M. S. Goldberg, F. Labreche, M. Dumont, R. Winqvist, K. Pylkas, A. Jukkola-

Vuorinen, M. Grip, H. Brauch, U. Hamann, T. Bruning, G. Network, P. Radice, P. Peterlongo, S. Manoukian, L. Bernard, N. V. Bogdanova, T. Dork, A. Mannermaa, V. Kataja, V. M. Kosma, J. M. Hartikainen, P. Devilee, R. A. Tollenaar, C. Seynaeve, C. J. Van Asperen, A. Jakubowska, J. Lubinski, K. Jaworska, T. Huzarski, S. Sangrajrang, V. Gaborieau, P. Brennan, J. McKay, S. Slager, A. E. Toland, C. B. Ambrosone, D. Yannoukakos, M. Kabisch, D. Torres, S. L. Neuhausen, H. Anton-Culver, C. Luccarini, C. Baynes, S. Ahmed, C. S. Healey, D. C. Tessier, D. Vincent, F. Bacot, G. Pita, M. R. Alonso, N. Alvarez, D. Herrero, J. Simard, P. P. Pharoah, P. Kraft, A. M. Dunning, G. Chenevix-Trench, P. Hall and D. F. Easton (2015). "Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer." Nat Genet **47**(4): 373-380.

Michel, B., G. Grompone, M. J. Flores and V. Bidnenko (2004). "Multiple pathways process stalled replication forks." Proc Natl Acad Sci U S A **101**(35): 12783-12788.

Mimori, T. and J. A. Hardin (1986). "Mechanism of interaction between Ku protein and DNA." J Biol Chem **261**(22): 10375-10379.

Miura, T., Y. Yamana, T. Usui, H. I. Ogawa, M. T. Yamamoto and K. Kusano (2012). "Homologous recombination via synthesis-dependent strand annealing in yeast requires the Irc20 and Srs2 DNA helicases." Genetics **191**(1): 65-78.

Mojica, F. J., C. Diez-Villasenor, J. Garcia-Martinez and C. Almendros (2009). "Short motif sequences determine the targets of the prokaryotic CRISPR defence system." Microbiology **155**(Pt 3): 733-740.

Mojica, F. J., C. Diez-Villasenor, J. Garcia-Martinez and E. Soria (2005). "Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements." J Mol Evol **60**(2): 174-182.

Mojica, F. J., C. Diez-Villasenor, E. Soria and G. Juez (2000). "Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria." Mol Microbiol **36**(1): 244-246.

Moore, J. K. and J. E. Haber (1996). "Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in Saccharomyces cerevisiae." Mol Cell Biol **16**(5): 2164-2173.

Moore, R., A. Spinhirne, M. J. Lai, S. Preisser, Y. Li, T. Kang and L. Bleris (2015). "CRISPR-based self-cleaving mechanism for controllable gene delivery in human cells." Nucleic Acids Res **43**(2): 1297-1303.

Morita, H., H. L. Rehm, A. Menesses, B. McDonough, A. E. Roberts, R. Kucherlapati, J. A. Towbin, J. G. Seidman and C. E. Seidman (2008). "Shared genetic causes of cardiac hypertrophy in children and adults." N Engl J Med **358**(18): 1899-1908.

Morrical, S. W. (2015). "DNA-pairing and annealing processes in homologous recombination and homology-directed repair." Cold Spring Harb Perspect Biol **7**(2): a016444.

Morrison, A. J., J. Highland, N. J. Krogan, A. Arbel-Eden, J. F. Greenblatt, J. E. Haber and X. Shen (2004). "INO80 and gamma-H2AX interaction links ATP-dependent chromatin remodeling to DNA damage repair." Cell **119**(6): 767-775.

Moscou, M. J. and A. J. Bogdanove (2009). "A simple cipher governs DNA recognition by TAL effectors." Science **326**(5959): 1501.

Motycka, T. A., T. Bessho, S. M. Post, P. Sung and A. E. Tomkinson (2004). "Physical and functional interaction between the XPF/ERCC1 endonuclease and hRad52." J Biol Chem **279**(14): 13634-13639.

Moynahan, M. E., A. J. Pierce and M. Jasin (2001). "BRCA2 is required for homology-directed repair of chromosomal breaks." Mol Cell **7**(2): 263-272.

