SMGroup

DNA Technology, Genes and Genetic Engineering

Pamela Mancha Agresti[1], Fillipe Luiz Rosa do Carmo[1], Rodrigo Dias de Oliveira Carvalho[1], Marcela Santiago Pacheco de Azevedo[1], Artur Silva[2], Vasco Azevedo[1]

[1]Universidade Federal de Minas Gerais, UFMG, Brazil

[2]Universidade Federal doPará, Brazil

***Corresponding author:** Vasco Azevedo, Universidade Federal de Minas Gerais/Instituto de CiênciasBiológicas, Minas Gerais, Brazil, Email: vasco@icb.ufmg.br

Published Date: February 15, 2015

ABSTRACT

The detailing physical and chemical bases of the DNA molecule in 1953 by Watson and Crick, followed by the *complete* description of *its structure and mechanisms of replication and gene expression,* allowed the development of recombinant DNA technology. The development of techniques used to transfer foreign DNA into new host cells proved successful and nowadays is commonly implemented in different areas of knowledge, such as microbiology, biochemistry and immunology. Moreover, recombinant DNA technology allowed the identification and sequencing of important genes, and provided a set of techniques developed with the purpose to manipulate producing hybrid organisms capable to express molecules useful for human beings health" with to manipulate organism and produce hybrids capable of expressing useful molecules for human health. One example is the genome editing and molecular cloning technique that are widely used in basic and applied research. Genetic engineering is a valuable tool for (i) the study of genomes, (ii) the construction of genetically modified organisms that can produce proteins of biotechnological interests or (iii) the development of transgenic organisms with new properties.

Keywords: Molecular cloning; restriction enzyme; genome editing; homologous

recombinations; heterologous protein

HISTORICAL PERSPECTIVE

In the 40s, Tatum and Beadle demonstrated that genes are responsible for coding proteins [1]. In the same decade Avery, McLeod and McCarty proved that the DNA is the material that causes bacterial transformation [2]. In the 50s, Hershey-Chase experiment helped to confirm that DNA is the genetic material [3]. Finally, Watson and Crick in 1953 established a three-dimensional structure model for the DNA [4], which yielded new insights into the genetic code and how proteins are synthesized. Throughout the 60s, Nirenberg and collaborators through various experiments depicted the genetic code [5] and Jacob and Monod studies elucidated the process of protein synthesis in bacterial cells giving rise to modern molecular biology [6].

The identification of both small circular DNA molecules, called plasmids, carrying antibiotic resistance genes and restriction endonucleases/DNA ligase enzymes founded in bacterial cells were the "ingredients" that allowed the process of cloning, as demonstrated in the work of Janet Mertz and Ronald Davis in 1972 [7]. Cutting a piece of DNA carrying a gene from one organism with the restriction endonucleases enzyme and inserting it into a plasmid that can be replicated by a host organism (i.e. bacteria) using the DNA ligase, it was possible to create a recombinant second organism expressing the protein of interest, recovered because of its capacity to grown in a medium containing an specific antibiotic marker. In the same year, Berg and collaborators performed a similar experiment where genes of *E. coli* and bacteriophages were inserted in the genome of the SV40 virus, which then were able to produce recombinant molecules [8]. The following year, in 1973, Boyer and S. Cohen cloned a DNA fragment of *Xenopuslaevis*in *Escherichia coli* using the pSC101 plasmid. Afterwards, the bacterium was able to synthesize some molecules normally produced by *Xenopuslaevis*. Molecular cloning allows the isolation of individual fragments, including the determination of nucleotide sequence. The development of sequencing techniques started with Sanger, in 1977, and made possible to sequence completely the genome of an organism [9]. Another important technique developed by Mullis in 1983, which was essential for the molecular biology development, is the polymerization chain reaction (PCR) [10]. This technique, turned possible to identify and amplify specific regions of the DNA from various sources, providing information about the organization and the expression of genes with medical and biotechnological application.

