UNIVERSITÀ
DEGLI STUDI
FIRENZE

DOTTORATO DI RICERCA IN
"SCIENZE BIOMEDICHE"

CICLO XXVII

COORDINATORE Prof. Persio Dello Sbarba

Transcription Activator Like-Effectors for dissection of the antibody diversification processes

Settore Scientifico Disciplinare MED/04

Dottorando
Dott. Riccardo Pecori

Tutore
Prof. Persio Dello Sbarba

Supervisor
Dott. Silvestro Conticello

Anni 2012/2014

UNIVERSITÀ
DEGLI STUDI
FIRENZE

DIPARTIMENTO DI
SCIENZE BIOMEDICHE
SPERIMENTALI E CLINICHE

Dottorato di Ricerca in
SCIENZE BIOMEDICHE
sede amministrativa: Dipartimento di
Scienze Biomediche Sperimentali e Cliniche
coordinatore: Prof. Persio Dello Sbarba

Dottorato di Ricerca in Scienze Biomediche

Presentazione del candidato: Riccardo Pecori
Curriculum: Oncologia Sperimentale e Clinica
Ciclo: XXVII
Titolo della tesi: Transcription Activator Like-Effectors for dissection of the antibody diversification processes

A conclusione del corso triennale del XXVII Ciclo del Dottorato di Ricerca in Scienze Biomediche (*curriculum* Oncologia Sperimentale e Clinica), il Collegio dei Docenti, facendo propria la relazione presentata dal Prof: Persio Dello Sbarba in qualità di *tutor*, circa l'attività di ricerca, l'operosità e l'assiduità del candidato, rilascia con parere unanime il seguente giudizio da presentare alla Commissione Giudicatrice ai fini dell'espletamento dell'esame finale.

Il Dott. Riccardo Pecori, nato a Bagno a Ripoli (FI) il 25/08/1986, laureato in Biologia Cellulare e Molecolare il 21/04/2011, discutendo una tesi dal titolo "Preparazione di costrutti plasmidici per la visualizzazione in tempo reale del Class Switch Recombination" con la votazione di 110/110 e lode, è stato ammesso, a partire dal 01/01/2012, al Dottorato di Ricerca in Scienze Biomediche *curriculum* Oncologia Sperimentale e Clinica (XXVII Ciclo), svolgendo la propria attività di ricerca presso l'Istituto Toscano Tumori, sotto il tutoraggio del Prof. Persio Dello Sbarba e la supervisione del Dott. Conticello.

Descrizione dell'attività di ricerca/Risultati ottenuti:

Activation Induced Deaminase (AID) è una deaminasi essenziale nella diversificazione secondaria degli anticorpi, svolgendo un ruolo primario nei processi di Somatic Hypermutation (SHM) e di Class Switch Recombination (CSR).
La capacità di AID di mutare il DNA lo rende anche, un elemento importante nella genesi di tumori nelle cellule B. Lo scopo del progetto di ricerca del candidato è stato lo sviluppo di un sistema che permetta di visualizzare il CSR in tempo reale nella linea cellulare CH12-F3, derivante da linfoma murino. L'approccio utilizzato è stato quello di inserire nel locus della catena pesante delle immunoglobuline mediante gene targeting un array attraverso cui seguire la progressione del processo di CSR. La potenzialità dell'approccio potrà permettere di superare i limiti delle tecniche esistenti e permettere uno studio dettagliato del coinvolgimento di AID e altri importanti cofattori.
Inoltre il Dott. Pecori ha lavorato per sfruttare la capacità mutagenica di AID e poter aumentare la frequenza di mutazione su sequenze specifiche. A tal fine ha creato una serie di chimere in cui AID è stato fuso con dei TALEs (Transcription Activator Like-Effectors) in modo da indirizzare la sua azione su geni di interesse,. Questo sistema potrà essere utilizzato per la creazione di modelli murini per studiare l'effetto di mutazioni su geni correlati allo sviluppo del cancro - protooncogeni e oncosoppressori, e studiarne gli effetti.

UNIVERSITÀ
DEGLI STUDI
FIRENZE

DIPARTIMENTO DI
SCIENZE BIOMEDICHE
SPERIMENTALI E CLINICHE

Pubblicazioni/Presentazione dati a congressi nazionali e internazionali:

2014 (24-27 September) - FISV XIII Congress, Pisa, Italy. Poster presentation "TALEs for every season: targeted DNA damage and live visualisation of DNA".

Seminari:

Dottorato anno I – Giugno 2012: "Visualization of Class Switch Recombination in live cells"

Dottorato anno II – Settembre 2013: "Visualization of Class Switch Recombination in live cells"

Dottorato anno III – Dicembre 2014: "TALEs for every season: targeted DNA damage and live visualisation of DNA"

Durante il corso di dottorato, il candidato ha seguito con il massimo impegno il programma didattico stabilito dal Collegio dei Docenti ed ha portato avanti con entusiasmo e determinazione le sue ricerche, dando prova di grande inventiva ed intraprendenza, nonché di notevole elasticità nella elaborazione dei dati sperimentali. Nel corso del triennio, il candidato ha inoltre maturato una buona cultura di base ed una vasta esperienza diretta in metodiche sperimentali.

Per quanto sopra, il Collegio dei Docenti unanime ritiene che il Dr. Riccardo Pecori possa meritatamente aspirare a conseguire il titolo di Dottore di Ricerca.

Firenze, 11/02/2015

Il Coordinatore del Corso

Prof. Persio Dello Sbarba

Table of abbreviations

A: Adenine

aa: Amino acid

Ab: Antibody

AD: Activation Domain

A-EJ: Alternative End Joining

AID: Activation Induced Deaminase

Ala or A: Alanine

Amp: Ampicillin

APE: Apurinic/apyrimidinic Endonuclease

APOBEC-1: APOlipoprotein B mRNA Editing enzyme Catalytic polypeptide 1

ARCUT: Artificial Restriction DNA Cutters

Arg or R: Arginine

Asn or N: Asparagine

Asp or D: Aspartic acid

BER: Base Excision Repair

bp: Base pair

BCR: B Cell Receptor

bsr: Blasticidin-S resistance gene

C: Cytosine

C-NHEJ: Classical Non-Homologous End Joining

C region: Constant region

C-terminal: Carboxyl-terminal

cDNA: Complementary DNA

CDR: Complementarity Determining Region

cds: coding sequence

C_H: Constant gene of Heavy chain

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeat

CRISPRi: Clustered Regularly Interspaced Short Palindromic Repeat Interference

crRNA: CRISPR RNA

CSR: Class Switch Recombination

CTS: C-Terminal Segment

Cys or C: Cysteine

dC: deoxycytidine

DMSO: Dimethylsulphoxide

DNA: Deoxyribonucleic Acid

dNTP: DeoxyNucleoside TriPhosphate

DSB: Double-Strand Break

DTA: Diphtheria Toxin A

E.coli: *Escherichia coli*

EGFP: Enhanced Green Fluorescent Protein

ES: Embryonic Stem

FACS: Fluorescence-Activated Cell Sorting

FBS: Fetal Bovine Serum

FIAU: Fialuridine

FLASH: Fast Ligation-based Automatable Solid-phage High-throughput

Flp: Flippase

FRAP: Fluorescence Recovery After Photobleaching

FRT: Flippase Recognition Target

G: Guanine

GFP: Green Fluorescent Protein

Glu or E: Glutamic Acid

Gly or G: Glycine

gRNA: Guide RNA

His or H: Histidine

HIGM2: Hyper-IgM Syndrome

HIV: Human Immunodeficiency Virus

HMA: Heteroduplex Mobility Assay

HR: Homologous Recombination

HRMA: High-Resolution Melting Analysis

HSP90: Heat Shock Protein 90

HSV-tk: Herpes Simplex Virus-thymidine kinase

ICA: Iterative Capped Assembly

iEμ: Intronic Enhancer

IFNγ: Interferon-γ

Ig: Immunoglobulin

IgA: Immunoglobulin A

IgC: Immunoglobulin Constant region

IgD: Immunoglobulin D

IgH: Immunoglobulin Heavy chain

IgL: Immunoglobulin Light chain

IgM: Immunoglobulin M

IgV: Immunoglobulin Variable region

Ile or I: Isoleucine

IL-4: Interleukin-4

iPS: Induced Pluripotent Stem cell

IPTG: Isopropyl-beta-D-Thiogalactopyranoside

IRES: Internal Ribosome Entry Site

Kana: Kanamycin

KRAB domain: Krueppel-Associated Box domain

L: Leucine

LIC: Ligation Independent Cloning

LSD1: Lysine-Specific Demethylase 1

Lys or K: Lysine

MMR: Mismatch Repair

MODC: Mouse Ornithine Decarboxylase degradation domain

MSH2: MutS Homolog 2

MSH6: MutS Homolog 6

mRNA: messenger RNA

N-terminal: Amino-terminal

neo^r: Neomycin resistance gene

NES: Nuclear Export Signal

NHEJ: Non-Homologous End-Joining

NF-kB: Nuclear Factor kB

NLS: Nuclear Localization Signal

NmCas9: *Neisseria meningitidis* Cas9

nt: Nucleotide

NTS: N-Terminal Segment

ORF: Open Reading Frame

PAM: Protospacer-Adjacent Motif

Pax5: Paired box 5

PKA: Protein Kinase A

PKA-Cα: Protein Kinase A Catalytic subunit-α

PTBP2: Polypyrimidine Tract-Binding Protein 2

Puro: Puromycin

R: Purine bases (A or G)

RAG: Recombination Activating Gene

RFLP: Restriction Fragment Length Polymorphism

RNA: Ribonucleic Acid

RNAi: RNA interference

RPA: Replication Protein A

RT-PCR: Reverse Transcription PCR

RVD: Repeat-Variable Di-residue

SHM: Somatic Hypermutation

scFokI: Single-chain FokI

SCID: Severe Combined ImmunoDeficiency

SpCas9: *Streptococcus pyogenes* Cas9

S regions: Switch regions

SSB: Single-Strand Break

ssDNA: Single-stranded DNA

T: Thymine

TB: Transformation Buffer

TGF-β: Transforming Growth Factor-β

TALE: Transcription Activator-Like Effector

TALEN: Transcription Activator-Like Effector Nuclease

TALER: TALE Recombinase

tracrRNA: Trans-activating crRNA

Trp or W: Tryptophan

U: Uracil

UNG: Uracil-N-Glycosylase

UTR: Untranslated Region

VDJ: Variable Diversity Joining

W: Weak bases (A or T)

WT: Wild-Type

XRCC4: X-ray Repair Cross-Complementing protein 4

Y: Pyrimidine bases (T or C)

YC: YFP C terminal part

YFP: Yellow Fluorescent Protein

YN: YFP N terminal part

ZF: Zinc Finger

ZFN: Zinc Finger Nuclease

Zn^{2+}: Zinc

3'-RR: 3'-Regulatory region

5mC: 5-methyl Cytosine

Introduction

Genome editing

Introduction and overview

GENE FUNCTION AND PHENOTYPE

Since a long time, the classical approach in modern genetics to study gene function and correlate it with phenotype. is termed “reverse genetics”, and allows us to understand the intricate relationship between genotype and phenotype through manipulation of the genetic information. These manipulations involve various biological approaches: knock-in, knockout, and gene expression modification (RNA interference, RNAi). For the past few decades, functional genes have been integrated into genomes and over-expressed through transposon-mediated modification, similar to T-DNA and p-elements (Chen et al., 2014). Moreover, gene targeting was performed by site-specific recombinase technology using the Cre/loxP system (Kilby et al., 1993), the Flp/FRT system (Dimecky, 1996), and the phiC31 integrase (Groth et al., 2004). RNAi was developed as a powerful tool to knock-down genes (Gonczy et al., 2000; Dietzel et al., 2007; Martin and Caplen, 2007).
Targeted gene knockdown by RNAi is a rapid, inexpensive and highly-efficient method compared to gene targeting (McManus et al., 2002). However, this approach presents disadvantages due to its time constrained-effect in knocked-down cells, to variability and reproducibility issues, to possible unpredictable off-target effects, and to a high background noise (Martin and Caplen, 2007).

Altering gene expression by transposons or stable transfection can generally be used to study overexpression of a given gene. Nonetheless, expression levels can be affected by the random insertion of the gene.

All in all, gene targeting has proven to be the most reliable tool to explore gene function. However this method has certain hindrances such as the low efficiency of targeting at the target site, the need for time-consuming and labor-intensive selection/screening strategies. Hence, traditional gene-targeting was only applied in few model systems, such as *Drosophila* (Venken and Bellen, 2007) and mouse (Capecchi, 2005), characterized by short generation times and easy inbreeding.

GENOME EDITING WITH SITE-SPECIFIC NUCLEASES

In recent years, a new approach has emerged to theoretically manipulate any desired gene in several different cell types and organisms. This technology is commonly called as genome editing and is based on three different systems: Zinc-Finger Nuclease (ZFN), Transcription Activator-Like Effector Nuclease (TALEN) and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR/Cas9 nuclease). While conceptually identical to gene targeting, these novel tools, due to the high targeting efficiency, makes it possible for scientists to modify the genome in an easy, efficient and inexpensive manner.

These systems can be designed and put into effect within few days using designing tools, traditional cloning or commercial kits, or commercial services (listed in Appendix C) (Gonzalez et al., 2010; Kim et al., 2011b; Li et al., 2011b; Miller et al., 2011; Reyon et al., 2012a,b; Ran et al., 2013).

ZFNs and TALENs are composed of sequence-specific DNA-binding domain fused with an unspecific endonuclease (Urnov et al., 2010; Carroll, 2011; Miller et al., 2011) while Cas9 is an RNA-guided DNA endonucleases. These proteins induce DNA double-strand breaks (DSBs) in the region of interest and hence stimulate the DNA repair mechanisms (Wyman and Kanaar, 2006). Error-prone non homologous end joining (NHEJ) can lead to gene inactivation, while homology recombination (HR) in presence of donor homologous DNA can lead to knock-in (FIG. 1).

Since their discovery, these proteins have been modified to increase their activity and specificity. For example, an highly-active FokI mutant called “Sharkey” was developed by directing evolution (Guo et al., 2010) and used in association with

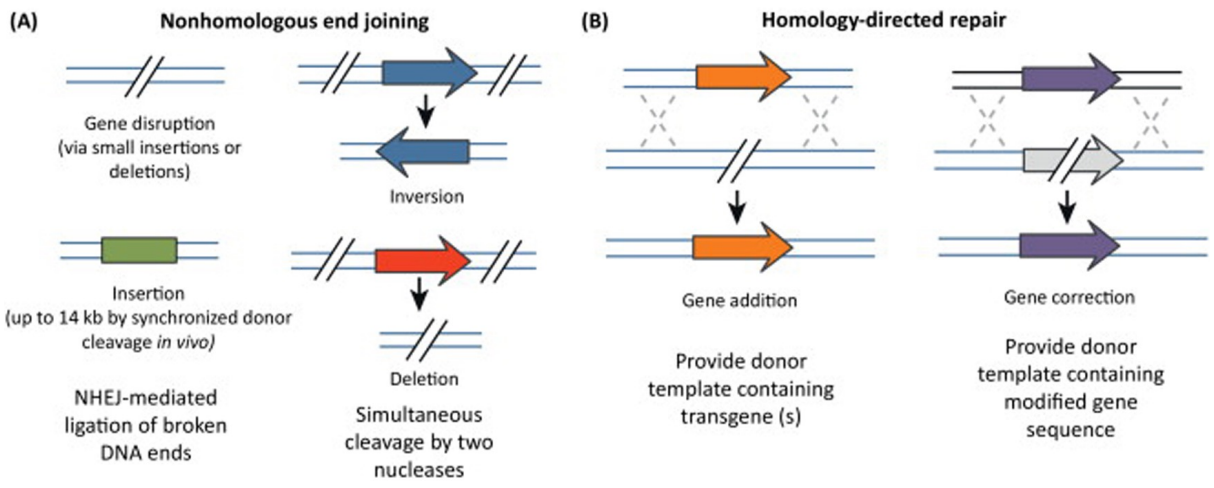

Figure 1

Nuclease-induced DNA double-strand breaks (DSBs) can be repaired by error-prone non-homologous end joining (NHEJ) or homologous recombination (HR). (A) In the presence of donor plasmid with homology arms, HR can introduce single or multiple transgenes to correct or replace the wild type gene. (B) In absence of donor plasmid, NHEJ-mediated repair induces small insertions or deletions that cause gene disruption. Simultaneous induction of two DSBs can lead to deletions, inversions and translocations of the intervening segment (Reproduced from Gaj et al., 2013).

either ZFNs (Ramalingam et al., 2011) or TALEs (Tesson et al., 2011; Cade et al., 2012) to increase their effectivity.

The heterodimers of FokI were developed, where only one of the two FokI protein is active; this results in a decreased off-target activity and genotoxicity of ZFNs and TALENs protein (Miller et al., 2007; Szczepek et al., 2007; Cade et al., 2012). These proteins were further improved by introducing a mutation (Asp_{450} to Ala) at the active site of the FokI domain in one of the subunits in order to develop DNA-nicking enzymes (nickases) (Sanders et al., 2009). The heterodimer nickases induce a single nick on one strand thus decreasing the chances of DSBs. These site-specific nicks can increase the frequency of HR without activating NHEJ (Kim et al., 2012; Ramirez et al., 2012; Wang et al. 2012).

In addition, simultaneous introduction of these endonucleases can result in multiple targeted mutagenesis (Wang et al., 2013), chromosomal deletions (Lee et al., 2010), duplication (Lee et al., 2012), inversions (Lee et al., 2012) or translocations (Piganeau et al., 2013).

Finally, there are alternative strategies to ZFNs, TALENs and CRISPR/Cas: engineered meganucleases have been used, but these proteins are refractory to customization for sequence specificity (Grizot et al. 2009; Munoz et al. 2011; Menoret

et al. 2013); Artificial Restriction DNA Cutters (ARCUT), a DNA-based DSB-inducing molecule is an other possibility for the same (Komiyama, 2013).

DNA repair and genome editing

DOUBLE-STRAND BREAKS

DNA double-strand breaks (DSBs) are a type of lesion that cells need to repair in order to maintain their genomic integrity and stability. These type of lesions are generated spontaneously or after exposure to exogenous DNA damaging agents. When unrepaired or improperly repaired, DSBs can lead to cell death and, in multicellular organisms, may promote tumorigenesis and formation of cancer (Van Gent et al., 2001).

As mentioned above, the two different pathways to repair these damages, HR and NHEJ, the DNA ends are aligned and joined using homologous sequence and non-homologous end joining which acquire the broken ends, bring these extremities together and rejoin them in the absence of long tracts of the homologous sequence.

NON-HOMOLOGOUS END JOINING

Non-Homologous End Joining is a process present in all organisms, from bacteria, to archea and eukaryotes, indicating its importance. NHEJ was initially identified in mammalian cells, demonstrating its pivotal contribution to the repair system of the cell. Contrarily to other organisms, such as *Saccharomyces cerevisiae* (Van Den Bosch et al., 2002), this repair machinery is the prevalent one for of DSB repair.

In NHEJ, the broken ends bind to specific factors to limit DNA degradation which may lead to loss of genetic information. Then, the DNA ends are juxtaposed through protein-protein interactions in which the end-binding factors bound to DNA ends interact with the bridging proteins. Alternatively, the end-binding proteins may work as the landing site for the bridging factors (Pfeiffer and Vielmetter, 1988; Thode et al., 1990). Sometimes, the juxtaposed DNA ends may directly be ligated by DNA ligases but often the DSBs generated ends are poor substrates for the ligases. In such cases, the extremities must be processed before the ligation. Several studies on break site repair by NHEJ, suggest that alignment of DNA through microhomology,

short complementary sequences (1-4 nucleotides), is involved to closely bring together the broken ends (Roth and Wilson, 1986; Kramer et al., 1994; Moore and Haber, 1996).

The microhomology-dependent alignments causes mismatches and possible next flaps and/or gaps on the juxtaposed DNA. The resulting gaps and flaps seeds to small insertions or deletions (indels), this process has been exploited to obtain a knock-out of specific gene.

HOMOLOGOUS RECOMBINATION

Homologous Recombination requires long stretches of homologous DNA and repairs DSBs utilizing the information on the correct sister-chromatid. As mentioned, this reparation system is predominant in yeast, conversely HR is a rare event in higher eukaryotes, probably due to the shorter relative persistence of the chromatids.

Initial investigations in Capecchi's laboratory showed evidences of HR in mammalians. They microinjected many copies of Herpes simplex virus thymidine kinase (*HSV-tk*) in mammalian cells deficient in thymidine kinase (tk^-) to obtain tk^+ cells. The *HSV-tk* genes were integrated in multiple copies and, even though integrated randomly in one or two points of genome, they were present as head-to-tail ordered concatamers (FIG. 2). Additionally, they demonstrated that the

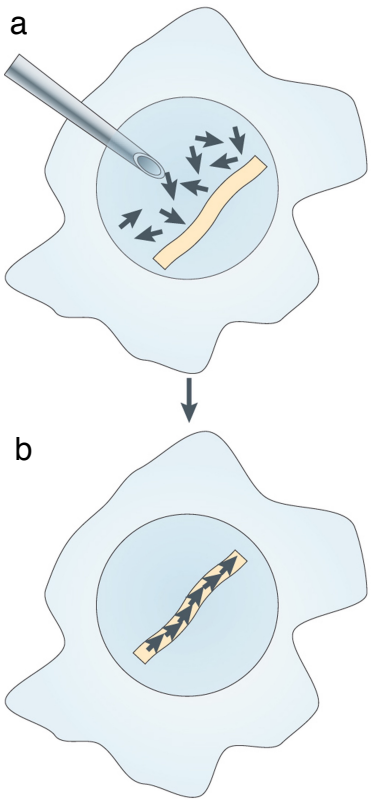

Figure 2

When multiple copies of a DNA sequence (arrows) are introduced into mammalian cells (a), they are efficiently integrated into one or a very few random site(s) within the host genome as concatamers (b).
The orientation of the sequences in the concatamer showed that the DNA copies are not randomly oriented within the concatamer, but are all oriented in the same direction (b) (Reproduced from Capecchi, 2005).

concatemers were produced by homologous recombination (Folger et al., 1982). This was the first demonstration that mammalian cells could mediate homologous recombination through the introduction of external plasmid DNA. This finding provides the basis for gene targeting and genome editing.

The efficiency of this machinery was evident, especially after the observation that microinjection of 100 molecules of *HSV-tk* plasmid per cell could cause the formation of concatemers composed from the 100 molecules. Capecchi wrote:

"*It was immediately clear to me that if we could harness this machinery to carry out homologous recombination between a newly introduced DNA molecule of our choice and the cognate sequence in the recipient cell, we could mutate or modify almost any gene in mammalian cells in any desired manner.*" (Capecchi, 2005).

The next logical step was to increase the low frequency of gene targeting in mammalian cells. To achieve this, Capecchi's laboratory developed a selection technique to eliminate cells which did not contain targeted homologous recombination products. They obtained a cell line with random insertion of a neomycin resistance gene (neo^r) containing an inactivating deletion. The cell line was then microinjected with a plasmid containing the neo^r gene inactivated by a single point mutation. HR between the targeting plasmid and the chromosomal region containing the defective neo^r gene could generate a functional neo^r gene, thus producing cells resistant to G418 (FIG. 3) (Thomas et al., 1986). Moreover, they exhibited that linear DNA molecules are preferred over circular DNA as substrates for HR (Thomas et al., 1986) and is cell-cycle dependent, with a peak of activity during early S phase (Wong and Capecchi, 1986).

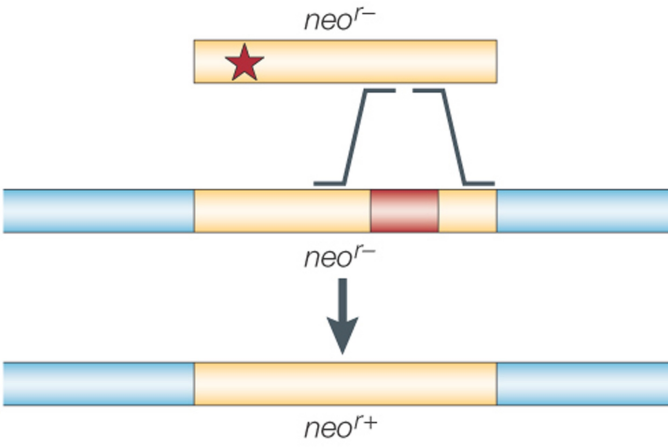

Figure 3

Generation of a functional neomycin resistance gene from two defective genes by gene targeting. The recipient mammalian cell line contains a defective neomycin resistance (neo^r) gene, with a small deletion (red block) at the 3' end of coding sequence, integrated randomly in the cellular genome. The targeting vector contains a 5′ point mutation (red star). The deletion in the integrated gene is corrected by homologous recombination through the exogenously added targeting vector with a frequency of approximately 1 in 1,000 cells (Reproduced from Capecchi, 2005).

Further examinations showed that

targeting efficiency was neither dependent on the concentration of the targeting vector nor on the number of targets present in the cell genome. Indeed, the introduction of either 100 molecules or a single molecule of plasmid per cell led to the same targeting frequency. This suggested that, the limiting step was the availability of the cellular machinery to direct recombination between the exogenous and the endogenous DNA. An important point from this work was that each chromosomal target position analyzed was equally accessible to HR, thus indicating that most of the genome, if not all, could be modified by gene targeting.

The low frequency of targeting proved the technique impractical for use directly in murine zygotes. The best option to obtain mice with specific mutations at a specific locus was to perform the gene targeting in cultured embryonic stem (ES) cells, and to select the targeted cells before introduction into a pre-implantation embryo. These cells would contribute to the formation of some tissues in the newborn mice, among them the germline.

Another improvement in the gene targeting technology came with the association of the targeting with a site-specific recombination system, such as Cre-loxP or Flp-FRT. This approach makes it possible to generate conditional mutants for evaluation of gene function in specific tissues and/or limited times (Gu et al., 1994).

In order to increase the efficiency of targeting and the accessibility to every laboratory, in 1986 Capecchi started using electroporation. Whereas less efficient compared to microinjection, electroporation could be applied to a higher number of cells (10^7), thus resulting in a higher number of targeted cells. Despite these advancements, the targeting ratio was still insufficient.

Capecchi's lab next strategized to enrich the targeting cells by a procedure called positive-negative selection. With such approach a positive selectable gene, such as neo^r, was used in conjunction with an element for negative selection such as *HSV-tk* located beyond the homologous arms, at the end of targeting vector (Mansour et al., 1988). Through positive selection in presence of G418, only cells that receive the targeting construct can grow regardless of the insertion, either through HR or NHEJ. The cells in which the targeting plasmid is inserted by NHEJ, would also express the *HSV-tk* gene, and thus could be negatively selected with FIAU (a drug that kills cells that express *HSV-tk* gene). The death of these cells would lead to the enrichment of targeted cells (FIG. 4).

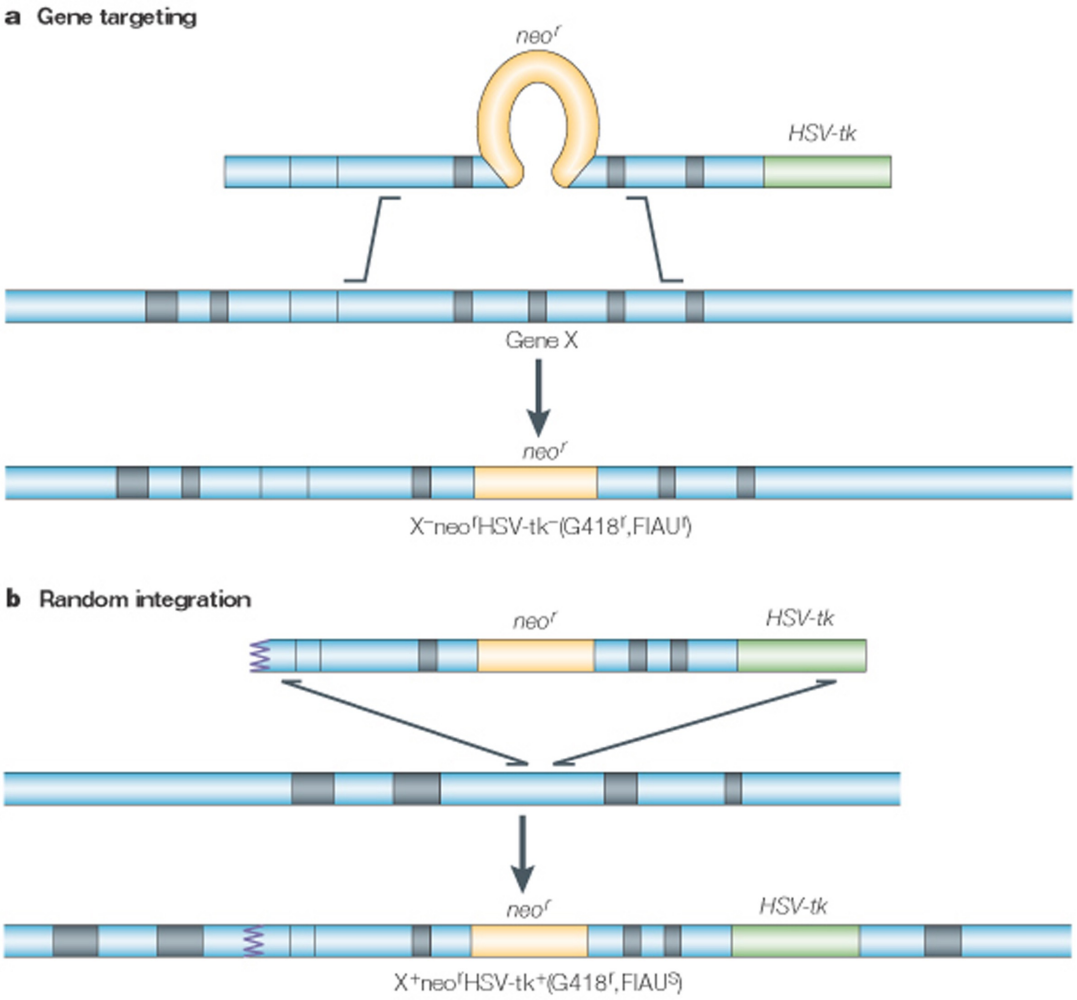

Figure 4

(a) The targeting vector contains a neomycin resistance (neo^r) expression cassette to be inserted in a gene X and an herpes virus thymidine kinase (*HSV-tk*) expression cassette at one end. Homologous recombination between the targeting vector and the chromosomal gene results in the disruption of one genomic copy of gene X and the loss of the vector *HSV-tk* gene. Cells in which this event has occurred will be X$^{+/-}$, neo^{r+} and *HSV-tk*$^-$, and will be resistant to both G418 and FIAU.
(b) Most frequently, the targeting vector will be integrated into the host cell genome at a random site, through non-homologous recombination. Because non-homologous insertion of exogenous DNA into the host cell chromosome occurs through the ends of the linearized targeting vector, the *HSV-tk* gene will remain linked to the neo^r gene. Cells derived from this type of recombination event will be X$^{+/+}$, neo^{r+} and *HSV-tk*$^+$, and therefore resistant to G418 but sensitive to FIAU (Reproduced from Capecchi, 2005).

Soon after, the diphtheria toxin A (DTA) has been utilized as marker to negative selection. This gene had several advantages, as drugs are not required to induce cell death and transient expression of toxin presented minimal toxicity. This approach exhibited an enrichment in targeted clones of 9- to 29-folds compared to positive selection alone (McCarrick et al., 1993).

Additionally, single-strand oligonucleotides (ssODNs) have been used, coupled with engineered endonucleases, as templates for HR-mediated DNA repair (Chen et al., 2011). Initial works with ssODNs had achieved corrections in reporter genes at frequencies above 0.7% (Wang et al., 2008; Radecke et al., 2010), but subsequently ssODNs have been used, achieving frequencies of correction ranging from 1-30% in absence of any antibiotic selection. Using this approach different types of gene targeting could be achieved: (1) targeted point mutation, (2) targeted deletion of small and large region of DNA and (3) simultaneous targeted deletion of large sequences and insertion of specific short DNA sequence (Chen et al., 2011).

Finally, the efficiency of targeting was shown to be improved using specific endonucleases, which stimulate the DNA repair machinery through the induction of targeted DSBs in eukaryotic cells (Rouet et al., 1994a,b; Bibikova et al., 2003; Porteus and Baltimore, 2003; Urnov et al., 2005; Moehle et al., 2007; Voytas, 2013).

Zinc-Finger Nucleases (ZFNs)

INTRODUCTION AND STRUCTURE

The Zinc-finger domain, originally discovered in *Xenopus* (Miller et al., 1985), is one of the most common DNA-binding motifs in eukaryotes and is amongst the most frequently encoded protein folds in humans. Each zinc-finger motif is composed of 30 amino acids in a conserved ββα configuration (Beerli and Barbas, 2002) (FIG. 5b). Amino acids on the α-helix bind 3 or 4bp in the major groove of DNA, with different levels of selectivity, while the side of conserved Cys and His residues form a complex with Zn^{2+} ion. The availability of the structure of a highly conserved linker sequence between Zinc-finger domains led to the possibility to fuse different zinc-finger domains together. Specific DNA-recognition protein were developed to recognize sequences which were 9-18bp in length (Liu et al., 1997) (FIG. 5a). The possibility to recognise specifically such length allowed binding to DNA targets in the genome, and represented the first method to increase targeting efficiency in the human genome (Beerli et al., 1998, 2000).

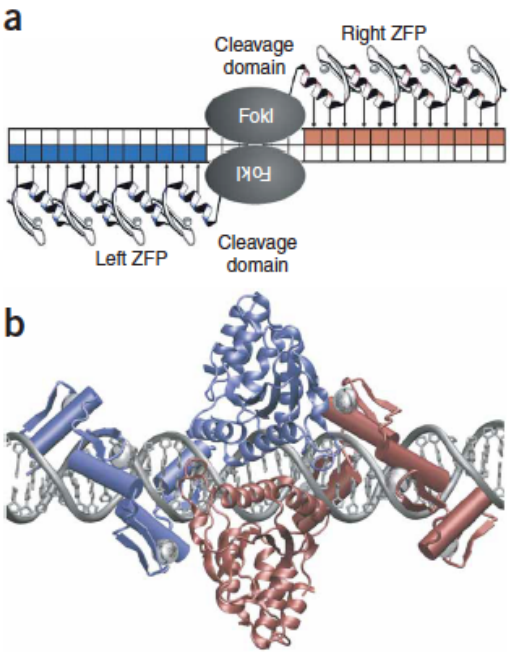

Figure 5

(a) Representation of zinc-finger nuclease (ZFN) dimer bound to DNA. ZFN target sites consist of two zinc-finger binding sites separated by a 5–7bp spacer sequence targeted by the FokI cleavage domain
(b) Model to the arrangement in (a). DNA is shown in gray, ZFNs are colored blue (left ZFN) or red (right ZFN), and reflective spheres denote zinc ions. Zinc-finger helices are represented by cylinders, whereas the central FokI are represented entirely as ribbon diagrams.
(Reproduced from Miller et al., 2007).

ASSEMBLY METHODS

Following these initial studies, several methods have been developed to construct zinc-finger proteins. The “modular assembly” used a preselected library of zinc-finger modules, to recognize nearly all of 64 possible nucleotide (nt) triplets, which were linked together to target a specific DNA sequence (Beerli et al., 1998, 2000; Beerli and Barbas, 2002; Gonzalez et al., 2010; Kim et al., 2011b; Bhakta et al., 2013).
An alternative possibility for the construction of ZFNs is a selection-based approach where an oligomerized pool engineering (OPEN) is used to select a new zinc-finger arrays from a random library, also considering the context-dependent interactions between close fingers (Maeder et al., 2008). Moreover, several approaches have been developed to combine the methods described above (Sander et al., 2011b; Gupta et al., 2012).

ADVANTAGES AND DISADVANTAGES

The main problem with ZFNs is their off-target activity. In order to reduce this effect and increase the specificity, several approaches have been developed such as FokI heterodimers (Miller et al., 2007; Szczepek et al., 2007) with optimized cleavage specificity and reduced toxicity, which can be used in combination with the “Sharkey” mutant (Guo et al., 2010) to further increase activity and specificity (Doyon et al., 2011). Moreover, evidences suggest that 4-6 zinc-finger domains for each half ZFN

enhance specificity and activity of protein (Guo et al., 2010; Perez-Pinera et al., 2012; Bhakta et al., 2013; Sood et al., 2013).

There are other methods to improve ZFN activity such as transient hypothermic culture conditions to increase nuclease expression levels (Doyon et al., 2010), co-transfection with DNA end-processing enzymes (Certo et al., 2012) and the use of fluorescent surrogate reporter vectors for enrichment of modified cells by ZFNs activity (Kim et al., 2011a).

In recent times, there have been advancements where zinc-finger nickases (ZFNickases) (Sanders et al., 2009; Kim et al., 2012; Ramirez et al., 2012; Wang et al., 2012a) were engineered to nick DNA, thus stimulating HR instead of NHEJ (McConnell et al., 2009). Consequently, this approach has fewer off-target activity than conventional ZFNs, however, the frequency of HR remains lower than traditional ZFNs.

DNA- and mRNA-based methods have been used to deliver ZFNs in the cells. These methods are closely dependent on cell types and exhibit undesirable effects; such as insertional mutagenesis, toxicity and low efficiency. Despite using nuclease pairs to increase its specificity, the transient co-delivery poses a technical limitation. Hence, monomer-type ZFNs have been established using single-chain FokI (scFokI), a tandem pair of FokI nuclease domain in one protein (Minczuk et al., 2008; Mino et al., 2009; Mori et al., 2009). These proteins have shown activity *in vitro* and in cultured cells, but there is still a need for additional analysis to get a clearer view on their efficiency. On the other hand, ZFs linked to I-TevI endonuclese can be used for genome editing (Kleinstiver et al., 2012).

To overcome the limitations, purified ZFN proteins were used directly onto the cells. This strategy exhibits insertional mutations sans any risks and has fewer off-target effects than conventional delivery systems (Gaj et al., 2012).

Transcription Activator-Like Effectors Nucleases (TALENs)

INTRODUCTION AND STRUCTURE

The Transcription Activator-Like Effectors (TALEs) are proteins found in plant pathogenic bacteria originating from the genus *Xanthomonas.* The native function of these proteins is to directly modulate host gene expression (Bogdanove and Voytas,

2011; Munoz Bodnar et al., 2013). Upon delivery into the host cells, TALEs enter the nucleus, through a Nuclear Localisation Signal (NLS), bind to effector-specific sequences in the host promoters and, using its activation domain (AD), activate transcription (Bogdanove et al., 2010). Its central DNA binding domain is composed of several tandem repeats. Each repeat is formed by 33-35 amino acids (aa) and recognizes a single nucleotide. The last repeat usually consists of a 20aa sequence, termed "half repeat". The specificity on the DNA is conferred by highly variable amino acids at positions 12 and 13, called repeat-variable di-residue (RVD); divergent RVDs bind preferentially to different nucleotides (Boch et al., 2009; Moscou and Bogdanove, 2009) (FIG. 6). The sides of the central domain comprise of a N-terminal segment (NTS) containing secretion signal peptides, a C-terminal segment (CTS) with the nuclear localization signal peptides and the transcription activator domain (White et al., 2009).

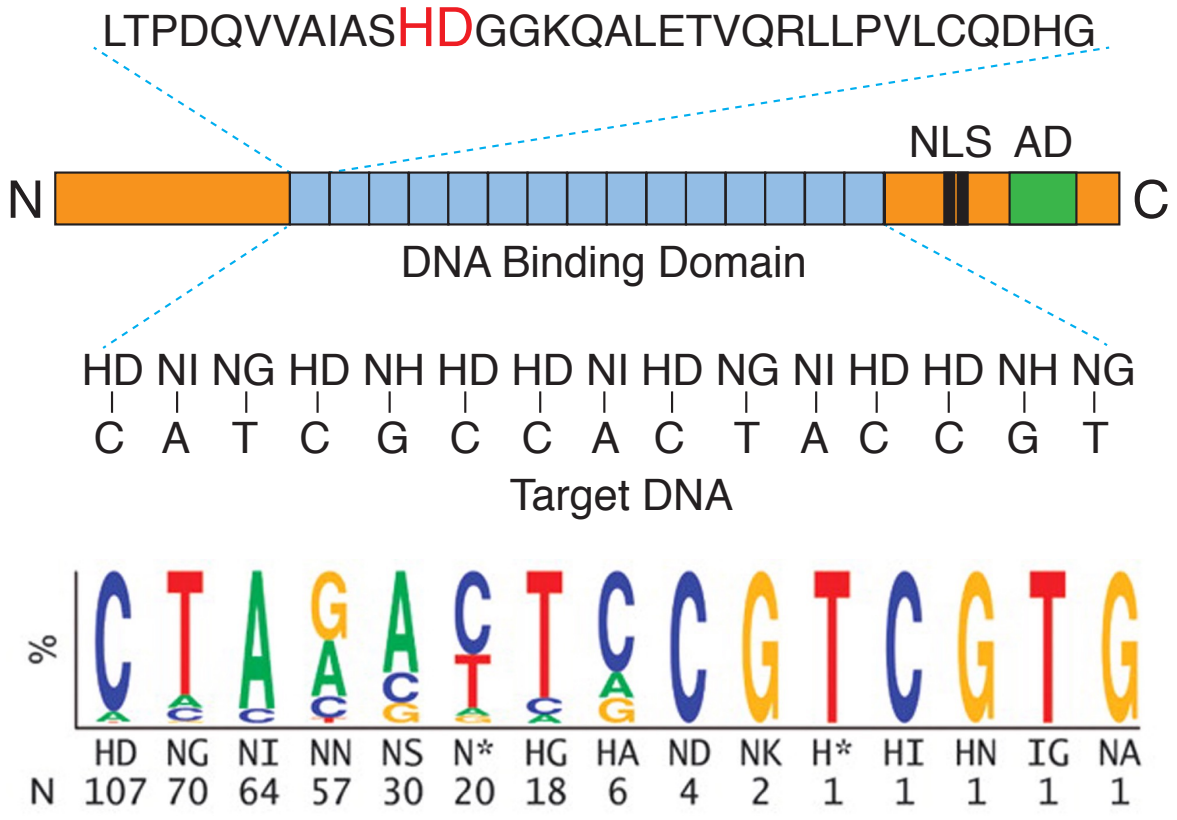

Figure 6

Up: Schematic representation of a TALE, with repeats (in blue) constituting the central DNA binding domain, the nuclear localization signal (NLS) and the activation domain (AD). An RVD is indicated in red (Adapted from Cermak et al., 2011)). Down: Frequencies of RVD-nucleotide associations in naturally occurring TALEs constitute a code that allows prediction and design. An asterisk indicates that the residue at position 13 is missing, resulting in a 33-amino acid repeat (Reproduced from Moscou and Bogdanove, 2009).

Similarly to ZFNs, Transcription Activator-Like Effectors Nucleases (TALENs) are TALE proteins fused with a non-specific FokI domain and acting as a pair. The TALEN structure was initially reported by fusing a natural TALE (AvrXa7) to the FokI domain. This protein was originally engineered comprising of the whole NTS and CTS (288aa and 295aa, respectively). Based on the structure, using yeast assay methodologies, the optimal spacer length between the two TALEN target sites was calculated in 16-31bp (Li et al., 2011a). Analysis of a TALEN based on a different TALEs (AvrBs3, PthXo1) found efficient DNA cleavage with a 13-30bp spacer (Christian et al., 2010). This scaffold was successfully used to obtain genome editing in various cell types and species, from yeast to mammalian cell lines, to zebrafish and transgenic mice (Hockemeyer et al., 2011; Huang et al., 2011; Sander et al., 2011a; Tesson et al., 2011; Wood et al., 2011; Cade et al., 2012; Lei et al., 2012; Moore et al., 2012; Reyon et al., 2012b). While usually transduced into cells encoded in plasmid constructs or as mRNA, TALENs have also been used in the form of purified proteins directly on human cells (Liu et al., 2014).

The next logical step regarded the optimization of TALEN architecture in order to obtain higher cleavage efficiency with minimal peptide length. Experiments using a series of TALEN with 136aa long NTS and varying lengths of CTS indicated that diverging TALEN scaffolds preferred different spacer lengths (Miller et al., 2011). Indeed, TALENs with a 63aa-CTS could lead to efficient gene editing in human cells when separated by 12-20bp spacers, while for a CTS with 28aa, a shorter separation range of 12-13bp was needed.

Additional work showed that shorter CTSs work efficiently with small spacers (Mussolino et al., 2011; Christian et al., 2012), highlighting tentative correlations between CTS length and the maximum spacer length, this may be due a short CTS producing less possible movement for FokI domains (TAB. 1).

