ISSN 2449-8955 European Journal of Biological Research *Review Article*

**DOI:** http://dx.doi.org/10.5281/zenodo.5527223

# A review on CRISPR-Cas9 and its role in cancer immunotherapy

**Rashi A. Bhavsar [1], Vishwa Maharajan [1], Evan Joseph [1], Salai S. Sumukhi [1], Akshatha Banadka [2], Kokila Srinivasa [3,*]**

[1] The Oxford College of Science, Department of Biotechnology, #32, 19th Main, 17th 'B' Cross, Sector – 4, HSR Layout, Bengaluru – 560 102, Karnataka, India

[2] Christ Deemed-to-be University, Department of Life Sciences, Hosur Road, Bhavani Nagar, S. G. Palya, Bengaluru – 560 029, Karnataka, India

[3] The Oxford College of Science, Department of Biochemistry, #32, 19th Main, 17th 'B' Cross, Sector – 4, HSR Layout, Bengaluru – 560 102, Karnataka, India

* Corresponding author: Email: kokila.s1281@gmail.com

**Received:** 19 June 2021; **Revised submission:** 10 August 2021; **Accepted:** 23 September 2021

**https://jbrodka.com/index.php/ejbr**



**ABSTRACT:** Since the discovery of CRISPR, the field of Molecular Genetics has revolutionized and has opened so many different doors to improve molecular techniques and interpret the early microbial life forms. The diversity found within the CRISPR-Cas systems has led to its application in various fields like diagnostics, medicine and also has given rise to an interesting field of genome engineering. The Nobel Prize in Chemistry was awarded to Emanuelle Charpentier and Jennifer Doudna for their work on CRISPR-Cas9 and its application as a genome engineering tool. Scientists have been using the CRISPR-Cas9 system to edit genomes and cure various genetic diseases associated with mutations in the human genome. One such application is the use of CRISPR-Cas9 in cancer immunotherapy. The entire world has been known to be affected by the rapidly dividing cellular disease of cancer. Since cancer cells have different morphology, they are attacked by our immune system. Cancer cells possess the ability to camouflage themselves and avoid these immune responses and thereby proliferate and metastasize to a much greater extent. Scientists have been able to genetically engineer T-cells with the help of CRISPR-Cas9 genome editing tool which has shown promising results in the course of immunotherapy. On the 4th of June 2021, in India, the first patient underwent CAR-T Cell therapy setting a milestone for future treatments. In this review, we aim to evaluate the potential and diversity of the profound CRISPR-Cas systems and the application of CRISPR-Cas9 in immunotherapy for refractory cancer.

**Keywords:** CRISPR-Cas9; Genome; Cancer; Genetically-modified T-cells.

## 1. HISTORY OF CRISPR-CAS SYSTEM

In 1987, Y. Ishino and team, from Osaka University, Japan, made the first description of CRISPRs while they were attempting to determine the nucleotide sequence of the iap gene which codes for alkaline phosphatase isozyme in *Escherichia coli*. An odd repeat sequence was detected in the 3'-end flanking region of the iap gene containing five homologous sequences of 29-nucleotide interspaced by 32-nucleotide

sequence. Initially, it was expected to be repetitive extragenic palindromic (REP) sequences commonly found in *Escherichia coli* and *Salmonella typhimurium* but no similarities were found between both these sequences, thus its existence in *E. coli* remained unexplained. Subsequently, similar repetitive sequences were identified in a few members of the Enterobacteriaceae such as *S. dysenteriae*, *S. Enterica*, and *Mycobacterium tuberculosis* as well as in other *E. coli* strains [1, 2].

A major advancement in its study came with the discovery of archaeal CRISPR repeats in 1993. During the investigation of understanding the regulatory mechanisms permitting halophilic archaea to adapt to highly saline conditions, Francisco Mojica identified regularly spaced repeats within a long DNA sequence in the genome of the archaeon *Haloferax mediterranei*. Nevertheless, the biological role of these repeats could not be explained. With the discovery of automated sequencing machines and new efficient procedures for DNA sequencing during the 20$^{th}$ century, scientists sequenced and analyzed genome sequences of several archaea and bacteria which showed the presence of these unusual sequences in most of the prokaryotes. In 2002, the term CRISPR was given to these sequences by Jansen and team.