MOLECULAR CLONING

In order to start the cloning procedure, DNA harboring the gene of interest must first be digested with specific restriction endonucleases enzymes that cut the DNA in a region flanking the gene. The cut DNA fragment binds to the cloning vector derived from a plasmid by using the DNA ligase enzyme. Both the cloning vector and the segment of DNA to be cloned must have been digested using the same restriction enzymes in order to create cohesive ends that can bind to each other. After the enzyme action, recombinant DNA molecule generated has to be inserted

into a host cell (i.e. bacteria). There are many methodologies described to ensure the entrance of the recombinant DNA into the cell, such physical methods like: electroporation, heat shock transformation, microinjection, microparticle intake and liposome fusion with the plasma membrane. As the introduced recombinant DNA molecule is not within the genome of the host cell, it replicates independently from the genomic DNA because cloning vectors has its own replication origin. With the intent to recover bacteria harboring the recombinant vector, antibiotics are added to the culture medium because almost all conventional plasmids use an antibiotic resistance gene as a selection markers. These transformed cells containing the recombinant DNA are multiplied in a suitable culture medium which allows the obtainment of large amounts of the cloned protein.

RESTRICTION ENZYMES

Restriction enzymes are produced by bacteria, viruses and act as a defense mechanism in digesting pieces of viral DNA. The first restriction enzyme (Hind III) was discovered by Hamilton in 1970 when working with *Haemophilus influence*. He noticed that the phage P22, when put in contact with this bacterium, had its DNA rapidly degraded. It was noticed that this degradation was linked to the action of an enzyme, Hind III, which was then isolated and studied [11]. Restriction enzymes move along the DNA strands, and when they reach specific target sequences from 4 to 8 base pairs, the DNA is distorted and the strands are cut off. These enzymes can be divided into several classes depending on the structure, cofactor requirements, activity, recognition and cleavage sites.

There are three types of restriction endonucleases, type I, II and III. Enzymes Type II, are of major importance in Recombinant DNA Technology as they presents a relative simple protein structure, and does not require ATP to cleave DNA [12], the recognition site for this type of enzyme is usually a palindromic sequence. Restriction enzymes can perform two types of cleavage, one that generate abrupt ends, and the other that generate cohesive ends.

DNA LIGASE

The DNA ligase plays an important role in DNA replication and repair, as it acts connecting the short pieces of lagging-strand DNA (Okazaki fragments) generated by replication of the late strand [12]. During cloning procedure, this enzyme promotes the binding of DNA inserts into the vector previously cleaved by the same restriction endonuclease that were used to isolate the DNA insert. This process was demonstrated for the first time *in vitro* in 1972 by Janet Mertz and Ronald Davis [7]. There are different DNA ligase enzymes, nevertheless, T4 DNA ligase, produced by bacteriophage T4, is one of the most frequently used as it connects both cohesive and blunt ends [12].

DNA POLYMERASE

DNA polymerase III has a very large and complex protein structure, consisting of several protein subunits. This enzyme is also very important for molecular biology as it acts adding

deoxyribonucleoside to the 3' end of a growing DNA chain. They are Commonly used to amplify DNA fragments *in vitro*. The first one to be isolated was a thermostable DNA polymerase (Taq DNA polymerase) originated from the thermophilic bacterium *Thermus aquaticus*, is largely used for PCR [13].

Several thermostable DNA polymerases were subsequently isolated, and many of these are commercially available for carrying out PCR technique. Some examples are the Tth isolated from *Thermus thermophilus*; BstI from *Bacillus stearothermophilus*; Psp "Deep Vent" isolated from *Pyrococcus sp*.; Pfu isolated from *Pyrococcus furiosus*, Pwo isolated *Pyrococcus woesei*, Tli "Vent" isolated from *Thermococcus litoralis*. These enzymes have different characteristics from each other, such as extension rate, processivity and fidelity, which are chosen according to the final purpose [14].

CLONING VECTORS

In order to amplify the genetic information of the gene of interest, the fragment must be inserted in a different DNA molecule; these DNA molecules are called molecular cloning vectors. These vectors basically have some essential components such as (i) the multiple cloning site (MCS) for insertion of the sequence of interest, (ii) a replication origin that allows replication of the cloning vector into the host cell and (iii) a gene coding that confers resistance to a certain antibiotic, enabling the identification and selection of cells harboring the recombinant vector [13].