Muller, O. J., F. Kaul, M. D. Weitzman, R. Pasqualini, W. Arap, J. A. Kleinschmidt and M. Trepel (2003). "Random peptide libraries displayed on adeno-associated virus to select for targeted gene therapy vectors." Nat Biotechnol **21**(9): 1040-1046.

Nam, K. H., C. Haitjema, X. Liu, F. Ding, H. Wang, M. P. DeLisa and A. Ke (2012). "Cas5d protein processes pre-crRNA and assembles into a cascade-like interference complex in subtype I-C/Dvulg CRISPR-Cas system." Structure **20**(9): 1574-1584.

Nathwani, A. C., U. M. Reiss, E. G. Tuddenham, C. Rosales, P. Chowdary, J. McIntosh, M. Della Peruta, E. Lheriteau, N. Patel, D. Raj, A. Riddell, J. Pie, S. Rangarajan, D. Bevan, M. Recht, Y. M. Shen, K. G. Halka, E. Basner-Tschakarjan, F. Mingozzi, K. A. High, J. Allay, M. A. Kay, C. Y. Ng, J. Zhou, M. Cancio, C. L. Morton, J. T. Gray, D. Srivastava, A. W. Nienhuis and A. M. Davidoff (2014). "Long-term safety and efficacy of factor IX gene therapy in hemophilia B." N Engl J Med **371**(21): 1994-2004.

Naumer, M., F. Sonntag, K. Schmidt, K. Nieto, C. Panke, N. E. Davey, R. Popa-Wagner and J. A. Kleinschmidt (2012). "Properties of the adeno-associated virus assembly-activating protein." J Virol **86**(23): 13038-13048.

Nelson, C. E., C. H. Hakim, D. G. Ousterout, P. I. Thakore, E. A. Moreb, R. M. Castellanos Rivera, S. Madhavan, X. Pan, F. A. Ran, W. X. Yan, A. Asokan, F. Zhang, D. Duan and C. A. Gersbach (2016). "In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy." Science **351**(6271): 403-407.

Nicolson, S. C. and R. J. Samulski (2014). "Recombinant adeno-associated virus utilizes host cell nuclear import machinery to enter the nucleus." J Virol **88**(8): 4132-4144.

Nishimasu, H., F. A. Ran, P. D. Hsu, S. Konermann, S. I. Shehata, N. Dohmae, R. Ishitani, F. Zhang and O. Nureki (2014). "Crystal structure of Cas9 in complex with guide RNA and target DNA." Cell **156**(5): 935-949.

Noguchi, S., Y. Yasui, J. Iwasaki, M. Kumazaki, N. Yamada, S. Naito and Y. Akao (2013). "Replacement treatment with microRNA-143 and -145 induces synergistic inhibition of the growth of human bladder cancer cells by regulating PI3K/Akt and MAPK signaling pathways." Cancer Lett **328**(2): 353-361.

Nonnenmacher, M. and T. Weber (2012). "Intracellular transport of recombinant adeno-associated virus vectors." Gene Ther **19**(6): 649-658.

Nunez, J. K., A. S. Lee, A. Engelman and J. A. Doudna (2015). "Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity." Nature **519**(7542): 193-198.

Ohashi, K., H. Nakai, L. B. Couto and M. A. Kay (2005). "Modified infusion procedures affect recombinant adeno-associated virus vector type 2 transduction in the liver." Hum Gene Ther **16**(3): 299-306.

Orom, U. A. and A. H. Lund (2007). "Isolation of microRNA targets using biotinylated synthetic microRNAs." Methods **43**(2): 162-165.

Orr-Weaver, T. L., J. W. Szostak and R. J. Rothstein (1981). "Yeast transformation: a model system for the study of recombination." Proc Natl Acad Sci U S A **78**(10): 6354-6358.

Parchem, R. J., N. Moore, J. L. Fish, J. G. Parchem, T. T. Braga, A. Shenoy, M. C. Oldham, J. L. Rubenstein, R. A. Schneider and R. Blelloch (2015). "miR-302 Is Required for Timing of Neural Differentiation, Neural Tube Closure, and Embryonic Viability." Cell Rep **12**(5): 760-773.