Thus, a systematic study with 10 different TALEN scaffolds with varying NTSs and CTSs have been tested against 10 substrates with different spacers in yeast in a 10 x 10 matrix. Two scaffolds with high activity were identified: the first one with 207aa NTS and 31aa CTS preferred target sites separated by 10-16bp spacers, and the second one comprising of a 207aa NTS and of a 63aa CTS could efficiently cut the target site with 14-32bp spacers. Moreover, it was seen that TALENs with 50aa NTS did not exhibit catalytic activity against any target (Sun et al., 2012).

This was further clarified when the availability of a TALE crystal structure showed the

Table 1 - Different TALEN scaffolds

NTS (aa)	CTS (aa)	Spacer (bp)	Reporter system	Refs
288	295	16-31	β-galactosidase assay in yeast	Li et al., 2011a
288	285	16[a]	Transient expression assay in tobacco leaves	Mahfouz et al., 2011
287	231	13-30	β-galactosidase assay in yeast	Christian et al., 2010
287	63	15[a]	Mutagenesis in medaka embryos	Ansai et al., 2013
207	63	14-32	β-galactosidase assay in yeast	Sun et al., 2012
207	31	10-16	β-galactosidase assay in yeast	Sun et al., 2012
153	47	12-21	dsEGFP assy in HEK293	Mussolino et al., 2011
153	17	12	dsEGFP assy in HEK293	Mussolino et al., 2011
136	63	12-20	Surveyor nuclease assay in K562	Miller et al., 2011
136	28	12-13	Surveyor nuclease assay in K562	Miller et al., 2011
136	18	13-16	β-galactosidase assay in yeast	Christian et al., 2012

[a] the spacer length was not optimized in this study (Reproduced from Sun and Zhao, 2013)

presence of an extended N-terminal DNA-binding region of 127aa immediately before the central DNA-binding domain, with its structural features being very similar. These 127aa form four repeats, containing two α-helices and an interposing loop (FIG. 7). Although this region did not seem to confer sequence specificity, it is essential for DNA binding and thus it explains why functional TALEN scaffolds need at least 127aa NTSs (Gao et al., 2012). A second generation of TALEN scaffolds was thus developed, named as GoldyTALEN (Bedell et al., 2012). The GoldyTALEN have the same scaffold comprising of a 136aa NTS and of a 63aa CTS (Miller et al., 2011), with nine aa substitutions in the NTS and five in the CTS. These proteins were used to target several loci in zebrafish with a 100% modification efficiency. This represented the first study to show HR-based genome editing in zebrafish using DNA single-strand for donor. GoldyTALENs were also used to obtain efficient gene knockout in livestocks (Carlson et al., 2012).

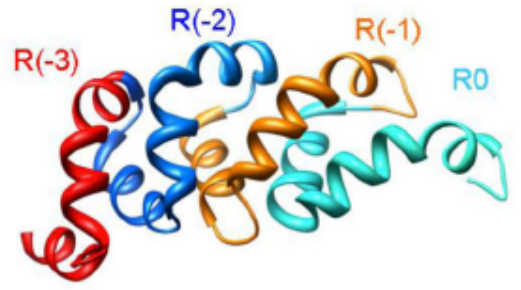

Figure 7

Crystal structures of TALEs NTS, characterized by four continuous repeats, which are important for DNA binding (PDB identification code: 4HPZ) (Reproduced from Sun and Zhao, 2013).

DNA RECOGNITION SPECIFICITY

The specificity of TALENs on DNA is conferred by RVDs at position 12 and 13 of each repeat. Through the alignment of natural TALEs more than 20 RVDs have been found where NI, NG, HD, NN and HG are the most common. The RVDs bind to the nucleotides A, T, C, G/A and T respectively (Boch et al., 2009; Moscou and Bogdanove, 2009).

Studies on crystal structures have shown that TALEs bind DNA as a right-handed superhelix (FIG. 8). Each module forms two helices that present the loop containing the RVD to the major groove of the DNA (Deng et al., 2012a; Mak et al., 2012). The residue 12, the first residue of RVDs, does not make contact with the DNA, while the side chain of Ala at position 8, forms a hydrogen bond stabilizing the local conformation of the RVD-containing loop. Sequence specific contacts of TALEs allow specific interaction of residue 13 (the second residue in each RVDs) in each repeat to the corresponding base on the 5'-3' strand of the DNA.

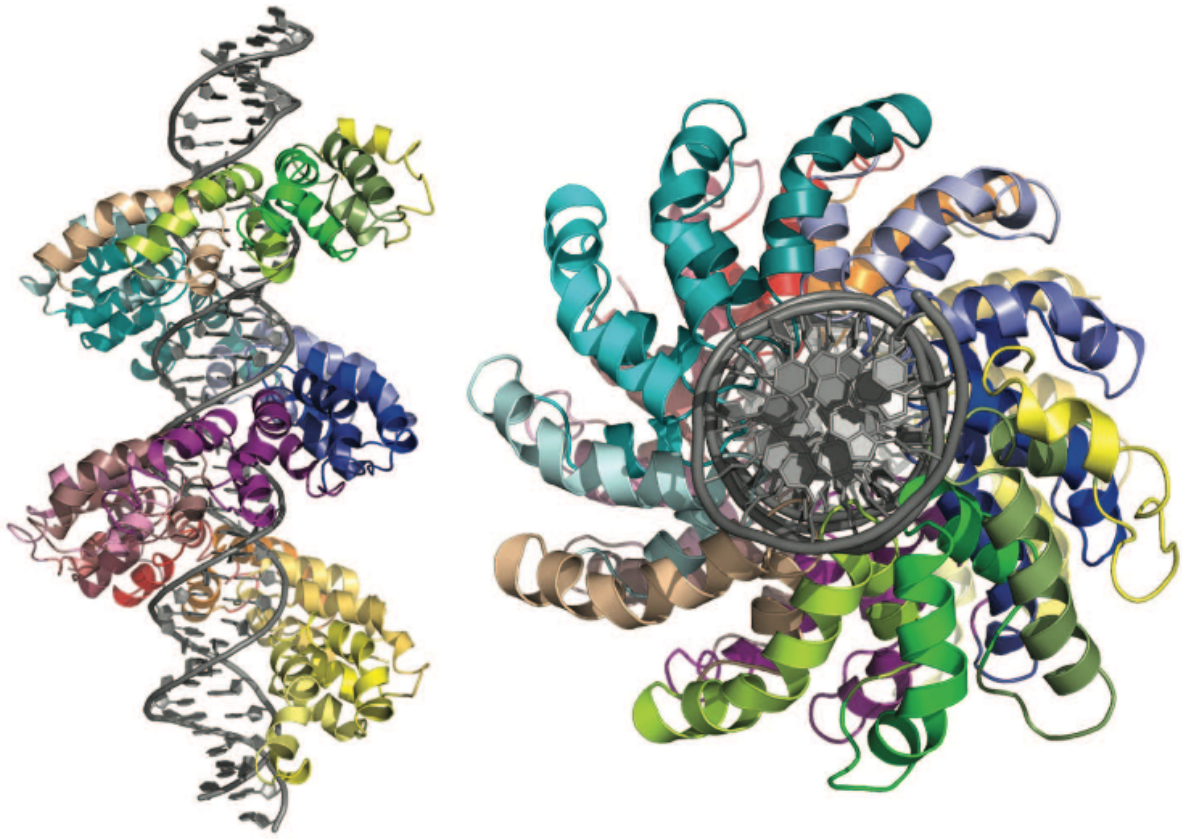

Figure 8

Structure of the PthXo1 DNA binding region in complex with its target site (Reproduced from Mak et al., 2012).

In the HD-RVD, which binds C specifically, the oxygen in the Asp[13] carboxylate forms an hydrogen bond with the amine group of the cytosine, while other interactions are excluded with different bases primarily due to physical or electrostatic hindrances (FIG. 9). NG- and HG-RVD are specific for T. In this case, the α-carbon of Gly[13] forms a non-polar van der Waals contact with the methyl group of thymine. In the NI-RVD, specific for A, the aliphatic chain of Ile[13] makes a non-polar van der Waals bond with C8 and N7 in adenine. NN-RVD is commonly used to bind G: the lateral chain of

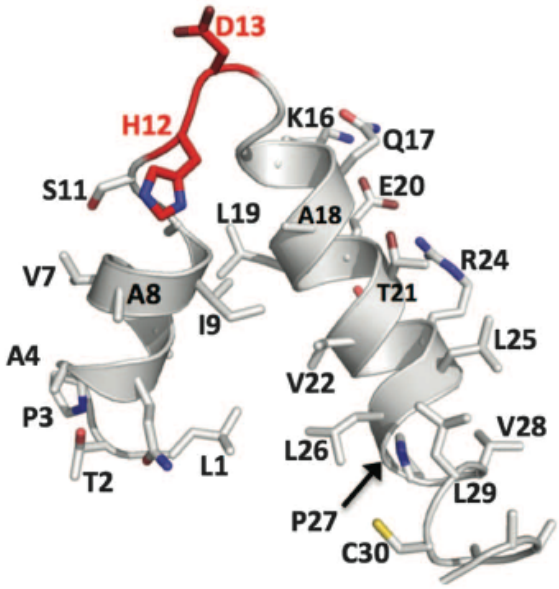

Figure 9

Structure of a single TAL effector repeat. Arrows indicate the start and the end of the crystallized protein.
The sequence and structure of a representative repeat is RVD residues (HD) that recognize cytosine are shown in red. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; G, Gly; H, His; I, Ile; K, Lys; L, Leu; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; and V, Val (Reproduced from Mak et al., 2012).

Asn[13] forms a hydrogen bond with N7 nitrogen of guanine. Nonetheless the same interaction can also partake with the N7 nitrogen of adenine. Thus, NN bind to both A and G with the same frequency (FIG. 10).

In conclusion, HD- and NN-RVDs recogniser the nucleobases through hydrogen bonds. On the other hand , NI-, NG- or HG-RVDs interact via van der Waals bonds, which are weaker than the hydrogen bonds. This is why it is recommended to include at least 3-4 strong RVDs during the design of the TALENs in order to increase their specificity (Streubel et al., 2012).

The non-specificity of NN represents an important limitation, and can generate off-target cleavages, inducing genomic instability and cytotoxicity. Some evidences indicate that NK-RVDs can bind to G nucleotides with stronger affinity than adenine, which represents a good premise for specific recognition of the G. Despite this, substitution of NN with NK has been shown to significantly reduce the activity of the TALENs in zebrafish embryo (Huang et al., 2011), plants and mammalian cells (Cong et al., 2012; Streubel et al., 2012). This is because the NN-RVDs are stronger than

the NK-RVDs. NH-RVD has been reported to have specific binding with G, higher than NK-RVD (Cong et al., 2012; Streubel et al., 2012). In fact, the imidazole ring of His^{13} in NH-RVD provides a compact base-stacking interaction with guanine (Cong et al., 2012) (FIG. 10).

Natural TALEs are unable to bind methylated DNA (Bultmann et al., 2012) as they fail to bind 5-methyl cytosine (5mC), the major epigenetic mark in fungi, plant and mammalian genome (Su et al., 2011). Moreover, 5mC has been identified within CpG islands, an important regulatory region of many promoters (Maunakea et al., 2010). On the other hand, it has been shown that NG-and N*-RVDs (the asterisk indicates the deletion of residue 13 in the repeat unit) binds 5mC efficiently *in vitro* and *in vivo* (Deng et al., 2012b; Valton et al., 2012). The 5mC is structurally similar to thymine, besides position 4, which is not involved in TALE bindings, thus the lack of side chain in Gly^{13} of NG provides space to insert the 5-methyl group of 5mC and form van der Waals interactions (FIG. 10). N* is equivalent to NG as the RVDs are followed by two conserved Glycines. These RVDs represented the first example of TALEN-mediated modification of a methylated locus in human cells (Deng et al., 2012b).

All natural TALEs target sites are preceded by a T at position 0 (5'-T), which initially seemed essential for TALEs function (Boch et al., 2009; Moscou and Bogdanove, 2009). The crystal structures of TALEs reveal two degenerate repeats before the central repeat domain. These repeats co-operate in order to bind the conserved thymine (Mak et al., 2012) through the indole ring of a Trp in repeat R(-1) to form van der Waals interactions with the methyl group of the thymine (FIG. 10). It has been reported that TALENs with short CTSs (31aa) recognize target sites with a T in position 0 with higher efficiency than sites preceded by A, C or G. However, TALENs with longer CTSs (63-117aa) are capable of binding target sequences preceded by any base (Sun et al., 2012). Further evidences suggest that thymine at position 0 is not rigorously required for TALEN activity (Miller et al., 2011; Yu et al., 2011; Briggs et al., 2012). Moreover, it has been shown that nine leucine zipper like heptad repeats closely linked to the C-terminus of the central repeat domain (Yang and Gabriel, 1995), may mediate binding on DNA and increase the strength of TALE/DNA interactions, making the 5'-T less of a requirement. The need of 5'-T can be mitigated using different TALEN scaffolds, allowing incredible flexibility to choose target sites in the genome.

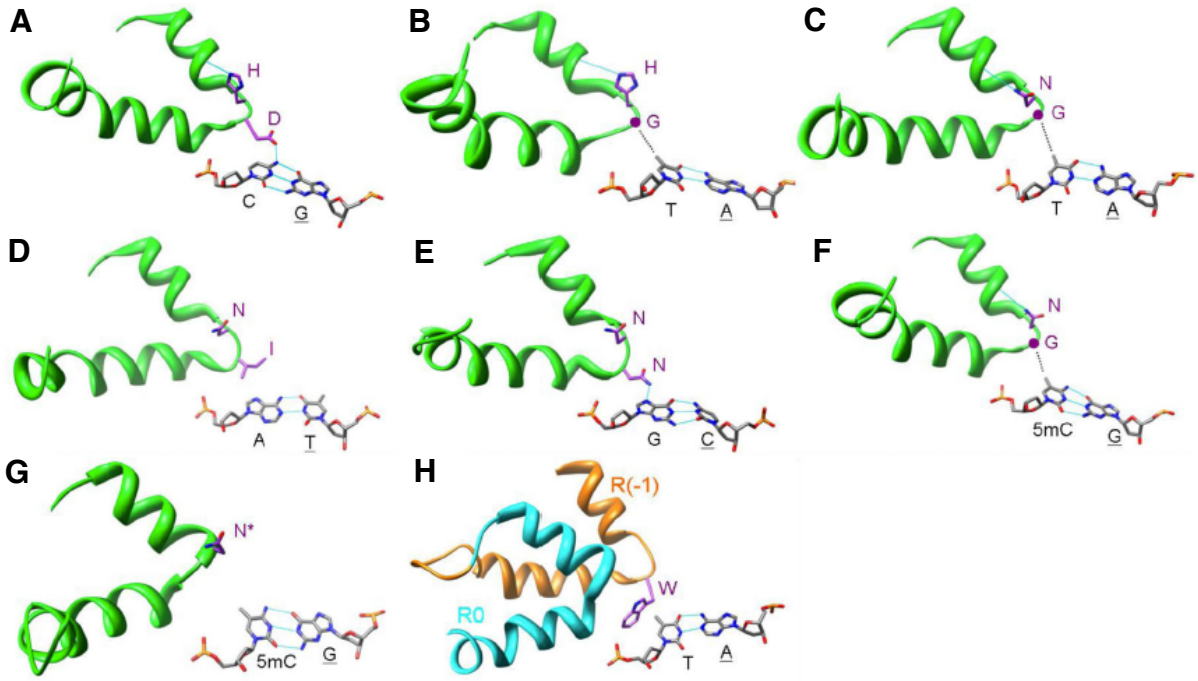

Figure 10

A: Interaction of RVD HD with cytosine (PDB identification code: 4HPZ). Hydrogen bonds are indicated by cyan lines. B: Interaction of RVD HG with thymine (PDB identification code: 3UGM). A non-polar van der Waals interaction is shown in a dotted line. C: Interaction of RVD NG with thymine (PDB identification code: 4HPZ). D: Interaction of RVD NI with adenine (PDB identification code: 3UGM). E: Interaction of RVD NN with guanine (PDB identification code: 3UGM). F: Interaction of RVD NG with 5-methyl cytosine (PDB identification code: 4GJR). G: Interaction of RVD N with 5-methyl cytosine based on a structural model. H: Interaction of the NTS with the 50-preceding thymine. (Adapted from MPDB identification code: 3UGM) (Reproduced from Sun and Zhao, 2013).

ASSEMBLY METHODS

Developing systems to obtain plasmids encoding TALEs for a specific sequence is a daunting task due to the high similarity of each repeat. Therefore numerous methods have been designed to assist TALEs construction (Joung and Sander, 2013).

Restriction enzyme and ligation (REAL) is a method where single TALE repeats are bound together using standard restriction digestion and ligation (Sander et al., 2011a). The single repeats obtained by DNA synthesis are inserted in plasmids, thus providing a library containing each repeat. These plasmids are digested and ligated together through several steps of cloning. This technique starts with association of two single TALE repeats, and arithmetically progresses to form paired repeats until it obtains a TALE array of the desired length. An improved version of this method called REAL-Fast was developed in order to achieve a rapid and less-laborious process (Reyon et al., 2012a).

The Golden Gate cloning system, is the most utilized system, and its development has enormously facilitated and accelerated the construction process of TALEs (Cermak et al., 2011; Geissler et al., 2011; Li et al., 2011b, 2012b; Morbitzer et al., 2011; Weber et al., 2011; Zhang et al., 2011; Sanjana et al., 2012). This method uses type IIS restriction endonucleases that cut outside of its recognition site and thus leaves 4bp overhanging bases at the 5' which are different for each plasmid containing the repeats. The overhangs are designed to allow binding in a specific order where the 3'-end of first repeat ligates only with the 5'-end of the second repeat and goes on in an ordered manner. This permits us to obtain two intermediate plasmid containing two halves of our TALE. The process is concluded after cloning of the TALE in the final plasmid containing the FokI endonuclease at its 3' end. The main advantage of this system is the absence in the final product of the restriction site. This allows the cleavage and ligation reactions in a single step, thereby increasing the cloning efficiency in presence of a *LacZ* and a toxic *ccdB* genes, which are used for blue/white screening and negative selection.

The requirement of two intermediate plasmids generated using this method yields the desired plasmid in 5 days. The use of preassembled TALE repeats - tetramers and trimers - which have been developed in a single step Golden Gate strategy can be used to generate TALENs that recognize 15bp targets in 2 days (Ding et al., 2013). The fast ligation-based automatable solid-phage high-throughput (FLASH), a solid-phase ligation strategy, was developed for large scale TALEN production, and was optimized for high-throughput and cost effectiveness (Briggs et al., 2012; Reyon et al., 2012b; Wang et al., 2012b). The system allows the assembly of approximately 100 TALEs in a single day. The solid-phase contains an immobilized DNA double strand adaptor with a single restriction site at the end, which specifically binds to TALE repeats until a TALE of the desired length is assembled. The final product is released at the end of the process by restriction digestion and it is cloned into the final backbone. This is the fastest assembly method, but is hindered by high start-up costs, sophisticated machinery and maintenance. In recent years there have been several advancements. Among these, the Iterative Capped Assembly (ICA) (Briggs et al., 2012) and Ligation Independent Cloning (LIC) (Schmid-Burgk et al., 2013) methods were developed.

OTHER APPLICATIONS

Several proteins were fused to TALEs to create novel chimeric proteins with specific functions; TALEs fused to transcription activator domain were obtained to induce transcription of specific genes in plants (Morbitzer et al., 2010) and human cells (Geissler et al., 2011; Zhang et al., 2011; Bultmann et al., 2012; Garg et al., 2012; Li et al., 2012d; Tremblay et al., 2012). TALEs, via fusion with repressor domains, have been used to repress specific genes in bacteria (Politz et al., 2012), yeast (Blount et al., 2012), plants (Mahfouz et al., 2012) and human cells (Cong et al., 2012; Li et al., 2012d). Furthermore, chimeric TALE recombinases (TALERs), obtained by fusing TALE and hyper-activated catalytic domain of DNA invertase Gin, have been used to recombine DNA in bacterial and mammalian cells in an approach alternative to the classical one (Mercer et al., 2012). Additionally, TALEs fused with lysine-specific demethylase 1 (LSD1) efficiently removed enhancer-associated chromatin modifications from target loci (Shi et al., 2004; Mendenhall et al., 2013). Finally, TALEs fused with fluorescent proteins could be used to visualize repetitive sequences in the genome (telomeric and centromeric regions) (Ma et al., 2013; Miyanari et al., 2013). Similarly, purified TALEs fused with fluorescent proteins can act as probes to detect repetitive sequences on fixed cells (Ma et al., 2013).

Clustered Interspaced Short Palindromic Repeats (CRISPRs)

INTRODUCTION AND STRUCTURE

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs), are part of an immune pathway present in bacteria to defend against viruses or plasmids (Barrangou et al., 2007; Horvath and Barrangou, 2010; Fineran and Charpentier, 2012; Wiedenheft et al., 2012). The CRISPR locus comprises of a series of repeat sequences (direct repeats) interspaced by non repetitive sequences called spacers and clustered CRISPR-associated (Cas) genes (FIG. 11A). CRISPR/Cas is present in almost 40% of available bacterial genomes and in almost 90% of archaeal ones (Mojica et al., 2000; Grissa et al., 2007). The CRISPR/Cas system is based on the ability of the Cas nucleases to process exogenous DNA in small fragments which are then incorporated into the CRISPR locus as spacers. During viral infection these

sequences act as transcriptional templates for producing CRISPR RNAs (crRNA). These particular RNAs guides Cas to target and cleave homologous nucleic acids (FIG. 11B).

The sequences and structures of Cas proteins have been used to classify three different types of CRISPR systems (Haft et al., 2005; Makarova et al., 2011b). The different CRISPR loci contain multiple Cas proteins to form complexes with crRNA (e.g: CASCADE complex for type I; Cmr or Csm RAMP complexes for type III) for the recognition and destruction of target nucleic acids (Brouns et al., 2008; Hale et al., 2009). Another CRISPR system, the type II, is different as it has lower number of Cas proteins.

Initial studies showed that CRISPR loci were transcribed, and archeal cells carrying specific spacers were resistant to corresponding viruses (Tang et al., 2002; Mojica et al., 2005). The breakthrough in understanding the functions of CRISPR locus happened in 2005, when varying independent spacer sequences analysis suggested their origin to be linked to phage-associated sequences (Bolotin et al., 2005; Mojica et al., 2005; Pourcel et al., 2005). It dates to 2007 the first experimental evidence linking the natural role of a type II CRISPR system to a bacterial adaptive immunity: CRISPR spacers define target specificity and Cas enzymes controlled spacer acquisition and phage defense (Barrangou et al., 2007). With regards to type I CRISPR locus in *E. coli*, it was shown that CRISPR arrays are transcribed and converted into small crRNAs, and individual spacers guide the activity of Cas nuclease (Brouns et al., 2008). In the same year, studies on *S. epidermidis* showed that the type III-A

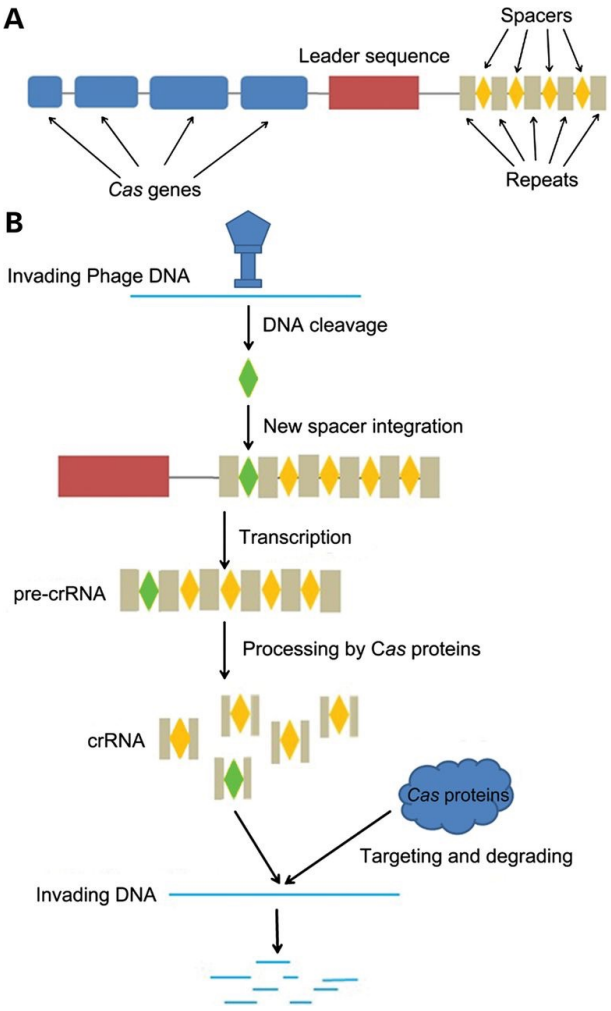

Figure 11

Overview of CRISPR/Cas bacterial immune system. (A) A typical structure of CRISPR locus; (B) illustration of new spacer acquisition and invading DNA cleavage (Reproduced from Zhang et al., 2014).

CRISPR system blocked plasmid acquisition by targeting DNA through the Cas enzymes (Marraffini and Sontheimer, 2008). Nonetheless later works found that a different type III-B system from *P. furiosus* could trigger crRNA-directed RNA cleavage (Hale et al., 2009, 2012).

It was initially speculated that protospacer-adjacent motifs (PAMs) might allow the type II Cas9 nuclease to cleave DNA (Bolotin et al., 2005). This was indeed demonstrated in 2008, when it was shown that phage genomes bearing mutations in PAMs avoided type II CRISPR defense activity (Deveau et al., 2008). With regards to type I and II, the absence of PAM within the CRISPR array prevents self-targeting, whereas in type III systems mismatches between the end of crRNA and DNA target were required for plasmid interference (Marraffini and Sontheimer, 2010).

Characterisation of type II CRISPR system showed that Cas9 (also known as: Cas5, Csn1 or Csx12) was the only enzyme in the cas gene cluster responsible for the cleavage of the target DNA (Garneau et al., 2010). A latter study revealed that a key molecule for crRNAs processing in type II CRISPR system is a noncoding trans-activating crRNA (tracrRNA) which hybridizes with crRNA to facilitate RNA-guided targeting through Cas9 (Deltcheva et al., 2011). These studies suggested that there are at least three essential components in the CRISPR system, namely: Cas9, mature crRNA, and tracrRNA for the type II CRISPR nuclease system.

The moment for Cas9 as a tool genome editing was nearly there. It was first demonstrated that type II CRISPR could be transferred among different bacterial strains (Sapranauskas et al., 2011), that the crRNA could guide Cas9 to cleave a target DNA *in vitro* (Gasiunas et al., 2012; Jinek et al., 2012), and that fusion of crRNA with tracrRNA forms a single guide RNA (sgRNA) which allows easy DNA cleavages *in vitro* (Jinek et al., 2012). Then, two studies simultaneously showed successful genome editing in mammalian cells using the type II CRISPR system (Cong et al., 2013; Mali et al., 2013a). This showed that a sgRNA can direct Cas9 for single or multiple cleavages in mammalian cells. Since then the usage of Cas9 in genome editing increased exponentially in many different experimental model systems (Sander and Joung, 2014).

Cas9 family proteins bear two nuclease domains, the RuvC and HNH, each known for its homology with already known nuclease domain structures. While the HNH comprises of a single domain, the RuvC domain consists of three subdomains, the RuvC I near the N-terminal region of Cas9, RuvC II and RuvC III, both flanking the

HNH domain close to the centre of protein. Recently, studies highllighted the structural mechanism of cleavage by Cas9. Single-particle electron microscopeic reconstructions of *S. pyogenes* Cas9 (SpCas9) revealed large structural modifications between Cas9 and Cas9 complexed with crRNA and tracrRNA. The association with nucleic acids forms a central channel to take up the RNA-DNA heteroduplex (Jinek et al., 2014). The high-resolution structure of SpCas9 bound to sgRNA and the complementary strand of the target DNA showed a domain organization with an α-helical recognition (REC) lobe, and a nuclease (NUC) lobe constituting of the HNH domain, the assembled RuvC subdomains and a PAM interacting (PI) C-terminal region (Nishimasu et al., 2014) (FIG. 12).

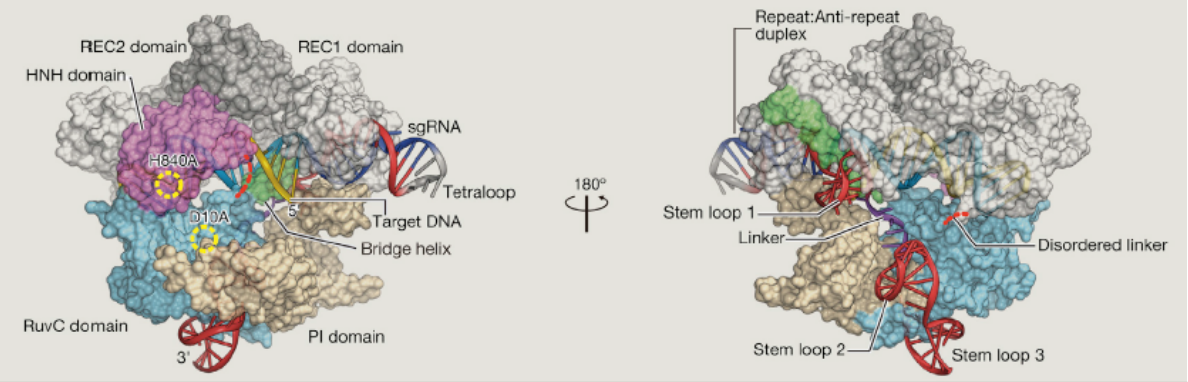

Figure 12

Crystal structure of *Streptococcus pyogenes* Cas9 in complex with sgRNA (in red) and target DNA (in blue). The domains of Cas9 are represented with different colors: HNH (in purple), RuvC (in light blue), PI (in brown), REC1 (in light grey) and REC2 (in dark grey) (Reproduced from Hsu et al., 2014).

These studies suggest that SpCas9 presents an autoinhibited conformation, where the HNH domain active site is blocked by the RuvC domain, further moved away from REC lobe (Jinek et al., 2014), thus rendering the protein unable to bind and cleave the DNA. The RNA-DNA heteroduplex binding results in alterations of this structure. Similar to the ribonucleoprotein complexes, the sgRNA act as a scaffold around which Cas9 folds and organizes its multiple domains (Nishimasu et al., 2014). The crystal structure of SpCas9 in complex along with an sgRNA and a target DNA showed the presence of an arginine-rich bridge helix (BH) within a REC lobe; serving an important point to contact with 8-12nt at the 3' of the RNA-DNA heteroduplex. This region is known as the "seed sequence" (Jinek et al., 2012; Cong et al., 2013; Fu et al., 2013; Hsu et al., 2013; Mali et al., 2013b; Pattanayak et al., 2013; Nishimasu et al., 2014). The availability of the structure provided an excellent platform for engineering of Cas9, where recombination or truncation of the REC2 domain, known

to be poorly conserved, represents an effective way to minimize the size of Cas9. The mutants lacking REC2 domain exhibit around 50% of wild-type cleavage activity, which may be due to its lower expression levels (Nishimasu et al., 2014).

PROTOSPACER ADJACENT MOTIF (PAM)

An essential part of the Cas9 system is the protospacer-adjacent motif (PAM), this sequence flanks the 3' end of the DNA target site and facilitates self/non-self discrimination by Cas9 (Shah et al., 2013), as PAM sites are not present in the direct repeats. Several studies hypothesize that PAM recognition is involved in transition between Cas9 target binding and cleavage conformations (Jinek et al., 2014; Nishimasu et al., 2014; Sternberg et al., 2014). Binding to the PAM and to the matching target activates the Cas9 nuclease activity by triggering the HNH and RuvC domains (Nishimasu et al., 2014). The complexity of the PAM sequence determines the frequency of the possible target sequences in a genome: e.g. the trinucleotide NGG, the PAM for SpCas9, is occurs on average every 8 bp in the human genome (Cong et al., 2013; Hsu et al., 2013). The fact that SpCas9 can also target sequences which flanked by NAG with low efficiency increases the versatility of this system (Hsu et al., 2013; Jiang et al., 2013). Other PAM sequences present an higher complexity. Thus Cas9 orthologs from *S. thermophilus* CRISPR1 is NNAGAAW, and Cas9 from *S. thermophilus* CRISPR3 is NGGNG (Deveau et al., 2008; Horvath et al., 2008). A alternative PAM, NNNNGATT, from *N. meningitidis* was recently used in human pluripotent stem cells (Hou et al., 2013; Zhang et al., 2013). These findings opened up numerous possibilities as Cas9s with different PAM requirements can be used together for simultaneous and differential genome engineering strategies. Thus independent transcriptional repression and nuclease activity was studied using NmCas9 and SpCas9 (Esvelt et al., 2013). Another interesting tool is based on the modification of PAM specificity by replacing PAM-interacting (PI) domains from different Cas9 orthologs. In the case of the PI from *S. thermophilus* CRISPR3 Cas9, it was substituted with the corresponding domain of *S. pyogenes* Cas9 and was successfully altered from NGGNG to NGG (Nishimasu et al., 2014). The advantage of the system is its inherent ability to effectively cleave several distinct target sequences simultaneously (Barrangou et al., 2007; Garneau et al., 2010; Deltcheva et al., 2011). This allows co-expression of CRISPR arrays containing spacers to target different genes (Cong et al., 2013), multiple sgRNAs (Mali et al., 2013a; Wang

et al., 2013) along with SpCas9 to obtain multiplex editing in mammalian cells. Indeed, CRISPR arrays containing direct repeats interspaced by designer spacers could be engineered in such a way that are cleaved by endogenous mammalian RNases and can mature in sgRNAs (Cong et al., 2013)

ADVANTAGES AND DISADVANTAGES

Genome editing leads to permanent modifications in the genome, hence the specificity of CRISPR system is crucial. Cas9 target recognition is defined by base-pairing interactions between the sgRNA and the target DNA. This allowed to elucidate how the number and position of mismatches between sgRNA and DNA affect the activity of Cas9. *S. pyogenes* Cas9 specificity has been characterized using mismatched sgRNA libraries, *in vitro* selection, and reporter assays by several researchers (Fu et al., 2013; Hsu et al., 2013; Mali et al., 2013b; Pattanayak et al., 2013). Initial studies suggested the presence of a "seed sequence" where the first 8–12nt PAM-proximal sequence determined its specificity (Jinek et al., 2012; Cong et al., 2013). Subsequent studies demonstrated how Cas9 tolerates mismatches throughout the guide sequence, and its sensitivity to the numbers, positions, and distribution of the mismatches (Fu et al., 2013; Hsu et al., 2013; Mali et al., 2013b; Pattanayak et al., 2013). PAM-distal bases are not distinct aspects for specificity, thus mismatches at these positions often do not abolish Cas9 activity. While Cas9 is still able to bind DNAs with low base-pairing with the sgRNA, its ability to cleave the DNA is almost entirely abolished. This highlights that, even in the presence of off-target binding sites, Cas9 cleaves only a small fraction of them (Wu et al., 2014). All in all, the amount of "real" Cas9 off-target activity is still an open question.

Enzymatic concentration represents an essential factor for the specificity of the system, as up to five mismatches within the target site can be tolerated by Cas9 at higher concentrations, thus leading to a higher off-target activity (Fu et al., 2013; Hsu et al., 2013; Pattanayak et al., 2013). On the other hand, low Cas9 concentrations are associated with improvements in the efficiency of on-target cleavages and reduction the off-target activity (Hsu et al., 2013). Meanwhile, the development of bioinformatic tools and the availability of complete genome sequences allows the modelling of off-target sites during the design of the CRISPR target sequence.

The activity of RuvC and HNH nuclease domains makes it possible to cleave to the DNA where each domain nicks one strand of DNA and they together generate a

blunt-ended DSB (FIG. 13). In order to reduce its off-target activity, SpCas9 has been converted into a DNA "nickase" to generate a single-strand break (SSB). This has been achieved through inactivation of either RuvC (D10A mutant) or HNH (N863A and H840A mutant) nuclease domains (Gasiunas et al., 2012; Jinek et al., 2012; Sapranauskas et al., 2011) (FIG. 14A). Since SSBs are repaired through the high-fidelity base excision repair (BER) pathway (Dianov and Hübscher, 2013), off-target effects from these molecules is greatly reduced. On-target cleavage is obtained by the association of a pair of sgRNAs driving the induction of a double-nick at the desired target, thus providing an approach similar to that used with dimeric ZFNs or TALENs (Hsu et al., 2013) (FIG. 14B). This approach increases the specificity by 1500x compared to wild type Cas9 (Ran et al., 2013). Two or three nt truncations in the sgRNAs have also been reported to increase SpCas9 targeting specificity, possibly due to higher mismatch sensitivity (Fu et al., 2014). Combination of these methods yields an improved and efficient method to reduce off-target mutagenesis (Fu et al.,

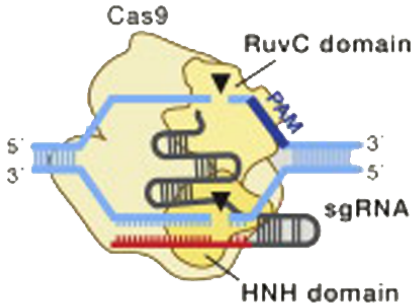

Figure 13

The figure shows a schematic drawing of the recognition by Cas9 of specific DNA sequences via the sg(red), directly base-pairing with the DNA target.
Cas9 cleaves the DNA on both strands through its domains RuvC and HNH, each of which cuts one of the strands. The PAM sequence is shown in blue (Reproduced from Hsu et al., 2014).

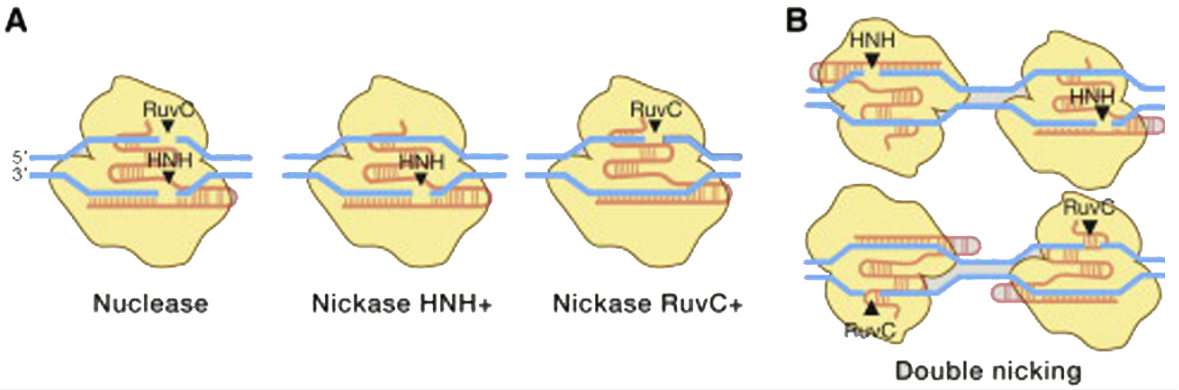

Figure 14

(A) The Cas9 nuclease cleaves DNA via its RuvC and HNH nuclease domains, each of which nicks a DNA strand to generate blunt-end DSBs. Either catalytic domain can be inactivated to generate nickase mutants that cause single-strand DNA breaks.
(B) Two Cas9 nickase complexes with appropriately spaced target sites can mimic targeted DSBs via cooperative nicks, doubling the length of target recognition without sacrificing cleavage efficiency (Reproduced from Hsu et al., 2014).

2014). A thorough in-depth targeting analysis and Cas9 engineering via rational design or directed evolution will likely improve Cas9 specificity.

OTHER APPLICATIONS

Wild-type Cas9 has been also converted through inactivation of both catalytic domains into an RNA-guided homing device (dCas9) which, fused with other effector domains, can be used in several different ways. For example, binding of dCas9 to DNA can repress transcription through sterical interference with the RNA polymerase machinery (Qi et al., 2013). This interference, known as CRISPRi, works well in prokaryotic genomes but is less effective in eukaryotic cells (Gilbert et al., 2013). This repressive function may be increased by fusing it with transcriptional repressor domains such as KRAB or SID effectors. Currently, only partial transcriptional knockdown has been achieved (Gilbert et al., 2013; Konermann et al., 2013).

Fusion of dCas9 with VP16/VP64 or p65 activation domains may activate transcription. Nonetheless, it was observed that targeting Cas9 activators through a single sgRNA induces just a modest transcriptional activation. On the other hand, multiple sgRNAs led to a strong increase in transcriptional activation (Maeder et al., 2013; Mali et al., 2013b; Perez-Pinera et al., 2013).

Recent advancements have made it possible to study the spatial organization and genome interaction in live-cell-imaging, where fluorescently tagged dCas9 were labelled for specific DNA loci (Chen et al., 2013a).

Comparison between ZFNs, TALENs and CRISPRs

ZFNs suffer from multiple limitations to their use, compared to the other techniques, as each ZF domain recognises only 3-nt targets, and the design and the assembly of the proteins is an expensive, laborious process. TALENs construction is easier and obtaining functional proteins is relatively easy. From the perspective of design, CRISPR are the easiest technique as the target DNA is defined only by the sgRNA, which can be easily cloned into the CRISPR plasmids using synthesized oligonucleotides (Mali et al., 2013a; Ran et al., 2013).

In contrast to ZFNs which present a context-dependent DNA binding, TALENs can target any given DNA sequence as it has a simple protein-DNA code with a modular

structure. TALENs also exhibit fewer off-target activity and cytotoxicity when compared to ZFNs (Mussolino et al., 2011; Ding et al., 2013).

TALENs and CRISPRs are two of the most effective tools for gene targeting technologies as they represent an approach relatively not expensive and easily accessible to any lab (TAB. 2).

ZFNs and TALENs work as juxtaposed pairs to activate FokI activity, this dimerization increases the specificity to cut only at the target locus, leading to a decrease in possible off-target activity. Concurrently, simultaneous use of two proteins to cut a single site may pose a problem for cell lines which are difficult to transfect into.

Table 2 - Comparisons between different genome editing technologies

Name	Components	Mechanism of action	Specificity/off-target effects	Possibility to rapidly generate large-scale libraries
Genome editing				
Zinc finger nucleases (ZFNs)	Fok1 restriction nuclease fused to multiple zinc finger peptides; each targeting 3 bp of genomic sequence	Induces double-strand breaks in target DNA	Can have off-target effects	No – requires customization of protein component for each gene
Transcription activator-like effector nucleases (TALENs)	Non-specific DNA-cleaving nuclease fused to a DNA-binding domain specific for a genomic locus	Induces double-strand breaks in target DNA	Highly specific	Feasible, but technically challenging (Reyon et al., 2012)
Homing meganucleases	Endonuclease with a large recognition site for DNA (12–40 base pairs)	Induces double-strand breaks in target DNA	Highly specific	No – limited target sequence specificity available
CRISPR/Cas	20 nt crRNA fused to tracrRNA and Cas9 endonuclease	Induces double-strand breaks in target DNA (wt Cas9) or single-strand DNA nicks (Cas9 nickase)	Some off-target effects that can be minimized by selection of unique crRNA sequences	Yes – requires simple adapter cloning of 20 nt Oligos targeting each gene into a plasmid

Reproduced from Heintze et al., 2013

Moreover, the fact that they work as pairs reduces the possibility for multiplex editing, to induce multiple targeted mutagenesis, chromosomal deletions, duplication, inversions or translocations (Lee et al., 2010; Lee et al., 2012; Piganeau et al., 2013; Wang et al., 2013). In these respects the CRISPR system is particularly versatile as only a single Cas9 nuclease is required to interact with any given number of sgRNAs. Considering the potential off-target effects, several reports have suggested potent off-target cleavages by CRISPRs in cultured cells (Cradick et al., 2013; Fu et al., 2013; Hsu et al., 2013; Mali et al., 2013b; Pattanayak et al., 2013).

There is a common consensus that CRISPR system may allow multiple mismatches in the target sequence, including PAM. Contrarily, TALENs are only active as dimers and the only limitations in design depends on the length of the spacers (Miller et al.,

2011), thus limiting its off-target effects. Meanwhile, single TALENs have exhibited higher specificity compared to the CRISPR/Cas9 (Mali et al. 2013b).
Taken together, all evidences points to the CRISPR system as the most flexible, and to the TALENs as the most accurate.