Cloning vectors are bacterial plasmids, genetic elements consisting of circular DNA molecules within a cell that is physically separated from a chromosomal DNA found in bacteria. They have the ability to replicate because of its own origin of replication and, in general, they contain genes that confer resistance to antibiotics, such as ampicillin and tetracycline. These characteristics, in addition to its small size (a few thousand of base pairs) turned them very suitable to be used as cloning vectors. pBR322 was one of the first recombinant plasmid to be constructed and it is formed with a number of other plasmids (pBR318, pBR320 pBR313) [15]. Very well known cloning vectors, such as pUC18 and pUC19, were developed from pBR322. pUC19, for instance, is a commonly used plasmid cloning vector in *E.coli*. It is a small molecule and has a high copy number, a multiple cloning site polylinker, and a gene that codes for ß-galactosidase (lac Z gene). Insertion of foreign DNA within the lac Z gene causes insertional inactivation and abolishes intra-allelic complementation with another protein, permease, encoded by host chromosome [12,16,17]. Therefore, the recombinant bacteria able to hydrolyze X-gal, the substrate for the active enzyme, giving rise to white colonies, which can be differentiated on culture media from non-recombinant cells, which are blue. Bacteriophage vectors are also used as cloning vectors, keeping the insert inside infectious phage particles. It is an effective vector for cloning DNA fragments no larger than 23kb [17]. For fragments or genes that are larger in size, between 35 to 40 kb, the cosmids are an excellent choice. They are hybrid vectors containing a DNA fragment from the lambda phage, including the cos site, and bacterial plasmid DNA. They are able to pack DNA and infect host cells

releasing large DNA molecules into these recipient cells. However, for very high inserts still the most popular is the bacterial artificial chromosome (BAC), that can pack inserts from 100 to 500 kb, and yeast artificial chromosomes (YACs), that can harbor inserts ranging from 250 to 1000 kb [12].

These cloning vectors that have the ability to carry large inserts were used for constructing genomic libraries, collection of the total genomic DNA from a single organism. Another type of genetic library that was much explored using these vectors was the one where the reverse transcriptase enzyme produced cDNA using mRNA as a template. The creation of cDNA library was very important in a large number of studies, such as the search for potential targets for antimicrobial and antitumor drugs [12,16,17].

RECOMBINANT DNA TECHNIQUES

In order to amplify the cloned insert using the PCR method is necessary to use two oligonucleotides, also named primers, commercially synthesized. They can bind in a sequence-specific manner to their respective complementary oligonucleotides in the DNA strand (forward 5'-3' and reverse 3'-5' strand) that is flanking the region of interest amplifying thus the fragment that was cloned. Therefore, these oligonucleotides Serve as primers for DNA synthesis *in vitro*. The automatized PCR reaction begins with the heating of the machine, which promotes the separation of double-stranded DNA. Then, it is cooled to allow the hybridization of the oligonucleotides to their respective complementary DNA sequences. The melting temperatures are in the range of 52 to 58°C. After this step, machine is warmed up again to an optimum temperature in which the enzyme DNA polymerase can extend the strands by adding deoxyribonucleotides (dATP, dCTP, dGTP and dTTP) once the primers have been annealed. This cycle is repeated around thirty times. At the end it is generated a large number of recombinant DNA molecules. The identification of the amplified sequences occurs through the gel electrophoresis method, which consists on the separation and analysis of DNA (also RNA and proteins) and their fragments, based on their size and charge. The concept of the technique is that ionized molecules placed in an electric field migrate according to their charges and molecular weights. If the charge is positive, they will migrate to the negative pole, conversely, if it is negative, migration will occur towards the positive pole. The amplified fragments migrate according to their charge, size and shape, forming a characteristic band in the agarose gel. This method allows to qualitatively identify the sequence of interest based on its size calculated in the amount of nitrogenous base pairs.

In order to determine whether the sequence was correctly amplified, the sequencing method described by Sanger and colleagues at the end of the 70s, has been widely used. This method of DNA sequencing is based on the selective incorporation of chain-terminating dideoxynucleotides stained with different fluorescent dyes by DNA polymerase during *in vitro* DNA replication [12,16,17]. Nowadays there are other techniques, besides cloning and sequencing, to determine gene sequence as well their biological function. For example, genome editing approaches has been

described and recent data shows that they represent a very good genomic techniques advances.