Pasquinelli, A. E., B. J. Reinhart, F. Slack, M. Q. Martindale, M. I. Kuroda, B. Maller, D. C. Hayward, E. E. Ball, B. Degnan, P. Muller, J. Spring, A. Srinivasan, M. Fishman, J. Finnerty, J. Corbo, M. Levine, P. Leahy, E. Davidson and G. Ruvkun (2000). "Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA." Nature **408**(6808): 86-89.

Paulk, N. K., K. Wursthorn, Z. Wang, M. J. Finegold, M. A. Kay and M. Grompe (2010). "Adeno-associated virus gene repair corrects a mouse model of hereditary tyrosinemia in vivo." Hepatology **51**(4): 1200-1208.

Perabo, L., J. Endell, S. King, K. Lux, D. Goldnau, M. Hallek and H. Buning (2006). "Combinatorial engineering of a gene therapy vector: directed evolution of adeno-associated virus." J Gene Med **8**(2): 155-162.

Pereira, D. M., P. M. Rodrigues, P. M. Borralho and C. M. Rodrigues (2013). "Delivering the promise of miRNA cancer therapeutics." Drug Discov Today **18**(5-6): 282-289.

Perez-Pinera, P., D. D. Kocak, C. M. Vockley, A. F. Adler, A. M. Kabadi, L. R. Polstein, P. I. Thakore, K. A. Glass, D. G. Ousterout, K. W. Leong, F. Guilak, G. E. Crawford, T. E. Reddy and C. A. Gersbach (2013). "RNA-guided gene activation by CRISPR-Cas9-based transcription factors." Nat Methods **10**(10): 973-976.

Persson, H., A. Kvist, N. Rego, J. Staaf, J. Vallon-Christersson, L. Luts, N. Loman, G. Jonsson, H. Naya, M. Hoglund, A. Borg and C. Rovira (2011). "Identification of new microRNAs in paired normal and tumor breast tissue suggests a dual role for the ERBB2/Her2 gene." Cancer Res **71**(1): 78-86.

Petersen-Jones, S. M., J. T. Bartoe, A. J. Fischer, M. Scott, S. L. Boye, V. Chiodo and W. W. Hauswirth (2009). "AAV retinal transduction in a large animal model species: comparison of a self-complementary AAV2/5 with a single-stranded AAV2/5 vector." Mol Vis **15**: 1835-1842.

Petrs-Silva, H., A. Dinculescu, Q. Li, S. H. Min, V. Chiodo, J. J. Pang, L. Zhong, S. Zolotukhin, A. Srivastava, A. S. Lewin and W. W. Hauswirth (2009). "High-efficiency transduction of the mouse retina by tyrosine-mutant AAV serotype vectors." Mol Ther **17**(3): 463-471.

Pilichou, K., A. Nava, C. Basso, G. Beffagna, B. Bauce, A. Lorenzon, G. Frigo, A. Vettori, M. Valente, J. Towbin, G. Thiene, G. A. Danieli and A. Rampazzo (2006). "Mutations in desmoglein-2 gene are associated with arrhythmogenic right ventricular cardiomyopathy." Circulation **113**(9): 1171-1179.

Pourcel, C., G. Salvignol and G. Vergnaud (2005). "CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies." Microbiology **151**(Pt 3): 653-663.

Qi, L. S., M. H. Larson, L. A. Gilbert, J. A. Doudna, J. S. Weissman, A. P. Arkin and W. A. Lim (2013). "Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression." Cell **152**(5): 1173-1183.

Ramakrishna, S., A. B. Kwaku Dad, J. Beloor, R. Gopalappa, S. K. Lee and H. Kim (2014). "Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA." Genome Res **24**(6): 1020-1027.

Rampazzo, A., A. Nava, S. Malacrida, G. Beffagna, B. Bauce, V. Rossi, R. Zimbello, B. Simionati, C. Basso, G. Thiene, J. A. Towbin and G. A. Danieli (2002). "Mutation in human desmoplakin domain binding to plakoglobin causes a dominant form of arrhythmogenic right ventricular cardiomyopathy." Am J Hum Genet **71**(5): 1200-1206.