Genome editing in cultured cells and therapeutic applications

The development of ZFNs, TALENs and CRISPRs has opened new frontiers for genome editing *in vivo*, and these molecules are revolutionising our approaches to cellular models *in vitro*.
In cultured cells, gene editing may be achieved by introducing plasmids to express the nucleases or *in vitro* transcribed RNAs. The advantage in using RNA rests in the avoidance of random integration of the vectors.
An alternative way to deliver nucleases to the cells is the direct transduction of purified proteins. Delivery of SFNs obtained from inclusion bodies is used to engineer HEK293T, CHO (Chinese hamster ovary), primary $CD4^+$ cells and other cell lines (Gaj et al., 2012), and TALENs and CRISPRs can be delivered in a similar manner (Liu et al., 2014; Ramakrishna et al., 2014b).
There were several instances of human genome engineering in cell lines to correct genetic disorders: ZFNs were used in K562 cells (human erythroleukemia cell line) and in primary CD4+ T-cells to disrupt and repair the IL2RG gene, which can causes severe combined immunodeficiency (SCID) (Urnov et al., 2005); the LAMB3 gene, responsible for Epidermolysis bullosa was inserted using ZFNs in epithelial stem cells (Coluccio et al., 2013); and dystrophin was corrected in skeletal myoblasts from patients affected by Duchenne muscular dystrophy (Ousterout et al., 2013).
ZFNs and CRISPR have been used to confer HIV-1 resistance in primary T-cells as well as CD34+ hematopoietic stem cells by destructing HIV co-receptor C-C chemokine receptor type 5 (CCR5) (Perez et al., 2008; Holt et al., 2010; Ebina et al., 2013).
All these approaches have generate considerable interest in correlation to their possible use in human induced pluripotent stem (iPS) cells, a promising frontier in regenerative medicine and cell-based therapies. These cell lines are easy to engineer and selection and genotyping are more feasible than in pluripotent stem

cells (Capecchi, 2005). Moreover, auxiliary systems have been developed to enrich the cellular population in which nucleases have been active (Kim et al., 2011a; Kim et al., 2013; Ramakrishna et al., 2014a). Despite these advantages, genome editing in these cells is still constrained by the possibility of off-target effects. Furthermore, sequential targeting or highly active nucleases to concurrently hit both alleles are necessary to obtain homozygous mutations. However, high-levels of cutting activity increases the chances of off-targeting effects. In order to allow secure human *in vivo* applications further development of these technologies is needed, to increase the nuclease specificity and to enhance mutation detection.

Genome editing in model organisms

ZFNs, TALENs and CRISPRs are important tools that have been extensively used to edit the genomes of several model organisms. Their introduction in the cells can happen either as plasmid DNAs, or as *in vitro* transcribed mRNA.

These editors have been used to target genes mainly for disruption, but other purposes such as knock-in, chromosomal deletions or inversions have been obtained (Gupta et al., 2013; Xiao et al., 2013). A summary of various genome modification obtained in animal models is presented in TAB. 3.

In plants, ZFNs, TALENs, and CRISPRs have been used for genome engineering: e.g. *Arabidopsis* (Zhang et al., 2010; Li et al., 2013a) and several crop species (Shukla et al., 2009; Townsend et al., 2009; Shan et al., 2013), were engineered to be resistant to diseases or herbicides (Shukla et al., 2009; Townsend et al., 2009; Li et al., 2012c; Xie and Yang 2013).

In animals, the screening methods to identify the targeted mutants are an essential aspect of the approach. Presently, there are several methods to screen mutants, based on detection of DNA hybrids, such as: restriction fragment length polymorphism (RFLP) (Ochiai et al., 2010; Ansai et al., 2013; Suzuki et al., 2013), Cel-I digestion (Guschin et al., 2010), high-resolution melting analysis (HRMA) (Dahlem et al., 2012), heteroduplex mobility assay (HMA) (Ota et al., 2013) and direct detection by standard or next-generation DNA sequencing. Recent analysis revealed that a combination of these methods helped to improve the efficiency of screening (Nakagawa et al., 2013).

The diversity of organisms engineered by these site-specific nucleases continue to expand everyday, showcasing the opportunities these proteins pose to explore the repertoire of model system for basic and applied research.

Table 3 - Examples of various genome editing in animals
{Sakuma and Woltjen, 2014}

	ZFNs	TALENs	CRISPRs/Cas9
Targeted mutagenesis	Various animals such as nematode, fly, zebrafish, frog, mouse and rat (all in 24)		
Multiple targeted mutagenesis		Frog[1]	Zebrafish[2], mouse[3] and rat[4]
Knock-in using ssODN	Mouse[5] and rat[6]	Nematode[7], zebrafish[8], newt[9] and mouse[10]	Fly[11], zebrafish[12] and mouse[3]
Knock-in using targeting vector	Fly[13], sea urchin[14], mouse[15, 25] and rat[6, 25]	Zebrafish[16] and mouse[17]	Nematode[18] and mouse[19]
Chromosomal deletion or inversion		Silkworm[20] and zebrafish[21,22]	Zebrafish[22] and mouse[23]

([1]Sakane et al., 2013; [2]Jao et al., 2013; [3]Wang et al., 2013; [4]Li et al., 2013b; [5]Meyer et al., 2012; [6]Brown et al., 2013; [7]Lo et al., 2013; [8]Bedell et al., 2012; [9]Hayashi et al., 2013; [10]Wefers et al., 2013; [11]Gratz et al., 2013; [12]Hwang et al., 2013; [13]Beumer et al., 2008; [14]Ochiai et al., 2012; [15]Meyer et al., 2010; [16]Zu et al., 2013; [17]Jones and Meisler, 2013; [18](Chen et al., 2013b; Dickinson et al., 2013); [19]Yang et al., 2013; [20]Ma et al., 2012; [21]Gupta et al., 2013; [22]Xiao et al., 2013; [23]Fujii et al., 2013; [24](Wood et al. 2011; Bibikova et al., 2002; Liu et al., 2012; Aryan et al., 2013; Smidler et al., 2013; Merlin et al., 2013; Watanabe et al., 2012; Ma et al., 2012; Ochiai et al., 2010; Kawai et al., 2012; Doyon et al., 2008; Meng et al., 2008; Hisano et al., 2013; Ansai et al., 2012, 2013; Young et al., 2011; Suzuki et al., 2013; Carbery et al., 2010; Sung et al., 2013; Mashimo et al., 2010, 2013; Hauschild et al., 2011; Carlson et al., 2012); [25](Cui et al., 2011).

Secondary Antibody Diversification Processes

The antibodies

At the core of vertebrate adaptive immunity are the antibodies, molecules that are produced in B-cells to recognize and bind the antigens. The B cell antigen receptor (BCR) is composed of immunoglobulin (Ig) heavy (IgH) and light (IgL) chains.

A huge repertoire of specific antibody molecules are produced through a series of processes that select antibodies specific for any given antigen. The portion of the antibody responsible for antigen recognition is encoded by the immunoglobulin genes (heavy and light chains), and is generated through gene rearrangement events during the early phases of B lymphocyte development (Tonegawa, 1983). This region is assembled through the recombination of variable (V), diversity (D, only in the genes encoding for the heavy chains) and joining (J) segments. V(D)J recombination is responsible for the generation of the primary repertoire of antibodies. Key enzymes in this process are the RAG1/RAG2 recombinases. While this primary repertoire has a complexity in the order of 10^{-6}, these antibodies, generated at random, do not possess the required affinity to recognise the wide range of antigens that the organisms will encounter (FIG. 15).

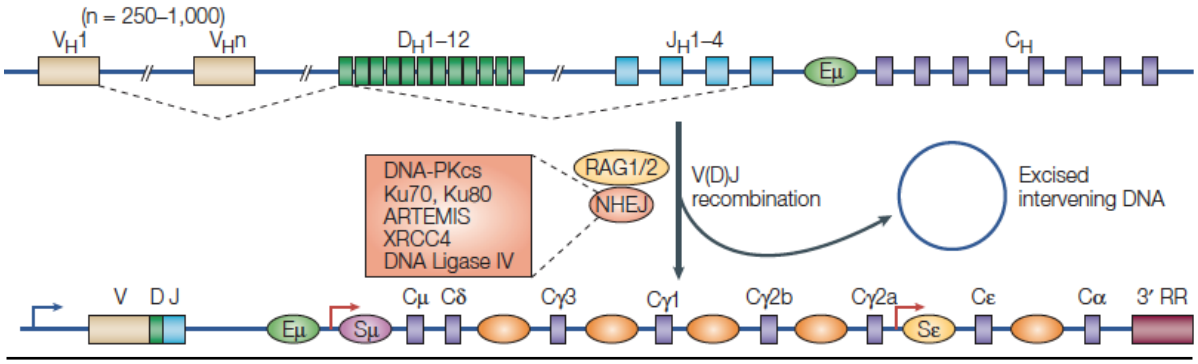

Figure 15

The variable region of the immunoglobulin heavy chain is assembled from variable (VH), diversity (DH), and joining (JH) gene segments by V(D)J recombination. The process is carried out by the RAG1–RAG2 complex. Joining of the DNA ends requires NHEJ proteins, including Ku70, Ku80, ARTEMIS, X-ray repair cross-complementing protein 4 (XRCC4), DNA ligase IV and the catalytic subunit of DNA-dependent protein kinase (DNA-PKcs). Transcription across the locus is driven by a promoter upstream of the rearranged VDJ segment (blue arrow) (Reproduced from Chaudhuri and Alt, 2004).

In order to increase the specificity for the antigen, another layer of diversification processes, the Somatic Hypermutation (SHM), has evolved in jawed vertebrates.
Upon antigen encounter, during SHM, mutations are inserted in the antigen binding region of the antibody gene, and cells expressing high-affinity antibodies are then selected in the germinal centers (Neuberger and Milstein, 1995; Milstein and Neuberger, 1996; Rajewsky, 1996).
Finally, through the Class Switch Recombination (CSR), constant regions in the antibody genes are replaced in order to obtain different effector functions depending on the isotype selected. Contrarily to SHM, the CSR process appears to have evolved later than the SHM, being present only in tetrapods (Flajnik, 2002).

Class Switch Recombination

The Class of the antibody, also termed Isotype, is the moiety defining the specific function of the antibody: constant regions allow the binding of the antibodies to the cell surface, and the various constant regions determine the type of cells that will be involved downstream to the antibody-antigen interaction, leading to the activation of macrophages, Natural-Killer cells or mast cells. The different isotypes of membrane-bound antibodies are diverse in their cytoplasmic domains, thus resulting in the activation of different intracellular signaling pathways (Martin and Goodnow, 2002; Horikawa et al., 2007; Waisman et al., 2007).
The Constant regions are encoded by the heavy chain constant genes (C_H) in the immunoglobulin gene. The different C_H regions form an array (Cμ, Cδ, $C\gamma_3$, $C\gamma_1$, $C\gamma_{2b}$, $C\gamma_{2a}$, Cε, Cα in human) downstream to the variable region and to the Intronic Enhancer (iEμ). Upstream to each C_H region there is a repetitive region, the Switch Region (Sμ, $S\gamma_3$, $S\gamma_1$, $S\gamma_{2b}$, $S\gamma_{2a}$, Sε, Sα). The exception is the Cδ region, in which IgD expression occurs by alternative splicing between the Cμ and the Cδ genes. Finally, located downstream of Cα there is a control region, 3'-regulatory region (3'-RR). S regions are composed by tandem repeats of short G-rich sequences (20-80 bp) that are different between each isotype. The length of each S region ranges from ~1 kb to 12 kb and CSR can happen anywhere inside or near the S regions (Dunnick et al., 1993; Min et al., 2005).

CSR occurs between the Sμ and one of the downstream ones, leading to a change from IgM and IgD expression in naive B cells to the expression of a different antibody isotype in memory- and plasma- cells.

CSR is initiated by AID, which targets the S regions and converts cytosines to uracil by deamination (Muramatsu et al., 2000; Revy et al., 2000; Petersen-Mahrt et al., 2002; Chaudhuri et al., 2003; Dickerson et al., 2003; Pham et al., 2003). Repair of these uracils lead to the formation of DSBs both at the Sμ donor region and at the S acceptor region to induce intrachromosomal DNA recombination (FIG. 16).

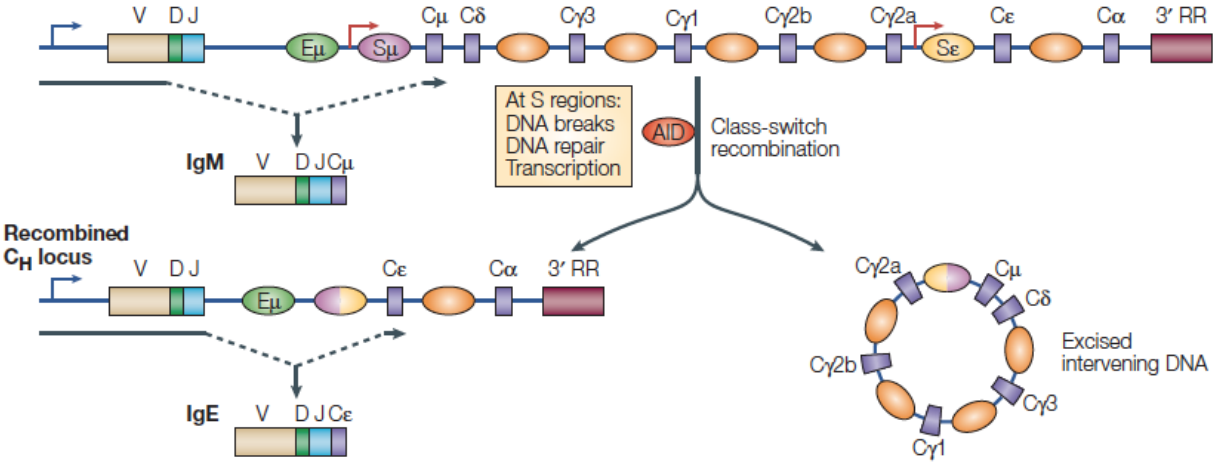

Figure 16

Secondary isotypes are produced by class-switch recombination (CSR), a process that exchanges the constant region of the heavy chain (C_H) with a set of downstream constant-region genes (CSR to IgE is shown). This recombination reaction, which requires AID, involves the generation of DNA breaks at switch (S) regions, which precede the constant-region genes, followed by the repair of DNA that leads to a rearranged C_H locus. Cytokines stimulate transcription (red arrows) through the C_H gene and determine the immunoglobulin isotype that the B cell will switch to. The Eμ and 3′-regulatory region (3′ RR) enhancers influence V(D)J recombination and CSR, respectively (Reproduced from Chaudhuri and Alt, 2004).

Closely related to the activity of AID are several other elements that determine the fate of CSR. The primary element determining the success of CSR, as well as that of SHM, is the transcriptional status of the Ig locus: only transcriptionally active S regions can undergo CSR (Pettersson et al., 1990; Yang and Schatz, 2007; Alt et al., 2013; Buerstedde et al., 2014; Storb, 2014). Such transcription is induced at levels sufficient for CSR only in the presence of both a fully functional IgH iEμ (Perlot et al., 2005; Perlot and Alt, 2008; Meng et al., 2014; Qian et al., 2014) and the locus control region located downstream of Cα (3'RR) (Pettersson et al., 1990; Dunnick et al., 2009). Highlighting the strong relation between transcription and CSR, CSR is impaired in cells deficient for proteins linked to transcription control and RNA processing, such as the transcription elongation factors SPT5 and SPT6,

polypyrimidine tract-binding protein 2 (PTBP2), a regulator of RNA splicing, and the RNA exosome, which has a function in RNA processing (Conticello et al., 2008; Pavri et al., 2010; Basu et al., 2011; Nowak et al., 2011; Okazaki et al., 2011). In addition transcription of S regions causes RNA polymerase II pausing or stalling (Rajagopal et al., 2009; Wang et al., 2009a; Storb, 2014), and its interaction with AID is mediated by SPT5, which, similar to RNA polymerase II, is enriched in S regions (Pavri et al., 2010). Due to the mutational pattern observed in S regions (more at the 5' end than downstream), researchers have proposed that AID is recruited by RNA polymerase II at the beginning of transcription within ~150 bp from transcription start (Longerich et al., 2005; Bransteitter et al., 2006; Longerich et al., 2006; Xue et al., 2006).

Activation Induced Deaminase (AID)

The main actor in SHM and CSR is Activation Induced Deaminase (AID), a 198 amino acid protein that targets deoxycytosines in the context of DNA, induces DNA damage on the immunoglobulin locus, thus recruiting the DNA repair machinery (Muramatsu et al., 2000). AID deficiency causes a form of primary immunodeficiency, the Hyper-IgM Syndrome (HIGM2), in which there are no antigen-driven antibody diversification processes (Revy et al., 2000). This demonstrates the essential role of AID in the secondary antibody diversification.

When AID was identified, the only protein with some degree of similarity was APOBEC1 (Muramatsu et al., 1999). This protein is an RNA editor that deaminates C6666 to U in the mRNA encoding the apolipoptein B, leading a premature stop codon and the synthesis of truncated form of apolipoptein B (Blanc and Davidson, 2010). This similarity initially suggested that AID could work in antibody diversification as an RNA editing enzyme for some unknown endonuclease (Muramatsu et al., 2000). Indeed, few years later it was demonstrated that AID acts directly on DNA by deaminating C residues within the Ig locus (Petersen-Mahrt et al., 2002; Bransteitter et al., 2003; Chaudhuri et al., 2003; Dickerson et al., 2003; Ramiro et al., 2003; Sohail et al., 2003).

AID targets single-stranded DNA (Bransteitter et al., 2003) with a preference for cytidines within the WRC motif (where W = A or T, and R = A or G) (Pham et al., 2003; Beale et al., 2004). On the Ig locus, AID physiologically shows a preference for

RGYW sequences (where Y = T or C) which represents overlapping WRC consensus sequences on opposite DNA strands. This might be an indication of AID acting as dimeric/multimeric complexes (Beale et al., 2004). The presence of these hotspots in the immunoglobulin locus are essential to proceed with the antibody diversification processes (Yang and Schatz, 2007). The non-transcribed strand (coding) is preferentially targeted by AID in bacteria. This could be due to the coding strand in the bubble formed by RNA polymerase being more available than the template strand, which is transiently bound to the nascent RNA (Chaudhuri et al., 2003; Pham et al., 2003; Martomo et al., 2005; Ramiro et al., 2003). However, the presence of the RNA exosome in the transcription bubble enables AID to deaminate cytosines on both strands (Basu et al., 2011; Pefanis et al., 2014). This might explain why both strands are equally mutated in the Ig locus in physiological conditions (Milstein et al., 1998; Longerich et al., 2005; Xue et al., 2006).

AID is predominantly localized in the cytoplasm even though it performs its function in nucleus (Rada et al., 2002). AID presents an N-terminal nuclear-localization signal (NLS) and a C-terminal nuclear-export signal (NES) (Ta et al., 2003; Brar et al., 2004; Ito et al., 2004b; McBride et al., 2004) through which it can enter and exit from the nucleus (FIG. 17).

The catalytic domain contains the residue E58, whose carboxylic acid group serves as a proton donor in the deamination reaction, and the H56, C87 and C90 residues that are responsible for the binding to the Zn^{2+}. The enzymatic activity of AID is virtually abolished by mutations of the R112 and R24 residues, within the APOBEC-like domain and the DNA-binding N-terminal region. These two positively charged residues are frequently mutated in patients with HIGM2 syndrome. In addition, R112 is located just outside the loop that determines the substrate specificity for AID activity (Conticello et al., 2007) (FIG. 17). Mutations in the C-terminal region (last ~30 aa) have been found in patients from hyper-IgM syndromes (Ta et al., 2003), suggesting the importance of this region for AID activity. The last 10 aa of this region are absolutely essential for CSR but seem dispensable for SHM (Ta et al., 2003; Barreto et al., 2003; Shinkura et al., 2004; Ranjit et al., 2011). Moreover, this region is rich in leucines, suggesting of a protein-protein interaction domain (FIG. 17).

Some proteins, such as MDM2 (a regulatory protein of p53) and DNA-PKcs (a protein normally implicated in the joining of broken DNA ends), have been identified to be able to interact with the C-terminal portion of AID both in yeast two-hybrid and pull-

down assays (MacDuff et al., 2006; Wu et al., 2005). Other proteins termed 14-3-3 proteins have an important role in CSR for their ability to bridge DNA ends with other factors from the DNA repair machinery (Morrison, 2009). These proteins play an essential role in CSR, in fact 14-3-3 adaptors are recruited to S regions, and inhibition of their binding to S regions results in diminished CSR. B cells deficient in 14-3-3γ are defective in CSR (Xu et al., 2010). 14-3-3 proteins interact directly with AID and protein kinase A catalytic subunit-α (PKA-Cα) and, mediating their interaction (Xu et al., 2010), they lead to phosphorylation of AID through PKA (Vuong et al., 2009). The interaction between 14-3-3 proteins and AID depend on the C-terminal region of AID (Xu et al., 2010) (FIG. 17).

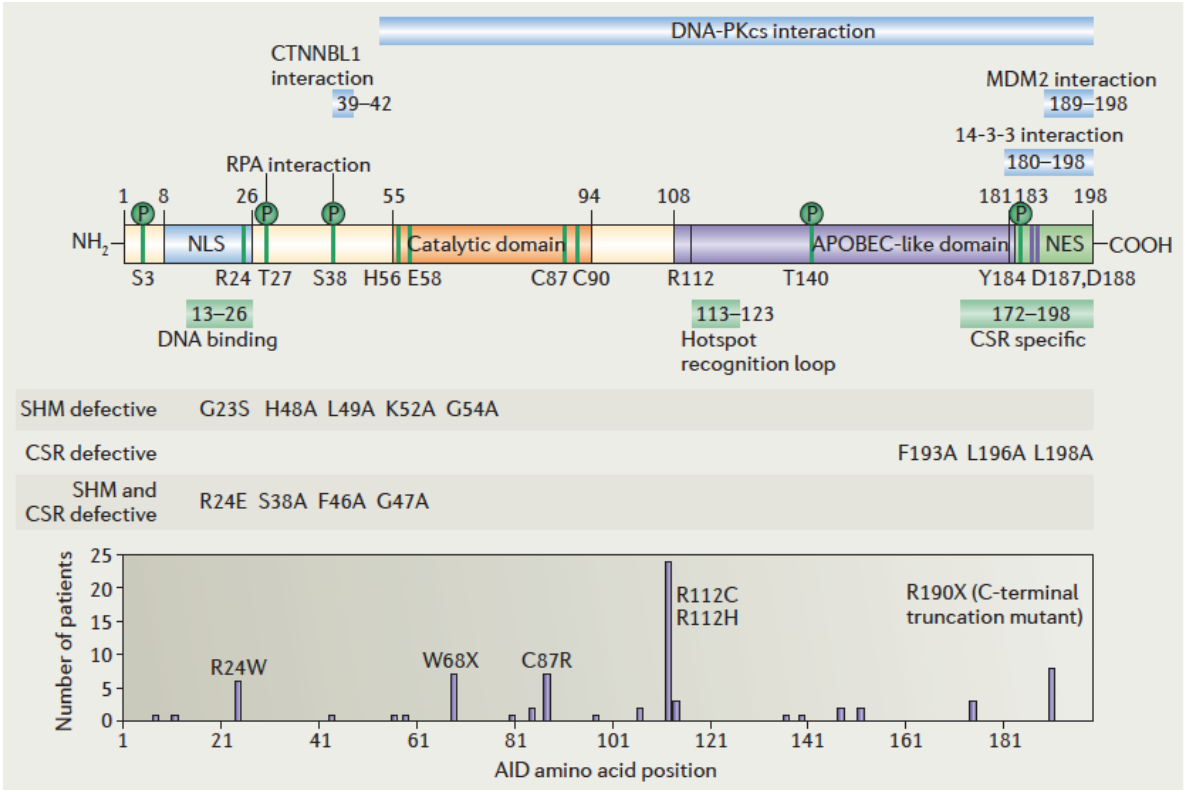

Figure 17

AID is a 198aa protein that initiates CSR and SHM. It shares a conserved catalytic domain with other members of the APOBEC family. The catalytic domain contains the amino acid residue E58, H56, C87 and C90, which bind to Zn^{2+}. The APOBEC-like domain of AID binds to the DNA surrounding dC and influences the substrate specificity. The C-terminal domain is essential to mediate CSR. Naturally occurring mutations in the AID gene are responsible for the autosomal recessive disorder hyper-IgM type 2 (HIGM2) syndrome causing defects in CSR and/or SHM. AID deamination activity and CSR are virtually abolished by mutation of R112 in the APOBEC-like domain and R24 in the DNA-binding N-terminal region; these two positively charged residues are frequently mutated in patients with HIGM2 syndrome. R112 is just outside the hotspot recognition loop (amino acids 113–123). The N-and C-terminal domains also function as the NLS and nuclear NES, respectively. AID is phosphorylated at T27 and S38 to create a binding site for replication protein A (RPA), and this promotes CSR. The regions in AID that interact with the spliceosome-associated factor CTNNBL1, 14-3-3 proteins, MDM2 and the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) are also depicted. (Reproduced from Xu et al., 2012).

Regulation of AID

Primary stimuli induce the expression of AID and other proteins important in CSR in B cells, through activation of transcription factors such as nuclear factor-κB (NF-κB) (He et al., 2004; Pone et al., 2012; Rawlings et al., 2012). On the other hand, secondary stimuli, while not directly related to AID expression, are required to direct the class switch recombination to specific classes of Ig (IgG, IgE or IgA). These stimuli involve interleukin-4 (IL-4), transforming growth factor-β (TGF-β), interferon-γ (IFNγ; in mice but not humans) and CD40 activation. These stimuli activate canonical and non-canonical NF-κB pathways (Zarnegar et al., 2004; Pone et al., 2012): the canonical pathway is rapidly activated to induce immediate but transient AID gene expression, by binding to the promoter of AID and to its upstream enhancer element (Tran et al., 2010). On the other hand, the kinetics of the induction of AID (peaking at 48–60 hours after stimulation) resembles that of the non-canonical pathway of NF-κB, which sustain gene expression to support cell proliferation (required for CSR) and differentiation (Smale, 2011). In addition NF-κB regulates many genes and transcription factors (HOXC4, STAT6, PAX5, E2A and others) that influence positively AID expression (Gonda et al., 2003; Sayegh et al., 2003; Dedeoglu et al., 2004; Xu et al., 2007; Park et al., 2009; Tran et al., 2010).

Being a DNA editor, AID expression needs a tight regulation to avoid mutations or chromosomal translocations and to maintain genomic integrity (Pasqualucci et al., 2008; Robbiani et al., 2009; Hasham et al., 2010). This is achieved through fine control of the AID gene and protein, at the transcriptional (Xu et al., 2007; Stavnezer, 2011), post-transcriptional (Delker et al., 2009), and post-translational level. These regulatory layers effect nuclear and cytoplasmic distribution of AID (Geisberger et al., 2012; Orthwein et al., 2010; Häsler et al., 2011; Orthwein and Di Noia, 2012) and stability (Aoufouchi et al., 2008; Uchimura et al., 2011) and enzymatic function (Chaudhuri et al., 2004; Basu et al., 2005; McBride et al., 2006; Li et al., 2012a).

The stability of AID is dependent on its subcellular localization, in fact cytoplasmic AID is more stable than nuclear AID (Aoufouchi et al., 2008). In particular it has been showed that Heat Shock Protein 90 (HSP90) interacts with AID, thus preventing its polyubiquitination and consequent proteasomal degradation (Orthwein et al., 2010). Contrarily, nuclear AID seems to be constantly targeted to the proteasome by ubiquitin-dependent and -independent pathways (Aoufouchi et al., 2008; Uchimura et

al., 2011). This suggests that while a cell can afford the presence of AID in its cytoplasm, unless needed, it avoids as much as possible AID presence in the nucleus.

Phosphorylation of AID by PKA at the S38 and T27 residues is another important regulatory step for CSR. Substitution of these aa by alanines eliminate the ability of AID to initiate CSR when transduced into *aid*$^{-/-}$ splenic B cells (Chaudhuri et al., 2004; Basu et al., 2005; McBride et al., 2006). The phosphorylation of AID is required for AID interaction replication protein A (RPA), allowing the contact and deamination of transcribed DNA (Chaudhuri et al., 2004; Basu et al., 2005; McBride et al., 2006).

Events downstream to the deamination

AID deaminates cytosines to uracils on single-stranded DNA (ssDNA) in both SHM and CSR (RawlingsDi Noia and Neuberger, 2007). Elimination of uracil (U) residues after deamination through the base excision repair (BER) pathway is essential for CSR (Di Noia and Neuberger, 2002; Rada et al., 2002; Schrader et al., 2005). BER is a highly active DNA repair system to eliminate oxidided and deaminated bases. These DNA lesions happen spontaneously hundreds of times per day in each cell by oxidation or spontaneous hydrolysis, even more often during inflammation (Christmann et al., 2003). In mammals, the BER system is composed by four enzymes that excise the uracils. These are the DNA glycosylases UNG, SMUG1, TDG and MBD4. UNG is the enzyme most tightly related to CSR. In fact, CSR is reduced by 95% in UNG deficient B cells. This has been shown both experimentally in chicken and mice, and in patients with inactivating mutations in UNG (Di Noia and Neuberger, 2002; Rada et al., 2002; Imai et al., 2003; Schrader et al., 2005). Thus, DSBs in S regions are reduced considerably in splenic B cells from ung$^{-/-}$ mice induced for CSR (Schrader et al., 2005). MBD4 and SMUG1 seem to not play a substantial role in CSR (Rada et al., 2002; Bardwell et al., 2003). Nonetheless SMUG1 overexpression can support low level CSR in ung$^{-/-}$ cells/ However, coexpression of UNG and SMUG1 inhibits CSR, suggesting that correction of AID-induced damage with the right timing, as UNG and SMUG1 are differentially expressed during cell cycle, is important (Di Noia et al., 2006).

The BER enzyme responsible for the repair of the abasic site left by UNG is the apurinic/apyrimidinic endonuclease (APE) that excises the abasic site and leaves a SSBs (Christmann et al., 2003). In mammals there are three different AP endonucleases: APE1 and APE2, very similar to each other (Hadi et al., 2002), and a third one, PALF/APLF/XIP-1 (Bekker-Jensen et al., 2007; Iles et al., 2007; Kanno et al., 2007). The main protein involved in BER is APE1 as its activity is essential for early embryonic development and human cells viability (Xanthoudakis et al., 1996; Fung and Demple, 2005). Induction of CSR in murine splenic B cells deficient for APE1 and/or APE2 demonstrate that both these proteins contribute to resolution of the CSR (Guikema et al., 2007). PALF/APLF/XIP-1 together with Artemis appear to be among the primary nucleases involved in NHEJ and responsible for most nucleolytic end processing in NHEJ (Pannunzio et al., 2014).

It is hypothesised that the activity of the AP endonucleases are responsible for the formation of a SSB, which can develop in DSB when close to each other on opposite DNA strands, or in the case that Mismatch Repair activity converts them to DSBs.

Physiologically the single strand gap left by UNG and APE is filled in by DNA polymerase β (Polβ) (Barnes and Lindahl, 2004; Beard and Wilson, 2006). We know that Polβ is recruited by APE1 (Barnes and Lindahl, 2004), but this would limit the efficiency of CSR, since resynthesis over the gap would avoid the formation of the DSBs crucial for CSR. It is possible that even in the case of Polβ recruitment by APE1, its levels or its activity might be inhibited. It is also possible that the number of AID-induced lesions at the S regions overload the BER machinery and, even though not inhibited during CSR, Polβ cannot repair all the damages; so in this case Polβ inhibits CSR when the number of SSBs are limited (Wu and Stavnezer, 2007).

In addition down-regulation of BER during switching is dangerous due to the large amount of reactive oxygen species produced during activation and proliferation of B cells (Ito et al., 2004a) thus it is more plausible thinking about BER overwhelmed than BER inhibition, in order to maintain the integrity of the genome.

Indeed, AID leads to many more mutations at the Sμ region in ung$^{-/-}$msh2$^{-/-}$ B cells compared to wild-type cells, probably because in wild type cells most of the damages are correctly repaired (Xue et al., 2006). Moreover, artificially introduced I-SceI sites in Sμ and Sγ$_1$ mediate switch from IgM to IgG$_1$, suggesting that a single DSB is sufficient in both donor and acceptor S regions to obtain CSR (Zarrin et al., 2007) (FIG. 18).

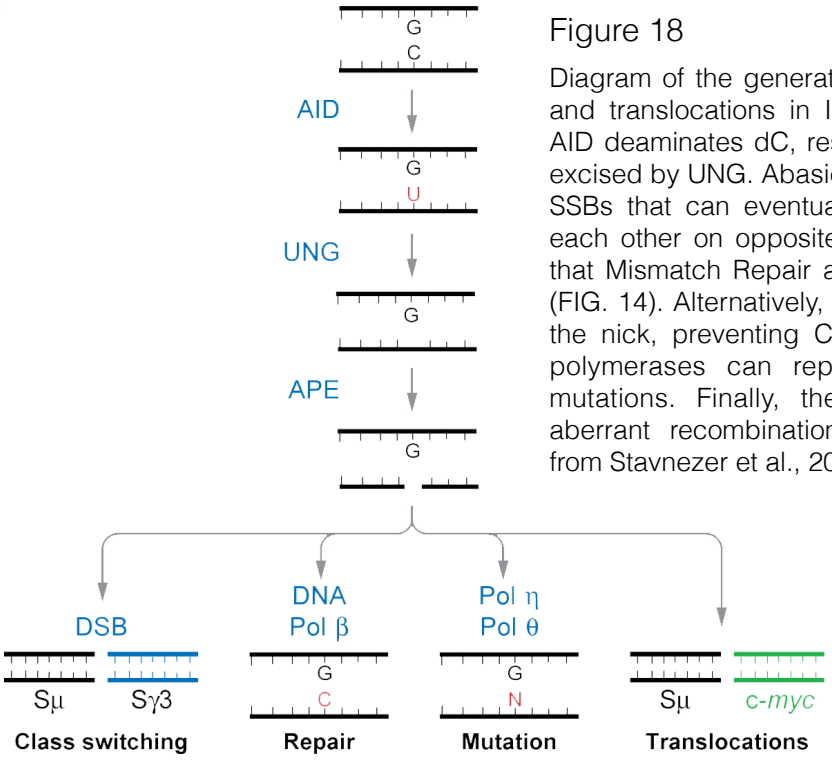

Figure 18

Diagram of the generation of DNA breaks, mutations, and translocations in IgS regions by AID-UNG-APE. AID deaminates dC, resulting in dU bases, which are excised by UNG. Abasic sites are cut by APE, creating SSBs that can eventually form DSBs when close to each other on opposite DNA strands, or in the case that Mismatch Repair activity converts them to DSBs (FIG. 14). Alternatively, DNA Pol β can correctly repair the nick, preventing CSR, or error-prone translesion polymerases can repair the nick by introducing mutations. Finally, the DNA breaks can lead to aberrant recombinations/translocations (Reproduced from Stavnezer et al., 2008).

Another repair pathway that contributes to CSR is Mismatch Repair (MMR). The role of MMR is to correct mis-incorporation of nucleotides during DNA synthesis (Kunkel and Erie, 2005), after recognition of the mismatch by the Msh2-Msh6 heterodimer (for short mismatches) or Msh2-Msh3 (for large ones). These heterodimers recruit the MIh1-Pms2 heterodimer, thus forming an heterotetramer, which recruits Replication factor C, the processivity factor proliferating cell nuclear antigen (PCNA) and exonuclease1 (Exo1). The complex excises the single strand segment flanking the mutated nucleotide (Genschel et al., 2002; Genschel and Modrich, 2003; Kunkel and Erie, 2005).

Some experiments demonstrate that knock-out for MMR genes in mice induces a relative reduction in CSR efficiency, depending on Ig isotype (Ehrenstein and Neuberger, 1999; Schrader et al., 1999; Ehrenstein et al., 2001; Bardwell et al., 2004; Li et al., 2004; Martomo et al., 2004). Indeed, MMR could be involved in CSR by converting SSBs generated by the AP endonucleases in DSBs (Stavnezer and Schrader, 2006; Schrader et al., 2007). In fact, while two nearby SSBs on opposite strands can spontaneously resolve in a DSB, when the SSBs are far away from each other, they are repaired more easily, without destabilising the region (FIG. 19). Several experiments confirm this idea: B cells with a deletion of the Sμ tandem repeats show a reduction of CSR, explainable with the reduction of AID targets (Luby et al., 2001).

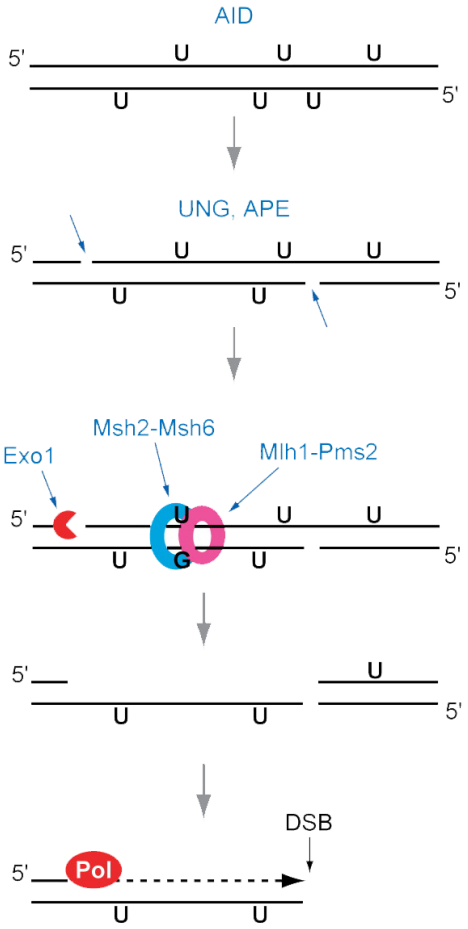

Figure 19

Model of the conversion from SSBs to DSBs by Mismatch Repair. AID introduces several dU residues in S regions during one cell cycle. Some of the dU residues are excised by UNG, and some of the abasic sites are nicked by APE. The U:G mismatches that remain can be substrates for Msh2-Msh6.
Msh2-Msh6, along with Mlh1-Pms2, recruit Exo1 (and accessory proteins) to a nearby 5' nick, from where Exo1 begins to excise the region towards the mismatch, thus creating a DSB with a 5' single-strand overhang, which can be filled in by DNA polymerases. Fill-in synthesis is probably performed by translesion polymerases owing to the presence of abasic sites (Reproduced from Stavnezer et al., 2008).

However when also MMR is abolished (such as in Msh2 or Mlh1 deficient mice) CSR is completely absent (Min et al., 2003; Schrader et al., 2007). Indeed, MMR-deficient B cells have a lower number of DSBs compared to wild-type cells (Schrader et al., 2007).

At the S-S junctions a number of mutations are present, both at G:C and A:T pairs (Dunnick et al., 1989; Schrader et al., 2003; Lahdesmaki et al., 2004). It has been hypothesized that A:T mutations are introduced by Pol η recruited by Msh2-Msh6 heterodimer (Wilson et al., 2005) both in not recombined Sμ and at S-S junctions after CSR (Faili et al., 2004; Delbos et al., 2007). On the other hand G:C mutations are mediated by DNA Pol θ (Masuda et al., 2005; Zan et al., 2005; Masuda et al., 2006) (FIG. 18).

Usually DSBs are produced in G2 phase or in S phase during DNA replication, they are repaired by homologous recombination in presence of homologous not damaged DNA. On the other hand CSR-related DSBs in S regions don't present sufficient homology for the HR repair pathway, and they are generated and resolved in G1 phase (Schrader et al., 2007) using NHEJ. Four proteins are essential for NHEJ,

Ku70 and Ku80, which bind together to DNA ends and improve the binding of the ligase complex XRCC4-ligase IV (Chen et al., 2000; Nick McElhinny et al., 2000; Costantini et al., 2007). Each of these proteins is essential for CSR (Casellas et al., 1998; Manis et al., 1998; Reina-San-Martin et al., 2003; Rooney et al., 2004; Ma et al., 2005; Pan-Hammarstrom et al., 2005; Sonoda et al., 2006; Soulas-Sprauel et al., 2007).

Ku deficient cells go in apoptosis after induction of CSR, and CSR is nearly ablated in this that survive (Reina-San-Martin et al., 2003). While XRCC4 or ligase IV deficiency impairs brain development, mice deficient in these genes have been created (Frank et al., 1998; Soulas-Sprauel et al., 2007; Yan et al., 2007) and people with hypomorphic mutations have been described. XRCC4 is important but not essential for CSR as showed in mice *xrcc4*$^{-/-}$ (Yan et al., 2007). Indeed mutations in ligase IV causes a lower number of blood cells in patients. Yet B cells can undergo CSR (Pan-Hammarstrom et al., 2005). Analysis of S-S junction in these mice and patients show an increased length of the junctional micro-homology (Pan-Hammarstrom et al., 2005; Yan et al., 2007), which is quite different from wild-type mice or humans in which S-S junctions show very little micro-homology (Dunnick et al., 1993). Thus, these data suggest that CSR uses NHEJ but can also occur by an alternative end joining (A-EJ) pathway that uses micro-homologies.

An important unresolved question regards the proximity of donor and acceptor S regions during CSR. Pre-association of the S regions could ideally increase the probability of correct S-S recombination. Transcription of these regions could pay a role in preparing this association. Indeed, during V(D)J recombination, V and J genes are brought closely each other through the loop formation in Ig locus (Roldan et al., 2005; Sayegh et al., 2005).

Other three important protein are Mre11, Rad50 and Nbs1, which constitute the MRN complex. This complex scans the DNA to find DSBs and consequently recruits the repair factors (Lee and Paull, 2005; Moreno-Herrero et al., 2005). Inactivation of in any of these proteins lead to aberrations in the chromosomes and chromosomal translocations (Reina-San-Martin et al., 2005). This complex acts downstream to Ku proteins (Lieber et al., 2003) and, a lower efficiency of CSR is present in splenic B cells from mouse with inactivation of the Nbs1 gene (Kracker et al., 2005; Reina-San-Martin et al., 2005). After binding of MRN complex ti DNA, through Nbs1, the kinase ataxia telangiectasia mutated (ATM) is activated, leading to accumulation of MRN

complexes, recruitment of other repair proteins and activation of cell-cycle checkpoints (FIG. 20) (Difilippantonio et al., 2005; Falck et al., 2005; Cerosaletti et al., 2006).