GENOME EDITING WITH SITE-SPECIFIC NUCLEASES

The implementation of large genome annotation projects, and new and accessible approaches for whole-genome sequencing, have been permitted to scientists to be self-confident, to modify basic science and personalized medicine. Homologous recombination (HR) has been the traditional and powerful approach to inactivate target genes, replacement, or addition, and also offers conclusive information for evaluating gene function [18]. Nevertheless, the low efficiency of this system in mammalian cells and model organisms, time-consuming, the possibility for hostile mutagenic effects, radically restricts the utility of this methodology [19]. An alternative to HR is RNAi process [20]. RNAi is a rapid and cheaper mechanism, that happens in response to the incorporation of RNA molecules into cells, RNA molecules inhibit or knockdown gene expression. Even so, RNAi has restrictions like inefficiency, incomplete knockdown gene, random off-target effects, time-consuming, labor intensive and supplies only temporary inhibition of gene function; and until now, traditional gene-targeting technology has only been capable to be applied in rare model systems, such as Drosophila [21] and mouse [22], categorized by short generation times and easy inbreeding. These limitations hamper the practical application of RNAi technology [19].

The chance of modifying DNA sequences within the cell in a controlled way permit the understanding of gene function, consequently, virtual handling of any gene in a varied range of cell types and organisms, usually is denominated as 'genome editing'. This approach was initially applied to *Drosophila melanogaster* [23,24] is founded on the use of engineered nucleases constituted of sequence-specific DNA-binding domains attached to a non specific DNA cleavage module [25,26]. This methods permits an efficient and precise genetic modification through the induction of a double-strand break (DSB) in a specific genomic target sequence, followed by the generation of desired modifications during subsequent DNA break repair [27].

In the last years, it has been obeserved observed an innovation in gene targeting technology. ZFN (Zinc Finger Nuclease) [24,28], TALEN (Transcription Activator-Like Effector Nuclease) [29] and CRISPR/Cas9 nuclease (Clustered Regularly Interspaced Short Palindromic Repeats) [30,31] systems, currently make it possible for scientists to easily, efficiently and inexpensively modify the genome. It is known that any gene can be mutated by these methods, in an efficiently way, sometimes overhead a 50% of frequency [32,33].

The advent and speedy development of those techniques has elevated and pronounced interest in their applications in either model or non-model organisms [34]. An underlying mechanism similar to this techniques-by inducing targeted DNA to generate DSB, followed-by DSBs being corrected by error-prone non-homologous end joining (NHEJ) [35] and HR [36]. In eukaryotic cells NHEJ and HR are the two keys of DNA repair mechanisms.

Each of these systems ZFN, TALEN and CRISPR/Cas9, has different way to recognize and operated on the target DNA [24,29,37,38]. Both ZFN and TALEs have been fused to enzymatic

domains. The ZFNs were the first engineered endonucleases developed for targeted genome modification. ZFs (Zinc Finger) composed of three Cys2His2 zinc fingers engineered to bind specific DNA sequences. As one ZF links to three bases, consequently joined ZF array can recognize 9–18 bp, per cleavage site [25]. ZFs are linked to a FokI nuclease domain, a non specific DNA cleavage domain [39,40]. A dimerization of FokI induces local endonuclease activity. This dimerization requires high specificity for DSB generation only at the target locus. Therefore, ZFNs are used in pairs with specificity to opposing DNA strands that assemble on both sides of the targeted cleavage site. The capability to regular the recognition specificity of zinc fingers, opens the prospect of directing cleavage to arbitrarily chosen chromosomal sites, without prior manipulation of the target [24].

The other alternative approach for introducing chromosomal breaks at selected sites of DNA is TALEN, was first described in 2007 [41]. Similar a ZFNs conception, consist of assembled DNA-binding motifs coupled to FokI nuclease [42-45]. TALEs proteins are produced by plant pathogens in the genus Xanthomonas, and consist of a central domain commissioner for DNA binding, a nuclear localization signal, and a domain that activates the target gene transcription [46]. Domains consist in tandem repeats monomers of 33–35-amino-acid residues; each of them binds one nucleotide in the target nucleotide sequence. The domains responsible for the recognition of a specific nucleotide are located at the 12 and 13 positions, and they are highly variable (repeat variable di-residue, RVD). Repeats with different RVDs recognize different DNA base pairs, and there is a one-to-one correspondence between the RVDs in the repeat domain and the nucleotides in the target DNA sequence [42,47]. As with ZFNs, TALEN-mediated double-strand breaks, also stimulated HR in human cells at levels that are similar to the levels achieved with ZFNs [48,49]. The skill to predict the DNA binding specificity of native or artificial TALEs suggests a variety of applications for these proteins in the targeted modification of genomes. These site-specific nucleases have modified a diversity of organism's genome, and they are in continuous growing, developing and expanding the repertoire of model systems for basic research [19].