Ran, F. A., P. D. Hsu, C. Y. Lin, J. S. Gootenberg, S. Konermann, A. E. Trevino, D. A. Scott, A. Inoue, S. Matoba, Y. Zhang and F. Zhang (2013). "Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity." Cell **154**(6): 1380-1389.

Reilly, M. E., G. McKoy, D. Mantle, T. J. Peters, G. Goldspink and V. R. Preedy (2000). "Protein and mRNA levels of the myosin heavy chain isoforms Ibeta, IIa, IIx and IIb in type I and type II fibre-predominant rat skeletal muscles in response to chronic alcohol feeding." J Muscle Res Cell Motil **21**(8): 763-773.

Reinhart, B. J., F. J. Slack, M. Basson, A. E. Pasquinelli, J. C. Bettinger, A. E. Rougvie, H. R. Horvitz and G. Ruvkun (2000). "The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans." Nature **403**(6772): 901-906.

Ren, J., P. Jin, E. Wang, F. M. Marincola and D. F. Stroncek (2009). "MicroRNA and gene expression patterns in the differentiation of human embryonic stem cells." J Transl Med **7**: 20.

Richard, P., P. Charron, L. Carrier, C. Ledeuil, T. Cheav, C. Pichereau, A. Benaiche, R. Isnard, O. Dubourg, M. Burban, J. P. Gueffet, A. Millaire, M. Desnos, K. Schwartz, B. Hainque, M. Komajda and E. H. F. Project (2003). "Hypertrophic cardiomyopathy: distribution of disease genes, spectrum of mutations, and implications for a molecular diagnosis strategy." Circulation **107**(17): 2227-2232.

Rodino-Klapac, L. R., L. G. Chicoine, B. K. Kaspar and J. R. Mendell (2007). "Gene therapy for duchenne muscular dystrophy: expectations and challenges." Arch Neurol **64**(9): 1236-1241.

Rutledge, E. A., C. L. Halbert and D. W. Russell (1998). "Infectious clones and vectors derived from adeno-associated virus (AAV) serotypes other than AAV type 2." J Virol **72**(1): 309-319.

Salmon, F., K. Grosios and H. Petry (2014). "Safety profile of recombinant adeno-associated viral vectors: focus on alipogene tiparvovec (Glybera(R))." Expert Rev Clin Pharmacol **7**(1): 53-65.

Samulski, R. J., K. I. Berns, M. Tan and N. Muzyczka (1982). "Cloning of adeno-associated virus into pBR322: rescue of intact virus from the recombinant plasmid in human cells." Proc Natl Acad Sci U S A **79**(6): 2077-2081.

San Filippo, J., P. Chi, M. G. Sehorn, J. Etchin, L. Krejci and P. Sung (2006). "Recombination mediator and Rad51 targeting activities of a human BRCA2 polypeptide." J Biol Chem **281**(17): 11649-11657.

Santamarina-Fojo, S., C. Haudenschild and M. Amar (1998). "The role of hepatic lipase in lipoprotein metabolism and atherosclerosis." Curr Opin Lipidol **9**(3): 211-219.

Sapranauskas, R., G. Gasiunas, C. Fremaux, R. Barrangou, P. Horvath and V. Siksnys (2011). "The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli." Nucleic Acids Res **39**(21): 9275-9282.

Schlehofer, J. R., M. Ehrbar and H. zur Hausen (1986). "Vaccinia virus, herpes simplex virus, and carcinogens induce DNA amplification in a human cell line and support replication of a helpervirus dependent parvovirus." Virology **152**(1): 110-117.

Schumann, K., S. Lin, E. Boyer, D. R. Simeonov, M. Subramaniam, R. E. Gate, G. E. Haliburton, C. J. Ye, J. A. Bluestone, J. A. Doudna and A. Marson (2015).

"Generation of knock-in primary human T cells using Cas9 ribonucleoproteins." Proc Natl Acad Sci U S A **112**(33): 10437-10442.

Semenova, E., M. M. Jore, K. A. Datsenko, A. Semenova, E. R. Westra, B. Wanner, J. van der Oost, S. J. Brouns and K. Severinov (2011). "Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence." Proc Natl Acad Sci U S A **108**(25): 10098-10103.

Senapathy, P., J. D. Tratschin and B. J. Carter (1984). "Replication of adeno-associated virus DNA. Complementation of naturally occurring rep- mutants by a wild-type genome or an ori- mutant and correction of terminal palindrome deletions." J Mol Biol **179**(1): 1-20.