One of the aberrant outcome in CSR is the onset of chromosomal alterations (see below). One of the most commonly seen in murine experimental models is the one involving the Sμ region and the *c-myc* gene (FIG. 18)(Ramiro et al., 2004; Unniraman et al., 2004; Franco et al., 2006). The *c-myc-Igh* translocation is AID dependent and

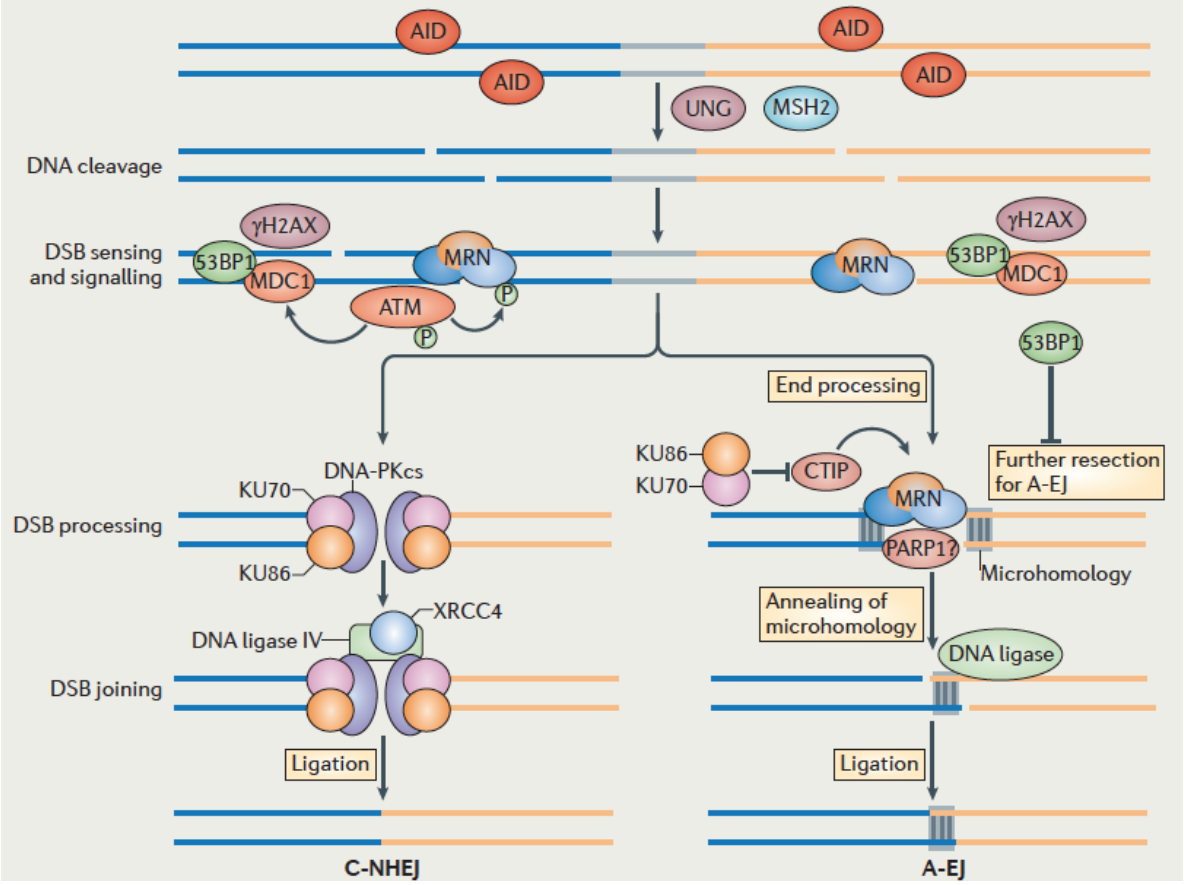

Figure 20

DNA damage sensors, adaptors and effectors in the DNA damage response have important roles in the DSBs resolution stage of CSR. The MRE11–RAD50–NBS1 (MRN) complex is an early sensor of AID-induced DSBs and localizes to damage sites, recruiting and activating the serine-protein kinase ATM. MRN also activates the γH2AX–MDC1–53BP1 pathway; phosphorylated histone H2AX (γH2AX) and p53-binding protein 1 (53BP1) mediate the synapsis of upstream and downstream DSBs. MRN can process DSB ends for both the classical non-homologous end joining (C-NHEJ) and the alternative end joining (A-EJ) pathway independently of the nuclease activity of MRE11 (a nuclease required for DSB repair mediated by HR). In C-NHEJ, DSBs are bound by KU70 and KU86, which form complexes with the DNA-PKcs and recruit other essential factors, such as the XRCC4–DNA ligase IV complex, to complete the end-joining process forming the junctions between the recombining switch (S) regions. In A-EJ, the DSB ends are processed by MRN and CTIP, generating microhomologies between the DSB ends. The choice of C-NHEJ is important for maintaining genome integrity, over A-EJ, is frequently associated with chromosomal translocations, depends on the expression of 53BP1, which protects DSB ends from resection for A-EJ, and on the presence of KU70 and KU86, which inhibit CTIP. Aberrant repair of DSBs can lead to chromosomal translocations and genomic instability. (Reproduced from Xu et al., 2012).

happens six times more frequently in Nbs1 hypomorph B cells than in wild-type B cells (Ramiro et al., 2004; Ramiro et al., 2006). These findings suggest that the MRN complex is involved in organizing the efficient and accurate S-S recombination.

ATM is a ser/thr protein kinase involved in DNA damage repair and cell-cycle control. It is recruited by MRN, and after accumulation in repair foci it repairs DSBs and initiates a cell-cycle check point (Bakkenist and Kastan, 2003; Downs et al., 2007). Indeed CSR is threefold reduced in splenic B cultured cells from mice *atm*$^{-/-}$ (Lumsden et al., 2004; Reina-San-Martin et al., 2004) and Sμ region translocates with *c-myc* eightfold more frequently than wild-type cells (Ramiro et al., 2006). Therefore ATM is important to repair DSBs and contributes to the correct positioning during CSR to obtain accurate S-S recombination, also blocking cells cycle in case of errors.

Another important factor for CSR is 53BP1, a transcriptional coactivator for p53 (Huang et al., 2007), which is involved in DSBs repair after ionizing radiation (Pryde et al., 2005). 53BP induces phosphorylation of ATM (Mochan et al., 2004), thus increasing MRN activity and its accumulation on DSBs. Cultured splenic B cells *53bp1*$^{-/-}$ display a reduction of 90% in CSR, although serum IgM levels are normal. This means that the reduction IN CSR is not due to decrease in cellular proliferation (Manis et al., 2004; Ward et al., 2004), and S-S junctions are normal. *atm*$^{-/-}$ cells deficient lacking 53BP1, while not showing an increase in chromosomal instability, display an higher number of aberrations at the IgH locus. This suggests a specific role for 53BP1 at this locus (Adams and Carpenter, 2006; Franco et al., 2006). In addition, differently from *atm*$^{-/-}$ cells, induction of CSR in *53bp1*$^{-/-}$ IgM hybridomas results in a three-fold increase in deletions within S regions compared to wild-type IgM hybridomas (Reina-San-Martin et al., 2007). Taken together, these results suggest a possible role of 53BP1 to bring close, or connect, Sμ and the downstream S regions (Manis et al., 2004; Adams and Carpenter, 2006). Yet, it still not known the possible mechanisms by which this could happen. Finally it is possible that 53BP1 foci induced by irradiation are mediated by RNA binding as they are disrupted by RNase treatment (Huang et al., 2007).

AID and cancer

Although in humans there is more or less the same number of B and T lymphocytes, more to 95% of lymphomas originate from B cells (Küppers, 2005). This could be related to aberrations deriving from the DNA damage induced by AID (Alt et al., 2013). In fact, whereas AID preferentially targets Ig light and heavy chain loci, it can also deaminates and induces DNA breaks in non-Ig genes (Liu et al., 2008; Robbiani et al., 2008; Hakim et al., 2012) as well as several oncogenes such as BCL6, Myc, MIR142, CD95 and others (Pasqualucci et al., 1998; Shen et al., 1998; Müschen et al., 2000; Tsai et al., 2008; Robbiani et al., 2009; Hasham et al., 2010; Chiarle et al., 2011; Klein et al., 2011; Hakim et al., 2012; Kato et al., 2012).

In addition ablation of AID significantly reduces formation of Ig-translocation and the development of B cell tumors in mice (Ramiro et al., 2004; Kovalchuk et al., 2007, 2012; Robbiani et al., 2008; Takizawa et al., 2008).

However the mechanisms through which AID target specific DNA sites. Two recent studies show that AID targets are present within regions containing super enhancers, characterized by chromatin accessibility and transcriptional activity an order of magnitude higher than other active sites (Parker et al., 2013; Whyte et al., 2013). Moreover this model could predict 91% of mouse AID targets, suggesting that some characteristics of these regions lead a nuclear microenvironment highly prepared to AID mediated deamination (Qian et al., 2014). According to this, another study found that many oncogene translocations in human B-cell lymphomas happen downstream of transcription start site, especially when the super enhancer regions lie within the transcribed regions (Meng et al., 2014). These data suggest that AID recruitment is mediated by SPT5, and RNA polymerase II stalling represents the link between AID activity and the transcriptional machinery (Peters and Storb, 1996). This, coupled with the ability of the RNA exosome that enables AID to deaminate both DNA strands (Basu et al., 2011), can explain how AID can generate mutations (Liu et al., 2008) and DSBs in many transcribed genes, throughout the genome (Staszewski et al., 2011; Hakim et al., 2012).

Open questions

Many questions are still unanswered with regards to AID activity and CSR. The main question, the one with the potential to explain the off-target activity of AID, regards the mechanisms and the factors responsible of the targeting of AID to the V and S regions.

Moreover, considering the involvement of AID in both SHM and CSR, it is not clear yet whether from a mechanistic perspective these processes occur simultaneously or specific factors are needed to determine if a cell undertakes CSR or SHM.

Finally, the C terminus of AID is important for CSR but not for SHM (Barreto et al., 2003; Ta et al., 2003; Shinkura et al., 2004). This seems to be related to its ability to recruit repair factors to the site of DNA damage. Is this domain of AD and its interactors related to the onset of aberrant outcomes in CSR?

Answering to these question could help understanding both the physiology of the antigen-driven antibody diversification processes, and - more importantly - the pathological consequences of AID disfunction.

Aims of the work

Visualization of Class Switch Recombination

AID is essential for the antigen-driven antibody diversification processes, as it initiates both Somatic Hypermutation (SHM) and Class Switch Recombination (CSR). Indeed AID deficiency causes Type 2 Autosomal Recessive form of the Hyper-IgM syndrome (Revy et al., 2000). However, despite its crucial physiological role, its mutagenic activity can lead to DNA damage and - eventually - cancer development (Okazaki et al., 2003; Ramiro et al., 2004; Rucci et al., 2006; Pasqualucci et al., 2008). More to 95% of lymphomas are B cell originated (Küppers, 2005) and these are often characterised by genetic alterations that can be traced back to DNA damage induced by AID (Alt et al., 2013): oncogenes usually mutated in these tumors feature a mutational signature similar to that observed in SHM, and most of the chromosomal translocations involve the IgH gene, thus suggesting a failure in the CSR process. Indeed there are hints that several layers of regulation limit AID mutagenic potential to the Ig locus, but these regulatory pathways are yet to be elucidated (Okazaki et al., 2003; Muto et al., 2006; Matsumoto et al., 2007; Shen et al., 2008).

In light of this, it is necessary to understand the factors that physiologically partake in the SHM and CSR process and elucidate whether failure of any element in this machinery can unleash the mutagenic potential of AID. While most of the factors involved downstream to the action of AID are linked to DNA repair, many of the factors involved in the targeting of AID to the IgH locus belong to pathways whose disruption is likely to heavily disrupt basic cellular processes such as transcription and mRNA processing. It is thus difficult to study the involvement of these factors through classical techniques such as reverse genetics. To overcome this problem we

aimed at developing a system to visualize the real time progress of CSR in live cells in order to be able to follow the action of the various factors involved.
To this aim we used CH12-F3 cells, a B-cell line from murine lymphoma, a cell line in which CSR from IgM to IgA can be induced through a specific stimulation cocktail (Nakamura et al., 1996). In this cell line we planned to visualise the IgH locus in order to follow it during CSR. We also aimed to follow transcription of the IgM transcript as a reporter for the progression of CSR. Thus, we have knocked-in an array of MS2-repeats (Rafalska-Metcalf and Janicki, 2007) in the 3' UTR of the Cμ region, which we tried to use in association with MS2-YFP to visualise IgM transcription.
On the other hand, in order to label the IgH locus, we are developing a pair of chimeric TALEs for bimolecular fluorescence complementation (Hu et al., 2002).

TALE-targeted Mutagenesis

Based on the ability of TALEs to recognize specific DNA sequences, we developed a TALE-AID chimera to target AID to specific sequences. With the development of such a system few opportunities would open up. On one hand, this approach could be used to to direct the evolution of specific genes by targeting only a selected region of interest. On the other hand, our approach could be used to generate cancer models to study the effects of mutations on specific oncogenes and tumor suppressors. Such approach would have few advantages over more classical approaches.
To this aim we have generated several chimeras in which a TALE designed to recognize the sequence of a reporter gene was fused to the AID and APOBEC1. We used this chimeras in conjunction with a construct encoding for a reporter gene: transfected cells would become GFP(+) upon targeted mutation of the reporter construct. Indeed we obtained a GFP(+) population when the reporter construct was cotransfected with the TALE-AID and TALE-APOBEC1, markedly higher than in the controls.

Enriching Genome Editing

Despite the power of the tools recently developed - TALENs and CRISPRs - genome editing is still hindered by the bottleneck of the selection of the targeted clones. Based on a recently developed approach (Kim et al., 2011a), we have modified it in order to obtain an enrichment of targeted clones without stable insertion of selection cassettes in the genome. To this aim we built a reporter cassette to select for cells in which genome editing tools are active. The reporter is based on a BlasticidinS-resistence gene (bsr) placed out-of-frame with an upstream mCherry coding sequence. The two genes are linked by a sequence homologous to the genomic target. Activation of the genome editing tool after transfection will target both the genomic target and the one present on the reporter construct. Targeting of this linker will bring the bsr in frame, thus providing a transient resistance to BlasticidinS. Treatment with this antibiotic will thus force the selection of the cells in which the TALEN/CRISPR have been active. We have tested our approach to obtain inactivation of the AID gene in CH12-F3 cells. Parallel experiments in presence or absence of a transient BlasticidinS treatment show that our approach can deliver up to 1/3 of the clones in which both alleles have been targeted.

Our approach appears feasible to improve the efficiency and the ease to obtain knock-in and knock-out using these genome editing tools.

Materials and Methods

Solutions and buffers

General standard buffers used in our work are listed here. Other specific buffers and solutions are explained in specific experiments.

Transformation buffer (TB)

Pipes:	10mM
$CaCl_2$:	15mM
KCl:	250mM
to pH 6.7 with 1M KOH	
and then $MnCl_2$:	55mM

10X TBE

Tris Base:	89mM
Boric Acid:	890mM
Na_2EDTA:	20mM g

Loading dye (6x)

Bromophenol blue:	0.25%
Glycerol:	30%

Church buffer

Sodium phosphate buffer:	0.25M
EDTA:	1mM
BSA:	1% (w/v)
SDS:	7% (w/v)

20x SSC

NaCl:	3M
Trisodium citrate pH 7.0:	300mM
to pH 7 with 1M HCl	
sterilize by autoclaving	

Denaturation buffer (Southern)

NaCl:	1.5M
NaOH:	0.5M g

Depurination buffer (Southern)

HCl:	0.25M

2X Cracking buffer

NaOH:	0.1mM
EDTA 0.5M	10mM
SDS 10%:	1%
Glycerol 100%:	10%

Trypan blue solution in PBS

Trypan blue	0.4%

General bacterial techniques

Bacterial strains

Bacterial strains used are listed in following table.

Strain	Genotype	Comments	References
DH5α	F- ϕ80d*lac*ZΔM15 Δ(lacZYA-argF) U169 endA1 recA1 sdR17($r_K^-m_K^+$) deoR thi-1 supE44 λ-gyrA96 relA1.	Host strain for general routine cloning.	(Invitrogen, 1986)
STBL3	F- glnV44 recA13 mcrB mrr hsdS20(rB-, mB-) ara-14 galK2 lacY1proA2 rpsL20 xyl-5 leu mtl-1.	Host strain for retroviral/lentiviral cloning.	Life Technologies
XL10-GOLD	endA1 glnV44 recA1 thi-1 gyrA96 relA1 lac Hte Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 tetR F'[proAB lacIqZΔM15 Tn10(TetR Amy CmR)]	Host strain for Quick Change II Site-Directed Mutagenesis Kit Stratagene	Stratagene
JM109	endA1 glnV44 thi-1 relA1 gyrA96 recA1 $mcrB^+$ Δ(lac-proAB) e14- [F' traD36 $proAB^+$ $lacI^q$ lacZΔM15] hsdR17($r_K^-m_K^+$)	Host strain for general routine cloning.	Promega

Media

Bacterial media LB (Luria-Bertani), SOB (LB medium + $MgSO_4$ 10%), SOC (SOB medium + Glucose 1M), and the corresponding plates were made following standard recipes (Sambrook and Russell, 2001). LB media was sterilized by autoclaving at 120°C for 20 minutes. SOB and SOC were prepared by adding the additional substance to the LB medium and next filtering by 0,2µm filters.

To select bacteria we used antibiotics to the following final concentrations:

Ampicillin	50-100 µg/ml
Kanamycin	50-100 µg/ml
Spectinomycin	50 µg/ml
Tetracycline	10 µg/ml

Preparation of chemically competent bacteria

Starting from an overnight culture in LB media, the bacteria were inoculated in 250ml of SOB at 37°C and grown until OD600=0.6 (measured by Biophotometer Eppendorf spectrometer). The bacteria were chilled on ice 10 minutes and spun down at 2500g for 10 minutes at 4°C. After resuspension in 80ml ice cold TB, they were chilled on ice for 10 minutes and again pelleted at 4°C 2500g for 10 minutes. They were then resuspended in 20ml of ice cold TB. DMSO was added to a final concentration of 7%. After incubation on ice for 10 minutes, bacteria were aliquoted in tubes, frozen in liquid nitrogen and then store at -80°C (Inoue et al., 1990).

Transformation of chemically competent bacteria

The chemically competent cells were mixed with 0.1-0.2µg plasmid DNA for 30 minutes to allow adsorption of the DNA to the cells. The mixture was underwent to a heat shock at 42°C for 45 seconds and immediately put on ice for 2 minutes. After addition of 1ml of SOC media the cells were grown at 37°C for one hour in order to let them express the antibiotic resistance. They were then gently pelleted and resuspended in 0.1-0.3ml of the remaining SOC, after removing the major part of the volume, and finally they were spread out on plates with antibiotic.

Storage of bacterial strains

An equal volume of freezing medium 2X (LB + 40% glycerol) was directly added to the bacterial culture. The bacterial strains were then stored at -80°C.

Blue/White colony screening

The blue/white bacteria colony screening was performed by adding 40µl of X-Gal solution (40 mg/ml) and 40µl of IPTG solution (100 mM) to each LB agar plate.

Cracking

This technique represents a very fast method to screen bacterial colonies without DNA extraction. Colonies were resuspended in 10µl of sterile ddH_2O. 5µl from the bacterial resuspension were added to 100µl of LB media or NaCl 0.9% (physiological solution) to maintain viable bacteria and eventually expand them later. The remaining 5 µl were added to 5µl of 2X cracking buffer to lyse the bacterial cells. Samples were then loaded on a 0.8% agarose gel not submerged in the running buffer. 1X TBE buffer was added up to the edge of the gel and gel was run. After few minutes, to let the samples in the gel, we completely covered the gel with buffer and the run continued. The DNA was then visualized at the end of the run on a transilluminator.

General Cell Culture techniques

Cell lines

The cell lines used in our work are listed in the following table.

Name	Derivation	References
HEK293T	Human Embryonic Kidney cell line	(DuBridge et al., 1987)
CH12-F3	Mouse B Lymphoma Cells cell line	(Nakamura et al., 1996)

HEK293T and NIH3T3 cell lines, were maintained in DMEM supplemented with 10% FBS, 2mM L-Glutamine and penicillin/streptomycin at 37°C with 5% CO_2.
CH12-F3 cells (kindly provided by Eva Severinson and Tasuku Honjo (Nakamura et al., 1996) were maintained in RPMI1640 supplemented with 10% FBS, 50µM β-mercaptoethanol, 2mM L-Glutamine, 1mM Sodium pyruvate and penicillin/streptomycin at 37°C in 5% CO_2.
Cell counting was performed using a Neubauer chamber cell counting.

Freezing and thawing of cells

Cells were harvested from healthy logarithmic cultures and spun down at 300g for 5 minutes. After removal of the supernatant, the pellet was resuspended in 1ml of sterile, ice cold freezing medium (90% FBS, 10% DMSO), the density of cells depended on the cell line. Cells were transferred into cryogenic tubes and frozen at -80°C. For long-term storage, the cryogenic tubes were transferred into liquid nitrogen. In order to thaw a cellular line, the vials were warmed to 37°C in a water-bath, thus the cells was transferred into a 15ml conical tube with 10ml of appropriate culture medium, were spun down at 300g for 3 minutes, resuspended and transferred into a culture flask or a Petri-dish with fresh tissue culture medium.

Class Switch Recombination Analysis

CH12-F3 cells at a density of $1x10^4$/ml were induced to perform CSR with TGF-β (2ng/ml), IL-4 (2µg/ml) and anti-CD40 antibody (0.5mg/ml). After 72 hours we analyzed the stimulated cells by FACS using the antibodies listed in the table:

Name	Source	Feature(s) and usage	Amount used
Anti-mouse IgM-FITC	Southern Biotech, Birmingham, AL	IgM Flow Cytometric Analysis	≤ 1µg/10^6 cells
Anti-mouse IgA-APC	Southern Biotech, Birmingham, AL	IgA Flow Cytometric Analysis	≤ 0.1µg/10^6 cells

Cre-recombination

CH12-F3 cells were plated at a density of 3-5x10^5/ml. We starved the cells overnight and then we added TAT-CRE (100µg/ml) (Peitz et al., 2002). After 1 hour we added chloroquine (100µM) for 30 minutes and, after washing the, we plated the cells in fresh media. We prepared a limiting dilution in order to obtain single clones to be tested for the success of the CRE treatment. In our case, we checked the loss of the puromycin resistance conferred by the antibiotic cassette. The TAT-CRE utilized in our experiments was a gift from Dr. Svend Petersen-Mahrt.

Transfection of eukaryotic cells

Transient transfections of HEK293T and NIH3T3 cells were performed using Lipofectamine LTX (Life Technologies) or Gene Juice (Novagen) following the procedures recommended by the manufacturer.

To electroporate CH12-F3 cells, 10^7 cells were washed, resuspended in 0,6ml of PBS and incubated with 25µg of DNA on ice for 10 minutes. The cells were the electroporated in a 0,4cm electroporation cuvette, in a Gene Pulser X-cells Total System (Biorad) using the following parameters: Voltage = 250V; Capacity = 500 or 950µF; Resistance = ∞. Finally the cells were transferred in a new flask with fresh media. Cells were plated in 96 wells in selection media after 48h.

Flow cytometry

About 2-5x10^5 cells were transferred into 5ml polystyrene round-bottom tubes, centrifuged at 1000g at 4°C, washed once in 2% FBS/PBS and spun down again. The cell pellet was resuspended in 0.1ml of 2% FBS/PBS containing the specific antibodies as required.

All buffers were equilibrated on ice before use, and incubations were carried out on ice, protecting the mix from light, to preserve the fluorescent reagents.

Following incubation for 30 minutes, cells were washed twice in PBS, re-suspended in 0.2-0.5ml PBS, and analyzed by FACS.

For the analysis of the reporter fluorescent gene expressions cells have been washed, resuspended in PBS and analyzed.

FACS analysis was performed on an Accuri C6 flow cytometer with a standard configuration (BD Biosciences, San Jose, CA) and data were analyzed by Accuri CFlowPlus analysis software.

DNA extraction from eukaryotic cells

DNA from eukaryotic cells was extracted by the Wizard Genomic Purification Kit (Promega) according to the manufacturer's instructions, and quantified with NanoDrop, UV spectrometer or by gel quantification.

RNA extraction from eukaryotic cells

RNA extraction was performed using TriPure isolation reagent (Roche) and RNA was purified from eventual DNA contaminations by the DNase I recombinant kit (Roche) according to manufacturer’s instructions.

Confocal microscopy

CH12-F3 were analyzed by confocal microscopy. After 3 PBS washes, cells were pulled down at 300g for 5 minutes and spotted on a cover slip, which was mounted on microscope slide.
The slides were analyzed using a Leica microscope SP2-AOBS, 516nm of laser excitation, 527-590nm acquisition window. Objective HCX PL APO lbd.BL 63.0x1.40 Oil, 512x512 scan size, voxel size (XYZ) 122x122x366nm.

PCR procedures

Oligonucleotides

Oligonucleotides, including PCR primers and sequencing primers, are listed in Appendix A and were synthesized by the company PRIMM Biotech Services.

Standard PCR

The following PCR protocol was used in all routine PCR, using AB analytical reagents.

Template DNA:	100 ng
10X ABtaq buffer:	5 µl
dNTPs mix (10mM each):	1 µ l
AB $MgCl_2$:	1,5 µl
Primer Forward (100 µM):	0,3 µl
Primer Reverse (100 µM):	0,3 µl
ABtaq (AB Analitica):	1 µl
ddH_2O:	to 50 µl

Cycling Step	Temperature	Time	Nº of Cycles
Initial denaturation	94° C	2 min	1
Denaturation	94° C	30 sec	35
Annealing	58° C	30 sec	
Extension	72° C	1 min/kb	
Final extension	72° C	5 min	1
Hold	4° C	∞	1

In order to screen bacterial colonies by PCR, the colonies were resuspended in 10µl of sterile ddH_2O. 5µl were added to 100µl of LB media or NaCl 0.9% (physiological solution) to maintain viable bacteria and eventually expand them later, while the others 2µl were used directly for the PCR reaction.
The initial PCR denaturation step was performed at 95°C for 5 minutes.

High-fidelity PCR

PCR amplification of fragments for cloning or sequencing were performed using Kod hot start DNA polymerase reagents (Novagen):

Template DNA:	50 ng
10X Buffer for KOD hot start DNA polymerase:	5 µl
dNTPs (2 mM each):	1 µl
25 mM $MgSO_4$:	3 µl
Primer Forward (100 µM):	0,15 µl
Primer Reverse (100 µM):	0,15 µl
Kod hot start DNA polymerase (Novagen):	1 µl
ddH_2O:	to 50 µl

Cycling Step	Temperature	Time	Nº of Cycles
Initial denaturation	95° C	2 min	1
Denaturation	95° C	20 sec	25
Annealing	57° C	10 sec	
Extension	70° C	15 sec/kb	
Final extension	70° C	5 min	1
Hold	4° C	∞	1

Touch-down PCR

To perform an amplification reaction to reduce non-specific products, we used a “touchdown” amplification program:

Cycling Step	Temperature	Time	Nº of Cycles
Initial denaturation	94° C	2 min	1
Denaturation	94° C	30 sec	10
Annealing	65° C (1°C each cycle)	30 sec	
Extension	72° C	1 min/kb	
Denaturation	94° C	30 sec	30
Annealing	55° C	30 sec	
Extension	72° C	1 min/kb	
Final extension	72° C	5 min	1
Hold	4° C	∞	1

Nested-PCR

Nested-PCR is an alternative PCR protocol used to reduce non-specific and to increase the efficiency of amplification, allowing detection of poorly represented targets in the sample. This technique involves two couples of primers, with a pair homologous for sequences internal to the other pair. Two rounds of PCR are performed, with the second reaction performed on 1 µl from the first PCR product.
The PCR screening for the knock-in of the MS2 array was performed by nested PCR using 45-46 primers for the first and the 47-48 primers for the second PCR. Forward primers were designed outside of the 5' arm and reverse primers on the puromycin cassette in order to obtain amplification only in the case of a targeted clone. These amplifications were performed using touchdown PCR protocol.

RT-PCR

Retrotranscription was performed using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystem).

Template RNA:	5 µg
10X RT buffer:	2 µl
25 X dNTP mix (100mM):	1 µl
10x RT random primers:	2 µl
RNase inhibitor:	1 µl
MultiScribe Reverse Transcriptase (100mM):	1 µl
ddH_2O:	to 20 µl

Cycling Step	Temperature	Time	Nº of Cycles
Step 1	25° C	10 min	1
Step 2	37° C	120 sec	1
Step 3	85° C	3 min	1
Hold	4° C	∞	1

The PCR to check the expression of MS2 cassette in 3' UTR of IgM mRNA, was performed using primers 53-54 on 2µl from the obtained cDNA.
In all these PCR techniques, both temperature and time of annealing can change depending on primers were used to amplify a specific target sequence.

Cross-over PCR

This method consists of two different PCRs to amplify the two DNA fragments that need to be joined or mutated. Special primers, containing 5' extensions complementary to each other, were used.
The products derived from the first pair of PCR reactions, were mixed to perform the second PCR. The second PCR was performed in absence of primers for the first five cycles (to extend the PCR products on each other). The external primers were then added for the remaining cycles (FIG. 21). The two PCRs were performed using the high-fidelity protocol and 18 PCR cycles for both PCR reactions.

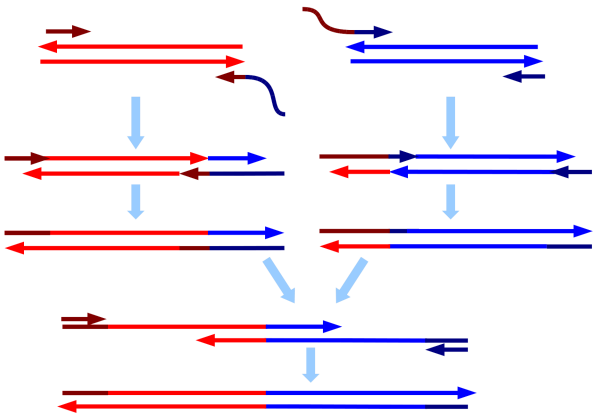

Figure 21

A graphical explanation of cross-over PCR techniques, to merge together two DNA sequences (red and blue).

Recombinant DNA techniques

Plasmid DNA preparation

Plasmids were prepared from bacterial cultures and in some experiments from transfected cell lines, using EuroGOLD Plasmid Miniprep Kit I (EuroClone) for

miniprep and High pure plasmid maxiprep kit (Life Technologies) according to manufacturer's instructions.

Phosphorylation

New England Biolabs T4 Polynucleotide Kinase (PNK) was used to phosphorylate annealed oligonucleotides according to manufacturer's instructions.

Restriction digestions

New England Biolabs endonucleases were used for restriction digestion, according to manufacturer's recommendations.
Screening of plasmids after minipreps were performed for 1 hour at 37°C.
Restriction digestions for cloning were performed overnight unless suggested differently by the manufacturer's recommendations.

Agarose gel electrophoresis

For analysis and purification of digested DNA fragments or PCR amplification products, samples were mixed with loading dye and loaded into agarose gel. Depending on the size of the DNA fragment to be analyzed, the gels contained 0.7% to 2.0% (w/v) agarose in 1x TBE buffer supplemented with EuroSafe Nucleic Acid Staining Solution (20,000x) (EuroClone), as recommended by the manufacturer, for the visualization of the DNA fragments. 5µl of 100-plus or 1kb or 1kb-plus ladder (New England Biolabs) were loaded in one lane of the gel for estimating the sizes and the concentration of DNA fragments.
Electrophoresis was carried out in a mini-gel apparatus (Biorad) in 1x TBE buffer, subsequently the gels were placed on a UV light-box, where the UV radiation at 254nm is absorbed by SYBR safe and re-emitted at 530nm in the green region of the visible spectrum, and photographed with a CCD camera.

DNA fragment and PCR purification

To clone DNA fragments, the bands were cut from gel and purified after visualization with blue-light source.
The DNA bands were excised from agarose gels using a clean scalpel, and the DNA was purified using Nucleospin Extract II kit (Macherey-Nagel), following manufacturer's instructions.
PCR samples were purified using the Nucleospin Extract II kit (Macherey-Nagel) as recommended from manufacturer.

Dephosphorylation

The phosphate group from 5' overhangs of cut vectors was removed by incubating the digested DNA with Shrimp Alkaline Phosphatase (Rapid DNA Dephos & Ligation Kit, Roche) according to manufacturer's instructions. The DNA was used directly for ligation reaction.

Ligation

For the ligation reaction we typically used 5 to 20ng of vector and the DNA insert in a 1:5 to 1:10 molar ratio with the vector. T4 DNA ligase (Rapid DNA Dephos & Ligation Kit, Roche) was used with the ligase buffer supplied in kit. Ligation reactions were incubated at room temperature for at least 30 minutes before bacterial transformation, according to manufacturer's instructions.

Blunting

To obtain blunt the ends of DNA fragments with 5'- or 3'- overhangs generated by restriction enzyme digestion we used Clone JET PCR Cloning Kit (Thermo Scientific) according to the manufacturer's instructions.

TOPO cloning

For direct cloning of PCR products amplified using Taq DNA polymerase the TOPO TA cloning kit (Life Technologies) was used. Otherwise Kod DNA polymerase PCR products were cloned by the Zero Blunt TOPO PCR cloning kit (Life Technologies). Both kits were used according to the recommendations of manufacturer.

DNA sequencing

DNA sequencing was performed by the following companies:
-PRIMM Biotech Services (www.primmbiotech.com)
-Macrogen (www.macrogen.com)

Southern blot

In this work we used southern blot analysis (Southern, 1975) to screen clones for gene targeting. The protocol used is as follows:

DAY 1

10 µg of genomic DNA, extracted by singles clones, was digested overnight with an appropriate restriction enzyme. The samples were run on agarose gel at 80 V for 5-6 h. After the run a photo of the gel was taken to keep track of the running distances of all marker bands. The gel was then incubated with gentle agitation in depurination buffer for 15-30 min. The gel was then incubated with denaturation buffer for other 30 min. The gel was then prepared for the blotting to the “tower” to transfer the DNA from the gel to nylon membrane overnight.

DAY 2

After disassembling the tower, the position of the wells on the membrane were marked and the membrane was cooked in a 3MM paper envelope at 80ºC for 2 h.

After a wash in 2X SSC buffer the membrane was prehybridized with 15-20 ml of Church buffer at 65ºC in the hybridizer (Techne) for 1 h.
The probe was prepared by PCR amplification using primers 43 and 44 on CH12-F3 genome and radiolabeled with ^{32}PCTP Amersham Ready-To-Go DNA Labelling Beads (-dCTP) (GE Healtcare). After purification with ProbeQuant G-50 MicroColumns (GE Healtcare) to eliminate not incorporated radioactive the probe was denatured.
The prehybridization buffer was substituted with new Church buffer and the radioactive probe was added. The hybridisation was performed overnight at 65ºC.

DAY 3

The radioactive buffer was removed and the membrane was washed:
2 washes with 2X SSC + 0,1% SDS for 10 min each
2 washes with 0.1X SSC + SDS 0.1% SDS for 20-30 min each.
Excess buffer was removed from the membrane using 3MM paper, and the membrane was exposed in the phosphoimager cassette after being enveloped in plastic film. The radioactive bands were detected with a phosphoimager Typhoon FLA 7000 (GE).

Plasmids and vectors

Plasmids and vectors used in this work are listed in Appendix B.

TALEs and TALENs

All TALEs and TALENs plasmids were generated by Golden Gate TALEN and TAL Effector Kit 2.0. The kit was a gift from Daniel Voytas and Adam Bogdanove (Addgene kit # 1000000024). In order to obtain several TALEs chimeras, the DNA binding domains, generated by the kit, were inserted in different final plasmid built by us (see Vectors and Plasmids construction section). For TALENs the final plasmid used was pCAGGS-TAL-NC2 (a gift from Takashi Yamamoto - Addgene plasmid # 43856). The TALEs and TALENs used in this work, are listed in the following table:

Name	Target sequence	Purpose
TALEN CH12 f	GTCTCTGTCACCTGCAG	to target the IgM gene in CH12-F3
TALEN CH12 r	CTGCTGTCCTTCCATGC	to target IgM gene in CH12-F3
TALE1	GGGATGGAGCTGGATCT	to bind V region in CH12-F3
TALE2	GCAGTTCCTGACAGGA	to bind V region in CH12-F3
TALE-AID	GAAAAACACGATAATACCG	to bring AID on IRES sequence
TALE-A1	GAAAAACACGATAATACCG	to bring A1 on IRES sequence
TALE-AID^{E58A}	GAAAAACACGATAATACCG	to bring AID^{E58A} on IRES sequence
TALE(EGFP)-AID	GCCCGTGCCCTGGCCCACC	to bring AID on EGFP sequence

CRISPRs

The CRISPRs plasmid used in our work was obtained by insertion of primers 49 and 50, annealed and phosphorylated, in pX330-U6-Chimeric_BB-CBh-hSpCas9 plasmid. This plasmid was a gift from Feng Zhang (Addgene plasmid # 42230). The sequence recognized in the AID gene is:

Name	Target sequence	Purpose
AID CRISPR	ACCATTTCAAAAATGTCCGC	to target the second exon of AID gene in CH12-F3 cells

Vectors and Plasmids construction

Targeting vector: The targeting vector to introduce MS2 array in the 3' UTR of the Cμ gene was planned to contain long homology arms, as suggested by previous works (Han and Yu, 2008; Han et al., 2010). We decided to use a 5' arm of 2.8 kb (292,907 - 290,084 on the murine chromosome 12, accession number: NT_114985) and a 3' arm of 6.8 kb (290,079 - 283,272). Another important element in the construct is the DTA cassette for negative selection of cells with non targeted integration (McCarrick et al., 1993) positioned externally to the 3' arm. Finally, a puromycin resistance cassette for positive selection was placed between the

homology arms, flanked by Lox sites. The 5' arm was obtained by PCR on CH12-F3 genomic DNA with primers 3 and 4, while the 3' arm was amplified using primers 1 and 2. DTA, puromycin and MS2 array cassettes was derived from the pROSA26, the pLoxPuro, and the p3216 plasmids respectively.

pLoxPuro 7bp, 10bp, 13bp, 16bp target plasmid: Target plasmids containing sequence to test TALE-half-YFPs proteins for the IgH visualization, were obtained by insertion of four couples of phosphorylated annealed oligonucleotides (Appendix A: 5-6, 7-8, 9-10, 11-12) in pLoxPuro plasmid, linearised by SpeI restriction enzyme. These plasmids contained four different lengths spacers, 7bp, 10bp, 13bp and 16bp, between TALE-half-YFPs target sequences.

pEGFP-TALE: In order to obtain visualize the IgH locus, we amplified from pCAGGS-TAL-NC2, the cassette for the last step of the Golden Gate TALEN and TAL Effector Kit 2.0, using primers 13 and 14 and we inserted it in pEGFP-N1 digested with SacII and BamHI.

pTALE-Jun-YN and pTALE-YN: From the pEGFP-TALE, we excised the EGFP through BamHI and BsrGI enzymatic digestion and inserted the Jun-YN and the YN fragments amplified from pBiFC-bJunYN using forward 15 or 16 and reverse 17 primers.

TALE1-Jun-YN and TALE1-YN: The plasmid obtained above were used in the last step of Golden Gate TALEN and TAL Effector Kit 2.0 to insert TALE1 central DNA binding domain (see TALEs and TALENs section).

pTALE-Fos-YC and pTALE-YC: Fos-YC and YC were amplified from pBiFC-bFosYC, using forward primer 18 or 19 and reverse primer 20, and inserted in the pEGFP-TALE cut by AgeI and AflII.

TALE2-Fos-YC and TALE2-YC: The plasmids above were used in the last step of Golden Gate TALEN and TAL Effector Kit 2.0 to insert TALE2 central DNA binding domain (see TALEs and TALENs section).

TALE2-Fos-YC-MODC: A 117bp (39aa) fragment encompassing the region 422-461 of the mouse ornithine decarboxylase degradation domain (MODC) (Li et al., 1998) was inserted in the TALE2-Fos-YC using the BsrGI site. MODC was amplified using primers 21 and 22 on genomic DNA from NIH3T3.

pBiFC-bFos: We modified pBiFC-bFosYC to shift YC out-of-frame by cutting (KpnI) blunting and religating.

pBiFC-bJunYN(Y66A): We mutagenized YN (Y66A) through crossover PCR techniques using 23 and 24 as internal primers, and 25 and 26 as external primers. The mutated YN was cloned in pBiFC-bJunYN substituting of the wt YN with the YN Y66A.

pTALE-AID and pTALE-A1: We excised the EGFP cds from pEGFP-AID and pEGFP-A1 by NheI and BsrGI enzyme. We then inserted in this plasmid a PCR-amplified cassette from the last step of Golden Gate TALEN and TAL Effector Kit 2.0 from pCAGGS-TAL-NC2, using 27 and 28 primers.

pTALE-AID^{E58A}: The AID mutant E58A was inserted in place of wt AID (HindIII and BamHI), by sub cloning it from pAID^{E58A}-Express Puro2 digested with HindIII and BglII. Finally, we the AID cds to obtain a control plasmid for our experiments.

TALE-AID, TALE-A1, TALE-AID^{E58A} and TALE(EGFP)-AID: The plasmids above were used in the last step of Golden Gate TALEN and TAL Effector Kit 2.0 to insert TALE(IRES) and TALE(EGFP) central DNA binding domains (see TALEs and TALENs section).

mCherry-IRES-EGFP and mCherry-IRES-EGFP WRC: This target plasmid, containing an array of ACG trinucleotides upstream to the mutated translation start site in the EGFP cds, was obtained in two steps. In the first step, using cross-over PCR, we inserted a unique BamHI restriction site just before the ATG of EGFP and we mutated this ATG into ACG. The cross-over PCR was performed using 29 and 30 as internal primers and 31 and 32 as external primers on pAID-Express Puro2. The PCR product was cloned in pEGFP-mCherry-EGFP cut by BsrGI. In the second step,

we inserted several pairs of oligonucleotides, annealed and phosphorylated, in the unique BamHI site. Primers 33 and 34 were used in order to obtain an array of ACG upstream and in-frame with ACG of EGFP. While, we used oligonucleotides (Appendix A: 35-36 and 37-38) with ACG interposed by GCA in the two possible orders to obtain arrays containing WRC motifs.

mCherry-IRES-EGFP w/o CMV and mCherry-IRES-EGFP w/o CMV-mCherry: The mCherry-IRES-EGFP was excised by NheI-BsmBI and by NdeI-BsmBI, and subsequently blunted and religated in order to obtain plasmids without the mCherry cds and the CMV promoter/mCherry cds, respectively.

mCherry-out-bsr: We inserted oligonucleotides 41-42 annealed and phosphorylated in the pBS-CMV-mCherry-EGFP digested with BsmBI enzyme to insert the target sequence for AID CRISPR (see CRISPRs section) and leaving the EGFP cds out-of-frame. Then we amplified the bsr gene from pLoxBsr using primers 39 and 40, and we cloned this fragment into pBS-CMV-mCherry-EGFP plasmid containing the CRISPRs target sequence, in place of the EGFP.

Results and Discussion

Visualization of Class Switch Recombination

AID is necessary to trigger the secondary diversification of the antibody genes, as it initiates both CSR and SHM (Muramatsu et al., 2000). Indeed its deficiency causes HIGM2 syndrome (Revy et al., 2000). However AID, due to its ability to induce DNA damage, is also well correlated with mutations and cancer (Okazaki et al., 2003; Ramiro et al., 2004; Rucci et al., 2006; Pasqualucci et al., 2008; Alt et al., 2013).

In CSR, AID deaminates cytosines in the S regions upstream to the constant regions that will undergo isotype switching. The uracils are eliminated by UNG, which leaves an abasic site, which in turn is excised by APE. Many factors are in the conversion of the lesion induced by AID to DSBs, as well as in the joining of Sμ donor and S acceptor regions to obtain change of isotype (Stavnezer et al., 2008).

Yet, many questions remain open, especially those regarding the targeting of AID and of the DNA repair pathways to the Ig locus. Moreover, most of the factors identified in recent years as involved in the targeting of AID to the IgH locus belong to pathways whose disruption is likely to heavily hamper basic cellular processes such as transcription and mRNA processing. It is thus difficult to study the involvement of these factors through classical techniques such as reverse genetics as their ablation could simply be lethal, or - on the other end - they can present an high degree of redundancy.

From a technical point of view, there are many methods to analyze CSR: PCR techniques, southern blot, or sequencing can assess the DNA/RNA level; western blot analysis, immunolabeling, and FACS to assess the protein level. These techniques give us information about quality and quantity of CSR, and they have been used to study the factors involved in CSR through evaluation of the changes in the levels of CSR linked to variation in these factors. However none of these

techniques can be used to address the exact timing of the CSR. Moreover, each of the assays not based on FACS imply the destruction of the sample in order to perform the analysis.

To overcome these issues we aimed at developing a system to visualize the real time progress of CSR in live cells in order to be able to follow the action of the various factors involved. In order to do so, we needed a reporter for the position of the IgH locus, as well as a marker to visualise the occurrence of the CSR.

To this aim we decided to use CH12-F3 cells, a B-cell line from murine lymphoma, a cell line in which CSR from IgM to IgA can be induced through a specific stimulation cocktail (Nakamura et al., 1996). This is the cell line from where AID was originally identified (Muramatsu et al., 2000; Revy et al., 2000).

We planned to use as marker for the CSR the transcription of the IgM as - upon completion of the CSR - the genomic region encoding for it would be excised. In order to visualise IgM transcription we planned to use an MS2 array in conjunction with a fluorescent MS2-Coat protein (Bertrand et al., 1998; Janicki et al., 2004; Shav-Tal et al., 2004). We originally planned to label the IgH locus with a Lac array that would be visualised by means of a fluorescent LacI chimera (Chuang et al., 2006; Thakar et al., 2006). Thanks to the availability of TALEs we switched to a plan involving the development of a fluorescent TALE to visualise the IgH locus.

Visualisation of IgM transcripts

One of the most commonly used techniques to visualise transcripts in live cells involves the use of the MS2 bacteriophage viral replicase translational operator. This sequence forms a 19 nucleotide RNA stem loop, which is recognized by the MS2 coat protein. The MS2 coat protein binds the stem loop as dimer tagging RNA (Beckett and Uhlenbeck, 1988). This specific pairing has been exploited using the RNA moiety together with a GFP-MS2 coat protein chimera (Bertrand et al., 1998; Janicki et al., 2004; Shav-Tal et al., 2004).

We thus decided to use this system to visualize the mRNA of IgM, and we aimed at inserting an MS2 array constituted by 24 MS2 repeats in the 3' UTR of the Cμ gene. This would have been followed by transfection with YFP-MS2 to visualize IgM mRNA.

The strategy for the knock-in of the MS2 array is depicted in (FIG. 22). The targeting vector to introduce MS2 array in the 3' UTR of the Cμ gene was planned to contain long homology arms, as suggested by previous works (Han and Yu, 2008; Han et al., 2010). We decided to use a 5' homology arm of 2.8 kb (292,907 - 290,084 on the murine chromosome 12, accession number: NT_114985) and a 3' homology arm of 6.8 kb (290,079 - 283,272). For the selection of the targeted clones we used a DTA cassette to reduce non targeted integration (McCarrick et al., 1993), and a puromycin resistance cassette for positive selection. The puromycin cassette was flanked by Lox sites, to allow its removal from the IgH locus.