The more recently described technique of genome editing is the CRISPR /Cas system, who has emerged as an alternative to ZFNs and TALENs for inducing targeted genetic alterations. The loci of CRISPR was discovery in 1987 in *E.coli* by Ishino and co-workers, but their function wasn't definite until 2007 by Barrangou and colleagues, who verified that *S. thermophilus* can acquire resistance against a bacteriophage by integrating a genome fragment of an infectious virus into its CRISPR locus [50]. This system has been described as an adaptive defense mechanism against various mobile genetic elements, in a wide range of prokaryotes, including the majority of Archaea and many Bacteria. CRISPR loci consist in a series of 24–47 bp repeated sequences separated by unique sequences of a similar length (known as spacers) that match the genomes of bacteriophages and other mobile genetic elements [51,52]. Upstream of the CRISPR locus is a leader region containing 20 to 534 bp, A-T reach region, the promoter of this loci, is in the sequence leader region [53,54]. Near the repeats array was describe a conserved family of

few CRISPR-associated (cas) genes. This family consisted of genes encoding proteins like DNA helicases (cas3) and exonucleases (cas4), which implied that Cas genes likely have function in DNA metabolism [54].

Three mains stages are involves in CRISPR mechanism. The first stage is *Adaptation*, phase with acquisition of memory, where a small fragment of foreign DNA that entered a bacterial cell is inserted into the CRISPR locus of the host genome, forming a new spacer. In prokaryotes genome, this fragment is present as a protospacer that is complementary to the spacer and flanked by a short (2–5 bp), conserved sequence called PAM (protospacer adjacent motif) [55,56]. Always the new spacer is inserted on the side of the leader sequence located before a CRISPR cassette. After that, take place the transcription, or biogenesis phase, in which the entire CRISPR locus is transcribed into a long pre-crRNA (poly-spacer precursor crRNA), which will be processing into mature crRNA, generate the guide RNA components (gRNA); the third phase is the interference of the invading cognate nucleic acids by ribonucleoprotein complexes consisting of Cas proteins and the guide RNAs. Cas9 is a dsDNA nuclease, enzyme specialized for cutting DNA, with two active cutting sites, one site for each strand of the DNA's double helix, that uses a crRNA guide to specify the site of cleavage [57,58].

ZFN, TALEN and CRISPR/Cas9, are the three new genome-editing techniques, which accomplish efficient and precise genome modification through similar mechanisms, like as by inducing targeted DNA to produce double strand breaks (DSB), followed-by DSBs being repaired by NHEJ [35,36].

ZFN and TALEN systems stimulate DSBs by a non-specific FokI nuclease domain fused to their binding domains; been ZFN a method that has limited target sites because of its 3-nucleotides recognizing model, and the system is also more expensive and difficult to assemble. For the other hands, TALEN and CRISPR/Cas9 techniques are pondered to be the ideal gene-targeting technologies, since they are easier to assemble, are more efficient, and have more abundant target-specific recognition sites and activations compared with a similar range of cell types and organisms [34]. It is important to note that in the last few years the CRISPR/Cas9 system has been successfully used to target genomic loci in the mammalian cell lines [30,33,59] and several species, including mice and rat [45,60].

Once the sequence of a target gene is studied through all these different methods mentioned above, the scientists might want to obtain the protein of interest for using it for different applications. One manner by which it is possible to produce a "biological factory" capable to produce in great quantities the protein of interest is to clone it into different biological expression systems.

EXPRESSION OF HETEROLOGOUS PROTEINS

The next step to produce the protein of interest is to insert the cloned DNA fragment into

expression vectors. They are very similar to the cloning vectors, however, they present transcription and translation signals necessary to regulate the expression of the cloned gene. These vectors are available for both prokaryotic and eukaryotic host cells. The gene to be expressed is inserted into the polylinker near the promoter of the gene, which can be inducible or constitutive. It also contains transcription terminator sequences to improve the stability and to establish the amount of the produced mRNA. Another component that is present is the ribosome binding site sequence that provides the signals necessary for the efficient translation of the mRNA. Prokaryotes and eukaryotes use different strategies to specify the site where the mRNA begins to be translated in protein. mRNA from procaryotes usually has a conserved sequence of six nucleotides, known as the Shine Dalgarno or linker site of the ribosome (RBS), which is always found few bases (5') above to the start codon. mRNAs from eukaryotic organisms has a sequence known as Kozak consensus sequence that also plays a major role in the initiation of the translation process[12,16,17].