Shen, S., K. D. Bryant, S. M. Brown, S. H. Randell and A. Asokan (2011). "Terminal N-linked galactose is the primary receptor for adeno-associated virus 9." J Biol Chem **286**(15): 13532-13540.

Son, M. Y. and P. Hasty (2018). "Homologous recombination defects and how they affect replication fork maintenance." AIMS Genet **5**(4): 192-211.

Sonntag, F., K. Schmidt and J. A. Kleinschmidt (2010). "A viral assembly factor promotes AAV2 capsid formation in the nucleolus." Proc Natl Acad Sci U S A **107**(22): 10220-10225.

Srivastava, A., E. W. Lusby and K. I. Berns (1983). "Nucleotide sequence and organization of the adeno-associated virus 2 genome." J Virol **45**(2): 555-564.

Stenvang, J. and S. Kauppinen (2008). "MicroRNAs as targets for antisense-based therapeutics." Expert Opin Biol Ther **8**(1): 59-81.

Sun, W., W. Ji, J. M. Hall, Q. Hu, C. Wang, C. L. Beisel and Z. Gu (2015). "Self-assembled DNA nanoclews for the efficient delivery of CRISPR-Cas9 for genome editing." Angew Chem Int Ed Engl **54**(41): 12029-12033.

Symington, L. S. (2002). "Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair." Microbiol Mol Biol Rev **66**(4): 630-670, table of contents.

Syrris, P., D. Ward, A. Evans, A. Asimaki, E. Gandjbakhch, S. Sen-Chowdhry and W. J. McKenna (2006). "Arrhythmogenic right ventricular dysplasia/cardiomyopathy associated with mutations in the desmosomal gene desmocollin-2." Am J Hum Genet **79**(5): 978-984.

Szostak, J. W., T. L. Orr-Weaver, R. J. Rothstein and F. W. Stahl (1983). "The double-strand-break repair model for recombination." Cell **33**(1): 25-35.

Taylor, E. M., S. M. Cecillon, A. Bonis, J. R. Chapman, L. F. Povirk and H. D. Lindsay (2010). "The Mre11/Rad50/Nbs1 complex functions in resection-based DNA end joining in Xenopus laevis." Nucleic Acids Res **38**(2): 441-454.

Thomas, K. R. and M. R. Capecchi (1987). "Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells." Cell **51**(3): 503-512.

Thomson, D. W., C. P. Bracken and G. J. Goodall (2011). "Experimental strategies for microRNA target identification." Nucleic Acids Res **39**(16): 6845-6853.

Tian, Y., Y. Liu, T. Wang, N. Zhou, J. Kong, L. Chen, M. Snitow, M. Morley, D. Li, N. Petrenko, S. Zhou, M. Lu, E. Gao, W. J. Koch, K. M. Stewart and E. E. Morrisey (2015). "A microRNA-Hippo pathway that promotes cardiomyocyte proliferation and cardiac regeneration in mice." Sci Transl Med **7**(279): 279ra238.

Tichy, E. D., R. Pillai, L. Deng, L. Liang, J. Tischfield, S. J. Schwemberger, G. F. Babcock and P. J. Stambrook (2010). "Mouse embryonic stem cells, but not somatic cells, predominantly use homologous recombination to repair double-strand DNA breaks." Stem Cells Dev **19**(11): 1699-1711.

Tran, N. D., X. Liu, Z. Yan, D. Abbote, Q. Jiang, E. B. Kmiec, C. D. Sigmund and J. F. Engelhardt (2003). "Efficiency of chimeraplast gene targeting by direct nuclear injection using a GFP recovery assay." Mol Ther **7**(2): 248-253.

Tratschin, J. D., I. L. Miller and B. J. Carter (1984). "Genetic analysis of adeno-associated virus: properties of deletion mutants constructed in vitro and evidence for an adeno-associated virus replication function." J Virol **51**(3): 611-619.

Tratschin, J. D., M. H. West, T. Sandbank and B. J. Carter (1984). "A human parvovirus, adeno-associated virus, as a eucaryotic vector: transient expression and encapsidation of the procaryotic gene for chloramphenicol acetyltransferase." Mol Cell Biol **4**(10): 2072-2081.