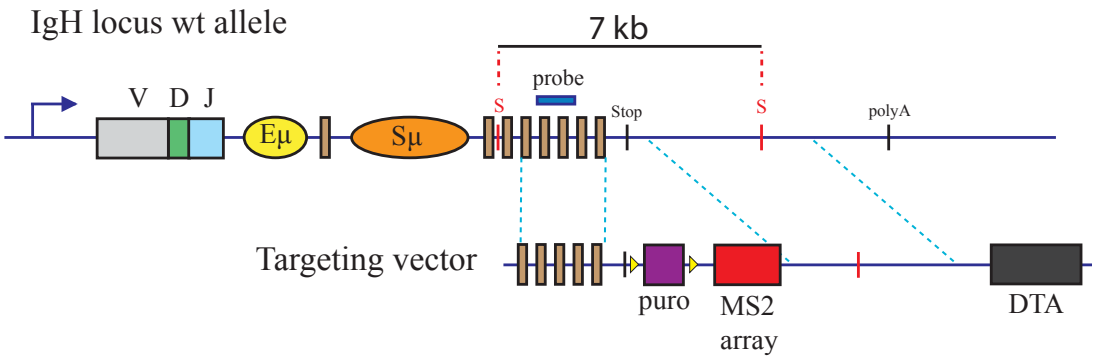

Figure 22

The schematic drawing represents the genomic organization of the murine IgH locus and the strategy for the knock-in. Brown boxes indicate the exons of Cμ and the small yellow triangles indicate lox P sites. DTA, diphtheria toxin A. Puro, puromycin resistance gene. Restriction enzyme sites: S, SacI. Using the probe (blue bar) we can visualize the expected DNA fragments by southern blot: a 7kb band for the wt allele and that a 11.6kb one for the targeted allele.

In each transfection 10 million CH12-F3 cells were electroporated with the linearized targeting vector. Transfected cells were related and treated with puromycin after 48h from transfection. To increase the frequency of targeting, we tried to change amounts of plasmid and cells, the electroporation protocol and also to use a PARP-1 inhibitor in order to boost Homologous Recombination (Bryant and Helleday, 2006). Puromycin resistant clones were picked in 10-15 days and after DNA extraction, were screened by PCR and southern blot analysis (see FIG. 23 for a representative blot). PCR screens and southern blotting were performed on ~120 clones. Unfortunately, no targeted clones could be found.

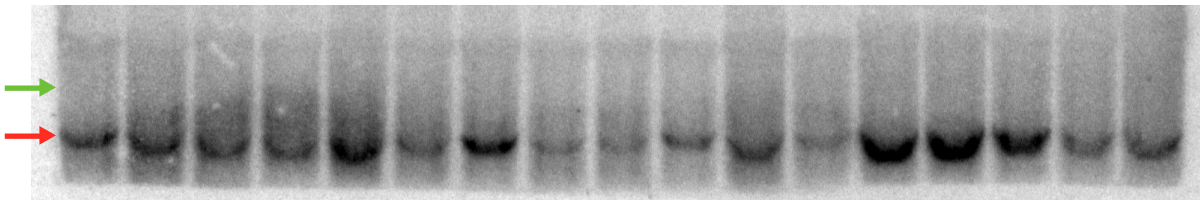

Figure 23

Representative Southern blot analysis of SacI-digested genomic DNA. The red arrow is positioned at 7kb and corresponds to wild-type allele. The green arrow indicates the expected position of the 11.6kb targeted allele.

Meanwhile, the seminal work on TALENs became available, together with the first kits to build specific TALENs (Cermak et al., 2011). We decided to use them to introduce a DSB in the 3' UTR of Cμ gene, as DSBs at a specific site had been demonstrated to increase the frequency of HR in presence of a donor DNA (Rouet et al., 1994a,b; Bibikova et al., 2003; Porteus and Baltimore, 2003; Urnov et al., 2005; Moehle et al., 2007; Voytas, 2013).

We thus designed and built a pair of TALENs (CH12 f and CH12 r) in order to target the 3' UTR of Cμ gene in a position internal to the homology arms of the targeting vector (to avoid targeting it). In the first experiment using the TALENs, we performed electroporation with 15μg of targeting vector and 5μg of each TALEN construct. We obtained 34 clones, but none targeted. Thus we decided to invert the proportion of the plasmids, 10μg for each TALEN construct and 5μg for targeting vector, under the assumption that the important step would have been the induction of the DSB at the targeted site. From this electroporation we obtained 18 clones, and six of them were found positive according to the nested PCR screen (FIG. 24).

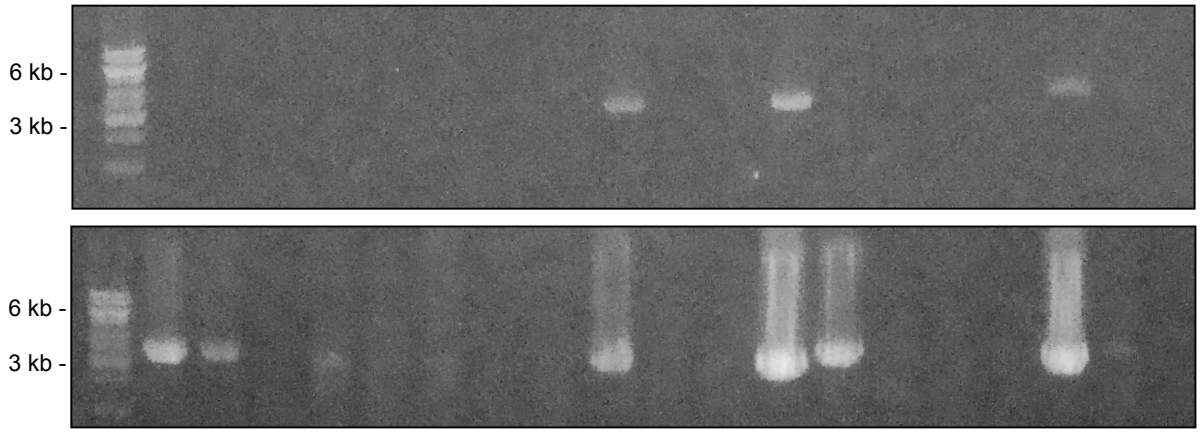

Figure 24

The photo shows the results of a nested PCR screen on selected CH12-F3 clones. The PCR using external primers evidenced three positive samples, at the expected size of ~3400bp (up). The nested PCR, using internal primers, showed six positive samples, at the expected size of 3170bp (down).

These six DNA fragments were extracted from gel and cloned in TOPO vector. A restriction digest revealed that only three of them displayed the expected pattern (clones 3, 4 and 6 in FIG. 25). The correct integration of targeting vector for these clones was confirmed by Sanger sequencing.

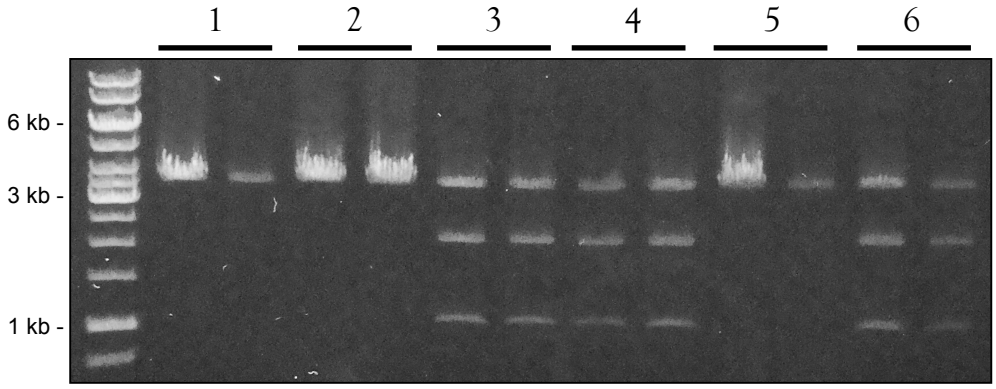

Figure 25

Enzymatic digestion (EcoRI and NheI) of the TOPO blunt plasmids containing the six bands obtained by nested PCR. The samples 3, 4 and 6 presented the expected pattern of digestion.

The next step was to analyze the targeted clones for the expression of the MS2 array. We thus performed RT-PCR to amplify the junction between the puromycin cassette and the last exon of the IgM transcript (we could not amplify MS2 due to its tandem repeats). The three clones expressed the MS2 cassette as shown in FIG. 26.

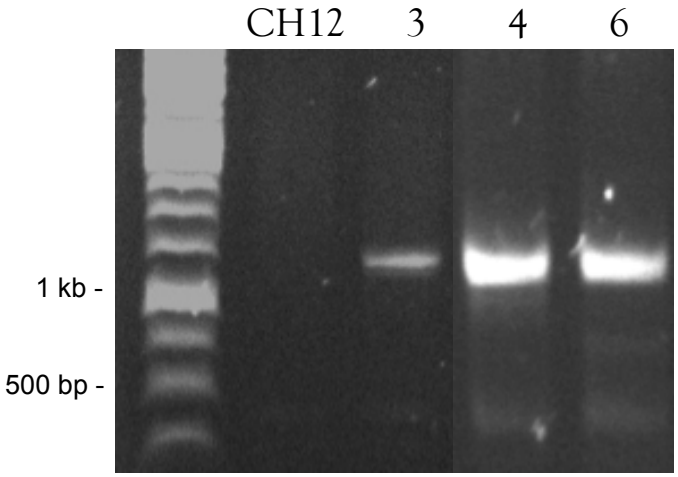

Figure 26

RT-PCR amplification confirmed the presence of our integrated cassette in the 3' UTR of the expressed IgM mRNA.

We then treated the clones with TAT-CRE recombinase protein to remove the puromycin cassette and we performed limiting dilutions in 96 well plates. Single clones were picked after 6-8 days and tested for resistance to puromycin. We selected clones from the original clones, which were then tested for their ability to perform CSR to make sure that the presence of the array did not hinder CSR. The

resulting levels of switching are showed in FIG. 27, suggesting that the ability to switch had been preserved.

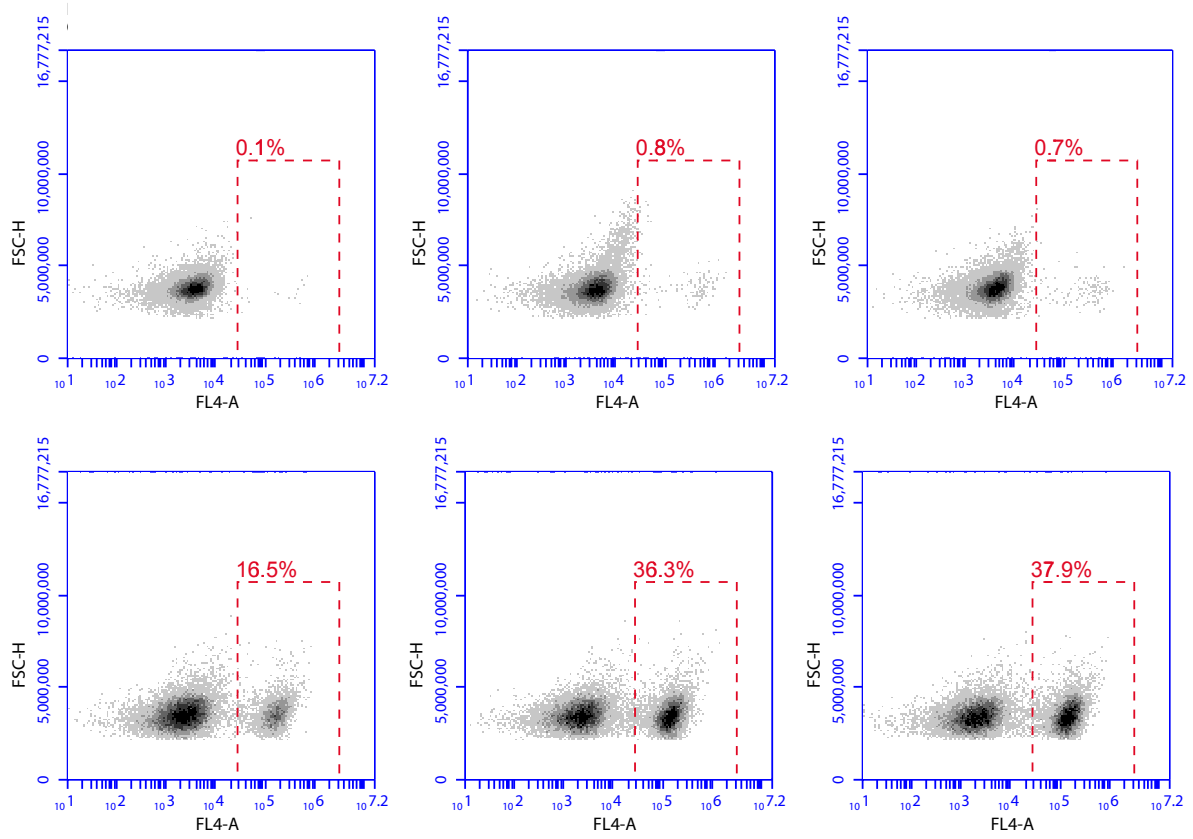

Figure 27

FACS analysis of selected CH12-F3 (MS2+CREtreated) clones that were stimulated in order to perform CSR. Upper blots show unstimulated clones. Lower blots show the same clones after stimulation.

We then chose the clone with the higher potential of CSR, and we transfected it with the CMV-YFP-MS2 plasmid. After selection of stable clones, we analyzed the selected clones by FACS for the expression of YFP-MS2 coat protein chimera. We found several clones that expressed the fluorescent protein, and we chose three clones displaying various levels of fluorescence, as showed in FIG. 28.

Finally we analyzed the clones by confocal microscopy in order to assess whether we were able to observe a nuclear yellow spot, indicating the expression of IgM mRNA. Unfortunately we could not find any, as the nuclei of the transfected cells resulted homogeneously yellow (FIG. 29). This could be due by two possible reasons. The first one is technical, we might have chosen clones expressing an excessive amount of YFP-MS2 coat protein. The second is intrinsic to the system: if transcription of the IgM locus is not sufficiently high, the mRNA could diffuse before a fluorescent spot could form.

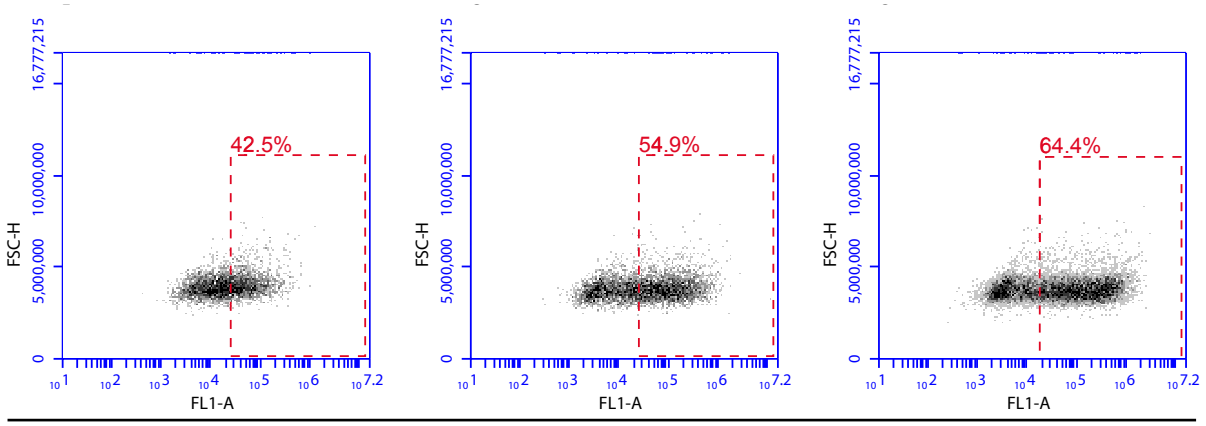

Figure 28

FACS analysis of CH12-F3 (MS2+CREtreated) clones stably transfected with CMV-YFP-MS2 plasmid, with different expression level of YFP-MS2 coat protein.

Using FRAP (Fluorescence Recovery After Photobleaching) analysis we observed a very rapid diffusion of the YFP-MS2 chimera. In fact, as showed in FIG. 30, the effect of photobleaching disappeared within 0.5-0.7s, suggesting that proteins were highly mobile, probably correlated with the high amount of protein in the cells nuclei. We cannot be sure of the specific reason for the failure of the assay. However, after analysis of clones with different expression levels of YFP-MS2 chimera, we are reasonably convinced that diffusion of IgM mRNA could be the main problem.

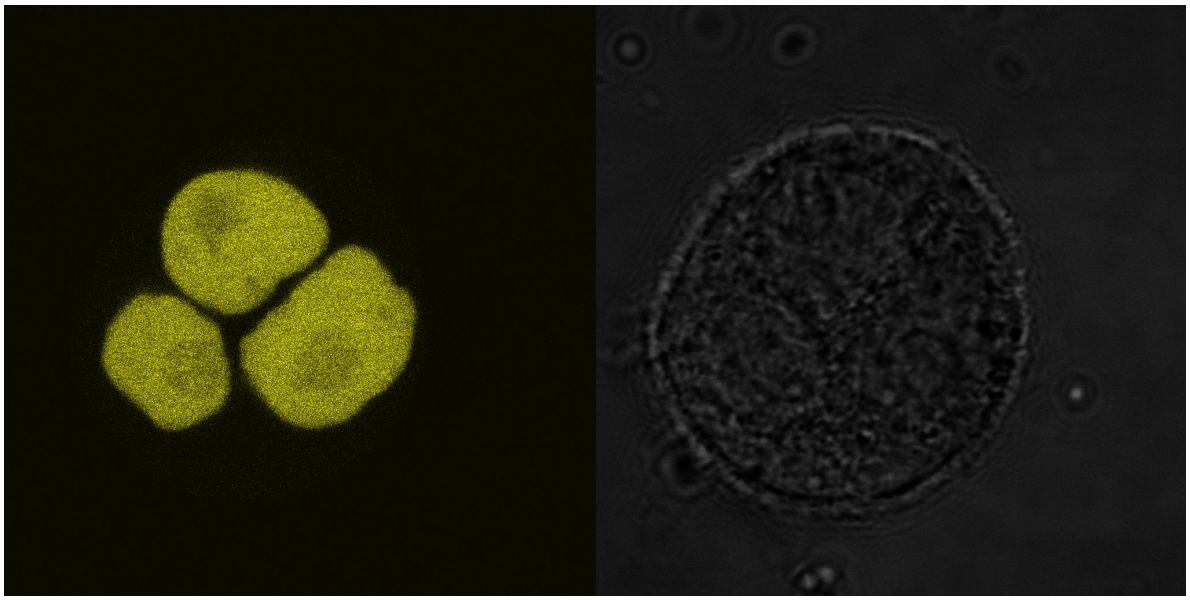

Figure 29

Fluorescence analysis of CH12-F3 clones by confocal microscopic shows cells with homogeneously fluorescent nuclei.

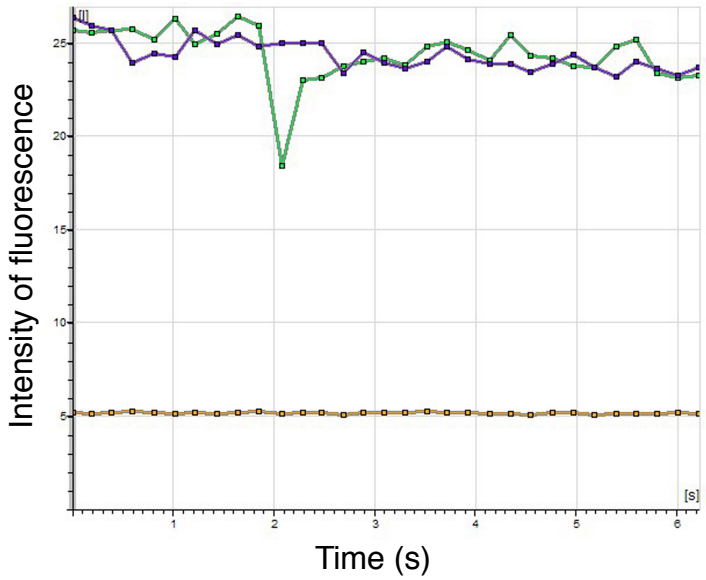

Figure 30

FRAP analysis shows the time needed to repopulate the photo bleached area. The short time needed reveals the high mobility of the fluorescent protein. The yellow line represent the fluorescent background; the purple line the fluorescent in any other point of the yellow cell, and the green represents the photo-bleaching event.

IgH locus visualization

The initial plan for the visualisation of the IgH locus involved the targeting of the locus with a Lac array. However, thanks to the availability of the TALEs, we decided to try a different approach involving the development of a fluorescent TALE.

To visualize the nuclear position of the IgH locus we decided to utilize two TALEs bound to two half-YFPs (Hu et al., 2002) that become fluorescent only upon dimerization. We originally designed the constructs to clone the N-terminal and C-terminal half of YFP at the C-terminus of the TALE cds (TALE1-YN and TALE2-YC). The Jun and Fos cds used in the original study were used as a linker between the TALE and the YFP cds, to facilitate dimerization (Hu et al., 2002). The two TALEs were designed to bind two different sequences located near each other on opposite strands of the IgH locus. In this way when the two TALE chimeras would bind to their target sequences, the two half-YFP could interact with each other and induce fluorescence at the IgH locus. In order to increase the signal intensity, TALEs were designed to recognize repeated sequences in promoters of the V regions upstream of V(D)J region rearranged in CH12-F3. We thus identified the V region used in CH12-F3 and we aligned, using Seaview (Gouy et al., 2010), all V regions left upstream to the rearranged one, locating a possible target sequence that was perfectly repeated eleven times and had partial conservation other seven ones. (though these repeats could have been of limited use, as some papers had shown

that a single point mutation could compromises DNA binding of the TALEs (Miyanari et al., 2013).

We initially aimed at validating such approach. We therefore prepared a plasmid construct containing the selected TALEs target sequences connected by linkers of different length. We then transiently transfected this plasmid together with the pair of TALE-Jun/Fos constructs in HEK293T cells. We analyzed the samples by FACS 48h after transfection and results are shown in FIG. 31. Indeed an increase in fluorescence was present in cells transfected with the target sequence compared to the cells transfected with a control plasmid. Yet, the presence of a fluorescent population also in absence of target sequence suggests that the TALE1-Jun-YN and TALE2-Fos-YC can reconstitute YFP spontaneously.

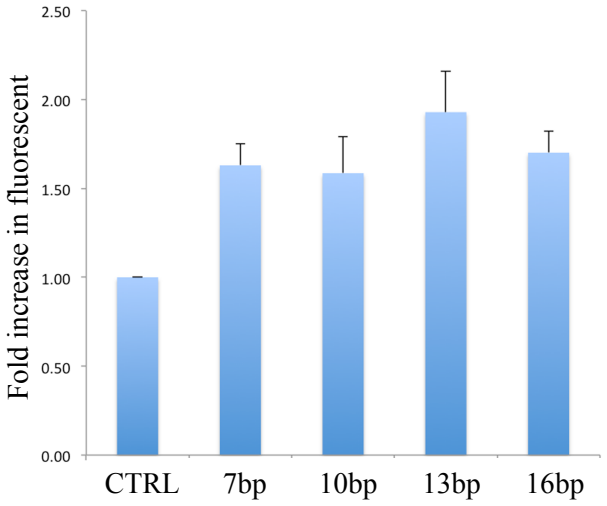

Figure 31

The bar diagram shows the differences in fluorescence in HEK293T cells transiently transfected using TALE1-Jun-YN and TALE2-Fos-YC together with plasmids containing the target sequences separated by spacers of different length. In the Control sample the target plasmid is an empty vector.

In order to reduce this background we removed the Jun and Fos cds from the TALE chimeras and we repeated the experiments. As shown in FIG. 32, we obtained a decrease in total signal but the background remained high. In addition we observed that analysis at 24h showed higher differences of levels of fluorescence than at 48h (FIG. 33), thus suggesting that the first hours after transfection were important to determine the different levels of fluorescence.

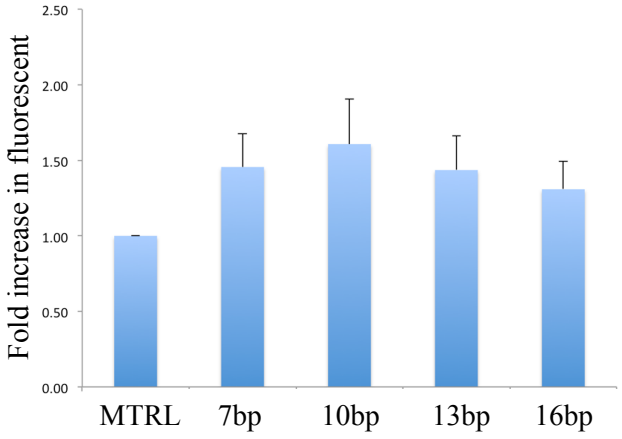

Figure 32

The bar graph depicts the differences in fluorescence in HEK293T cells transiently transfected using TALE1-YN and TALE2-YC together with plasmids containing the target sequences separated by spacers of different length. In the CTRL sample the target plasmid is an empty vector.

To better understand this point, we

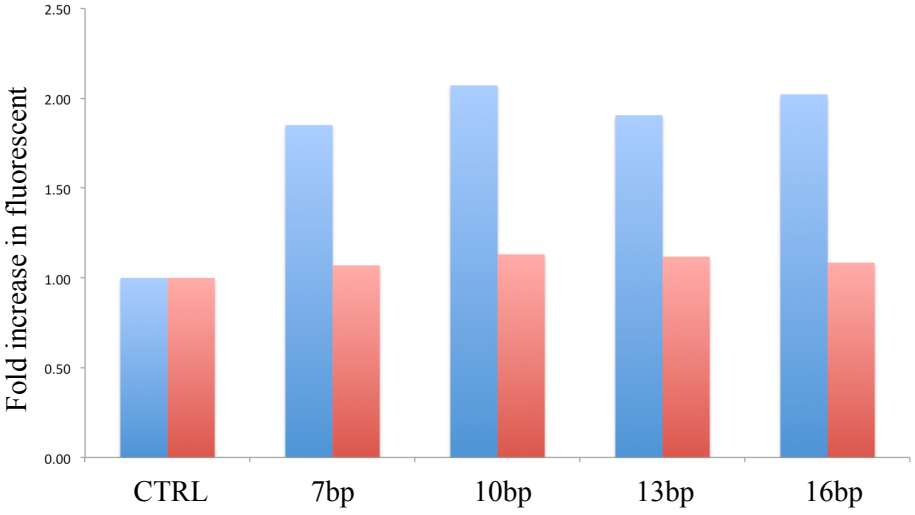

Figure 33

The histogram shows the differences in fluorescence in HEK293T cells transiently transfected using TALE1-Jun-YN and TALE2-Fos-YC together with plasmids containing the target sequences separated by spacers of different length. The coloured bars indicate the time from the transfection, 24h (in blue) and 48h (in red). In the CTRL sample the target plasmid is an empty vector.

performed the experiment in two steps: we started by transfecting HEK293T cells with the plasmids encoding the TALE chimeras, and we performed a second transfection with plasmid containing the target sequences after 24h. FACS analysis was performed after 6h, 12h, and 24h but the results were essentially the same. As showed in FIG. 34 there are almost no differences between cells transfected the second time and cells transfected only with half-YFPs. This confirmed our idea that the differences we observed in the initial experiments were due to an initial stage, in which dimerization of the few synthesized TALE chimeras was facilitated by the presence of the target sequence. At later stages, when higher levels of TALE chimeras were present, the signal from spontaneous dimerization would overnight that specific signal.

We attempted to decrease the background using two other approaches. In the first one, we saturated the aspecific signal with non-fluorescent interacting molecules such as the bare Fos or a Jun-YN with an inactivated chromofore (Y66A). We also tried to add the region 422-461 of the mouse ornithine decarboxylase degradation domain (MODC) (Li et al., 1998) to the TALE-Fos-YC/YC plasmids in order to reduce the half-life of these chimeras. Unfortunately there were no changes, and the background noise remained too high.

Our data suggest that the essential problems in our approach, causing the high background, depends on three factors: (a) the TALE chimeras are able to demerit

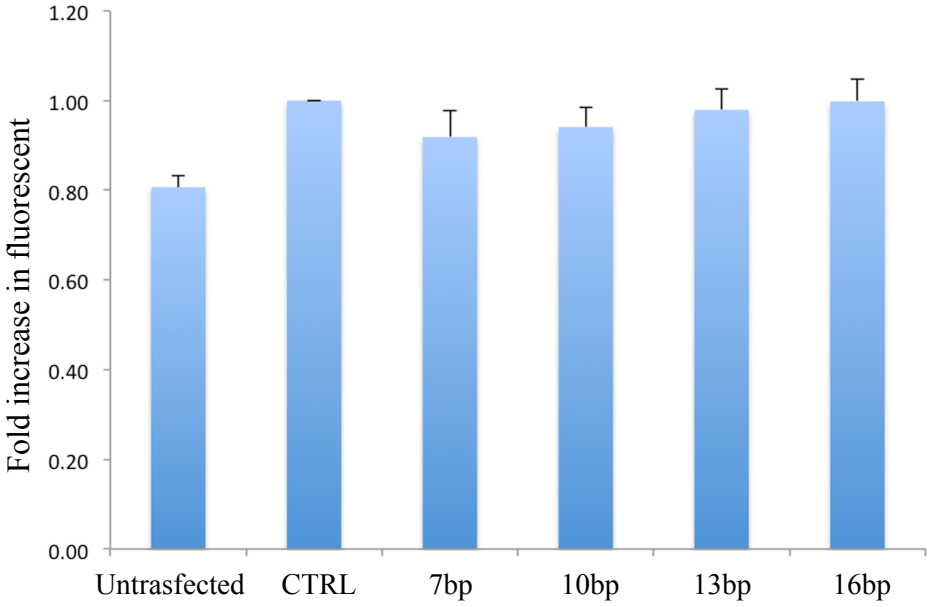

Figure 34

The histogram shows the results of the two-step experiment using transiently transfeced HEK293T cells. In the first step we transfected the TALE1-YN and TALE2-YC, After 24h the cells were retransfected whit the target plasmids. In the CTRL sample the target plasmid is an empty vector (transfected in the second round), while the untrasfected sample derives from cells transfected only in the first round.

spontaneously; (b) the binding of the TALE to their targets is dynamic, thus leading to (c) an further increase in the background due to the stability of the YFP heterodimer. All this suggests that far-reaching changes to the approach will be needed in order to to obtain an usable system. One possible change could target the stability of the YFP heterodimer.

TALE-targeted Mutagenesis

Based on the ability of TALEs to recognize specific DNA sequences, we aimed at developing a tool to target mutagenesis to specific regions of the genome. Such a system could be used both as a biotechnological tool, to direct the evolution of specific genes of interest and as a tool to model cancer.

To this aim we merged the mutagenic activity of AID and DNA specificity of TALEs through the construction of a chimeric TALE-AID. The TALE domain of our chimera was designed to bind the last 19bp of an IRES sequence inserted within a modified mCherry-IRES-EGFP cassette in a target plasmid. The cassette was modified by mutagenizing the start codon of the EGFP cds from ATG to ACG. We further inserted, upstream to the ACG, an array of ACGs, in frame with the EGFP cds, just downstream to the target sequence of the TALE. The ACG array had two purposes: (a) to provide redundancy to the possible sites that could be mutated by AID and (b) to accommodate for position of the TALE-AID on the DNA, as we could not model the steric hindrance of the TALE-AID chimera. The mCherry cds would serve as a reporter for the transfection efficiency, whereas only mutations at any of the ACGs would restore the ATG start codon and let the cell synthesis EGFP.

We thus co-transfected the target construct with or without the TALE-AID in HEK293T, and the cellular fluorescence was analysed from the time ranging between 24 hours to 96 hours after transfection by FACS . As shown in FIG. 35, the presence of TALE-AID induced a marked increase in GFP(+) cells, ~1% of transfected cells, as measured by mCherry fluorescence, or ~0.5% over the entire cell population.

It is noteworthy that even in the mock-transfected cells there was a small percentage of with a very low GFP(+) phenotype. This background fluorescence could probably be due to a couple of CTG codons downstream to the EGFP canonical start codon.

In order to further our analysis, we prepared a chimera replacing AID with APOBEC1 (TALE-A1), another AID/APOBEC member which ca target DNA as well (Harris et al., 2002; Petersen-Mahrt and Neuberger, 2003; Severi et al., 2011; Saraconi et al., 2014).

We repeated the experiments adding both TALE-A1 as well as untagged AID and APOBEC1, which should not be targeted specifically to the sequence. As expected, the untagged AID/APOBEC1 were not able to induce an increase in GFP(+) cells. On

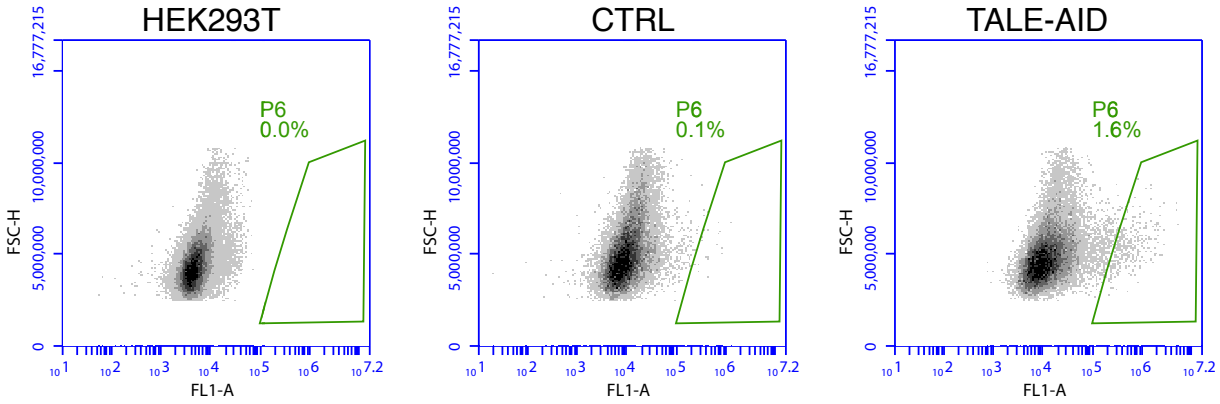

Figure 35

FACS analysis of HEK293T cells transiently transfected with mCherry-IRES-EGFP and TALE-AID. The CTRL sample was transfected with mCherry-IRES-EGFP and an empty vector.

the other hand TALE-A1 induced just a moderate effect (FIG. 36). The result could point to the importance of the preference context for the activity of the DNA mutators. Whereas APOBEC1 seems more mutagenic than AID (Harris et al., 2002; Saraconi et al. 2014), the preference context for the two deaminases is different: AID preferentially mutates cytosines within a WRC context (W=A/T; R=G/A; C=targeted base), on the other hand APOBEC1 prefers a TC context while avoiding AC dinucleotides. Since the ACG array contains Cs in a context is closer to that preferred by AID, this could explain the result.

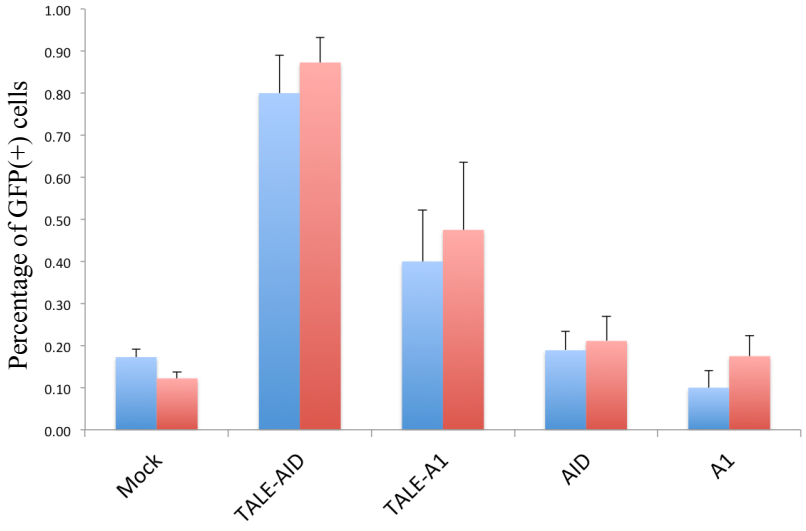

Figure 36

The bar graph shows the percentage of GFP(+) HEK293T transiently transfected using the mCherry-IRES-EGFP and plasmids encoding for TALE-AID, TALE-A1, or untagged APOBECs. The color indicates the time after transfection: 72h (blue) and 96h (red).

In addition we repeated the transfections with scalar amounts of target plasmid, paired to either TALE-AID or TALE-A1. As shown in FIG. 37 the background was reduced. On the other hand, the increment in the GFP(+) cells in presence of either AID or A1 was proportional to the quantity of the target plasmid. This suggests that the amount of target plasmid is the limiting factor.

We then attempted to sequence the transfected plasmids in order to identify the mutations. We thus prepared minipreps to obtain the plasmid from the transfected cells, and after PCR amplification of the IRES-EGFP fragment we sequenced it using the Illumina miSeq technology. Unfortunately the run failed and the samples could not be analysed.

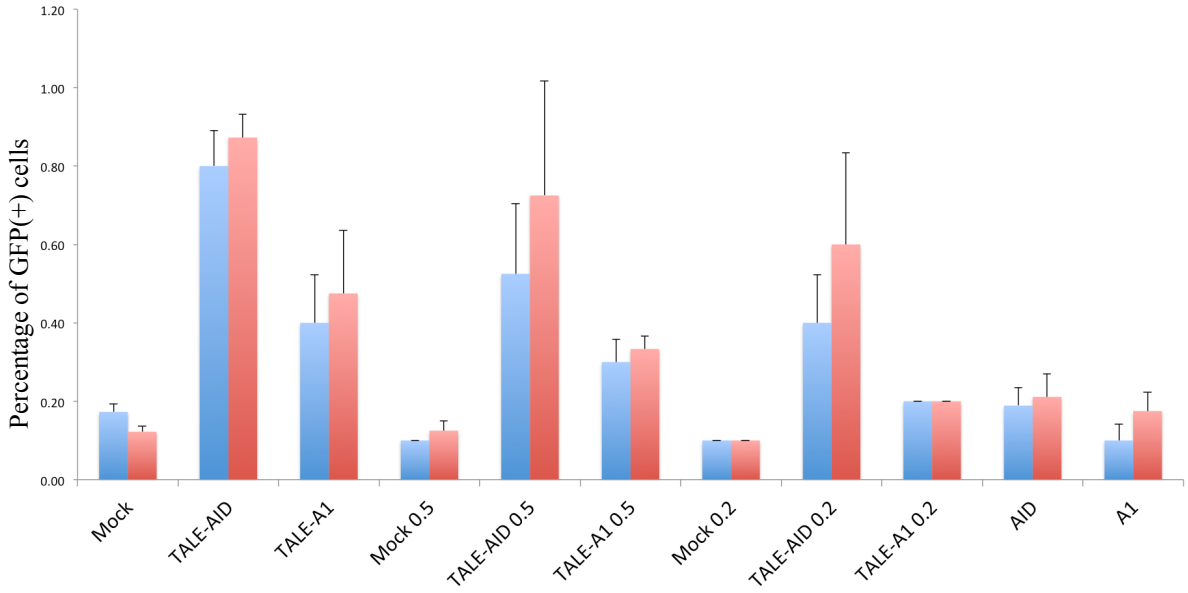

Figure 37

The histogram shows the percentage of GFP(+) cells in of HEK293T transiently transfected using different amounts of mCherry-IRES-EGFP together with plasmids encoding for TALE-AID and TALE-A1, or untagged APOBECs at 72h (blue) and 96h (red).

To make sure that the observed effect was mediated by the mutagenic activity of TALE-AID, we repeated the experiment with a chimera TALE-AID^{E58A}, an AID mutant in which the glutamate acting as a proton donor in the deamination reaction was mutated. Surprisingly, as shown in FIG. 38**,** TALE-AID^{E58A} induced an increment in the GFP(+) population similar to that observed with the TALE-AID. We initially thought that the effect might be due to the TALE itself, but this was clearly not the case (FIG. 38).

The result obtained is puzzling, as it seems that AID can trigger expression of the EGFP through a non-enzymatic non-mutagenic route.

There is a strong correlation between AID and transcription. Indeed It has been shown that AID is targeted to genes that are highly transcribed (Yang and Schatz, 2007; Alt et al., 2013; Buerstedde et al., 2014; Storb, 2014).

It could be possible that it could also be the opposite, with the presence of AID inducing transcription. Indeed, the GFP(+) population lied within the cells which showed higher than average transcription of the mCherry promoter (FIG. 39).

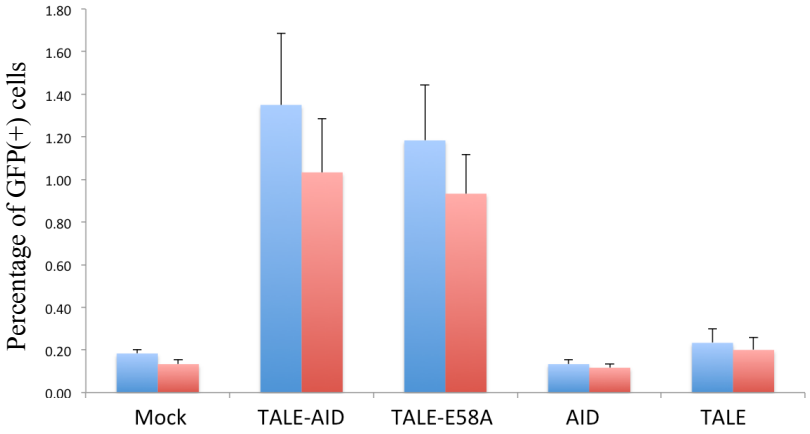

Figure 38

The histogram shows the percentage of GFP(+) HEK293T transiently transfected with the mCherry-IRES-EGFP and with plasmids encoding for TALE-AID, TALE-AID^{E58A}, untagged AID, and TALE alone at 72h (blue) and 96h (red). Previous experiments had shown a peak of the signal at 96h, whereas these one peak at 72h. This is due to a change in the transfection procedure (starvation of cells throughout the transfection vs starvation only before transfection). This modification in the protocol led to a change in the timing of plasmid expression.

To test whether the presence of AID could modulate/drive the transcription of the EGFP we prepared constructs in which the promoter was removed from the mCherry, or a wild-type EGFP cds was cloned in absence of a promoter or juxtaposed to the IRES sequence. No transcription of the EGFP could be observed either with TAL-AID, TALE-AID^{E58A} or any control.

On the other hand we tested whether the position of the targeting was relevant for

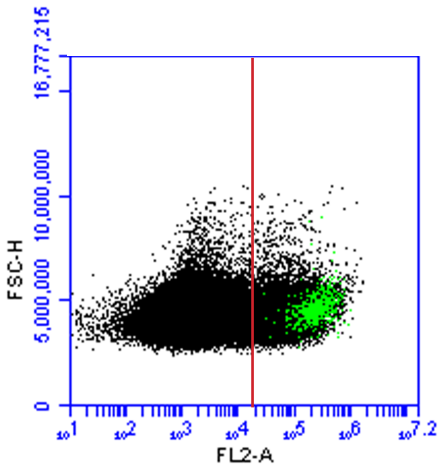

Figure 39

The FACS plot shows the position of the GFP(+) cells within the mCherry(+) cells that showed higher than average transcription of the mCherry promoter. The red bar indicates the threshold for the green fluorescence.

the effect. We thus prepared new TALEs specific for a central region of the EGFP cds, at least 110bp downstream either the canonical start codon or of the potential CTG codons. In this case the effect was restored, suggesting that the position of the target did not matter (FIG. 40).