The most frequently used plasmids for the expression of recombinant proteins contains the Lac promoter (inducible) or the T7 promoter (constitutive). The lac promoter is repressed by lacI repressor protein in the absence of an inducer (lactose or isopropyl-β-d- thiogalactopyranoside - IPTG) and induced in its presence in a growing medium lacking glucose. The T7 promoter is originated from the Bacteriophage T7. In molecular biology they are very expored due to its extremely high affinity for T7 RNA polymerase leading to a high level of protein expression. One of the most common expression vectors that uses the T7 promoter are the popular pET expression system (Novagem, 2005).

Vectors designed to express the protein of interest in mammalian cells needs to contain sequences to allow the production of the protein in eukaryotic cells. However, the vector needs to propagate in bacteria because of the cloning steps. Therefore they have origin of replication for both gram positive and negative bacteria as well as a selection marker for selection of clones harboring the vector [17]. They have to present promoters recognized by the eukaryotic cell transcription machinery to allow the expression of the protein. The most used one is the cytomegalovirus promoter (pCMV). Another component is the polyadenylation signal, originated from the Bovine Growth Hormone (BGH) gene to stabilize the mRNA recently sythetized.

The cell that can host the cloning or expression vector can be weither a virus or bacteria. The bacterium that was first and most widely used for this purpose is *Escherichia coli* because it has well-described genetic characteristics, can grow easier in culture medium that are not so expensive and are able to express very high amounts of proteins. Despite its widespread use, expression of recombinant proteins in *E. coli* system has some disadvantages when compared to other systems such as viral or gram positive bacteria. It has lipopolyssacharides (LPS) on its outer membrane which is considered to be an endotoxin that can contaminate the final product. Moreover, recombinant protein can be included inside inclusion bodies hampering its purification [61,62]. The system also does not allow post-translational modifications in the protein of interest,

which ends up limiting its use. Another disadvantage is that some combinations of codons are found at low frequency in *E. coli* preventing protein translation. To overcome this problem many codon usage indices has been studied, such as ENC (effective number of codons), P (codon preference plot), frequency of optimal codon, codon bias index and codon adaptation index which can establish which codon should be used for E. coli or other bacterial strain, therefore, allowing some authors to develop several softwares to analyse such indexes, which justifies the requirement of indices calculation for the research in recombinant protein expression in order to obtain high levels of the targeted proteins [63,64]. Actually, there are many strategies that have been used and being developed in order to improve both the expression of heterologous proteins in prokaryotic host cells (i.e. efficient expression vectors, construction of genetically modified bacterial strains lacking extracellular proteases).

Expression vectors that works only in eukaryotic cells are necessary only when the protein needs to be expressed on its native form as these mammalian cells provide post-translational modifications, such as acetylation, methylation, phosphorylation, especially N-linked glycosylation in the recombinant protein[65]. The first recombinant product produced on a large scale in mammalian cells was the tissue plasminogen activator (tPA), the product was licensed in 1987 and is produced on large scale in chinese ovary hamster CHO cell lines [66].