Trempe, J. P. and B. J. Carter (1988). "Alternate mRNA splicing is required for synthesis of adeno-associated virus VP1 capsid protein." J Virol **62**(9): 3356-3363.

Trempe, J. P. and B. J. Carter (1988). "Regulation of adeno-associated virus gene expression in 293 cells: control of mRNA abundance and translation." J Virol **62**(1): 68-74.

van Spaendonck-Zwarts, K. Y., A. Posafalvi, M. P. van den Berg, D. Hilfiker-Kleiner, I. A. Bollen, K. Sliwa, M. Alders, R. Almomani, I. M. van Langen, P. van der Meer, R. J. Sinke, J. van der Velden, D. J. Van Veldhuisen, J. P. van Tintelen and J. D. Jongbloed (2014). "Titin gene mutations are common in families with both peripartum cardiomyopathy and dilated cardiomyopathy." Eur Heart J **35**(32): 2165-2173.

Vatta, M., B. Mohapatra, S. Jimenez, X. Sanchez, G. Faulkner, Z. Perles, G. Sinagra, J. H. Lin, T. M. Vu, Q. Zhou, K. R. Bowles, A. Di Lenarda, L. Schimmenti, M. Fox, M. A. Chrisco, R. T. Murphy, W. McKenna, P. Elliott, N. E. Bowles, J. Chen, G. Valle and J. A. Towbin (2003). "Mutations in Cypher/ZASP in patients with dilated

cardiomyopathy and left ventricular non-compaction." J Am Coll Cardiol **42**(11): 2014-2027.

Velu, C. S. and H. L. Grimes (2012). "Utilizing antagomiR (antisense microRNA) to knock down microRNA in murine bone marrow cells." Methods Mol Biol **928**: 185-195.

Vignier, N., S. Schlossarek, B. Fraysse, G. Mearini, E. Kramer, H. Pointu, N. Mougenot, J. Guiard, R. Reimer, H. Hohenberg, K. Schwartz, M. Vernet, T. Eschenhagen and L. Carrier (2009). "Nonsense-mediated mRNA decay and ubiquitin-proteasome system regulate cardiac myosin-binding protein C mutant levels in cardiomyopathic mice." Circ Res **105**(3): 239-248.

Wang, H., H. Yang, C. S. Shivalila, M. M. Dawlaty, A. W. Cheng, F. Zhang and R. Jaenisch (2013). "One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering." Cell **153**(4): 910-918.

Wang, Y., D. Cortez, P. Yazdi, N. Neff, S. J. Elledge and J. Qin (2000). "BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures." Genes Dev **14**(8): 927-939.

Wang, Z., L. Lisowski, M. J. Finegold, H. Nakai, M. A. Kay and M. Grompe (2012). "AAV vectors containing rDNA homology display increased chromosomal integration and transgene persistence." Mol Ther **20**(10): 1902-1911.

Ward, H. A., T. Norat, K. Overvad, C. C. Dahm, H. B. Bueno-de-Mesquita, M. Jenab, V. Fedirko, F. J. van Duijnhoven, G. Skeie, D. Romaguera-Bosch, A. Tjonneland, A. Olsen, F. Carbonnel, A. Affret, M. C. Boutron-Ruault, V. Katzke, T. Kuhn, K. Aleksandrova, H. Boeing, A. Trichopoulou, P. Lagiou, C. Bamia, D. Palli, S. Sieri, R. Tumino, A. Naccarati, A. Mattiello, P. H. Peeters, E. Weiderpass, L. A. Asli, P. Jakszyn, J. Ramon Quiros, M. J. Sanchez, M. Dorronsoro, J. M. Huerta, A. Barricarte, K. Jirstrom, U. Ericson, I. Johansson, B. Gylling, K. E. Bradbury, K. T. Khaw, N. J. Wareham, M. Stepien, H. Freisling, N. Murphy, A. J. Cross and E. Riboli (2016). "Pre-diagnostic meat and fibre intakes in relation to colorectal cancer survival in the European Prospective Investigation into Cancer and Nutrition." Br J Nutr **116**(2): 316-325.