All these experiments suggest that, while transcription is needed in order for TALE-AID to affect the GFP(+) population, the transcription itself does not seem to be

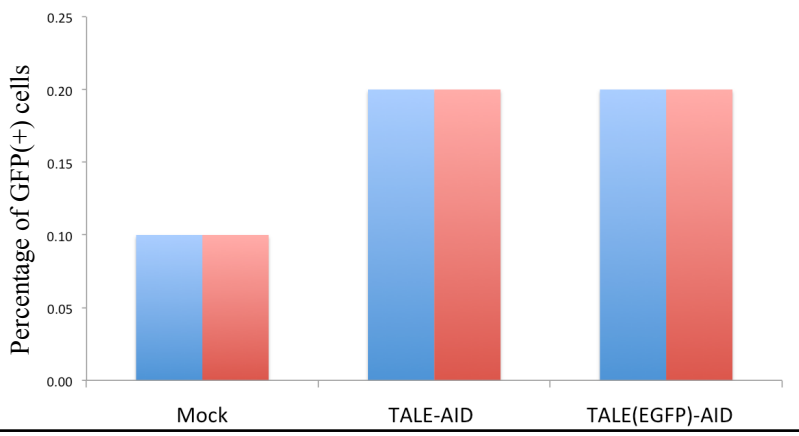

Figure 40

The histogram shows the percentage of GFP(+) HEK293T transiently transfeced using the mCherry-IRES-EGFP and plasmids encoding TALE-AID and TALE(EGFP)-AID at 72h (blue) and 96h (red).

modulated, nor the assembly of the translation machinery seems to be involved. The observation that the effect is visible regardless of the position of the target sequence might suggest that the observed effect could be due to the interaction of AID with the transcript itself.

Finally, we tested whether we could increase a mutagenic activity of AID on the target site by modifying the ACG array. We therefore replaced the ACG array with one in which the ACG trinucleotides were placed in a canonical AID context (GCAACG). Also in this case an effect due to TALE-AID^{E58A} was visible, yet the signal for TALE-AID was slightly higher, suggesting that this difference could be due to enzymatic deamination by AID (FIG. 41).

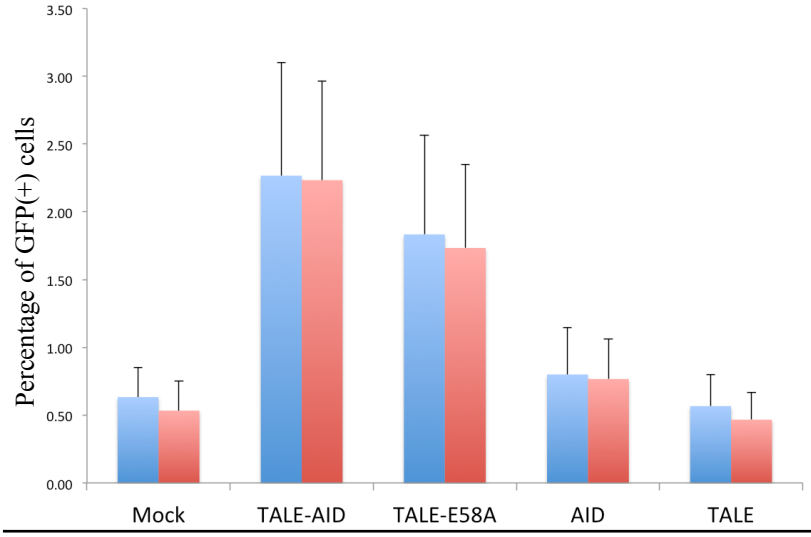

Figure 41

The histogram shows the percentage of GFP(+) HEK293T cells transiently transfected using the mCherry-IRES-EGFP WRC target plasmid and plasmids encoding for TALE-AID, TALE-AID^{E58A} , untagged AID orTALE alone at 72h (blue) and 96h (red).

Enriching Genome Editing

Gene targeting is the most direct way to dissect the role of genes in the cellular/ organismal environment (Capecchi, 2005). Development of nucleases - TALENs and CRISPRs - to target specific DNA sequence and thus induce the cellular DNA repair mechanisms, including NHEJ and HDR, has greatly increased the ease to obtain knock-out and knock-in in cellular and animal models. Despite the power of these tools, genome editing is still hindered by the bottleneck of selection of the targeted clones, and targeting efficiency is sometimes too low. Based on a recently developed approach (Kim et al., 2011a), we have modified it in order to obtain an enrichment of targeted clones without stable insertion of selection cassettes in the genome.

To this aim we built a reporter cassette to select for cells in which genome editing tools are active. The reporter is based on a BlasticidinS-resistance gene (bsr) placed out-of-frame with an upstream mCherry coding sequence. The two genes are linked by a sequence homologous to the genomic target. Activation of the genome editing tool after transfection will target both the genomic target and the one present on the reporter construct: deletions at the endogenous locus will lead to gene inactivation, deletion on the reporter construct will lead to the correction of the frameshift. This will provide a transient resistance to BlasticidinS. Treatment with this antibiotic will thus force the selection of the cells in which the TALEN/CRISPR have been active.

Based on our interest in antibody diversification, we decided to test our approach by targeting the AID gene in the CH12-F3 cell line. We chose to target the second exon of the AID gene. We thus designed the sgRNA to be used with the CRISPR/Cas9 system as well as the target to be inserted in the mCherry-out-bsr construct.

Twenty million CH12-F3 cells were used for the electroporation, and transfections were performed using either the CRISPR/Cas9 construct alone or in association with our reporter construct. Limiting dilutions were performed with regards to the control cells. On the other hand, cells transfected with our reporter construct were treated with Blasticidin-S for 72 hours before removing the antibiotic and plating them in 96 well plates. After approximately 2 weeks we obtained 52 clones from the bsr-treated cells and 78 from the control ones.

Considering that AID is necessary for Class Switch Recombination, we first selected the clones for their ability to switch from surface IgM to IgA. We thus induced CSR in each clone and we analyzed the percentage of switching by FACS. At the end of this analysis we obtained three typologies of clones: clones with a percentage of switch au pair with that of wild type CH12-F3, clones with a decreased ability to switch, and clones with absolutely no switch (FIG. 42). Interestingly, the number of clones in each category originating from the two samples was different (FIG. 43). An higher percentage of Blasticidin-S-treated clones showed an impaired CSR (~67% in treated clones vs ~30% in the control ones) (FIG. 43). More interestingly, whereas CSR was absent in a small percentage (~4%) of control cells, CSR was completely absent in ~38% of the Blasticidin-S-treated clones (FIG. 43).

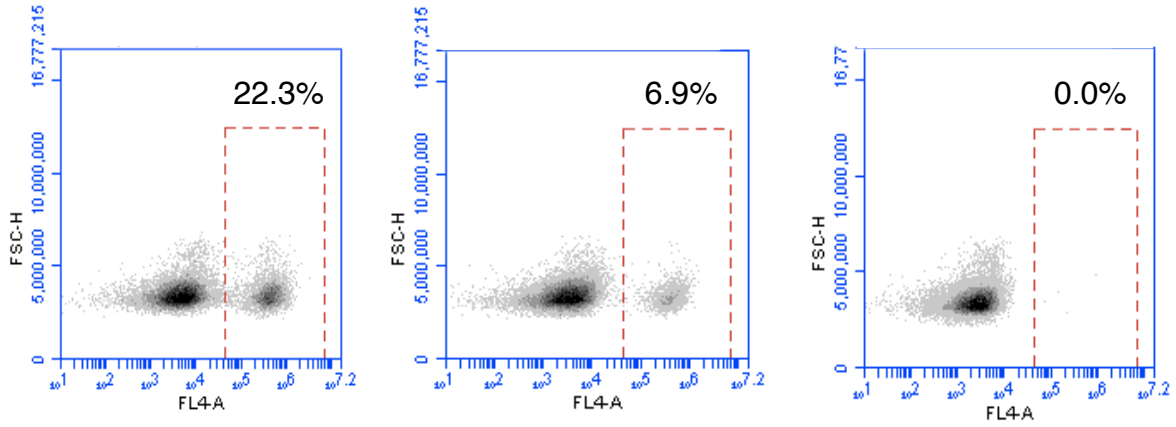

Figure 42

FACS analysis of stimulated CH12-F3 clones obtained after selection for gene targeting using the AID CRISPR and selection for 72h in bsr. The clones were stimulated and CSR visualised after 72h. Three different typologies of clones were obtained: clones with normal CSR (left), reduced switch (middle), or clones in which CSR was absent altogether (right).

We then extracted the genomic DNA from the clones with decreased or absent CSR , and we performed a touch-down PCR (Appendix A: 51-52) to amplify the second exon of AID. We initially cloned the fragments in plasmids in order to obtain sequence information from both AID alleles. Yet, we soon realised that most of the clones with absent CSR were homozygous for targeted deletions at the second exon. Each clone displayed its own deletion - typically 5-10bp -, suggesting that the selected clones were not originated from a clonal population (FIG. 44). On the other hand all the clones with a decreased ability to undergo were heterozygous at the site, with the wild type allele associated to a deleted allele. This observation is in line with the fact that AID is haploinsufficient (Sernández et al., 2008).

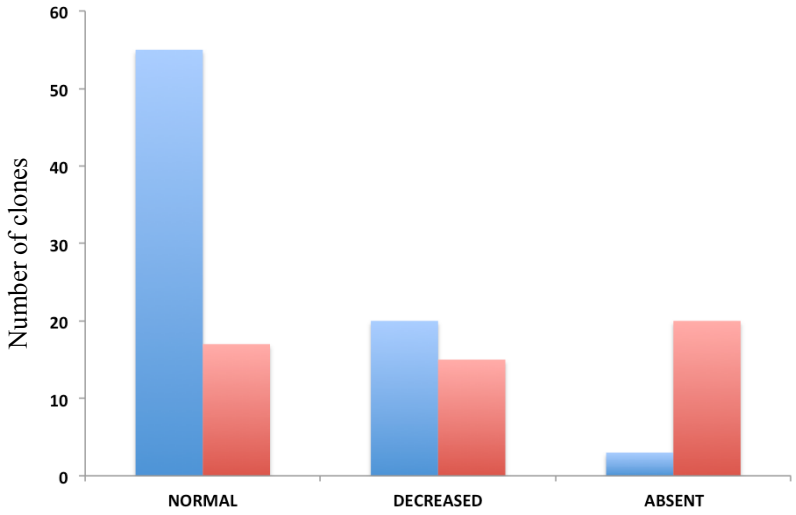

Figure 43

The histogram shows the number of CH12-F3 clones in each CSR-efficiency group.The color of the bars indicate whether the cell underwent bsr selection (red), or not (blue).

The observation that most AID knock-out clones were homozygous for the deletion, suggests that the deletion of the alleles occurred sequentially, with the targeted allele acting as a template for the repair of the second allele. This could be due to the fact that the deletion on the targeted allele impairs the ability of the CRISPR/Cas9 to retarget it. On the other hand, the wild-type allele will still be targeted, and - if homologous recombination is attempted - the deleted allele will be used as a template. Intriguingly, we found that most heterozygous clones from bsr-treated cells bore a ~200bp clonal deletion that encompassed the CRISPR target site and the flanking regions. We could not find though any homozygous clone bearing such deletion. This could be due to the fact that the deletion was too large to let the allele be used as a template for the second allele.

In our approach we used the CRISPR/Cas9 system. In this system the target sequence is lost when NHEJ repairs the DSB leaving a deletion. On the other hand, TALENs maintain the ability to target the DNA, as the DSBs induced by the FokI domain occur in the sequence flanking the TALEN targets. This suggest that with regards to gene targeting of multiple alleles, the CRISPR/Cas9 system might represent a better strategy. Of course, such an high frequency in HR could be specific to the cell line used, but - if the same behaviour is observed in other cells - this could make the targeting of both alleles much easier.

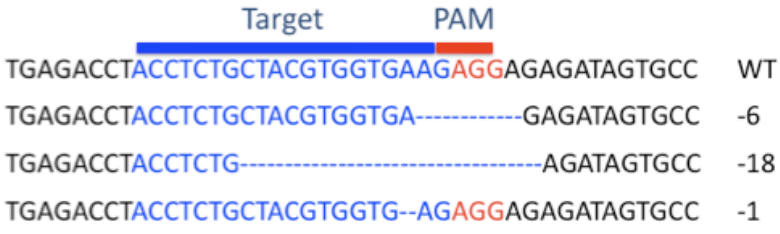

Figure 44

The image shows some of deletions found in our clones. In blue is represented the target sequence and in red the PAM sequence.

In conclusion we have proved that our mCherry-bsr reporter construct represents to date a very efficient approach to obtain specific gene inactivation, and more importantly, leads to a relevant increase in the percentage of homozygous clones.

Conclusions and future prospects

Visualization of Class Switch Recombination

As discussed in the Results chapter, the main problem we faced with the visualisation of the IgM transcript via an MS2 array, was the absence of a definite spot in the cells, indicating the source of the transcription. This could be either to the background fluorescence from the unbound YFP-MS2 or - more probably - to the pervasiveness of IgM transcripts in the cell. One possible route to test this could be to block transcription in the cell with Actinomycin D (Kleeff et al., 2000; Narita et al., 2000) to reduce the amount of IgM transcripts. Upon removal of the drug, we should be able to follow the IgM transcript as transcription at the IgH gene restarts. Of course, we cannot envision to use such a 'drugged' approach to follow the progress of CSR, yet this would give us an insight on the potential of the system. Moreover we could use such approach to follow the progression of a physiologically expressed gene. A different project, but it could put the cell we generated to a good use.

Another possibility to test our system lies on the addition of 422-461aa fragment from the mouse ornithine decarboxylase degradation domain (MODC) to the C-terminal of the YFP-MS2 coat protein, to induce a reduction in its half-life and obtain a lower fluorescent background (Li et al., 1998).

In the case of the visualization of IgH locus, our data suggest that the elemental problem lies in the capability of the half-YFP-TALEs to maintain the dimerisation even in absence of a suitable DNA target, thus increasing the background. Thus, to obtain a usable system, we would have to modify the dimerisation characteristics of the half-YFPs. An effective way could be the production of a library of mutants and a screening method to test the variation of the features in order to the screen for the mutants.

On the other hand, with the available tools, today there might be better approaches in order to visualize the progress of CSR: one possibility could be to use FRET by substituting the half-YFPs with fluorescent epitopes capable of FRET (e.g. EGFP and Tag-RFP fluorescent proteins). Another possibility could be to design fluorescent TALEs - either with or without the MODC degradation domain - to recognize repeat sequences at the IgH locus. Ideally we could target the Switch Regions as the AID-targeting repeats would represent thousands of targets for the TALEs: the Sμ region to keep track of the locus (part of it would be maintained even after CSR), and any of the S regions between the Sμ and the Sα regions, as they would be lost during CSR.

TALE-targeted Mutagenesis

With regards to this part of my work, there are two issues that need to be clarified, both revolving around our observation that expression of a TALE-AID can induce an increase in the fluorescence due to its targeting to an inactivated EGFP cds. The first one is whether this fluorescence is induced through back-mutation of the EGFP cds. The second one regards the modality through which a similar result is obtained through a TALE-AID^{E58A}, which is catalytically inactive.

One way to address this is to test whether the background fluorescence observed in untreated samples is due to use of the downstream CTGs as start codons for the EGFP. If this is the case, it might explain the results obtained with the TALE-AID^{E58A} as an effect of AID on the mRNA. It would also provide us a clean tool to test again for a targeted mutagenesis driven by the TALE-AID. This can be easily tested by mutagenizing the CTGs in the EGFP cds.

Moreover, in case we observe a limited mutagenic effect by AID, we could use upmutants of AID, such as those described previously described (Wang et al., 2009b) or one in which the C-terminal domain is ablated (Barreto et al., 2003; Ta et al., 2003; Shinkura et al., 2004). In alternative, we could use other APOBECS, which present different mutagenic capabilities.

It might also be possible that steric hindrance or spatial configuration of the chimera with respects to the DNA could affect the access of AID to the target DNA. This could

be solved either by testing more flexible/longer linker sequences to connect AID to the TALE and, eventually, by trying to a TALE targeting the opposite strand.

Enriching Genome Editing

Our approach to enrich for targeted clones seems reasonably robust, as it is being used by others member in the lab and by other labs to target other genes of interest. Our approach has been further validated by two very recent studies, in which a similar approach is described (Kim et al., 2013; Ramakrishna et al., 2014a). The most interesting result in our method is the high percentage of perfectly homozygous clones obtained. This is useful *per se*, and intriguing as it could provide insight on the machinery underlying the repair of this sort of DNA damage. At the moment we are refining the protocol to use our construct to increase the efficiency for knock-in targeting, either as single base or as longer DNA fragments.

References

Adams MM and Carpenter PB (2006): Tying the loose ends together in DNA double strand break repair with 53BP1. Cell Div. 1:19.

Alt FW, Zhang Y, Meng FL, Guo C, Schwer B (2013): Mechanisms of programmed DNA lesions and genomic instability in the immune system. Cell 152, 417–429.

Ansai S, Ochiai H, Kanie Y, Kamei Y, Gou Y, Kitano T, Yamamoto T, Kinoshita M (2012): Targeted disruption of exogenous EGFP gene in medaka using zinc-finger nucleases. Dev. Growth Differ. 54, 546–556.

Ansai S, Sakuma T, Yamamoto T, Ariga H, Uemura N, Takahashi R, Kinoshita M (2013): Efficient targeted mutagenesis in medaka using custom-designed transcription activator-like effector nucleases. Genetics 193, 739–749.

Aoufouchi S, Faili A, Zober C, D'Orlando O, Weller S, Weill J-C, Reynaud CA (2008): Proteasomal degradation restricts the nuclear lifespan of AID. The Journal of Experimental Medicine 205:1357–68.

Arakawa H, Lodygin D, Buerstedde JM (2001): Mutant loxP vectors for selectable marker recycle and conditional knock-outs. BMC Biotechnol. 1:7.

Arakawa H, Saribasak H, Buerstedde JM (2004): Activation-induced cytidine deaminase initiates immunoglobulin gene conversion and hypermutation by a common intermediate. PLoS Biol. 2(7):E179.

Aryan A, Anderson MA, Myles KM, Adelman ZN (2013): TALEN-based gene disruption in the dengue vector Aedes aegypti. PLoS ONE 8, e60082.

Bakkenist CJ and Kastan MB (2003): DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. Nature 421:499–506.

Bardwell PD, Martin A,Wong E, Li Z, Edelmann W, Scharff MD (2003): Cutting edge: the G-U mismatch glycosylase methyl-CpG binding domain 4 is dispensable for somatic hypermutation and class switch recombination. J. Immunol. 170:1620–24.

Bardwell PD, Woo CJ, WeiK, Li Z, Martin A, Sack SZ, Parris T, Edelmann W, Scharff MD (2004): Altered somatic hypermutation and reduced class-switch recombination in exonuclease 1-mutant mice. Nat. Immunol. 5:224–29.

Barnes DE and Lindahl T (2004): Repair and genetic consequences of endogenous DNA base damage in mammalian cells. Annu. Rev. Genet. 38:445–76.

Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P (2007): CRISPR provides acquired resistance against viruses in prokaryotes. Science, 315:1709–1712.

Barreto V, Reina-San-Martin B, Ramiro AR, McBride KM, Nussenzweig MC (2003): C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. Mol Cell. 12(2):501-8.

Basu U, Chaudhuri J, Alpert C, Dutt S, Ranganath S, Li G, Schrum JP, Manis JP, Alt FW (2005): The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation. Nature 438:508–11.

Basu U, Meng FL, Keim C, Grinstein V, Pefanis E, Eccleston J, Zhang T, Myers D, Wasserman CR, Wesemann DR, Januszyk K, Gregory RI, Deng H, Lima CD, Alt FW (2011): The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. Cell 144, 353–363.

Baxter SK, Jacoby K, Ryu BY, Kiem HP, Gouble A, Paques F, Rawlings DJ, Scharenberg AM (2012): Coupling endonucleases with DNA end-processing enzymes to drive gene disruption. Nat Methods, 9(10):973-5.

Beale RC, Petersen-Mahrt SK, Watt IN, Harris RS, Rada C, Neuberger MS (2004): Comparison of the differential context-dependence of DNA deamination by APOBEC enzymes: correlation with mutation spectra in vivo. J Mol Biol. 337:585–96.

Beard WA and Wilson SH (2006): Structure and mechanism of DNA polymerase β. Chem. Rev. 106:361–82.

Beckett D and Uhlenbeck OC (1988): Ribonucleoprotein complexes of R17 coat protein and a translational operator analog. J. Mol. Biol. 204, 927–938.

Bedell VM, Wang Y, Campbell JM, Poshusta TL, Starker CG, Krug RG 2nd, Tan W, Penheiter SG, Ma AC, Leung AY, Fahrenkrug SC, Carlson DF, Voytas DF, Clark KJ, Essner JJ, Ekker SC (2012): In vivo genome editing using a high-efficiency TALEN system. Nature 491, 114–118.

Beerli RR, Segal DJ, Dreier B, Barbas CF 3rd (1998): Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. Proc Natl Acad Sci U S A, 95(25):14628-33.

Beerli RR, Dreier B, Barbas CF 3rd (2000): Positive and negative regulation of endogenous genes by designed transcription factors. Proc Natl Acad Sci U S A, 97(4):1495-500.

Beerli RR and Barbas CF 3rd (2002): Engineering polydactyl zinc-finger transcription factors. Nat Biotechnol 20(2):135-41.

Bekker-Jensen S, Fugger K, Danielsen JR, Gromova I, Sehested M, Celis J, Bartek J, Lukas J, Mailand N (2007): Human xip1 (c2orf13) is a novel regulator of cellular responses to DNA strand breaks. J. Biol. Chem. 282:19638–43.

Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, Long RM (1998): Localization of ASH1 mRNA particles in living yeast. Mol. Cell 2, 437–445.

Beumer KJ, Trautman JK, Bozas A, Liu JL, Rutter J, Gall JG, Carroll D (2008): Efficient gene targeting in Drosophila by direct embryo injection with zinc-finger nucleases. Proc. Natl Acad. Sci. USA 105, 19821–19826.

Bhakta MS, Henry IM, Ousterout DG, Das KT, Lockwood SH, Meckler JF, Wallen MC, Zycovich A, Yu Y, Leo H, Xu L, Gersbach CA, Segal DJ (2013): Highly active zinc-finger nucleases by extended modular assembly. Genome Res, 23(3):530-8.

Bibikova M, Golic M, Golic KG, Carroll D (2002): Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. Genetics 161, 1169–1175.

Bibikova M, Beumer K, Trautman JK, Carrol D (2003): Enhancing gene targeting with designed zinc finger nucleases. Science 2;300(5620):764.

Blanc V and Davidson NO (2010): APOBEC-1-mediated RNA editing. Wiley Interdiscip Rev Syst Biol Med. 2(5):594-602.

Blount BA, Weenink T, Vasylechko S, Ellis T (2012): Rational diversification of a promoter providing fine-tuned expression and orthogonal regulation for synthetic biology. PLoS ONE 7:e33279.

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

Bogdanove AJ, Schornack S, Lahaye T (2010): TAL effectors: finding plant genes for disease and defense. Curr Opin Plant Biol. 13(4):394-401.

Bogdanove AJ and Voytas DF (2011): TAL effectors: Customizable proteins for DNA targeting. Science 333:1843–1846.

Bolotin A, Quinquis B, Sorokin A, Ehrlich SD (2005): Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology 151, 2551–2561.

Bransteitter R, Pham P, Scharff MD, Goodman MF (2003): Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. Proc. Natl Acad. Sci. USA 100, 4102–4107.

Bransteitter R, Sneeden JL, Allen S, Pham P, Goodman MF (2006): First AID (activation induced cytidine deaminase) is needed to produce high affinity isotype-switched antibodies. J. Biol. Chem. 281:16833–36.

Branzei D and Foiani M (2008): Regulation of DNA repair throughout the cell cycle. Nat Rev Mol Cell Biol. 9(4):297-308.

Brar SS, Watson M, Diaz M (2004): Activation-induced cytosine deaminase (AID) is actively exported out of the nucleus but retained by the induction of DNA breaks. J Biol Chem. 279(25):26395-401.

Briggs AW, Rios X, Chari R, Yang L, Zhang F, Mali P, Church GM (2012): Iterative capped assembly: Rapid and scalable synthesis of repeat module DNA such as TAL effectors from individual monomers. Nucleic Acids Res 40:e117.

Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, Dickman MJ, Makarova KS, Koonin EV, Van Der Oost J (2008): Small CRISPR RNAs guide antiviral defense in prokaryotes. Science 321, 960–964.

Brown AJ, Fisher DA, Kouranova E, McCoy A, Forbes K, Wu Y, Henry R, Ji D, Chambers A, Warren J, Shu W, Weinstein EJ, Cui X (2013): Whole-rat conditional gene knockout via genome editing. Nat. Methods 10, 638–640.

Bryant HE and Helleday T (2006): Inhibition of poly (ADP-ribose) polymerase activates ATM which is required for subsequent homologous recombination repair. Nucleic Acids Res. 23;34(6):1685-91.

Buerstedde JM, Alinikula J, Arakawa H, McDonald JJ, Schatz DG (2014): Targeting of somatic hypermutation by immunoglobulin enhancer and enhancer-like sequences. PLoS Biol. 12, e1001831.

Bultmann S, Morbitzer R, Schmidt CS, Thanisch K, Spada F, Elsaesser J, Lahaye T, Leonhardt H (2012): Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. Nucleic Acids Res 40:5368–5377.

Cade L, Reyon D, Hwang WY, Tsai SQ, Patel S, Khayter C, Joung JK, Sander JD, Peterson RT, Yeh JR (2012): Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs. Nucleic Acids Res. 40, 8001–8010.

Capecchi MR (2005): Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. Nat. Rev. Genet. 6, 507–512.

Carbery ID, Ji D, Harrington A, Brown V, Weinstein EJ, Liaw L, Cui X (2010): Targeted genome modification in mice using zinc-finger nucleases. Genetics 186, 451–459.

Carlson DF, Tan W, Lillico SG, Stverakova D, Proudfoot C, Christian M, Voytas DF, Long CR, Whitelaw CB, Fahrenkrug SC (2012): Efficient TALEN-mediated gene knockout in livestock. Proc. Natl Acad. Sci. USA 109, 17382–17387.

Carroll, D. (2011): Genome engineering with zinc-finger nucleases. Genetics 188, 773–782.

Casellas R, Nussenzweig A, Wuerffel R, Pelanda R, Reichlin A, Suh H, Qin XF, Besmer E, Kenter A, Rajewsky K, Nussenzweig MC (1998): Ku80 is required for immunoglobulin isotype switching. EMBO J. 17:2404–11.

Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF (2011): Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39:e82.

Cerosaletti K, Wright J, Concannon P (2006): Active role for nibrin in the kinetics of atm activation. Mol. Cell. Biol. 26:1691–99.

Certo MT, Gwiazda KS, Kuhar R, Sather B, Curinga G, Mandt T, Brault M, Lambert AR,

Chandra A, Van Maldegem F, Andrews S, Neuberger MS, Rada C (2013): Deficiency in spliceosome-associated factor CTNNBL1 does not affect ongoing cell cycling but delays exit from quiescence and results in embryonic lethality in mice. Cell Cycle 1;12(5):732-42.

Chaudhuri J, Tian M, Khuong C, Chua K, Pinaud E, Alt FW (2003): Transcription-targeted DNA deamination by the AID antibody diversification enzyme. Nature 17;422(6933):726-30.

Chaudhuri J and Alt FW (2004): Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. Nat Rev Immunol. 4(7):541-52.

Chaudhuri J, Khuong C, Alt FW (2004): Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. Nature 430:992–98.

Chen L, Trujillo K, Sung P, Tomkinson AE (2000): Interactions of the DNA ligase IVXRCC4 complex with DNA ends and the DNA-dependent protein kinase. J. Biol. Chem. 275:26196–205.

Chen F, Pruett-Miller SM, Huang Y, Gjoka M, Duda K, Taunton J, Collingwood TN, Frodin M, Davis GD (2011): High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. Nat. Methods 8, 753–755.

Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, Park J, Blackburn EH, Weissman JS, Qi LS, Huang B (2013a): Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell 155, 1479–1491.

Chen C, Fenk LA, de Bono M (2013b): Efficient genome editing in Caenorhabditis elegans by CRISPR-targeted homologous recombination. Nucleic Acids Res. 41, e193.

Chen L, Tang L, Xiang H, Jin L, Li Q, Dong Y, Wang W, Zhang G (2014): Advances in genome editing technology and its promising application in evolutionary and ecological studies. GigaScience, 3:24.

Chiarle R, Zhang Y, Frock RL, Lewis SM, Molinie B, Ho YJ, Myers DR, Choi VW, Compagno M, Malkin DJ, Neuberg D, Monti S, Giallourakis CC, Gostissa M, Alt FW (2011): Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. Cell 147, 107–119.

Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, Bogdanove AJ, Voytas DF (2010): Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 186:757–761.

Christian ML, Demorest ZL, Starker CG, Osborn MJ, Nyquist MD, Zhang Y, Carlson DF, Bradley P, Bogdanove AJ, Voytas DF (2012): Targeting G with TAL effectors: A comparison of activities of TALENs constructed with NN and NK repeat variable di-residues. PLoS ONE 7:e45383.

Christmann M, Tomicic MT, Roos WP, Kaina B (2003): Mechanisms of human DNA repair: an update. Toxicology 193:3–34.

Chuang CH, Carpenter AE, Fuchsova B, Johnson T, de Lanerolle P, Belmont AS (2006): Long-range directional movement of an interphase chromosome site. Curr. Biol. 16, 825-831.

Coluccio A, Miselli F, Lombardo A, Marconi A, Malagoli Tagliazucchi G, Goncalves MA, Pincelli C, Maruggi G, Del Rio M, Naldini L, Larcher F, Mavilio F, Recchia A (2013): Targeted gene addition in human epithelial stem cells by zinc-finger nuclease-mediated homologous recombination. Mol. Ther. 21, 1695–1704.

Cong L, Zhou R, Kuo YC, Cunniff M, Zhang F (2012): Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. Nat Commun 3:968.

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013): Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819–823.

Conticello SG, Langlois MA, Yang Z, Neuberger MS (2007): DNA deamination in immunity: AID in the context of its APOBEC relatives. Adv Immunol. 94:37-73.

Conticello SG, Ganesh K, Xue K, Lu M, Rada C, Neuberger MS (2008): Interaction between antibody-diversification enzyme AID and spliceosome-associated factor CTNNBL1. Mol Cell. 31(4):474-84.

Cormack BP, Valdivia RH, Falkow S (1996): FACS-optimized mutants of the green fluorescent protein (GFP). Gene. 173(1):33-8.

Costantini S, Woodbine L, Andreoli L, Jeggo PA, Vindigni A (2007): Interaction of the Ku heterodimer with the DNA ligase IV/Xrcc4 complex and its regulation by DNA-PK. DNA Repair 6:712–22.

Cradick TJ, Fine EJ, Antico CJ, Bao G (2013): CRISPR/Cas9 systems targeting b-globin and CCR5 genes have substantial off-target activity. Nucleic Acids Res. 41, 9584–9592.

Crosetto N, Mitra A, Silva MJ, Bienko M, Dojer N, Wang Q, Karaca E, Chiarle R, Skrzypczak M, Ginalski K., Pasero P, Rowicka M, Dikic I (2013): Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. Nat. Methods 10, 361–365.

Cui X, Ji D, Fisher DA, Wu Y, Briner DM, Weinstein EJ (2011): Targeted integration in rat and mouse embryos with zinc-finger nucleases. Nat. Biotechnol. 29, 64–67.

Dahlem TJ, Hoshijima K, Jurynec MJ, Gunther D, Starker CG, Locke AS, Weis AM, Voytas DF, Grunwald DJ (2012): Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. PLoS Genet. 8, e1002861.

Dedeoglu F, Horwitz B, Chaudhuri J, Alt FW, Geha RS (2004): Induction of activation induced cytidine deaminase gene expression by IL-4 and CD40 ligation is dependent on STAT6 and NFkappaB. International Immunology 16: 395-404.

Delbos F, Aoufouchi S, Faili A,Weill JC, Reynaud CA (2007): DNA polymerase η is the sole contributor of A/T modifications during immunoglobulin gene hypermutation in the mouse. J. Exp. Med. 204:17–23.

Delker RK, Fugmann SD, Papavasiliou FN (2009): A coming-of-age story: activation-induced cytidine deaminase turns 10. Nat Immunol.10(11):1147-53.

Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu JK, Shi Y, Yan N (2012a): Structural basis for sequence-specific recognition of DNA by TAL effectors. Science 335:720–723.

Deng D, Yin P, Yan C, Pan X, Gong X, Qi S, Xie T, Mahfouz M, Zhu JK, Yan N, Shi Y. (2012b): Recognition of methylated DNA by TAL effectors. Cell Res 22:1502–1504.

Deveau H, Barrangou R, Garneau JE, Labonté J, Fremaux C, Boyaval P, Romero DA, Horvath P, Moineau S (2008): Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. J. Bacteriol. 190, 1390–1400.

Di Noia J and Neuberger MS (2002): Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. Nature. 419(6902):43-8.

Di Noia JM, Rada C, Neuberger MS (2006): SMUG1 is able to excise uracil from immunoglobulin genes: insight into mutation vs repair. EMBO J. 25:585–95.

Di Noia JM and Neuberger MS (2007): Molecular mechanisms of antibody somatic hypermutation. Annu. Rev. Biochem. 76, 1–22.

Dianov GL and Hübscher U (2013): Mammalian base excision repair: the forgotten archangel. Nucleic Acids Res. 41, 3483–3490.

Dickerson SK, Market E, Besmer E, Papavasiliou FN (2003): AID mediates hypermutation by deaminating single stranded DNA. J Exp Med. 19;197(10):1291-6.

Dickinson DJ., Ward JD, Reiner DJ, Goldstein B (2013): Engineering the Caenorhabditis elegans genome using Cas9-triggered homologous recombination. Nat. Methods 10, 1028–1034.

Dietzl G, Chen D, Schnorrer F, Su KC, Barinova Y, Fellner M, Gasser B, Kinsey K, Oppel S, Scheiblauer S, Couto A, Marra V, Keleman K, Dickson BJ (2007): A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. Nature, 448(7150):151–156.

Difilippantonio S, Celeste A, Fernandez-Capetillo O, Chen HT, Reina San Martin B, Van Laethem F, Yang YP, Petukhova GV, Eckhaus M, Feigenbaum L, Manova K, Kruhlak M, Camerini-Otero RD, Sharan S, Nussenzweig M, Nussenzweig A (2005): Role of Nbs1 in the activation of the Atm kinase revealed in humanized mouse models. Nat. Cell Biol. 7:675–85.

Ding Q, Lee YK, Schaefer EA, Peters DT, Veres A, Kim K, Kuperwasser N, Motola DL, Meissner TB, Hendriks WT, Trevisan M, Gupta RM, Moisan A, Banks E, Friesen M, Schinzel RT, Xia F, Tang A, Xia Y, Figueroa E, Wann A, Ahfeldt T, Daheron L, Zhang F, Rubin LL, Peng LF, Chung RT, Musunuru K, Cowan CA (2013): A TALEN genome editing system for generating human stem cell-based disease models. Cell Stem Cell 12:238–251.

Downs JA, Nussenzweig MC, Nussenzweig A (2007): Chromatin dynamics and the preservation of genetic information. Nature 447:951–58.

Doyon Y, McCammon JM, Miller JC, Faraji F, Ngo C, Katibah GE, Amora R, Hocking TD, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Amacher SL (2008): Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. Nat. Biotechnol. 26, 702–708.

Doyon Y, Choi VM, Xia DF, Vo TD, Gregory PD, Holmes MC (2010): Transient cold shock enhances zinc-finger nuclease-mediated gene disruption. Nat Methods, 7(6):459-60.

Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, Miller JC, Urnov FD, Gregory PD, Holmes MC (2011): Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. Nat Methods, 8(1):74-9.

DuBridge RB, Tang P, Hsia HC, Leong PM, Miller JH, Calos MP (1987): Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. Mol Cell Biol. 7(1):379-87

Dunnick W, Wilson M, Stavnezer J (1989): Mutations, duplications, and deletion of recombined switch regions suggest a role for DNA replication in the immunoglobulin heavy chain switch. Mol. Cell. Biol. 9:1850–56.

Dunnick W, Hertz GZ, Scappino L, Gritzmacher C (1993): DNA sequences at immunoglobulin switch region recombination sites. Nucleic Acids Res. 21:365–72.

Dunnick WA, Collins JT, Shi J, Westfield G, Fontaine C, Hakimpour P, Papavasiliou FN (2009): Switch recombination and somatic hypermutation are controlled by the heavy chain 3' enhancer region. J Exp Med. 206(12):2613-23.

Dymecki SM (1996): Flp recombinase promotes site-specific DNA recombination in embryonic stem cells and transgenic mice. Proc Natl Acad Sci U S A, 93(12):6191–6196.

Ebina H, Misawa N, Kanemura Y, Koyanagi Y (2013): Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. Sci. Rep. 3, 2510.

Ehrenstein MR and Neuberger MS (1999): Deficiency in Msh2 affects the efficiency and local sequence specificity of immunoglobulin class-switch recombination: parallels with somatic hypermutation. EMBO J. 18:3484–90.

Ehrenstein MR, Rada C, Jones AM, Milstein C, Neuberger MS (2001): Switch junction sequences in PMS2-deficient mice reveal a microhomology-mediated mechanism of Ig class switch recombination. Proc. Natl. Acad. Sci. USA 98:14553–58.

Esvelt KM, Mali P, Braff JL, Moosburner M, Yaung SJ, Church GM (2013): Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. Nat. Methods 10, 1116–1121.

Faili A, Aoufouchi S, Weller S, Vuillier F, Stary A, Sarasin A, Reynaud CA, Weill JC (2004): DNA polymerase η is involved in hypermutation occurring during immunoglobulin class switch recombination. J. Exp. Med. 199:265–70.

Falck J, Coates J, Jackson SP (2005): Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. Nature 434:605–11.

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

Flajnik MF (2002): Comparative analyses of immunoglobulin genes: surprises and portents. Nat Rev Immunol. 2(9):688-98.

Folger KR, Wong EA, Wahl G, Capecchi MR (1982):Patterns of integration of DNA microinjected into cultured mammalian cells: evidence for homologous recombination between injected plasmid DNA molecules. Mol. Cell. Biol. 2, 1372–1387.

Forrest KM and Gavis ER (2003): Live imaging of endogenous RNA reveals a diffusion and entrapment mechanism for nanos mRNA localization in Drosophila. Curr. Biol. 13, 1159–1168.

Franco S, Gostissa M, Zha S, Lombard DB, Murphy MM, Zarrin AA, Yan C, Tepsuporn S, Morales JC, Adams MM, Lou Z, Bassing CH, Manis JP, Chen J, Carpenter PB, Alt FW (2006): H2AX prevents DNA breaks from progressing to chromosome breaks and translocations. Mol. Cell 21:201–14.

Frank KM, Sekiguchi JM, Seidl KJ, Swat W, Rathbun GA, Cheng HL, Davidson L, Kangaloo L, Alt FW (1998): Late embryonic lethality and impaired V(D)J recombination in mice lacking DNA ligase IV. Nature 396:173–77.

Friedrich G and Soriano P (1991): Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. Genes Dev. 5, 1513–1523.

Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD (2013): High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat. Biotechnol. 31, 822–826.

Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK (2014): Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat. Biotechnol. 32, 279–284.

Fujii W, Kawasaki K, Sugiura K, Naito K (2013): Efficient generation of large-scale genome-modified mice using gRNA and CAS9 endonuclease. Nucleic Acids Res. 41, e187.

Fung H and Demple B (2005): A vital role for Ape1/Ref1 protein in repairing spontaneous DNA damage in human cells. Mol. Cell 17:463–70.

Fusco D, Accornero N, Lavoie B, Shenoy SM, Blanchard JM, Singer RH, Bertrand E (2003): Single mRNA molecules demonstrate probabilistic movement in living Mammalian cells. Curr. Biol. 13, 161–167.

Gaj T, Guo J, Kato Y, Sirk SJ, Barbas CF 3rd (2012): Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. Nat. Methods 9, 805–807.

Gaj T, Gersbach CA, Barbas III CF (2013): ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Cell Press, Vol. 31, No. 7: 397-405.

Gao H, Wu X, Chai J, Han Z (2012): Crystal structure of a TALE protein reveals an extended N-terminal DNA binding region. Cell Res 22:1716–1720.

Garg A, Lohmueller JJ, Silver PA, Armel TZ (2012): Engineering synthetic TAL effectors with orthogonal target sites. Nucleic Acids Res 40:7584–7595.

Garneau JE, Dupuis ME, Villion M, Romero DA, Barrangou R, Boyaval P, Fremaux C, Horvath P, Magadán AH, Moineau S (2010): The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 468, 67–71.

Gasiunas G, Barrangou R, Horvath P, Siksnys V (2012): Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. USA 109, E2579–E2586.

Geisberger R, Huemer M, Gassner FJ, Zaborsky N, Egle A, Greil R (2012): Lysine residue at position 22 of the AID protein regulates its class switch activity. PLoS One. 7(2):e30667.

Geissler R, Scholze H, Hahn S, Streubel J, Bonas U, Behrens SE, Boch J (2011): Transcriptional activators of human genes with programmable DNA-specificity. PLoS ONE 6:e19509.

Genschel J, Bazemore LR, Modrich P (2002): Human exonuclease I is required for 5 and 3 mismatch repair. J. Biol. Chem. 277:13302–11.

Genschel J and Modrich P (2003): Mechanism of 5-directed excision in human mismatch repair. Mol. Cell 12:1077–86.

Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, Lim WA, Weissman JS, Qi LS (2013): CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 154, 442–451.

Gonczy P, Echeverri C, Oegema K, Coulson A, Jones SJM, Copley RR, Duperon J, Oegema J, Brehm M, Cassin E, Hannak E, Kirkham M, Pichler S, Flohrs K, Goessen A, Leidel S, Alleaume AM, Martin C, Ozlu N, Bork P, Hyman AA (2000): Functional genomic analysis of cell division in C. elegans using RNAi of genes on chromosome III. Nature, 408(6810):331–336.

Gonda H, Sugai M, Nambu Y, Katakai T, Agata Y, Mori KJ, Yokota Y, Shimizu A (2003): The balance between Pax5 and Id2 activities is the key to AID gene expression. J. Exp. Med. 198:1427–37.

Gonzalez B, Schwimmer LJ, Fuller RP, Ye Y, Asawapornmongkol L, Barbas CF 3rd (2010): Modular system for the construction of zinc-finger libraries and proteins. Nat Protocols, 5(4):791–810.

Gouy M, Guindon S, Gascuel O (2010): SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol. 27:221-224.

Gratz SJ, Cummings AM, Nguyen JN, Hamm DC, Donohue LK, Harrison MM, Wildonger J, O'Connor-Giles KM (2013): Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. Genetics 194, 1029–1035.

Grissa I, Vergnaud G, Pourcel C (2007): The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. BMC Bioinform., 8, 172.

Grizot S, Smith J, Daboussi F, Prieto J, Redondo P, Merino N, Villate M, Thomas S, Lemaire L, Montoya G, Blanco FJ, Paques F, Duchateau P (2009): Efficient targeting of a SCID gene by an engineered single-chain homing endonuclease. Nucleic Acids Res. 37, 5405–5419.

Groth AC, Fish M, Nusse R, Calos MP (2004): Construction of transgenic Drosophila by using the site-specific integrase from phage ϕC31. Genetics, 166(4):1775–1782.

Gu H, Marth JD, Orban PC, Mossmann H, Rajewsky K (1994): Deletion of a DNA polymerase β-gene segment in T cells using cell type-specific gene targeting. Science 265, 103–106.

Guikema JEJ, Linehan EK, Tsuchimoto D, Nakabeppu Y, Strauss PR, Stavnezer J, Schrader CE (2007): APE1 and APE2 dependent DNA breaks in immunoglobulin class switch recombination. J. Exp. Med. 204:3017–26.

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

Gupta A, Christensen RG, Rayla AL, Lakshmanan A, Stormo GD, Wolfe SA (2012): An optimized two-finger archive for ZFN-mediated gene targeting. Nature Methods, 9(6), 588–590.

Gupta A, Hall VL, Kok FO, Shin M, McNulty JC, Lawson ND, Wolfe SA (2013): Targeted chromosomal deletions and inversions in zebrafish. Genome Res. 23, 1008–1017.

Guschin DY, Waite AJ, Katibah GE, Miller JC, Holmes MC, Rebar EJ (2010): A rapid and general assay for monitoring endogenous gene modification. Methods Mol. Biol. 649, 247–256.

Hadi MZ, Ginalski K, Nguyen LH, Wilson DM 3rd (2002): Determinants in nuclease specificity of Ape1 and Ape2, human homologues of Escherichia coli exonuclease III. J. Mol. Biol. 316:853–66.

Haft DH, Selengut J, Mongodin EF, Nelson KE (2005): A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. PLoS Comput. Biol. 1, e60.

Hakim O, Resch W, Yamane A, Klein I, Kieffer-Kwon KR, Jankovic M, Oliveira T, Bothmer A, Voss TC, Ansarah-Sobrinho C, Mathe E, Liang G, Cobell J, Nakahashi H, Robbiani DF, Nussenzweig A, Hager GL, Nussenzweig MC, Casellas R (2012): DNA damage defines sites of recurrent chromosomal translocations in B lymphocytes. Nature 484, 69–74.