References

1. Beadle GW, Tatum EL. Genetic Control of Biochemical Reactions in Neurospora. ProcNatlAcadSci U S A. 1941; 27: 499-506.
2. Oswald T Avery, Colin M MacLeod, and Maclyn McCarty. Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. J Exp Med. 1944; 79: 137–158.
3. Hershey AD, Chase M. Independent functions of viral protein and nucleic acid in growth of bacteriophage. J Gen Physiol. 1952; 36: 39-56.
4. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature. 1953; 171: 737-738.
5. Matthaei JH, Jones OW, Martin RG, Nirenberg MW. Characteristics and composition of RNA coding units. ProcNatlAcadSci U S A. 1962; 48: 666-677.
6. Jacob F, Perrin D, Sánchez C, Monod J, Edelstein S. [The operon: a group of genes with expression coordinated by an operator. C.R.Acad. Sci. Paris 250 (1960) 1727-1729]. C R Biol. 2005; 328: 514-520.
7. Mertz JE, Davis RW. Cleavage of DNA by R 1 restriction endonuclease generates cohesive ends. ProcNatlAcadSci U S A. 1972; 69: 3370-3374.
8. Berg P, Mertz JE. Personal reflections on the origins and emergence of recombinant DNA technology. Genetics. 2010; 184: 9-17.
9. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. ProcNatlAcadSci U S A. 1977; 74: 5463-5467.
10. Saiki RK,Gelfand DH, Stoffel S, Scharf SJ, Higuchi R. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science. 1988; 239: 487-491.
11. Smith HO, Wilcox KW. A restriction enzyme from Hemophilusinfluenzae. I. Purification and general properties. J Mol Biol. 1970; 51: 379-391.
12. Griffiths AJF, Sean B Carroll, John Doebley. Introduction to genetic analysis. 10thedn. New York: W.H. Freeman and Co. 2012.
13. James D Watson, Richard M Myers, Amy A Caudy, Jan A Witkowski. Recombinant DNA: genes and genomes: a short course. 3rdedn. New York: W.H. Freeman. 2007.
14. Showalter AK,Lamarche BJ, Bakhtina M, Su MI, Tang KH. Mechanistic comparison of high-fidelity and error-prone DNA polymerases and ligases involved in DNA repair. Chem Rev. 2006; 106: 340-360.

15. Watson N. A new revision of the sequence of plasmid pBR322. Gene. 1988; 70: 399-403.

16. Alberts B, Wilson JH, Hunt T. Molecular biology of the cell. 5thedn. New York: Garland Science. 2008.

17. James D Watson, Richard M Myers, Amy A Caudy, Jan A Witkowski. Recombinant DNA: genes and genomes: a short course. 3rdedn. New York: W.H. Freeman. 2007.

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

19. Gaj T,Gersbach CA, Barbas CF 3rd. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. 2013; 31: 397-405.

20. McManus MT, Sharp PA. Gene silencing in mammals by small interfering RNAs. Nat Rev Genet. 2002; 3: 737-747.

21. Venken KJ,Bellen HJ.Transgenesis upgrades for Drosophila melanogaster. Development. 2007; 134: 3571-3584.

22. Capecchi M. The first transgenic mice: an interview with Mario Capecchi. Interview by Kristin Kain. Dis Model Mech. 2008; 1: 197-201.

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

24. Bibikova M,Beumer K, Trautman JK, Carroll D. Enhancing gene targeting with designed zinc finger nucleases. Science. 2003; 300: 764.

25. Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. Genome editing with engineered zinc finger nucleases. Nat Rev Genet. 2010; 11: 636-646.

26. Carroll D. Genome engineering with zinc-finger nucleases. Genetics. 2011; 188: 773-782.

27. Urnov FD. Biological techniques: Edit the genome to understand it. Nature. 2014; 513: 40-41.

28. Durai S, Mani M, Kandavelou K, Wu J, Porteus MH. Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. Nucleic Acids Res. 2005; 33: 5978-5990.

29. Bogdanove AJ,Voytas DF. TAL effectors: customizable proteins for DNA targeting. Science. 2011; 333: 1843-1846.

30. Cong L, Ran FA, Cox D, Lin S, Barretto R. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013; 339: 819-823.

31. Mali P, Yang L, Esvelt KM, Aach J, Guell M. RNA-guided human genome engineering via Cas9. Science. 2013; 339: 823-826.

32. Dahlem TJ,Hoshijima K, Jurynec MJ, Gunther D, Starker CG. Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. PLoS Genet. 2012; 8: e1002861.

33. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol. 2013; 31: 227-229.

34. Chen L, Tang L2, Xiang H, Jin L3, Li Q3. Advances in genome editing technology and its promising application in evolutionary and ecological studies. Gigascience. 2014; 3: 24.

35. Hefferin ML,Tomkinson AE. Mechanism of DNA double-strand break repair by non-homologous end joining. DNA Repair (Amst). 2005; 4: 639-648.

36. van den Bosch M,Lohman PH, Pastink A. DNA double-strand break repair by homologous recombination. Biol Chem. 2002; 383: 873-892.

37. Pabo CO,Peisach E, Grant RA. Design and selection of novel Cys2His2 zinc finger proteins. Annu Rev Biochem. 2001; 70: 313-340.

38. Nishimasu H, Ran FA2, Hsu PD2, Konermann S3, Shehata SI3. Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell. 2014; 156: 935-949.

39. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. ProcNatlAcadSci U S A. 1996; 93: 1156-1160.

40. Vanamee ES,Santagata S, Aggarwal AK.FokI requires two specific DNA sites for cleavage. J Mol Biol. 2001; 309: 69-78.

41. Römer P, Hahn S, Jordan T, Strauss T, Bonas U. Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. Science. 2007; 318: 645-648.

42. Boch J,Scholze H, Schornack S, Landgraf A, Hahn S. Breaking the code of DNA binding specificity of TAL-type III effectors. Science. 2009; 326: 1509-1512.

43. Miller JC, Tan S, Qiao G, Barlow KA, Wang J.A TALE nuclease architecture for efficient genome editing. Nat Biotechnol. 2011; 29: 143-148.

44. Christian M,Cermak T, Doyle EL, Schmidt C, Zhang F. Targeting DNA double-strand breaks with TAL effector nucleases. Genetics. 2010; 186: 757-761.

45. Li D, Qiu Z, Shao Y, Chen Y, Guan Y. Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. Nat Biotechnol. 2013; 31: 681-683.

46. Schornack S, Meyer A, Römer P, Jordan T, Lahaye T. Gene-for-gene-mediated recognition of nuclear-targeted AvrBs3-like bacterial effector proteins. J Plant Physiol. 2006; 163: 256-272.

47. Moscou MJ,Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. Science. 2009; 326: 1501.

48. Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q. Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol. 2011; 29: 731-734.

49. Mahfouz MM, Li L, Shamimuzzaman M, Wibowo A, Fang X. De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. ProcNatlAcadSci U S A. 2011; 108: 2623-2628.

50. Barrangou R,Fremaux C, Deveau H, Richards M, Boyaval P. CRISPR provides acquired resistance against viruses in prokaryotes. Science. 2007; 315: 1709-1712.

51. Nakata A,Amemura M, Makino K. Unusual nucleotide arrangement with repeated sequences in the Escherichia coli K-12 chromosome. J Bacteriol. 1989; 171: 3553-3556.

52. Mojica FJ, Díez-Villaseñor C, Soria E, Juez G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. MolMicrobiol. 2000; 36: 244-246.

53. Horvath P, Romero DA, Coûté-Monvoisin AC, Richards M, Deveau H. Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus. J Bacteriol. 2008; 190: 1401-1412.

54. Jansen R,Embden JD, Gaastra W, Schouls LM. Identification of genes that are associated with DNA repeats in prokaryotes. MolMicrobiol. 2002; 43: 1565-1575.

55. Mojica FJ,Díez-Villaseñor C, García-Martínez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology. 2009; 155: 733-740.

56. Swarts DC,Mosterd C, van Passel MW, Brouns SJ. CRISPR interference directs strand specific spacer acquisition. PLoS One. 2012; 7: e35888.

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

58. Gasiunas G,Barrangou R, Horvath P, Siksnys V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. ProcNatlAcadSci U S A. 2012; 109: E2579-2586.

59. Cho SW, Kim S, Kim JM, Kim JS. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat Biotechnol. 2013; 31: 230-232.

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

61. Gomes FR, Maluenda AC, Tápias JO, Oliveira FLS, Sá-Rocha LC, et al. Expression of recombinant human mutant granulocyte colony stimulating factor (Nartograstim) in Escherichia coli. World J MicrobiolBiotechnol. 2012; 28:2593–2600.

62. MM Basso A. Biopharmaceutical and Biosimilar Products in Brazil: From Political To Biotechnological Overview. J Bioequivalence Bioavailab. 2013; 5: 60-66.

63. Behura SK, Severson DW. Codon usage bias: causative factors, quantification methods and genome-wide patterns: with emphasis on insect genomes. Biol Rev CambPhilos Soc. 2013; 88: 49-61.

64. Plotkin JB,Kudla G. Synonymous but not the same: the causes and consequences of codon bias. Nat Rev Genet. 2011; 12: 32-42.

65. Berlec A,Strukelj B. Current state and recent advances in biopharmaceutical production in Escherichia coli, yeasts and mammalian cells. J IndMicrobiolBiotechnol. 2013; 40: 257-274.

66. Altamirano C, Berrios J, Vergara M, Becerra S. Advances in improving mammalian cells metabolism for recombinant protein production. Electron J Biotechnol.2013.