Ware, J. S., J. G. Seidman and Z. Arany (2016). "Shared Genetic Predisposition in Peripartum and Dilated Cardiomyopathies." N Engl J Med **374**(26): 2601-2602.

Westra, E. R., P. B. van Erp, T. Kunne, S. P. Wong, R. H. Staals, C. L. Seegers, S. Bollen, M. M. Jore, E. Semenova, K. Severinov, W. M. de Vos, R. T. Dame, R. de Vries, S. J. Brouns and J. van der Oost (2012). "CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3." Mol Cell **46**(5): 595-605.

White, J. and S. Dalton (2005). "Cell cycle control of embryonic stem cells." Stem Cell Rev **1**(2): 131-138.

Wiedenheft, B., G. C. Lander, K. Zhou, M. M. Jore, S. J. Brouns, J. van der Oost, J. A. Doudna and E. Nogales (2011). "Structures of the RNA-guided surveillance complex from a bacterial immune system." Nature **477**(7365): 486-489.

Witzgall, R., E. O'Leary, A. Leaf, D. Onaldi and J. V. Bonventre (1994). "The Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression." Proc Natl Acad Sci U S A **91**(10): 4514-4518.

Worman, H. J. (2018). "Cell signaling abnormalities in cardiomyopathy caused by lamin A/C gene mutations." Biochem Soc Trans **46**(1): 37-42.

Wu, Z., A. Asokan and R. J. Samulski (2006). "Adeno-associated virus serotypes: vector toolkit for human gene therapy." Mol Ther **14**(3): 316-327.

Wu, Z., H. Yang and P. Colosi (2010). "Effect of genome size on AAV vector packaging." Mol Ther **18**(1): 80-86.

Xiao, P. J. and R. J. Samulski (2012). "Cytoplasmic trafficking, endosomal escape, and perinuclear accumulation of adeno-associated virus type 2 particles are facilitated by microtubule network." J Virol **86**(19): 10462-10473.

Xiao, W., K. H. Warrington, Jr., P. Hearing, J. Hughes and N. Muzyczka (2002). "Adenovirus-facilitated nuclear translocation of adeno-associated virus type 2." J Virol **76**(22): 11505-11517.

Xie, C., Y. P. Zhang, L. Song, J. Luo, W. Qi, J. Hu, D. Lu, Z. Yang, J. Zhang, J. Xiao, B. Zhou, J. L. Du, N. Jing, Y. Liu, Y. Wang, B. L. Li, B. L. Song and Y. Yan (2016). "Genome editing with CRISPR/Cas9 in postnatal mice corrects PRKAG2 cardiac syndrome." Cell Res **26**(10): 1099-1111.

Xu, F., J. Yang, J. Shang, F. Lan, M. Li, L. Shi, L. Shen, Y. Wang and J. Ge (2019). "MicroRNA-302d promotes the proliferation of human pluripotent stem cell-derived cardiomyocytes by inhibiting LATS2 in the Hippo pathway." Clin Sci (Lond) **133**(13): 1387-1399.

Yacoub, M. H. (2014). "Decade in review--cardiomyopathies: Cardiomyopathy on the move." Nat Rev Cardiol **11**(11): 628-629.

Yan, H., T. Toczylowski, J. McCane, C. Chen and S. Liao (2011). "Replication protein A promotes 5'-->3' end processing during homology-dependent DNA double-strand break repair." J Cell Biol **192**(2): 251-261.

Yan, X., X. Chen, H. Liang, T. Deng, W. Chen, S. Zhang, M. Liu, X. Gao, Y. Liu, C. Zhao, X. Wang, N. Wang, J. Li, R. Liu, K. Zen, C. Y. Zhang, B. Liu and Y. Ba (2014). "miR-143 and miR-145 synergistically regulate ERBB3 to suppress cell proliferation and invasion in breast cancer." Mol Cancer **13**: 220.

Yannone, S. M., I. S. Khan, R. Z. Zhou, T. Zhou, K. Valerie and L. F. Povirk (2008). "Coordinate 5' and 3' endonucleolytic trimming of terminally blocked blunt DNA

double-strand break ends by Artemis nuclease and DNA-dependent protein kinase." Nucleic Acids Res **36**(10): 3354-3365.