Hakim O, Resch W, Yamane A, Klein I, Kieffer-Kwon KR, Jankovic M, Oliveira T, Bothmer A, Voss TC, Ansarah-Sobrinho C, Mathe E, Liang G, Cobell J, Nakahashi H, Robbiani DF, Nussenzweig A, Hager GL, Nussenzweig MC, Casellas R (2012): DNA damage defines sites of recurrent chromosomal translocations in B lymphocytes. Nature. 484(7392):69-74.

Hale CR, Zhao P, Olson S, Duff MO, Graveley BR, Wells L, Terns RM, Terns MP (2009): RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. Cell 139, 945–956.

Hale CR, Majumdar S, Elmore J, Pfister N, Compton M, Olson S, Resch AM, Glover CV 3rd, Graveley BR, Terns RM, Terns MP (2012): Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. Mol. Cell 45, 292–302.

Han L and Yu K (2008): Altered kinetics of nonhomologous end joining and class switch recombination in ligase IV–deficient B cells. The Journal of Experimental Medicine, 205(12), 2745–2753.

Han L, Masani S, Yu K (2010): CTNNBL1 is dispensable for immunoglobulin class switch recombination. Journal of Immunology (Baltimore, Md. : 1950), 185(3), 1379–1381.

Harris RS, Petersen-Mahrt SK, Neuberger MS (2002): RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. Mol Cell. 10(5):1247-53.

Hasham MG, Donghia NM, Coffey E, Maynard J, Snow KJ, Ames J, Wilpan RY, He Y, King BL, Mills KD (2010): Widespread genomic breaks generated by activation-induced cytidine deaminase are prevented by homologous recombination. Nat Immunol. 11(9):820-6.

Häsler J, Rada C, Neuberger MS (2011): Cytoplasmic activation-induced cytidine deaminase (AID) exists in stoichiometric complex with translation elongation factor 1α (eEF1A). Proc Natl Acad Sci U S A. 108(45):18366-71.

Hauschild J, Petersen B, Santiago Y, Queisser AL, Carnwath JW, Lucas-Hahn A, Zhang L, Meng X, Gregory PD, Schwinzer R, Cost GJ Niemann, H(2011): Efficient generation of a biallelic knockout in pigs using zinc-finger nucleases. Proc. Natl Acad. Sci. USA 108, 12013–12017.

Hayashi T, Sakamoto K, Sakuma T, Yokotani N, Inoue T, Kawaguchi E, Agata K, Yamamoto T, Takeuchi T (2013): TALENs efficiently disrupt the target gene in Iberian ribbed newts (Pleurodeles waltl), an experimental model animal for regeneration. Dev. Growth Differ 56, 115–121.

He B, Qiao X, Cerutti A (2004): CpG DNA induces IgG class switch DNA recombination by activating human B cells through an innate pathway that requires TLR9 and cooperates with IL-10. J Immunol. 173(7):4479-91.

Heintze J, Luft C, Ketteler R (2013): A CRISPR CASe for high-throughput silencing. Front Genet. 4:193.

Hisano Y, Ota S, Arakawa K, Muraki M, Kono N, Oshita K, Sakuma T, Tomita M, Yamamoto T, Okada Y, Kawahara A (2013): Quantitative assay for TALEN activity at endogenous genomic loci. Biol. Open 2, 363–367.

Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, Cost GJ, Zhang L, Santiago Y, Miller JC, Zeitler B, Cherone JM, Meng X, Hinkley SJ, Rebar EJ, Gregory PD, Urnov FD, Jaenisch R (2011): Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol 29:731–734.

Holt N, Wang J, Kim K, Friedman G, Wang X, Taupin V, Crooks GM, Kohn DB, Gregory PD, Holmes MC, Cannon PM (2010): Human hematopoietic stem/progenitor cells modified by zinc-finger nuclease targeted to CCR5 control HIV-1 in vivo. Nat. Biotechnol. 28, 839–847.

Horikawa K, Martin SW, Pogue SL, Silver K, Peng K, Takatsu K, Goodnow CC (2007): Enhancement and suppression of signaling by the conserved tail of IgG memory-type B cell antigen receptors. J. Exp. Med. 204:759–69.

Horvath P, Romero DA, Coûté-Monvoisin AC, Richards M, Deveau H, Moineau S, Boyaval P, Fremaux C, Barrangou R (2008): Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus. J. Bacteriol. 190, 1401–1412.

Horvath P and Barrangou R (2010): CRISPR/Cas, the immune system of bacteria and archaea. Science, 327:167–170.

Hou Z, Zhang Y, Propson NE, Howden SE, Chu LF, Sontheimer EJ, Thomson JA (2013): Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. Proc. Natl. Acad. Sci. USA 110, 15644–15649.

Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X., Shalem O, Cradick TJ, Marraffini LA, Bao G, Zhang F (2013): DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. 31, 827–832.

Hsu PD, Lander ES, Zhang F (2014): Development and applications of CRISPR-Cas9 for genome engineering. Cell. 157(6):1262-78.

Hu CD, Chinenov Y, Kerppola TK (2002): Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. Mol Cell. 9(4):789-98.

Huang J, Sengupta R, Espejo B, Lee MG, Dorsey JA, Richter M, Opravil S, Shiekhattar R, Bedford MT, Jenuwein T, Berger SL (2007): p53 is regulated by the lysine demethylase LSD1. Nature 449:105–8.

Huang P, Xiao A, Zhou M, Zhu Z, Lin S, Zhang B (2011): Heritable gene targeting in zebrafish using customized TALENs. Nat Biotechnol 29:699–700.

Hwang WY, Fu Y, Reyon D, Maeder ML, Kaini P, Sander JD, Joung JK, Peterson RT, Yeh JR (2013): Heritable and precise zebrafish genome editing using a CRISPR-Cas system. PLoS ONE 8, e68708.

Iles N, Rulten S, El-Khamisy SF, Caldecott KW (2007): APLF (C2orf13) is a novel human protein involved in the cellular response to chromosomal DNA strand breaks. Mol. Cell. Biol. 27:3793–803.

Imai K, Slupphaug G, Lee WI, Revy P, Nonoyama S, Catalan N, Yel L, Forveille M, Kavli B, Krokan HE, Ochs HD, Fischer A, Durandy A (2003): Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination. Nat. Immunol. 4:1023–28.

Inoue H, Nojima H, Okayama H (1990): High efficiency transformation of Escherichia coli with plasmids. Gene 30;96(1):23-8.

Invitrogen (1986): New frozen competent puc host e. coli dh5α. Focus, 8(2):9.252.

Ito K, Hirao A, Arai F, Matsuoka S, Takubo K, Hamaguchi I, Nomiyama K, Hosokawa K, Sakurada K, Nakagata N, Ikeda Y, Mak TW, Suda T (2004a): Regulation of oxidative stress by ATM is required for self-renewal of haematopoietic stem cells. Nature 431:997–1002.

Ito S, Nagaoka H, Shinkura R, Begum N, Muramatsu M, Nakata M, Honjo T (2004b): Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. Proc Natl Acad Sci U S A 101:1975-1980.

Janicki SM, Tsukamoto T, Salghetti SE, Tansey WP, Sachidanandam R, Prasanth KV, Ried T, Shav-Tal Y, Bertrand E, Singer RH, Spector DL (2004): From silencing to gene expression: real-time analysis in single cells. Cell 116, 683-698.

Jao LE, Wente SR, Chen W (2013): Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. Proc. Natl Acad. Sci. USA 110, 13904–13909.

Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA (2013): RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat. Biotechnol. 31, 233–239.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012): A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816–821.

Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, Anders C, Hauer M, Zhou K, Lin S, Kaplan M, Iavarone AT, Charpentier E, Nogales E, Doudna JA (2014): Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science 343, 1247997.

Jones JM and Meisler MH (2013): Modeling human epilepsy by TALEN targeting of mouse sodium channel Scn8a. Genesis (in press, doi: 10.1002/dvg.22731)

Joung JK, Sander JD (2013): TALENs: A widely applicable technology for targeted genome editing. Nat Rev Mol Cell Biol 14:49–55.

Kanno S, Kuzuoka H, Sasao S, Hong Z, Lan L, Nakajima S, Yasui A. (2007): A novel human AP endonuclease with conserved zinc-finger-like motifs involved in DNA strand break responses. EMBO J. 26:2094–103.

Kato L, Begum NA, Burroughs AM, Doi T, Kawai J, Daub CO, Kawaguchi T, Matsuda F, Hayashizaki Y, Honjo T (2012): Nonimmunoglobulin target loci of activation-induced cytidine deaminase (AID) share unique features with immunoglobulin genes. Proc. Natl. Acad. Sci. USA 109, 2479–2484.

Kawai N, Ochiai H, Sakuma T, Yamada L, Sawada H, Yamamoto T, Sasakura Y (2012): Efficient targeted mutagenesis of the chordate Ciona intestinalis genome with zincfinger nucleases. Dev. Growth Differ. 54, 535–545.

Kilby NJ, Snaith MR, Murray JA (1993): Site-specific recombinases: tools for genome engineering. Trends in Genet: TIG, 9(12):413–421.

Kim H, Um E, Cho SR, Jung C, Kim H, Kim JS (2011a): Surrogate reporters for enrichment of cells with nuclease-induced mutations. Nat Methods. 8(11):941-3.

Kim S, Lee MJ, Kim H, Kang M, Kim JS (2011b): Preassembled zinc-finger arrays for rapid construction of ZFNs. Nat Methods, 8(1):7.

Kim E, Kim S, Kim DH, Choi BS, Choi IY, Kim JS (2012): Precision genome engineering with programmable DNA-nicking enzymes. Genome Res. 22, 1327–1333.

Kim H, Kim MS, Wee G, Lee CI, Kim H, Kim JS (2013): Magnetic separation and antibiotics selection enable enrichment of cells with ZFN/TALEN-induced mutations. PLoS One 8(2):e56476.

Kleeff J, Kornmann M, Sawhney H, Korc M (2000): Actinomycin D induces apoptosis and inhibits growth of pancreatic cancer cells. Int J Cancer. 1;86(3):399-407.

Klein U, Dalla-Favera R (2008): Germinal centres: role in B-cell physiology and malignancy. Nat. Rev. Immunol. 8, 22–33.

Kleinstiver BP, Wolfs JM, Kolaczyk T, Roberts AK, Hu SX, Edgell DR (2012): Monomeric site-specific nucleases for genome editing. Proc. Natl Acad. Sci. USA, 109, 8061–8066.

Kohli RM, Maul RW, Guminski AF, McClure RL, Gajula KS, Saribasak H, McMahon MA, Siliciano RF, Gearhart PJ, Stivers JT (2010): Local sequence targeting in the AID/APOBEC family differentially impacts retroviral restriction and antibody diversification. J Biol Chem. 24;285(52):40956-64.

Koller BH and Smithies O (1992): Altering genes in animals by gene targeting. Annu. Rev. Immunol. 10, 705–730.

Komiyama M (2013): Cut-and-Paste of DNA Using an Artificial Restriction DNA Cutter. Int. J. Mol. Sci. 14, 3343–3357.

Konermann S, Brigham MD, Trevino AE, Hsu PD, Heidenreich M, Cong L, Platt RJ, Scott DA, Church GM, Zhang F (2013): Optical control of mammalian endogenous transcription and epigenetic states. Nature 500, 472–476.

Kovalchuk AL, duBois W, Mushinski E, McNeil NE, Hirt C, Qi CF, Li Z, Janz S, Honjo T, Muramatsu M, Ried T, Behrens T, Potter M (2007): AID-deficient Bcl-xL transgenic mice develop delayed atypical plasma cell tumors with unusual Ig/Myc chromosomal rearrangements. J. Exp. Med. 204, 2989–3001.

Kovalchuk AL, Ansarah-Sobrinho C, Hakim O, Resch W, Tolarová H, Dubois W, Yamane A, Takizawa M, Klein I, Hager GL, Morse HC 3rd, Potter M, Nussenzweig MC, Casellas R (2012): Mouse model of endemic Burkitt translocations reveals the long-range boundaries of Ig-mediated oncogene deregulation. Proc. Natl. Acad. Sci. USA 109, 10972–10977.

Kracker S, Bergmann Y, Demuth I, Frappart PO, Hildebrand G, Christine R, Wang ZQ, Sperling K, Digweed M, Radbruch A (2005): Nibrin functions in Ig class-switch recombination. Proc. Natl. Acad. Sci. USA 102:1584–89.

Kramer KM, Brock JA, Bloom K, Moore JK, Haber JE (1994): Two different types of double-strand breaks in Saccharomyces cerevisiae are repaired by similar RAD52-independent, non-homologous recombination events, Mol. Cell. Biol. 14 (2) 1293–1301.

Kunkel T and Erie D (2005): DNA mismatch repair. Annu. Rev. Biochem. 74:681–710.

Küppers R (2005): Mechanisms of B-cell lymphoma pathogenesis. Nat. Rev. Cancer 5, 251–262.

Lahdesmaki A, Taylor AM, Chrzanowska KH, Pan-Hammarstrom Q (2004): Delineation of the role of the Mre11 complex in class switch recombination. J. Biol. Chem. 279:16479–87.

Lee JH and Paull TT (2005): ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex. Science 308:551–54.

Lee HJ, Kim E, Kim JS (2010): Targeted chromosomal deletions in human cells using zinc finger nucleases. Genome Res. 20, 81–89.

Lee HJ, Kweon J, Kim E, Kim S, Kim JS (2012): Targeted chromosomal duplications and inversions in the human genome using zinc finger nucleases. Genome Res 22, 539–548.

Lei Y, Guo X, Liu Y, Cao Y, Deng Y, Chen X, Cheng CH, Dawid IB, Chen Y, Zhao H (2012): Efficient targeted gene disruption in Xenopus embryos using engineered transcription activator-like effector nucleases (TALENs). Proc Natl Acad Sci USA 109:17484–17489.

Li X, Zhao X, Fang Y, Jiang X, Duong T, Fan C, Huang CC, Kain SR (1998): Generation of destabilized green fluorescent protein as a transcription reporter. J Biol Chem. 25;273(52):34970-5.

Li Z, Scherer SJ, Ronai D, Iglesias-Ussel MD, Peled JU, Bardwell PD, Zhuang M, Lee K, Martin A, Edelmann W, Scharff MD (2004): Examination of Msh6- and Msh3-deficient mice in class switching reveals overlapping and distinct roles of MutS homologues in antibody diversification. J. Exp. Med. 200:47–59.

Li T, Huang S, Jiang WZ, Wright D, Spalding MH, Weeks DP, Yang B (2011a): TAL nucleases (TALNs): Hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. Nucleic Acids Res 39:359–372.

Li T, Huang S, Zhao X, Wright DA, Carpenter S, Spalding MH, Weeks DP, Yang B (2011b): Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. Nucleic Acids Res, 39(14):6315–6325.

Li G, Pone EJ, Tran DC, Patel PJ, Dao L, Xu Z, Casali P (2012a): Iron inhibits activation-induced cytidine deaminase enzymatic activity and modulates immunoglobulin class switch DNA recombination. J Biol Chem. 287(25):21520-9.

Li L, Piatek MJ, Atef A, Piatek A, Wibowo A, Fang X, Sabir JS, Zhu JK, Mahfouz MM (2012b): Rapid and highly efficient construction of TALE based transcriptional regulators and nucleases for genome modification. Plant Mol Biol 78:407–416.

Li T, Liu B, Spalding MH, Weeks DP, Yang B (2012c): High-efficiency TALEN-based gene editing produces disease-resistant rice. Nat. Biotechnol. 30, 390–392.

Li Y, Moore R, Guinn M, Bleris L (2012d): Transcription activator-like effector hybrids for conditional control and rewiring of chromosomal transgene expression. Sci Rep 2:897.

Li JF, Norville JE, Aach J, McCormack M, Zhang D, Bush J, Church GM, Sheen J (2013a): Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. Nat Biotechnol, 31:688–691

Li W, Teng F, Li T, Zhou Q (2013b): Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. Nat. Biotechnol. 31, 684–686.

Lieber MR, Ma Y, Pannicke U, Schwarz K (2003): Mechanism and regulation of human non homologous DNA end-joining. Nat. Rev. Mol. Cell Biol. 4:712–20.

Liu Q, Segal DJ, Ghiara JB, Barbas CF 3rd (1997): Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. Proc Natl Acad Sci U S A, 94(11):5525-30.

Liu M, Duke JL, Richter DJ, Vinuesa CG, Goodnow CC, Kleinstein SH, Schatz DG (2008). Two levels of protection for the B cell genome during somatic hypermutation. Nature 451, 841–845.

Liu J, Li C, Yu Z, Huang P, Wu H, Wei C, Zhu N, Shen Y, Chen Y, Zhang B, Deng WM, Jiao R (2012): Efficient and specific modifications of the Drosophila genome by means of an easy TALEN strategy. J. Genet. Genomics 39, 209–215.

Liu J, Gaj T, Patterson JT, Sirk SJ, Barbas III CF (2014): Cell-Penetrating Peptide-Mediated Delivery of TALEN Proteins via Bioconjugation for Genome Engineering. PLoS ONE, 9(1), e85755.

Lo TW, Pickle CS, Lin S, Ralston EJ, Gurling M, Schartner CM, Bian Q, Doudna JA, Meyer BJ (2013): Precise and heritable genome editing in evolutionarily diverse nematodes using TALENs and CRISPR/Cas9 to engineer insertions and deletions. Genetics 195, 331–348.

Longerich S, Tanaka A, Bozek G, Nicolae D, Storb U (2005): The very 5 end and the constant region of Ig genes are spared from somatic mutation because AID does not access these regions. J. Exp. Med. 202:1443–54.

Longerich S, Basu U, Alt F, Storb U (2006): AID in somatic hypermutation and class switch recombination. Curr. Opin. Immunol. 18:164–74.
Luby TM, Schrader CE, Stavnezer J, Selsing E (2001): The μ switch region tandem repeats are important, but not required, for antibody class switch recombination. J. Exp. Med. 193:159–68.

Lumsden JM, McCarty T, Petiniot LK, Shen R, Barlow C, Wynn TA, Morse HC 3rd, Gearhart PJ, Wynshaw-Boris A, Max EE, Hodes RJ (2004): Immunoglobulin class switch recombination is impaired in Atm-deficient mice. J. Exp. Med. 200:1111–21.

Ma Y, Lu H, Schwarz K, Lieber MR (2005): Repair of double-strand DNA breaks by the human non homologous DNA end joining pathway: the iterative processing model. Cell Cycle 4:1193–200.

Ma S, Zhang S, Wang F, Liu Y, Liu Y, Xu H, Liu C, Lin Y, Zhao P, Xia Q (2012): Highly efficient and specific genome editing in silkworm using custom TALENs. PLoS ONE 7, e45035.

Ma H, Reyes-Gutierrez P, Pederson, T (2013): Visualization of repetitive DNA sequences in human chromosomes with transcription activator-like effectors. Proceedings of the National Academy of Sciences of the United States of America, 110(52), 21048 21053. doi:10.1073/pnas.1319097110

MacDuff DA, Neuberger MS, Harris RS (2006): MDM2 can interact with the C-terminus of AID but it is inessential for antibody diversification in DT40 B cells. Mol Immunol. 43(8):1099-108.

MacGinnitie AJ, Anant S, Davidson NO (1995): Mutagenesis of apobec-1, the catalytic subunit of the mammalian apolipoprotein B mRNA editing enzyme, reveals distinct domains that mediate cytosine nucleoside deaminase, RNA binding, and RNA editing activity. J Biol Chem. 270(24):14768-75.

Maeder M L, Thibodeau-Beganny S, Osiak A, Wright DA, Anthony RM, Eichtinger M, Jiang T, Foley JE, Winfrey RJ, Townsend JA, Unger-Wallace E, Sander JD, Müller-Lerch F, Fu F, Pearlberg J, Göbel C, Dassie JP, Pruett-Miller SM, Porteus MH, Sgroi DC, Lafrate AJ, Dobbs D, McCray PB Jr, Cathomen T, Voytas DF, Joung JK (2008): Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol Cell. 31(2):294-301.

Maeder ML, Linder SJ, Cascio VM, Fu Y, Ho QH, Joung JK (2013): CRISPR RNA-guided activation of endogenous human genes. Nat. Methods 10, 977–979.

Mahfouz MM, Li L, Shamimuzzaman M, Wibowo A, Fang X, Zhu JK (2011): De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. Proc Natl Acad Sci USA 108:2623–2628.

Mahfouz MM, Li L, Piatek M, Fang X, Mansour H, Bangarusamy DK, Zhu JK (2012): Targeted transcriptional repression using a chimeric TALESRDX repressor protein. Plant Mol Biol 78:311–321.

Mak AN, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL (2012): The crystal structure of TAL effector PthXo1 bound to its DNA target. Science 335:716–719.

Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, Moineau S, Mojica FJ, Wolf YI, Yakunin AF, Van Der Oost J, Koonin EV (2011b): Evolution and classification of the CRISPR-Cas systems. Nat. Rev. Microbiol. 9, 467–477.

Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM (2013a): RNA-guided human genome engineering via Cas9. Science 339, 823–826.

Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, Yang L, Church GM (2013b): CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat. Biotechnol. 31, 833–838.

Manis JP, Gu Y, Lansford R, Sonoda E, Ferrini R, Davidson L, Rajewsky K, Alt FW (1998): Ku70 is required for late B cell development and immunoglobulin heavy chain switching. J. Exp. Med. 187:2081–89.

Manis JP, Morales JC, Xia Z, Kutok JL, Alt FW, Carpenter PB (2004): 53BP1 links DNA damage-response pathways to immunoglobulin heavy chain class-switch recombination. Nat. Immunol. 5:481–87.

Mansour SL, Thomas KR, Capecchi MR (1988): Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. Nature 336, 348–352.

Marraffini LA and Sontheimer EJ (2008): CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science 322, 1843–1845.

Marraffini LA, and Sontheimer EJ (2010): Self versus non-self discrimination during CRISPR RNA-directed immunity. Nature 463, 568–571.

Martin A, Li Z, Lin DP, Bardwell PD, Iglesias-Ussel MD, Edelmann W, Scharff MD (2003): Msh2 ATPase activity is essential for somatic hypermutation at A-T basepairs and for efficient class switch recombination. J. Exp. Med. 198:1171–78.

Martin SW and Goodnow CC (2002): Burst-enhancing role of the IgG membrane tail as a molecular determinant of memory. Nat. Immunol. 3:182–88.

Martin SE, Caplen NJ (2007): Applications of RNA interference in mammalian systems. Annu Rev Genomics Hum Genet, 8:81–108.

Martomo SA, Yang WW, Gearhart PJ (2004): A role for Msh6 but not Msh3 in somatic hypermutation and class switch recombination. J. Exp. Med. 200:61–68.

Martomo SA, Fu D, Yang WW, Joshi NS, Gearhart PJ (2005): Deoxyuridine is generated preferentially in the nontranscribed strand of DNA from cells expressing activation induced cytidine deaminase. J. Immunol. 174:7787–91.

Mashimo T, Takizawa A, Voigt B, Yoshimi K, Hiai H, Kuramoto T, Serikawa T (2010): Generation of knockout rats with X-linked severe combined immunodeficiency (X-SCID) using zinc-finger nucleases. PLoS ONE 5, e8870.

Mashimo T, Kaneko T, Sakuma T, Kobayashi J, Kunihiro Y, Voigt B, Yamamoto T, Serikawa T (2013): Efficient gene targeting by TAL effector nucleases coinjected with exonucleases in zygotes. Sci. Rep. 3, 1253.

Masuda K, Ouchida R, Takeuchi A, Saito T, Koseki H, Kawamura K, Tagawa M, Tokuhisa T, Azuma T, O-Wang J (2005): DNA polymerase θ contributes to the generation of C/G mutations during somatic hypermutation of Ig genes. Proc. Natl. Acad. Sci. USA 102:13986–91.

Masuda K, Ouchida R, Hikida M, Nakayama M, Ohara O, Kurosaki T, O-Wang J (2006): Absence of DNA polymerase θ results in decreased somatic hypermutation frequency and altered mutation patterns in Ig genes. DNA Repair 5:1384–91.

Matsumoto Y, Marusawa H, Kinoshita K, Endo Y, Kou T, Morisawa T, Azuma T, Okazaki IM, Honjo T, Chiba T (2007): Helicobacter pylori infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. Nat Med. 13(4):470-6.

Matthews AJ, Zheng S, DiMenna LJ, Chaudhuri J (2014): Regulation of immunoglobulin class-switch recombination: choreography of noncoding transcription, targeted DNA deamination, and long-range DNA repair. Adv. Immunol. 122, 1–57.

Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, Johnson BE, Hong C, Nielsen C, Zhao Y, Turecki G, Delaney A, Varhol R, Thiessen N, Shchors K, Heine VM, Rowitch DH, Xing X, Fiore C, Schillebeeckx M, Jones SJ, Haussler D, Marra MA, Hirst M, Wang T, Costello JF (2010): Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature 466:253–257.

McBride KM, Barreto V, Ramiro AR, Stavropoulos P, Nussenzweig MC (2004): Somatic hypermutation is limited by CRM1-dependent nuclear export of activation-induced deaminase. J Exp Med. 199(9): 1235-44.

McBride KM, Gazumyan A,Woo EM, Barreto VM, Robbiani DV, Chait BT, Nussenzweig MC (2006): Regulation of hypermutation by activation-induced cytidine deaminase phosphorylation. Proc. Natl. Acad. Sci. USA 103:8798–803.

McCarrick JW, Parnes JR, Seong RH, Solter D, Knowles BB (1993): Positive-negative selection gene targeting with the diphteria toxin A-chain gene in mouse embryonic stem cells. Transgenic Research 2, 183-190.

McConnell Smith A, Takeuchi R, Pellenz S, Davis L, Maizels N, Monnat RJ Jr, Stoddard BL (2009): Generation of a nicking enzyme that stimulates site-specific gene conversion from the I-AniI LAGLIDADG homing endonuclease. Proc Natl Acad Sci U S A, 106(13):5099-104.

McManus MT and Sharp PA (2002): Gene silencing in mammals by small interfering RNAs. Nat. Rev. Genet. 3, 737–747.

Mendenhall EM, Williamson KE, Reyon D, Zou JY, Ram O, Joung JK, Bernstein BE (2013): Locus-specific editing of histone modifications at endogenous enhancers using programmable TALE-LSD1 fusions. Nature Biotechnology, 31(12), 10.1038/nbt.2701. doi:10.1038/nbt.2701

Meng X, Noyes MB, Zhu LJ, Lawson ND, Wolfe SA (2008): Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. Nat. Biotechnol. 26, 695–701.

Meng FL, Du Z, Federation A, Hu J, Wang Q, Kieffer-Kwon KR, Meyers RM, Amor C, Wasserman CR, Neuberg D, Casellas R, Nussenzweig MC, Bradner JE, Liu XS, Alt FW (2014): Convergent Transcription at Intragenic Super-Enhancers Targets AID-Initiated Genomic Instability. Cell 18;159(7): 1538-48.

Menoret S, Fontaniere S, Jantz D, Tesson L, Thinard R, Remy S, Usal C, Ouisse LH, Fraichard A, Anegon I (2013): Generation of Rag1-knockout immunodeficient rats and mice using engineered meganucleases. FASEB J. 27, 703–711.

Mercer AC, Gaj T, Fuller RP, Barbas CF III (2012): Chimeric TALE recombinases with programmable DNA sequence specificity. Nucleic Acids Res 40:11163–11172.

Merlin C, Beaver LE, Taylor OR, Wolfe SA, Reppert SM (2013): Efficient targeted mutagenesis in the monarch butterfly using zinc-finger nucleases. Genome Res. 23, 159–168.

Meyer M, de Angelis MH, Wurst W, Kühn R (2010): Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. Proc. Natl Acad. Sci. USA 107, 15022–15026.

Meyer M, Ortiz O, Hrabe de Angelis M, Wurst W, Kühn R (2012): Modeling disease mutations by gene targeting in one cell mouse embryos. Proc. Natl Acad. Sci. USA 109, 9354–9359.

Miller J, McLachlan AD, Klug A (1985): Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. The EMBO Journal, 4(6), 1609–1614.

Miller JC, Holmes MC, Wang J, Guschin DY, Lee YL, Rupniewski I, Beausejour CM, Waite AJ, Wang NS, Kim KA, Gregory PD, Pabo CO, Rebar EJ (2007): An improved zinc-finger nuclease architecture for highly specific genome editing. Nat. Biotechnol. 25, 778–785.

Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ, Dulay GP, Hua KL, Ankoudinova I, Cost GJ, Urnov FD, Zhang HS, Holmes MC, Zhang L, Gregory PD, Rebar EJ (2011): A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29:143–148.

Milstein C and Neuberger MS (1996): Maturation of the immune response. Adv Protein Chem. 49:451-85.

Milstein C, Neuberger MS, Staden R (1998): Both DNA strands of antibody genes are hypermutation targets. Proc. Natl. Acad. Sci. USA 95:8791–94.

Min I, Schrader C, Vardo J, D'Avirro N, Luby T, D'Avirro N, Stavnezer J, Selsing E (2003): The Sμ tandem repeat region is critical for isotype switching in the absence of Msh2. Immunity 19:515–24.

Min I, Rothlein L, Schrader C, Stavnezer J, Selsing E (2005): Shifts in targeting of class switch recombination sites in mice that lack µ switch region tandem repeats or Msh2. J. Exp. Med. 201:1885–90.

Minczuk M, Papworth MA, Miller JC, Murphy MP, Klug A (2008): Development of a single-chain, quasi-dimeric zinc-finger nuclease for the selective degradation of mutated human mitochondrial DNA. Nucleic Acids Res, 36, 3926–3938.

Mino T, Aoyama Y, Sera T (2009): Efficient double-stranded DNA cleavage by artificial zinc-finger nucleases composed of one zinc-finger protein and a single-chain FokI dimer. J. Biotechnol, 140, 156–161.

Miyanari Y, Ziegler-Birling C, Torres-Padilla ME (2013): Live visualization of chromatin dynamics with fluorescent TALEs. Nat Struct Mol Biol 20(11):1321-4. doi: 10.1038/nsmb.2680. Epub 2013 Oct 6.

Mochan TA, Venere M, DiTullio RA Jr, Halazonetis TD (2004): 53BP1, an activator of ATM in response to DNA damage. DNA Repair 3:945–52.

Moehle EA, Rock JM, Lee YL, Jouvenot Y, DeKelver RC, Gregory PD, Urnov FD, Holmes MC (2007): Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. Proc Natl Acad Sci U S A 104(9):3055-60.

Mojica FJ, Diez-Villaseñor C, Soria E, Juez G (2000): Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. Mol. Microbiol. 36, 244–246.

Mojica FJ, Diez-Villaseñor C, Garcia-Martinez, J, Soria E (2005): Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J. Mol. Evol. 60, 174–182.

Moore JK, Haber JE (1996): Cell cycle and genetic requirements of two pathways of non-homologous end joining repair of double-strand breaks in Saccharomyces cerevisiae, Mol. Cell. Biol. 16 (5) 2164–2173.

Moore FE, Reyon D, Sander JD, Martinez SA, Blackburn JS, Khayter C, Ramirez CL, Joung JK, Langenau DM (2012): Improved somatic mutagenesis in zebrafish using transcription activator-like effector nucleases (TALENs). PLoS ONE 7:e37877.

Morbitzer R, Romer P, Boch J, Lahaye T (2010): Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. Proc Natl Acad Sci USA 107:21617–21622

Morbitzer R, Elsaesser J, Hausner J, Lahaye T (2011): Assembly of custom TALE-type DNA binding domains by modular cloning. Nucleic Acids Res 39:5790–5799.

Moreno-Herrero F, de Jager M, Dekker NH, Kanaar R, Wyman C, Dekker C (2005): Mesoscale conformational changes in the DNA-repair complex Rad50/Mre11/Nbs1 upon binding DNA. Nature 437:440–43.

Mori T, Kagatsume I, Shinomiya K, Aoyama Y, Sera T (2009): Sandwiched zinc-finger nucleases harboring a single chain FokI dimer as a DNA-cleavage domain. Biochem. Biophys. Res. Commun. 390, 694–697.

Morrison DK (2009): The 14-3-3 proteins: integrators of diverse signaling cues that impact cell fate and cancer development. Trends Cell Biol. 19(1):16-23.

Moscou MJ and Bogdanove AJ (2009): A simple cipher governs DNA recognition by TAL effectors. Science 326:1501.

Munoz IG, Prieto J, Subramanian S, Coloma J, Redondo P, Villate M, Merino N, Marenchino M, D'Abramo M, Gervasio FL, Grizot S, Daboussi F, Smith J, Chion-Sotinel I, Paques F, Duchateau P, Alibes A, Stricher F, Serrano L, Blanco FJ, Montoya G (2011): Molecular basis of engineered meganuclease targeting of the endogenous human RAG1 locus. Nucleic Acids Res. 39, 729–743.

Munoz Bodnar A, Bernal A, Szurek B, Lopez CE (2013): Tell me a tale of TALEs. Mol Biotechnol 53:228–2235.

Muramatsu M, Sankaranand VS, Anant S, Sugai M, Kinoshita K, Davidson NO, Honjo T (1999): Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. J Biol Chem. 25;274(26):18470-6.

Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T (2000): Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell 1;102(5):553-63.

Müschen M, Re D, Jungnickel B, Diehl V, Rajewsky K, Küppers R (2000): Somatic mutation of the CD95 gene in human B cells as a side-effect of the germinal center reaction. J. Exp. Med. 192, 1833–1840.

Muto T, Okazaki IM, Yamada S, Tanaka Y, Kinoshita K, Muramatsu M, Nagaoka H, Honjo T (2006): Negative regulation of activation-induced cytidine deaminase in B cells. Proc Natl Acad Sci U S A. 103(8):2752-7.

Nakagawa Y, Yamamoto T, Suzuki K, Araki K, Takeda N, Ohmuraya M, Sakuma T (2013): Screening methods to identify TALEN-mediated knockout mice. Exp. Anim. 63, 79–84 (in press).

Nakamura M, Kondo S, Sugai M, Nazarea M, Imamura S, Honjo T (1996): High frequency class switching of an IgM+ B lymphoma clone CH12F3 to IgA+ cells. Int Immunol. 8(2):193-201.

Narita Y, Asai A, Kuchino Y, Kirino T (2000): Actinomycin D and staurosporine, potent apoptosis inducers in vitro, are potentially effective chemotherapeutic agents against glioblastoma multiforme. Cancer Chemother Pharmacol. 45(2):149-56.

Navaratnam N, Bhattacharya S, Fujino T, Patel D, Jarmuz AL, Scott J (1995): Evolutionary origins of apoB mRNA editing: catalysis by a cytidine deaminase that has acquired a novel RNA-binding motif at its active site. Cell. 81(2):187-95.

Neuberger MS and Milstein C (1995): Somatic hypermutation. Curr Opin Immunol. 7(2):248-54.

Nick McElhinny SA, Snowden CM, McCarville J, Ramsden DA (2000): Ku recruits the XRCC4-ligase IV complex to DNA ends. Mol. Cell. Biol. 20:2996–3003.

Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, Ishitani R, Zhang F, Nureki O (2014): Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell 156, 935–949.

Nowak U, Matthews AJ, Zheng S, Chaudhuri J (2011): The splicing regulator PTBP2 interacts with the cytidine deaminase AID and promotes binding of AID to switch-region DNA. Nat Immunol. 12(2):160-6.

Ochiai H, Fujita K, Suzuki K, Nishikawa M, Shibata T, Sakamoto N, Yamamoto T (2010): Targeted mutagenesis in the sea urchin embryo using zinc-finger nucleases. Genes Cells 15, 875–885.

Ochiai H, Sakamoto N, Fujita K, Nishikawa M, Suzuki K, Matsuura S, Miyamoto T, Sakuma T, Shibata T, Yamamoto T (2012): Zinc-finger nuclease-mediated targeted insertion of reporter genes for quantitative imaging of gene expression in sea urchin embryos. Proc. Natl Acad. Sci. USA 109, 10915–10920.

Okazaki IM, Hiai H, Kakazu N, Yamada S, Muramatsu M, Kinoshita K, Honjo T (2003): Constitutive expression of AID leads to tumorigenesis. J Exp Med. 5;197(9):1173-81.

Okazaki IM, Okawa K, Kobayashi M, Yoshikawa K, Kawamoto S, Nagaoka H, Shinkura R, Kitawaki Y, Taniguchi H, Natsume T, Iemura S, Honjo T (2011): Histone chaperone Spt6 is required for class switch recombination but not somatic hypermutation. Proc Natl Acad Sci U S A. 108(19):7920-5.

Olsen PA, Solhaug A, Booth JA, Gelazauskaite M, Krauss S (2009): Cellular responses to targeted genomic sequence modification using single-stranded oligonucleotides and zinc-finger nucleases. DNA Repair (Amst) 8(3):298-308.

Orthwein A, Patenaude A-M, Affar EB, Lamarre A, Young JC, Di Noia JM (2010): Regulation of activation-induced deaminase stability and antibody gene diversification by Hsp9 . The Journal of Experimental Medicine 207:2751–65.

Orthwein A and Di Noia JM (2012): Activation induced deaminase: how much and where? Semin Immunol. 24(4):246-54.

Ota S, Hisano Y, Muraki M, Hoshijima K, Dahlem TJ, Grunwald DJ, Okada Y, Kawahara A (2013): Efficient identification of TALEN-mediated genome modifications using heteroduplex mobility assays. Genes Cells 18, 450–458.

Ousterout DG, Perez-Pinera P, Thakore PI, Kabadi AM, Brown MT, Qin X, Fedrigo O, Mouly V, Tremblay JP, Gersbach CA (2013): Reading frame correction by targeted genome editing restores dystrophin expression in cells from duchenne muscular dystrophy patients. Mol. Ther. 21, 1718–1726.

Pan-Hammarstrom Q, Jones AM, Lahdesmaki A, Zhou W, Gatti RA, Hammarström L, Gennery AR, Ehrenstein MR (2005): Impact of DNA ligase IV on non homologous end joining pathways during class switch recombination in human cells. J. Exp. Med. 201:189–94.

Pannunzio NR, Li S, Watanabe G, Lieber MR (2014): Non-homologous end joining often uses microhomology: implications for alternative end joining. DNA Repair (Amst). 17:74-80.

Park SR, Zan H, Pal Z, Zhang J, Al-Qahtani A, Pone EJ, Xu Z, Mai T, Casali P (2009): HoxC4 binds to the promoter of the cytidine deaminase AID gene to induce AID expression, class-switch DNA recombination and somatic hypermutation. Nat Immunol. 10(5):540-50.

Parker SC, Stitzel ML, Taylor DL, Orozco JM, Erdos MR, Akiyama JA, Van Bueren KL, Chines PS, Narisu N, Black BL, Visel A, Pennacchio LA, Collins FS; NISC Comparative Sequencing Program; National Institutes of Health Intramural Sequencing Center Comparative Sequencing Program Authors; NISC Comparative Sequencing Program Authors (2013): Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. Proc. Natl. Acad. Sci. USA 110, 17921–17926.

Pasqualucci L, Migliazza A, Fracchiolla N, William C, Neri A, Baldini L, Chaganti RS, Klein U, Küppers R, Rajewsky K, Dalla-Favera R (1998): BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. Proc. Natl. Acad. Sci. USA 95, 11816–11821.

Pasqualucci L, Bhagat G, Jankovic M, Compagno M, Smith P, Muramatsu M, Honjo T, Morse HC 3rd, Nussenzweig MC, Dalla-Favera R (2008): AID is required for germinal center-derived lymphomagenesis. Nat Genet. 40(1):108-12.

Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR (2013): High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat. Biotechnol. 31, 839–843.

Pavri R, Gazumyan A, Jankovic M, Di Virgilio M, Klein I, Ansarah-Sobrinho C, Resch W, Yamane A, Reina San-Martin B, Barreto V, Nieland TJ, Root DE, Casellas R, Nussenzweig MC (2010): Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. Cell 143, 122–133.

Pefanis E, Wang J, Rothschild G, Lim J, Chao J, Rabadan R, Economides AN, Basu U (2014): Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. Nature 514, 389–393.

Peitz M, Pfannkuche K, Rajewsky K, Edenhofer F (2002): Ability of the hydrophobic FGF and basic TAT peptides to promote cellular uptake of recombinant Cre recombinase: a tool for efficient genetic engineering of mammalian genomes. Proc Natl Acad Sci U S A. 99(7):4489-94.

Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O, Wang N, Lee G, Bartsevich VV, Lee YL, Guschin DY, Rupniewski I, Waite AJ, Carpenito C, Carroll RG, Orange JS, Urnov FD, Rebar EJ, Ando D, Gregory PD, Riley JL, Holmes MC, June CH (2008): Establishment of HIV-1 resistance in CD4 + T cells by genome editing using zinc-finger nucleases. Nat. Biotechnol. 26, 808–816.

Perez-Pinera P, Ousterout DG, Brown MT, Gersbach CA (2012): Gene targeting to the ROSA26 locus directed by engineered zinc finger nucleases. Nucleic Acids Res. 40(8):3741-52.

Perez-Pinera P, Kocak DD, Vockley CM, Adler AF, Kabadi AM, Polstein LR, Thakore PI, Glass KA, Ousterout DG, Leong KW, Guilak F, Crawford GE, Reddy TE, Gersbach CA (2013): RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nat. Methods 10, 973–976.

Perlot T, Alt FW, Bassing CH, Suh H, Pinaud E (2005): Elucidation of IgH intronic enhancer functions via germ-line deletion. Proc Natl Acad Sci U S A. 102(40):14362-7.

Perlot T and Alt FW (2008): Cis-regulatory elements and epigenetic changes control genomic rearrangements of the IgH locus. Adv Immunol. 99:1-32.

Peters A and Storb U (1996): Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. Immunity. 4(1):57-65.

Petersen-Mahrt SK, Harris RS, Neuberger MS (2002): AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. Nature 4;418(6893):99-103.

Petersen-Mahrt SK and Neuberger MS (2003): In vitro deamination of cytosine to uracil in single-stranded DNA by apolipoprotein B editing complex catalytic subunit 1 (APOBEC1). J Biol Chem. 30;278(22):19583-6.

Pettersson S, Cook GP, Brüggemann M, Williams GT, Neuberger MS (1990): A second B cell-specific enhancer 3' of the immunoglobulin heavy-chain locus. Nature. 344(6262):165-8.

Pfeiffer P., Vielmetter W. (1988): Joining of non-homologous DNA double strand breaks in vitro, Nucl. Acids Res. 16 (3) 907–924.

Pham P, Bransteitter R, Petruska J, Goodman MF (2003): Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. Nature 3;424(6944):103-7.

Piganeau M, Ghezraoui H, De Cian A, Guittat L, Tomishima M, Perrouault L, Rene O, Katibah GE, Zhang L, Holmes MC, Doyon Y, Concordet JP, Giovannangeli C, Jasin M, Brunet E (2013): Cancer translocations in human cells induced by zinc finger and TALE nucleases. Genome Res. 23, 1182–1193.

Politz MC, Copeland MF, Pfleger BF (2012): Artificial repressors for controlling gene expression in bacteria. Chem Commun (Camb), DOI: 10.1039/c2cc37107c

Pone EJ, Zhang J, Mai T, White CA, Li G, Sakakura JK, Patel PJ, Al-Qahtani A, Zan H, Xu Z, Casali P (2012): BCR-signalling synergizes with TLR-signalling for induction of AID and immunoglobulin class-switching through the non-canonical NF-κB pathway. Nat Commun. 3;3:767.

Porteus MK, Baltimore D (2003): Chimeric nucleases stimulate gene targeting in human cells. Science 2;300(5620):763.

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

Pryde F, Khalili S, Robertson K, Selfridge J, Ritchie AM, Melton DW, Jullien D, Adachi Y (2005): 53BP1 exchanges slowly at the sites of DNA damage and appears to require RNA for its association with chromatin. J. Cell Sci. 118:2043–55.

Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA (2013): Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152, 1173–1183.

Qian J, Wang Q, Dose M, Pruett N, Kieffer-Kwon KR, Resch W, Liang G, Tang Z, Mathé E, Benner C, Dubois W, Nelson S, Vian L, Oliveira TY, Jankovic M, Hakim O, Gazumyan A, Pavri R, Awasthi P, Song B, Liu G, Chen L, Zhu S, Feigenbaum L, Staudt L, Murre C, Ruan Y, Robbiani DF, Pan-Hammarström Q, Nussenzweig MC, Casellas R (2014): B Cell Super-Enhancers and Regulatory Clusters Recruit AID Tumorigenic Activity. Cell 18;159(7):1524-37.

Rada C, Jarvis JM, Milstein C (2002): AID-GFP chimeric protein increases hypermutation of Ig genes with no evidence of nuclear localization. Proc Natl Acad Sci U S A 99:7003-7008.