Yin, H., C. Q. Song, J. R. Dorkin, L. J. Zhu, Y. Li, Q. Wu, A. Park, J. Yang, S. Suresh, A. Bizhanova, A. Gupta, M. F. Bolukbasi, S. Walsh, R. L. Bogorad, G. Gao, Z. Weng, Y. Dong, V. Koteliansky, S. A. Wolfe, R. Langer, W. Xue and D. G. Anderson (2016). "Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo." Nat Biotechnol **34**(3): 328-333.

Yin, H., W. Xue, S. Chen, R. L. Bogorad, E. Benedetti, M. Grompe, V. Koteliansky, P. A. Sharp, T. Jacks and D. G. Anderson (2014). "Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype." Nat Biotechnol **32**(6): 551-553.

Yosef, I., M. G. Goren and U. Qimron (2012). "Proteins and DNA elements essential for the CRISPR adaptation process in Escherichia coli." Nucleic Acids Res **40**(12): 5569-5576.

Young, S. M., Jr., D. M. McCarty, N. Degtyareva and R. J. Samulski (2000). "Roles of adeno-associated virus Rep protein and human chromosome 19 in site-specific recombination." J Virol **74**(9): 3953-3966.

Yu, W., S. Mookherjee, V. Chaitankar, S. Hiriyanna, J. W. Kim, M. Brooks, Y. Ataeijannati, X. Sun, L. Dong, T. Li, A. Swaroop and Z. Wu (2017). "Nrl knockdown by AAV-delivered CRISPR/Cas9 prevents retinal degeneration in mice." Nat Commun **8**: 14716.

Zalatan, J. G., M. E. Lee, R. Almeida, L. A. Gilbert, E. H. Whitehead, M. La Russa, J. C. Tsai, J. S. Weissman, J. E. Dueber, L. S. Qi and W. A. Lim (2015). "Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds." Cell **160**(1-2): 339-350.

Zetsche, B., J. S. Gootenberg, O. O. Abudayyeh, I. M. Slaymaker, K. S. Makarova, P. Essletzbichler, S. E. Volz, J. Joung, J. van der Oost, A. Regev, E. V. Koonin and F. Zhang (2015). "Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system." Cell **163**(3): 759-771.

Zha, S., C. Boboila and F. W. Alt (2009). "Mre11: roles in DNA repair beyond homologous recombination." Nat Struct Mol Biol **16**(8): 798-800.

Zhang, J., Z. Ma, A. Treszezamsky and S. N. Powell (2005). "MDC1 interacts with Rad51 and facilitates homologous recombination." Nat Struct Mol Biol **12**(10): 902-909.

Zhang, J., Q. Sun, Z. Zhang, S. Ge, Z. G. Han and W. T. Chen (2013). "Loss of microRNA-143/145 disturbs cellular growth and apoptosis of human epithelial cancers by impairing the MDM2-p53 feedback loop." Oncogene **32**(1): 61-69.

Zhang, Y., C. Long, H. Li, J. R. McAnally, K. K. Baskin, J. M. Shelton, R. Bassel-Duby and E. N. Olson (2017). "CRISPR-Cpf1 correction of muscular dystrophy mutations in human cardiomyocytes and mice." Sci Adv **3**(4): e1602814.

Zhong, L., B. Li, C. S. Mah, L. Govindasamy, M. Agbandje-McKenna, M. Cooper, R. W. Herzog, I. Zolotukhin, K. H. Warrington, Jr., K. A. Weigel-Van Aken, J. A. Hobbs, S. Zolotukhin, N. Muzyczka and A. Srivastava (2008). "Next generation of adeno-associated virus 2 vectors: point mutations in tyrosines lead to high-efficiency transduction at lower doses." Proc Natl Acad Sci U S A **105**(22): 7827-7832.

Zhu, J., X. Huang and Y. Yang (2009). "The TLR9-MyD88 pathway is critical for adaptive immune responses to adeno-associated virus gene therapy vectors in mice." J Clin Invest **119**(8): 2388-2398.

Zolotukhin, S., B. J. Byrne, E. Mason, I. Zolotukhin, M. Potter, K. Chesnut, C. Summerford, R. J. Samulski and N. Muzyczka (1999). "Recombinant adeno-associated virus purification using novel methods improves infectious titer and yield." Gene Ther **6**(6): 973-985.