Radecke S, Radecke F, Cathomen T, Schwarz K (2010): Zinc-finger nuclease-induced gene repair with oligodeoxynucleotides: wanted and unwanted target locus modifications. Mol Ther 18(4):743-53.

Rafalska-Metcalf IU and Janicki SM (2007): Show and tell: visualizing gene expression in living cells. J Cell Sci. 15;120(Pt 14):2301-7.

Rajagopal D, Maul RW, Ghosh A, Chakraborty T, Khamlichi AA, Sen R, Gearhart PJ (2009): Immunoglobulin switch mu sequence causes RNA polymerase II accumulation and reduces dA hypermutation. J. Exp. Med. 206, 1237–1244.

Rajewsky K (1996): Clonal selection and learning in the antibody system. Nature 381, 751–758.

Ramakrishna S, Cho SW, Kim S, Song M, Gopalappa R, Kim JS, Kim H (2014a): Surrogate reporter-based enrichment of cells containing RNA-guided Cas9 nuclease-induced mutations. Nat Commun 26;5:3378.

Ramakrishna S, Kwaku Dad AB, Beloor J, Gopalappa R, Lee SK, Kim H (2014b): Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. Genome Res; 24(6): 1020-7.

Ramalingam S, Kandavelou K, Rajenderan R, Chandrasegaran S (2011): Creating designed zinc-finger nucleases with minimal cytotoxicity. J. Mol. Biol. 405, 630–641.

Ramirez CL, Certo MT, Mussolino C, Goodwin MJ, Cradick TJ, McCaffrey AP, Cathomen T, Scharenberg AM, Joung JK (2012): Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. Nucleic Acids Res. 40, 5560–5568.

Ramiro AR, Stavropoulos P, Jankovic M, Nussenzweig MC (2003): Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. Nat Immunol. 4(5):452-6.

Ramiro AR, Jankovic M, Eisenreich T, Difilippantonio S, Chen-Kiang S, Muramatsu M, Honjo T, Nussenzweig A, Nussenzweig MC (2004): AID is required for c-myc/IgH chromosome translocations in vivo. Cell 118:431–38.

Ramiro AR, Jankovic M, Callen E, Difilippantonio S, Chen HT, McBride KM, Eisenreich TR, Chen J, Dickins RA, Lowe SW, Nussenzweig A, Nussenzweig MC (2006): Role of genomic instability and p53 in AID-induced c-myc-Igh translocations. Nature 440:105–9.

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F (2013): Genome engineering using the CRISPR-Cas9 system. Nat Protocols, 8(11):2281–2308.

Ranjit S, Khair L, Linehan EK, Ucher AJ, Chakrabarti M, Schrader CE, Stavnezer J (2011): AID recruits UNG and Msh2 to Ig switch regions dependent upon the AID C terminus. J Immunol. 187(5):2464-75.

Rawlings DJ, Schwartz MA, Jackson SW, Meyer-Bahlburg A (2012): Integration of B cell responses through Toll-like receptors and antigen receptors. Nat Rev Immunol. 12(4):282-94.

Reina-San-Martin B, Difilippantonio S, Hanitsch L, Masilamani RF, Nussenzweig A, Nussenzweig MC (2003): H2AX is required for recombination between immunoglobulin switch regions but not for intraswitch region recombination or somatic hypermutation. J. Exp. Med. 197:1767–78.

Reina-San-Martin B, Chen HT, Nussenzweig A, Nussenzweig MC (2004): ATM is required for efficient recombination between immunoglobulin switch regions. J. Exp. Med. 200:1103–10.

Reina-San-Martin B, Nussenzweig MC, Nussenzweig A, Difilippantonio S (2005): Genomic instability, endoreduplication, and diminished Ig class-switch recombination in B cells lacking Nbs1. Proc. Natl. Acad. Sci. USA 102:1590–95.

Reina-San-Martin B, Chen J, Nussenzweig A, Nussenzweig MC (2007): Enhanced intraswitch region recombination during immunoglobulin class switch recombination in 53BP1–/– B cells. Eur. J. Immunol. 37:235–39.

Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, Catalan N, Forveille M, Dufourcq-Labelouse R, Gennery A, Tezcan I, Ersoy F, Kayserili H, Ugazio AG, Brousse N, Muramatsu M, Notarangelo LD, Kinoshita K, Honjo T, Fischer A, Durandy A (2000): Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). Cell 1;102(5):565-75.

Reyon D, Khayter C, Regan MR, Joung JK, Sander JD (2012a): Engineering designer transcription activator-like effector nucleases (TALENs) by REAL or REAL-Fast assembly. Curr Protoc Mol Biol Chapter 12:Unit 12.15.

Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK (2012b): FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol, 30(5):460–465.

Robbiani DF, Bothmer A, Callen E, Reina San-Martin B, Dorsett Y, Difilippantonio S, Bolland DJ, Chen HT, Corcoran AE, Nussenzweig A, Nussenzweig MC (2008): AID is required for the chromosomal translocations in c-myc that lead to c-myc/IgH translocations. Cell 135, 1028–1038.

Robbiani DF, Bunting S, Feldhahn N, Bothmer A, Camps J, Deroubaix S, McBride KM, Klein IA, Stone G, Eisenreich TR, Ried T, Nussenzweig A, Nussenzweig MC (2009): AID produces DNA double-strand breaks in non-Ig genes and mature B cell lymphomas with reciprocal chromosome translocations. Mol. Cell 36, 631–641.

Roldan E, Fuxa M, ChongW,Martinez D, Novatchkova M, Busslinger M, Skok JA (2005): Locus ‘decontraction’ and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. Nat. Immunol. 6:31–41.

Rooney S, Chaudhuri J, Alt FW (2004): The role of the non homologous end-joining pathway in lymphocyte development. Immunol. Rev. 200:115–31.

Roth DB, Wilson JH (1986): Non-homologous recombination in mammalian cells: role for short sequence homologies in the joining reaction, Mol. Cell. Biol. 6 (12) 4295–4304.

Rouet P, Smih F, Jasin M (1994a): Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proc Natl Acad. Sci. USA, 91, 6064–6068.

Rouet P, Smih F, Jasin M (1994b): Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Mol Cell Biol 14(12) 8096-106.

Rucci F, Cattaneo L, Marrella V, Sacco MG, Sobacchi C, Lucchini F, Nicola S, Della Bella S, Villa ML, Imberti L, Gentili F, Montagna C, Tiveron C, Tatangelo L, Facchetti F, Vezzoni P, Villa A (2006): Tissue-specific sensitivity to AID expression in transgenic mouse models. Gene 1;377:150-8.

Sakane Y, Sakuma T, Kashiwagi K, Kashiwagi A, Yamamoto T, Suzuki K (2013): Targeted mutagenesis of multiple and paralogous genes in Xenopus laevis using two pairs of transcription activator-like effector nucleases. Dev. Growth Differ 56, 108–114.

Sakuma T, Hosoi S, Woltjen K, Suzuki K, Kashiwagi K, Wada H, Ochiai H, Miyamoto T, Kawai N, Sasakura Y, Matsuura S, Okada Y, Kawahara A, Hayashi S, Yamamoto T (2013): Efficient TALEN construction and evaluation methods for human cell and animal applications. Genes Cells 18, 315–326.

Sakuma T and Woltjen K (2014): Nuclease-mediated genome editing: At the front-line of functional genomics technology. Dev Growth Differ 56(1):2-13.

Sambrook J and Russell D (2001): Molecular cloning: A laboratory manual. Cold spring harbor laboratory press. 66, 75.

Sander JD, Cade L, Khayter C, Reyon D, Peterson RT, Joung JK, Yeh JR (2011a): Targeted gene disruption in somatic zebrafish cells using engineered TALENs. Nat Biotechnol 29:697–698.

Sander JD, Dahlborg EJ, Goodwin MJ, Cade L, Zhang F, Cifuentes D, Curtin SJ, Blackburn JS, Thibodeau-Beganny S, Qi Y, Pierick CJ, Hoffman Em, Maeder ML, Khayter C, Reyon D, Doobs D, Langenau DM, STupar RM, Giraldez AJ, Voytas DF, Peterson RT, Yeh JR, Joung, J. K. (2011b): Selection-Free Zinc-Finger Nuclease Engineering by Context-Dependent Assembly (CoDA). Nature Methods, 8(1), 67–69.

Sander JD, Joung JK (2014): CRISPR-Cas systems for editing, regulating and targeting genomes. Nat. Biotechnol. 32, 347–355.

Sanders KL, Catto LE, Bellamy SR, Halford SE (2009): Targeting individual subunits of the FokI restriction endonuclease to specific DNA strands. Nucleic Acids Res 37: 2105–2115.

Sanjana NE, Cong L, Zhou Y, Cunniff MM, Feng G, Zhang F (2012): A transcription activator-like effector toolbox for genome engineering. Nat Protoc 7:171–192.

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

Saraconi G, Severi F, Sala C, Mattiuz G, Conticello SG (2014): The RNA editing enzyme APOBEC1 induces somatic mutations and a compatible mutational signature is present in esophageal adenocarcinomas. Genome Biol. 31;15(7):417.

Sayegh CE, Quong MW, Agata Y, Murre C (2003): E-proteins directly regulate expression of activation-induced deaminase in mature B cells. Nat Immunol. 4(6):586-93.

Sayegh C, Jhunjhunwala S, Riblet R, Murre C (2005): Visualization of looping involving the immunoglobulin heavy-chain locus in developing B cells. Genes. Dev. 19:322–27.

Schmid-Burgk JL, Schmidt T, Kaiser V, Honing K, Hornung V (2013): A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. Nat Biotechnol 31:76–81.

Schrader CE, Edelmann W, Kucherlapati R, Stavnezer J (1999): Reduced isotype switching in splenic B cells from mice deficient in mismatch repair enzymes. J. Exp. Med. 190:323–30.

Schrader CE, Bradley SP, Vardo J, Mochegova SN, Flanagan E, Stavnezer J (2003): Mutations occur in the Ig Sμ region but rarely in Sγ regions prior to class switch recombination. EMBO J. 22:5893–903.

Schrader CE, Linehan EK, Mochegova SN,Woodland RT, Stavnezer J (2005): Inducible DNA breaks in Ig S regions are dependent upon AID and UNG. J. Exp. Med. 202:561–68. Sci. U.S.A. 107, 12028–12033.

Schrader CE, Guikema JEJ, Linehan EK, Selsing E, Stavnezer J (2007): AID-dependent DNA breaks in class switch recombination occur during G1 phase and are mismatch repair-dependent. J. Immunol. 179:6064–71.

Sena-Esteves M, Saeki Y, Camp SM, Chiocca EA, Breakefield XO (1999): Single-step conversion of cells to retrovirus vector producers with herpes simplex virus-Epstein-Barr virus hybrid amplicons. J Virol. 73(12):10426-39.

Sernández IV, de Yébenes VG, Dorsett Y, Ramiro AR (2008): Haploinsufficiency of activation-induced deaminase for antibody diversification and chromosome translocations both in vitro and in vivo. PLoS One. 3(12):e3927.

Severi F, Chicca A, Conticello SG (2011): Analysis of reptilian APOBEC1 suggests that RNA editing may not be its ancestral function. Mol Biol Evol. 28(3):1125-9.

Shah SA, Erdmann S, Mojica FJ, Garrett RA (2013): Protospacer recognition motifs: mixed identities and functional diversity. RNA Biol. 10, 891–899.

Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, Zhang K, Liu J, Xi JJ, Qiu JL, Gao C (2013): Targeted genome modification of crop plants using a CRISPR-Cas system. Nat Biotechnol, 31:686–688.

Shav-Tal Y, Darzacq X, Shenoy SM, Fusco D, Janicki SM, Spector DL, Singer RH (2004): Dynamics of single mRNPs in nuclei of living cells. Science304, 1797-1800.

Shen HM, Peters A, Baron B, Zhu X, Storb U (1998): Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. Science 280, 1750–1752.

Shen HM, Bozek G, Pinkert CA, McBride K, Wang L, Kenter A, Storb U (2008): Expression of AID transgene is regulated in activated B cells but not in resting B cells and kidney. Mol Immunol. 45(7): 1883-92.

Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA, Shi Y (2004): Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell 29;119(7):941-53.

Shinkura R, Ito S, Begum NA, Nagaoka H, Muramatsu M, Kinoshita K, Sakakibara Y, Hijikata H, Honjo T (2004): Separate domains of AID are required for somatic hypermutation and class-switch recombination. Nat Immunol. 5(7):707-12.

Shukla VK, Doyon Y, Miller JC, DeKelver RC, Moehle EA, Worden SE, Mitchell JC, Arnold NL, Gopalan S, Meng X, Choi VM, Rock JM, Wu YY, Katibah GE, Zhifang G, McCaskill D, Simpson MA, Blakeslee B, Greenwalt SA, Butler HJ, Hinkley SJ, Zhang L, Rebar EJ, Gregory PD, Urnov FD (2009): Precise genome modification in the crop species Zea mays using zinc-finger nucleases. Nature 459, 437–441.

Smale ST (2011): Hierarchies of NF-κB target-gene regulation. Nat Immunol. 12(8):689-94.

Smidler AL, Terenzi O, Soichot J, Levashina EA, Marois E (2013): Targeted mutagenesis in the malaria mosquito using TALE nucleases. PLoS ONE 8, e74511.

Sohail A, Klapacz J, Samaranayake M, Ullah A, Bhagwat AS (2003): Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations. Nucleic Acids Res. 15;31(12):2990-4.

Sonoda E, Hochegger H, Saberi A, Taniguchi Y, Takeda S (2006): Differential usage of non homologous end-joining and homologous recombination in double strand break repair. DNA Repair 5:1021–29.

Sood R, Carrington B, Bishop K, Jones M, Rissone A, Candotti F, Chandrasekharappa SC, Liu P (2013): Efficient methods for targeted mutagenesis in zebrafish using zinc-finger nucleases: data from targeting of nine genes using CompoZr or CoDA ZFNs. PLoS One, 8(2):e57239.

Soriano P (1999): Generalized lacZ expression with the ROSA26 Cre reporter strain. Nat. Genet., 21, 70–71.

Soulas-Sprauel P, Guyader GL, Rivera-Munoz P, Abramowski V, Olivier-Martin C, Goujet-Zalc C, Charneau P, de Villartay JP (2007): Role for DNA repair factor XRCC4 in immunoglobulin class switch recombination. J. Exp. Med. 204:1717–27.

Southern EM (1975): Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol. 98(3):503-17.

Staszewski O, Baker RE, Ucher AJ, Martier R, Stavnezer J, Guikema JE (2011): Activation-induced cytidine deaminase induces reproducible DNA breaks at many non-Ig Loci in activated B cells. Mol Cell. 41(2):232-42.

Stavnezer J and Schrader CE (2006): Mismatch repair converts AID-instigated nicks to double-strand breaks for antibody class-switch recombination. Trends Genet. 22:23–28.

Stavnezer J, Guikema JE, Schrader CE (2008): Mechanism and Regulation of Class Switch Recombination. Annual Review of Immunology 26, 261–292.

Stavnezer J (2011): Complex regulation and function of activation-induced cytidine deaminase. Trends Immunol. 32(5):194-201.

Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA (2014): DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature 507, 62–67.

Storb U (2014): Why does somatic hypermutation by AID require transcription of its target genes? Adv. Immunol. 122, 253–277.

Storici F, Durham CL, Gordenin DA, Resnick MA (2003): Chromosomal site-specific double-strand breaks are efficiently targeted for repair by oligonucleotides in yeast. Proc Natl Acad Sci U S A 100(25):14994-9.

Streubel J, Blucher C, Landgraf A, Boch J (2012): TAL effector RVD specificities and efficiencies. Nat Biotechnol 30:593–595.

Su Z, Han L, Zhao Z (2011): Conservation and divergence of DNA methylation in eukaryotes: New insights from single base-resolution DNA methylomes. Epigenetics 6:134–140.

Sun N, Liang J, Abil Z, Zhao H (2012): Optimized TAL effector nucleases (TALENs) for use in treatment of sickle cell disease. Mol Biosyst 8:1255–1263.

Sun N and Zhao H (2013): Transcription activator-like effector nucleases (TALENs): a highly efficient and versatile tool for genome editing. Biotechnol Bioeng. 110(7):1811-21.

Sung YH, Baek IJ, Kim DH, Jeon J, Lee J, Lee K, Jeong D, Kim JS, Lee HW (2013): Knockout mice created by TALEN-mediated gene targeting. Nat. Biotechnol. 31, 23–24.

Suzuki KT, Isoyama Y, Kashiwagi K, Sakuma T, Ochiai H, Sakamoto N, Furuno N, Kashiwagi A, Yamamoto T (2013): High efficiency TALENs enable F0 functional analysis by targeted gene disruption in Xenopus laevis embryos. Biol. Open 2, 448–452.

Szczepek M, Brondani V, Büchel J, Serrano L, Segal DJ, Cathomen T (2007): Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. Nat Biotechnol, 25(7):786-93.

Ta VT, Nagaoka H, Catalan N, Durandy A, Fischer A, Imai K, Nonoyama S, Tashiro J, Ikegawa M, Ito S, Kinoshita K, Muramatsu M, Honjo T (2003): AID mutant analyses indicate requirement for class-switch-specific cofactors. Nat Immunol. 4(9):843-8.

Tang TH, Bachellerie JP, Rozhdestvensky T, Bortolin ML, Huber H, Drungowski M, Elge T, Brosius J, Hüttenhofer, A (2002): Identification of 86 candidates for small non-messenger RNAs from the archaeon Archaeoglobus fulgidus. Proc. Natl. Acad. Sci. USA 99, 7536–7541.

Tesson L, Usal C, Menoret S, Leung E, Niles BJ, Remy S, Santiago Y, Vincent AI, Meng X, Zhang L, Gregory PD, Anegon I, Cost GJ (2011): Knockout rats generated by embryo microinjection of TALENs. Nat. Biotechnol. 29, 695–696.

Thakar R, Gordon G, Csink AK (2006): Dynamics and anchoring of heterochromatic loci during development. J. Cell Sci. 119, 4165-4175.

Thode S., Schäfer A, Pfeiffer P, Vielmetter W (1990): A novel pathway of DNA end-to-end joining, Cell 60 (6) 921–928.

Thomas KR, Folger KR, Capecchi MR (1986): High frequency targeting of genes to specific sites in the mammalian genome. Cell 44, 419–428.

Todaro GJ and Green H (1963): Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines. J Cell Biol. 17:299-313.

Tonegawa S (1983): Somatic generation of antibody diversity. Nature 302(5909):575-81.

Townsend JA, Wright DA, Winfrey RG, Fu F, Maeder ML, Joung JK, Voytas DF (2009): High-frequency modification of plant genes using engineered zinc-finger nucleases. Nature 459, 442–445.

Tran TH, Nakata M, Suzuki K, Begum NA, Shinkura R, Fagarasan S, Honjo T, Nagaoka H (2010): B cell-specific and stimulation-responsive enhancers derepress Aicda by overcoming the effects of silencers. Nat Immunol. 11(2):148-54.

Tremblay JP, Chapdelaine P, Coulombe Z, Rousseau J (2012): Transcription activator-like effector proteins induce the expression of the frataxin gene. Hum Gene Ther 23:883–890.

Tsai AG, Lu H, Raghavan SC, Muschen M, Hsieh CL, Lieber MR (2008): Human chromosomal translocations at CpG sites and a theoretical basis for their lineage and stage specificity. Cell 135, 1130–1142.

Uchimura Y, Barton LF, Rada C, Neuberger MS (2011): REG-gamma associates with and modulates the abundance of nuclear activation-induced deaminase. The Journal of Experimental Medicine 208:2385–91.

Unniraman S, Zhou S, Schatz DG (2004): Identification of an AID-independent pathway for chromosomal translocations between the Igh switch region and Myc. Nat. Immunol.5:1117–23.

Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, Jamieson AC, Porteus MH, Gregory PD, Holmes MC (2005): Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 2; 435(7042):646-51.

Urnov, FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD (2010): Genome editing with engineered zinc finger nucleases. Nat. Rev. Genet. 11, 636–646.

Valton J, Dupuy A, Daboussi F, Thomas S, Marechal A, Macmaster R, Melliand K, Juillerat A, Duchateau P (2012): Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. J Biol Chem 287:38427–38432.

Van Den Bosch M, Lohman PHM, Pastink A (2002): DNA Double-Strand Break Repair by Homologous Recombination. Biol. Chem., Vol. 383, pp. 873 – 892.

Van Gent DC, Hoeijmakers JH, Kanaar R (2001): Chromosomal stability and the DNA double-stranded break connection. Nat Rev Genet.;2(3):196-206.

Venken KJ, Bellen HJ (2007): Transgenesis upgrades for Drosophila melanogaster. Development, 134(20):3571–3584.

Victora GD and Nussenzweig MC (2012): Germinal centers. Annu. Rev. Immunol. 30, 429–457.

Voytas, DF (2013): Plant genome engineering with sequence specific nucleases. Annu. Rev. Plant Biol. 64, 327–350.

Vuong BQ, Lee M, Kabir S, Irimia C, Macchiarulo S, McKnight GS, Chaudhuri J (2009): Specific recruitment of protein kinase A to the immunoglobulin locus regulates class-switch recombination. Nat Immunol. 10(4):420-6.

Waisman A, Kraus M, Seagal J, Ghosh S, Melamed D, Song J, Sasaki Y, Classen S, Lutz C, Brombacher F, Nitschke L, Rajewsky K (2007): IgG1 B cell receptor signaling is inhibited by CD22 and promotes the development of B cells whose survival is less dependent on Ig α/β. J. Exp. Med. 204:747–58.

Wang Z, Zhou ZJ, Liu DP, Huang JD (2008): Double-stranded break can be repaired by single-stranded oligonucleotides via the ATM/ATR pathway in mammalian cells. Oligonucleotides 18(1): 21-32.

Wang L, Wuerffel R, Feldman S, Khamlichi AA, Kenter AL (2009a): S region sequence, RNA polymerase II, and histone modifications create chromatin accessibility during class switch recombination. J. Exp. Med. 206, 1817–1830.

Wang M, Yang Z, Rada C, Neuberger MS (2009b): AID upmutants isolated using a high-throughput screen highlight the immunity/cancer balance limiting DNA deaminase activity. Nat Struct Mol Biol. 16(7):769-76.

Wang J, Friedman G, Doyon Y, Wang NS, Li CJ, Miller JC, Hua KL, Yan JJ, Babiarz JE, Gregory PD, Holmes MC (2012a): Targeted gene addition to a predetermined site in the human genome using a ZFN-based nicking enzyme. Genome Res. 22, 1316–1326.

Wang Z, Li J, Huang H, Wang G, Jiang M, Yin S, Sun C, Zhang H, Zhuang F, Xi JJ (2012b): An integrated chip for the high-throughput synthesis of transcription activator-like effectors. Angew Chem Int Ed Engl 51: 8505–8508.

Ward IM, Reina-San-Martin B, Olaru A, Minn K, Tamada K, Lau JS, Cascalho M, Chen L, Nussenzweig A, Livak F, Nussenzweig MC, Chen J (2004): 53BP1 is required for class switch recombination. J. Cell Biol. 165:459–64.

Watanabe T, Ochiai H, Sakuma T, Horch HW, Hamaguchi N, Nakamura T, Bando T, Ohuchi H, Yamamoto T, Noji S, Mito T (2012): Non-transgenic genome modifications in a hemimetabolous insect using zinc-finger and TAL effector nucleases. Nat. Commun. 3, 1017.

Weber E, Gruetzner R, Werner S, Engler C, Marillonnet S (2011): Assembly of designer TAL effectors by Golden Gate cloning. PLoS ONE 6: e19722.

Wefers B, Meyer M, Ortiz O, Hrabe de Angelis M, Hansen J, Wurst W, Kühn, R (2013): Direct production of mouse disease models by embryo microinjection of TALENs and oligodeoxynucleotides. Proc. Natl Acad. Sci. USA 110, 3782–3787.

White FF, Potnis N, Jones JB, Koebnik R (2009): The type III effectors of Xanthomonas. Mol Plant Pathol 10:749–7766.

Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA (2013): Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell 153, 307–319.

Wiedenheft B, Sternberg SH, Doudna JA (2012): RNA-guided genetic silencing systems in bacteria and archaea. Nature, 482:331–338.

Wilson TM, Vaisman A, Martomo SA, Sullivan P, Lan L, Hanaoka F, Yasui A, Woodgate R, Gearhart PJ (2005): MSH2-MSH6 stimulates DNA polymerase η, suggesting a role for A:T mutations in antibody genes. J. Exp. Med. 201:637–45.

Wong EA, Capecchi MR (1986): Analysis of homologous recombination in cultured mammalian cells in transient expression and stable transformation assays. Somat. Cell Mol. Genet. 12, 63–72.

Wood AJ, Lo TW, Zeitler B, Pickle CS, Ralston EJ, Lee AH, Amora R, Miller JC, Leung E, Meng X, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Meyer BJ (2011): Targeted genome editing across species using ZFNs and TALENs. Science 333:307.

Wu X, Geraldes P, Platt JL, Cascalho M (2005): The double-edged sword of activation-induced cytidine deaminase. J Immunol. 15;174(2):934-41.

Wu X and Stavnezer J (2007): DNA polymerase β is able to repair breaks in switch regions and plays an inhibitory role during immunoglobulin class switch recombination. J. Exp. Med. 204:1677–89.

Wyman C and Kanaar R (2006): DNA double-strand break repair: all's well that ends well. Annu. Rev. Genet. 40, 363–383.

Xanthoudakis S, Smeyne RJ,Wallace JD, Curran T (1996): The redox/DNA repair protein, Ref-1, is essential for early embryonic development in mice. Proc. Natl. Acad. Sci. USA 93:8919–23.

Xiao A, Wang Z, Hu Y, Wu Y, Luo Z, Yang Z, Zu Y, Li W, Huang P, Tong X, Zhu Z, Lin S, Zhang B (2013): Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. Nucleic Acids Res. 41, e141.

Xie K and Yang Y (2013): RNA-guided genome editing in plants using a CRISPR-Cas system. Mol Plant. doi:10.1093/mp/sst119.

Xu Z, Pone EJ, Al-Qahtani A, Park SR, Zan H, Casali P (2007): Regulation of aicda expression and AID activity: relevance to somatic hypermutation and class switch DNA recombination. Crit Rev Immunol. 27(4):367-97.

Xu Z, Fulop Z, Wu G, Pone EJ, Zhang J, Mai T, Thomas LM, Al-Qahtani A, White CA, Park SR, Steinacker P, Li Z, Yates J 3rd, Herron B, Otto M, Zan H, Fu H, Casali P (2010): 14-3-3 adaptor proteins recruit AID to 5'-AGCT-3'-rich switch regions for class switch recombination. Nat Struct Mol Biol. 17(9):1124-35.

Xu Z, Zan H, Pone EJ, Mai T, Casali P (2012): Immunoglobulin class-switch DNA recombination: induction, targeting and beyond. Nat Rev Immunol. 12(7):517-31.

Xue K, Rada C, Neuberger MS (2006): The in vivo pattern of AID targeting to immunoglobulin switch regions deduced from mutation spectra in msh2–/– ung–/– mice. J. Exp. Med. 203:2085–94.

Yan CT, Boboila C, Souza EK, Franco S, Hickernell TR, Murphy M, Gumaste S, Geyer M, Zarrin AA, Manis JP, Rajewsky K, Alt FW (2007): IgH class switching and translocations use a robust nonclassical end-joining pathway. Nature. 449:478–82.

Yang Y and Gabriel DW (1995): Xanthomonas avirulence/pathogenicity gene family encodes functional plant nuclear targeting signals. Mol Plant Microbe Interact 8:627–631.

Yang SY and Schatz DG (2007): Targeting of AID-mediated sequence diversification by cis-acting determinants. Adv Immunol. 94:109-125.

Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R (2013): One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. Cell 154, 1370–1379.

Young JJ, Cherone JM, Doyon Y, Ankoudinova I, Faraji FM, Lee AH, Ngo C, Guschin DY, Paschon DE, Miller JC, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Harland RM, Zeitler B (2011): Efficient targeted gene disruption in the soma and germ line of the frog Xenopus tropicalis using engineered zinc-finger nucleases. Proc. Natl Acad. Sci. USA 108, 7052–7057.

Yu Y, Streubel J, Balzergue S, Champion A, Boch J, Koebnik R, Feng J, Verdier V, Szurek B (2011): Colonization of rice leaf blades by an African strain of Xanthomonas oryzae pv. oryzae depends on a new TAL effector that induces the rice nodulin-3 Os11N3 gene. Mol Plant Microbe Interact 24:1102–1113.

Zambrowicz BP, Imamoto A, Fiering S, Herzenberg LA, Kerr WG, Soriano P (1997): Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. Proc. Natl Acad. Sci. USA, 94, 3789–3794.

Zan H, Shima N, Xu Z, Al-Qahtani A, Evinger AJ Iii, Zhong Y, Schimenti JC, Casali P (2005): The translesion DNA polymerase θ plays a dominant role in immunoglobulin gene somatic hypermutation. EMBO J. 24:3757–69.

Zarnegar B, He JQ, Oganesyan G, Hoffmann A, Baltimore D, Cheng G (2004): Unique CD40-mediated biological program in B cell activation requires both type 1 and type 2 NF-kappaB activation pathways. Proc Natl Acad Sci U S A. 101(21):8108-13.

Zarrin AA, Del Vecchio C, Tseng E, Gleason M, Zarin P, Tian M, Alt FW (2007): Antibody class switching mediated by yeast endonuclease-generated DNA breaks. Science 315:377–81.

Zhang F, Maeder ML, Unger-Wallace E, Hoshaw JP, Reyon D, Christian M, Li X, Pierick CJ Dobbs D, Peterson T, Joung JK, Voytas DF (2010): High frequency targeted mutagenesis in Arabidopsis thaliana using zinc finger nucleases. Proc. Natl. Acad.

Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P (2011): Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol 29:149–153.

Zhang Y, Heidrich N, Ampattu BJ, Gunderson CW, Seifert HS, Schoen C, Vogel J, Sontheimer EJ (2013): Processing-independent CRISPR RNAs limit natural transformation in Neisseria meningitidis. Mol. Cell 50, 488–503.

Zhang F, Wen Y, Guo X (2014): CRISPR/Cas9 for genome editing: progress, implications and challenges. Hum Mol Genet. 23(R1):R40-6.

Zu Y, Tong X, Wang Z, Liu D, Pan R, Li Z, Hu Y, Luo Z, Huang P, Wu Q, Zhu Z, Zhang B, Lin S (2013): TALEN mediated precise genome modification by homologous recombination in zebrafish. Nat. Methods 10, 329–331.

Appendices

A: Oligonucleotides and Primers

	Sequence	Enzyme(s)	Purpose(s)
1	AAAAGCGGCCGCTCTCAGCATGGAAGGACAGCA	NotI	to clone 3' homology arm in targeting vector
2	AAAGGATCCAGCTGGCAGGGGATAAAGGAAAGA	BamHI	to clone 3' homology arm in targeting vector
3	AAACTCGAGTCCTTTGCCGACATCTTCCT	XhoI	to clone 5' homology arm in targeting vector
4	AAAGGATCCTCATTTCACCTGCAGGTGACA	BamHI	to clone 5' homology arm in targeting vector
5	CTAGTTGGGATGGAGCTGGATCTTAGATCTTCCTGTC AGGAACTGCAGG		to clone 7 bp spacer
6	CTAGCCTGCAGTTCCTGACAGGAAGATCTAAGATCCA GCTCCATCCCAA		to clone 7 bp spacer
7	CTAGTTGGGATGGAGCTGGATCTTTAGATCTTTTCCT GTCAGGAACTGCAGG		to clone 10 bp spacer
8	CTAGCCTGCAGTTCCTGACAGGAAAAGATCTAAAGAT CCAGCTCCATCCCAA		to clone 10 bp spacer
9	CTAGTTGGGATGGAGCTGGATCTTCTTAGATCTTTCT CCTGTCAGGAACTGCAGG		to clone 13 bp spacer
10	CTAGCCTGCAGTTCCTGACAGGAGAAAGATCTAAGAA GATCCAGCTCCATCCCAA		to clone 13 bp spacer
11	CTAGTTGGGATGGAGCTGGATCTTTCTTAGATCTTTCA ATCCTGTCAGGAACTGCAGG		to clone 16 bp spacer
12	CTAGCCTGCAGTTCCTGACAGGATTGAAAGATCTAAG AAAGATCCAGCTCCATCCCAA		to clone 16 bp spacer
13	AAACCGCGGCCTCTGCTAACCATGTTCATGCCT	SacII	to clone cassette for last step Golden Gate
14	AAAGGATCCAGCTGGGATCTGATCAATTCCG	BamHI	to clone cassette for last step Golden Gate
15	AAAAGATCTCATGGACTACAAAGACGATGACG	BglII	to clone Jun-YN

	Sequence	Enzyme(s)	Purpose(s)
16	AAAAGATCTTTCTAGAGGATCCAGATCCATCGC	BglII	to clone YN
17	AAATGTACACCCTAGGCCATGATATAGACGTTG	BsrGI	to clone JunYN and YN
18	AAAACCGGTCACCATGTACCCATACGATGTTCC	AgeI	to clone Fos-YC
19	AAAACCGGTTACGCTCTTATGGCCATGGA	AgeI	to clone YC
20	AAACTTAAGCTAGCATGCCTGCAGATCGACT	AflII	to clone Fos-YC and YC
21	AAATGTACAAGGCTGCCCAAATTTGGAGA	BsrGI	to clone MODC
22	AAATGTACACATTGATCCTAGCAGAAGCAC	BsrGI	to clone MODC
23	GACCACCTTCGGCGCCGGCCTGCAGTGC		internal primer for cross-over PCR Y66H
24	GCACTGCAGGCCGGCGCCGAAGGTGGTC		internal primer for cross-over PCR Y66H
25	CGTGGATAGCGGTTTGACTC		external primer for cross-over PCR Y66H
26	AAACATATGTAATACGACTCACTATAGGG	NdeI	external primer for cross-over PCR Y66H
27	AAAGCTAGCCCTCTGCTAACCATGTTCATGCCT	NheI	to clone cassette for last step Golden Gate
28	AAATGTACAGCTGGGATCTGATCAATTCCG	BsrGI	to clone cassette for last step Golden Gate
29	GAAAAACACGATAATACCGGATCCAACGGTGAGCAAG GGCGAG		internal primer for cross-over PCR, to mutate ATG of EGFP
30	CTCGCCCTTGCTCACCGTTGGATCCGGTATTATCGTG TTTTTC		internal primer for cross-over PCR , to mutate ATG of EGFP
31	AAATGTACACGAGACGCATTTCGTACTTTGGGA	BsrGI	external primer for cross-over PCR, to mutate ATG of EGFP
32	TTTGTCGACGGCTAGTCCAGATCCAGACA	SalI	external primer for cross-over PCR, to mutate ATG of EGFP
33	GATCTACGACGACGACGACGACGACGGAT		to insert ACG array
34	GATCATCCGTCGTCGTCGTCGTCGTCGTA		to insert ACG array
35	GATCTACGGCAACGGCAACGGCAACGGAT		to insert WRC ACG array

36	GATCATCCGTTGCCGTTGCCGTTGCCGTA		to insert WRC ACG array
37	GATCTGCAACGGCAACGGCAACGGCAGAT		to insert WRC ACG array
38	GATCATCTGCCGTTGCCGTTGCCGTTGCA		to insert WRC ACG array
39	AAAGAATTCTGCTGGTTATTGTGCTGTCTC	EcoRI	to clone bsr gene
40	AAAGAATTCCCACAACTAGAATGCAGTGA	EcoRI	to clone bsr gene
41	GTCGACCTCTGCTACGTGGTGAAGAGG		to clone CRISPR AID target sequence in mCherry-out-bsr plasmid
42	CAGACCTCTTCACCACGTAGCAGAGGT		to clone CRISPR AID target sequence mCherry-out-bsr plasmid
43	CCAAACCTACAAGGTCATAAGC		to amplify MS2 southern probe
44	TGAAGGTTAGGATGTCTGTGGA		to amplify MS2 southern probe
45	AATGCCTAGCCCTCCCAGATTA		to screen MS2 targeted clones, external PCR
46	CCCACCGACTCTAGAGGATCATAA		to screen MS2 targeted clones, external PCR
47	ATGCCACCCAGACATGGTCATT		to screen MS2 targeted clones, internal PCR
48	CTGCATTCTAGTTGTGGTTTGTCC		to screen MS2 targeted clones, internal PCR
49	CACCGACCATTTCAAAAATGTCCGC		to insert in pX330
50	AAACGCGGACATTTTTGAAATGGTC		to insert in pX330
51	AGGGTGGGCAGGGAAGGATTTTAAAG		to sequence the second exon of AID gene
52	GAAGGTGGCCGAAGTCCAGTGA		to sequence the second exon of AID gene
53	GTATCAGAACCTTCCCAACAC		to check expression of MS2 cassette by RT-PCR
54	CTGCATTCTAGTTGTGGTTTGTCC		to check expression of MS2 cassette by RT-PCR

B: Plasmids and Vectors

Plasmid	Feature(s) and usage	Marker(s)	Source
pCR-BluntII-TOPO	Routine cloning of blunt PCR products.	Kan	Life Technologies
pCR 2.1 TOPO-TA	Routine cloning of Taq PCR products.	Amp/Kan	Life Technologies
pAID-Express Puro2	Eukaryotic expression vector for AID	Amp/Puro	(Arakawa et al., 2004)
pLoxPuro	Eukaryotic expression vector	Amp/Puro	(Arakawa et al., 2001)
pLoxBsr	Eukaryotic expression vector	Amp/Bsr	(Arakawa et al., 2001)
pEGFP-N1	Eukaryotic expression vector for EGFP	Kan/G418	(Cormack et al., 1996)
pROSA26	ROSA26 locus targeting vector	Amp/G418	(Soriano, 1999)
CMV-YFP-MS2	Eukaryotic expression vector for chimera YFP-MS2 coat protein	Kan/G418	(Janicki et al., 2004)
p3216PECSKLMS2β	Eukaryotic expression vector	Amp/HygroB	(Janicki et al., 2004)
pBiFC-bJunYN	Eukaryotic expression vector for chimera Jun-YN	Amp	(Hu et al., 2002)
pBiFC-bFosYC	Eukaryotic expression vector for chimera Fos-YC	Amp	(Hu et al., 2002)
pCAGGS-TAL-NC2	Eukaryotic expression vector for last step Golden Gate TALEN/TAL Effector Kit 2.0	Amp	(Sakuma et al., 2013)Addgene: #43856
pX330-U6-Chimeric_BB-CBh-hSpCas9	Eukaryotic expression vector	Amp	(Cong et al., 2013)Addgene: #42230
pEGFP-AID	Eukaryotic expression vector for chimera EGFP-AID	Kan/G418	
pEGFP-A1	Eukaryotic expression vector for chimera EGFP-A1	Kan/G418	
pAID^{E58A}-Express Puro2	Eukaryotic expression vector for AID^{E58A}	Amp/Puro	
pEGFP-mCherry-EGFP	Eukaryotic expression vector for chimera mCherry-EGFP	Kan/G418	
pBS-CMV-mCherry-EGFP	Eukaryotic expression vector for chimera mCherry-EGFP	Amp	

C: TALENs and CRISPRs tools and commercial services

Online tools for TALENs

Tool	Website
TAL Effector Nucleotide Targeter 2.0	https://tale-nt.cac.cornell.edu/
E-TALEN	http://www.e-talen.org/E-TALEN/
TALEN designer	http://www.talen-design.de/
TALEN™ Hit	http://talen-hit.cellectis-bioresearch.com/search
Mojo Hand	http://www.talendesign.org/
TALE Toolbox	http://taleffectors.com/tools/
TAL Plasmids Sequence Assembly Tool	http://baolab.bme.gatech.edu/Research/BioinformaticTools/assembleTALSequences.html
Emily Talen	http://www.planetizen.com/topthinkers/talen
ZiFiT	http://zifit.partners.org/ZiFiT/
PROGNOS	http://baolab.bme.gatech.edu/cgi-bin/prognos/prognos.cgi
CHOPCHOP	https://chopchop.rc.fas.harvard.edu/
idTALE	http://idtale.kaust.edu.sa/index.html
TALENoffer/TALENgetter/TALENgetterLong	http://galaxy2.informatik.uni-halle.de:8976/
SAPTA	http://baolab.bme.gatech.edu/Research/BioinformaticTools/TAL_targeter.html
LIC TALE gene Assembler Version 1.0	http://www.hornunglab.de/TALEN.html

Online tools for CRISPRs

Tool	Website
CRISPR Genome Engineering	http://www.med.umn.edu/starrlab/
ZiFiT	http://zifit.partners.org/ZiFiT/
CRISPR DESIGN TOOL	http://www.broadinstitute.org/mpg/crispr_design/
Jack Lin's CRISPR/Cas9 gRNA finder	http://spot.colorado.edu/~slin/cas9.html
E-CRISP	http://www.e-crisp.org/E-CRISP/
Zhang's Lab CRISPR Design	http://crispr.mit.edu/
Cas9 Design Platform	http://cas9.cbi.pku.edu.cn/
CRISPR Target	http://bioanalysis.otago.ac.nz/CRISPRTarget/crispr_analysis.html
sgRNAcas9	http://www.biootools.com/
fly CRISPR	http://flycrispr.molbio.wisc.edu/
DRSC tool	http://www.flyrnai.org/crispr/, Drosophila
CasOT	http://eendb.zfgenetics.org/casot/index.php
Cas-OFFinder	http://sourceforge.net/projects/cas-offinder/
COD (Cas9 & Off-target Designer)	http://cas9.wicp.net/
CHOPCHOP	https://chopchop.rc.fas.harvard.edu/
CRISPRdirect	http://crispr.dbcls.jp/
CRISPR gRNA Design tool	https://www.dna20.com/eCommerce/cas9/input
CRISPOR	http://tefor.net/crispor/crispor.cgi
SSFinder	https://code.google.com/p/ssfinder/
Cas9 Target Finder	http://www.shigen.nig.ac.jp/fly/nigfly/cas9/cas9TargetFinder.jsp
CRISPRi	http://qi.ucsf.edu/CRISPR_transcription
CRISPRscreen	http://slave03.molbiol.ox.ac.uk/CRISPR/cgi-bin/CRISPR.cgi
CRISPR-PLANT	http://www.genome.arizona.edu/crispr/

Commercial services for TALENs

Website	Location
http://www.bioon.com.cn/server/Show_product.asp?id=8713	China
http://www.v-solid.com/service/TALE-TALEN-TALEA.aspx	China
http://www.biomart.cn/infosupply/9979783.htm	China
http://www.sangon.com/sangon_detail.aspx?newsID=659	China
http://www.sidansai.com/cn/	China
http://minimouse.sciencenet.cn/?uid-585947-action-viewcompany-itemid-20871	China
http://www.cellectis-bioresearch.com/talen-solutions	China
http://www.bioon.com.cn/show/index.asp?id=197083	China
http://www.genechem.com.cn/Pro_show.aspx?plb=791	China
http://www.ennovationlifesciences.com/product.aspx?SId=10	India
http://zgenebio-ko.weebly.com/	Taiwan
http://www.tebu-bio.com/	UK
http://transposagenbio.com/gene-modification-tools/xtn-talens/	USA
http://www.genecopoeia.com/product/talen-tal-effector/	USA
http://www.geneticservices.com/injection/talen-and-crispr-injections/	USA
http://pnabio.com/products/TALEN.htm	USA
http://www.cores.utah.edu/?page_id=5987	USA
http://www.umassmed.edu/Content.aspx?id=174126	USA
http://www.unmc.edu/genetics/custom_talens.htm	USA
http://www.systembio.com/	USA
https://hopecenter.wustl.edu/	USA

Commercial services for CRISPRs

Website	Location
http://www.biomart.cn/infosupply/10980185.html	China
http://qy.bio1000.com/njfish/promotion/itemid-37.shtml	China
http://www.nbri-nju.com/service-view-CRISPR	China
http://www.bioon.com.cn/server/Show_product.asp?id=10525	China
http://www.biomart.cn/infosupply/14975959.htm	China
http://www.biomart.cn/infosupply/14593490.htm	China
https://hopecenter.wustl.edu/	China
http://www.ennovationlifesciences.com/product.aspx?SId=10	India
http://zgenebio-ko.weebly.com/crispr.html	Taiwan
http://transposagenbio.com/crisprcas/	USA
http://www.genecopoeia.com/product/crispr-cas9/	USA
http://www.geneticservices.com/injection/talen-and-crispr-injections/	USA
http://pnabio.com/products/RGEN.htm	USA
http://www.cores.utah.edu/?page_id=5987	USA
http://www.umassmed.edu/Content.aspx?id=174126	USA
http://www.sigmaaldrich.com/	USA
http://www.systembio.com/	USA
http://www.blueheronbio.com/Services/Genome-Editing.aspx	USA