By comparing the CRISPR regions in the genome of many organisms, four conserved genes (termed cas (CRISPR-associated genes) 1 through 4 or cas1 to cas4) were discovered regularly present adjacent to the CRISPR regions. The discovery of several clusters of genes similar to *cas* genes in the genomes of hyperthermophilic archaea and the absence of the same in mesophilic archaea and moderate thermophilic bacteria by Makarova and colleagues lead to the prediction that these proteins could be part of a peculiar uncharacterized DNA repair system specific to thermophilic organisms. A crucial breakthrough was achieved by two groups independently - Christine Pourcel and team in Orsay, France and Francisco Mojica and her colleagues in Alicante, Spain. They observed that the host strains containing similar spacer sequences in the CRISPR were immune to the infection by certain bacteriophages and conjugative plasmids. Thus, they proposed that the CRISPR sequences function as a biological defense system to protect the host from such foreign extrachromosomal elements and also could incorporate pieces of the foreign invading DNA into the CRISPRs providing a memory of past aggressions [2, 3]. Using the lactic acid bacterium, *Streptococcus thermophilus,* their studies were experimentally proven in 2007 by Barrangou and team. Thus, the CRISPR-Cas system was identified to function as an acquired immune system in prokaryotes [2, 4]. In the later years, different CRISPR-Cas systems and their corresponding components were discovered and their role as immune defense systems was studied extensively in different hosts. Some of the prominent discoveries include matured crRNA guide interference for Type I systems, Type III-A system targets DNA, Type III-B Cmr complex cleaves ssRNA, CRISPR-Cas systems classified into three types and discovery of tracrRNA [3]. In August 2012, a major breakthrough in developing the CRISPR-Cas system as a tool for genome editing was made by researcher Jennifer Doudna, from the University of California, Berkeley with Emmanuelle Charpentier of the Hannover Medical School in Germany. They discovered that Cas9 (SpCas9) protein from the type II-A system of *Streptococcus pyogenes* could be used as crRNA guided DNA endonuclease to produce double stranded break at specific positions in the DNA. The study on another bacterium, *Streptococcus thermophilus* by Virginijus Siksnys and team further supported the former discovery. These two studies uncovered essential characteristics of the CRISPR- Cas9 system but could not explain how it could be used as a tool for genome editing in eukaryotic cells. Later in January 2013, four independent studies by Feng Zhang's group, George Church's group, Jenniffer Doudna's group and Jin-Soo Kim's group showed that genomic sites in human cells could be efficiently cleaved by CRISPR and chimeric gRNA (guideRNA) can be substituted to replace the tracrRNA-crRNA complex. In the subsequent months, genome editing with

CRISPR-Cas systems were successfully reported in different species [5, 6]. After this phenomenal advancement, the CRISPR-Cas9 system became the most widely and extensively used genome editing tool with its application varying from repairing genetic defects to developing genetically modified mouse models for human disease. In the following years, different CRISPR-Cas system types and further subtypes were identified in various prokaryotic organisms. Although the CRISPR-Cas9 system was capable of gene editing in eukaryotes, several obstacles, such as off-target mutations in the host genome were present to perform successful gene editing. Currently, scientists are working to fill the knowledge gaps in understanding the diverse CRISPR-Cas systems and modifying them to apply in various fields [5].

## 2. DIFFERENT TYPES OF CRISPR-CAS SYSTEMS

In most archaea and many bacteria, CRISPR-Cas systems were established to function as an adaptive immune system. The operation of these systems involves three principal phases: (1) incorporating the foreign DNA (protospacers) into the CRISPR array or adaptation, (2) CRISPR array transcription, maturation of crRNA and formation of transRNA-crRNA complex to guide the Cas protein, (3) cleavage of the foreign DNA or RNA, (4) regulatory and other CRISPR associated functions. To perform these operations, the CRISPR-Cas system constitutes two principal modules - adaptation module, effector complex (expression and interference module) and ancillary module. The integration of spacers into the CRISPR cassettes is executed by the adaptor module while the effector complex is responsible for processing and crRNA maturation, identification of the target foreign DNA or RNA, and cleavage of the foreign genome. The ancillary module consists of different proteins and domains performing regulatory functions [7, 8]. These defense systems are constantly engaged in an endless battle against foreign extrachromosomal elements like viruses, plasmids etc., which results in rapid evolution of the *cas* genes and thus leads to diversification in the mechanism and structure of the systems. This huge variance of CRISPR-Cas systems found in prokaryotes have been classified on the criteria of signature *cas* genes, organization of genes in the CRISPR-cas loci, analogous sequences between multiple shared Cas proteins, structure of CRISPR and phylogeny of Cas1 protein [9]. All the identified CRISPR-Cas systems have been divided into two distinct classes 1 and 2, on the basis of different effector proteins encoded by the case genes [2].

Class 1 systems include the presence of multi-subunit crRNA-effector complexes. Class 1 systems are classified into three different types, I, III, and IV, on the criteria of different architecture of the effector complex. Type I and III systems mostly occur in archaea and are less frequent in bacteria while the rare Type IV system occurs in bacteria [7-9].

### 2.1. Type I CRISPR-Cas systems

The principal gene for these systems is cas3 encoding for a ssDNA stimulated helicase capable of unwinding dsDNA and RNA-DNA duplexes. The HD family endonuclease domain is fused with the helicase domain and is involved in the cleavage of DNA. The domains are encoded by two genes, cas3″ and cas3′ respectively, on the same loci. Cascade or CRISPR-associated complex for antiviral defense is the effector complex in type I systems. The multi-subunit effector complex consists of paralogous RAMP's Cas5 and Cas7, Cas6, a large subunit and a small subunit. The core fold of the RAMP (Repeat-Associated Mysterious Proteins) is a nucleic acid-binding domain called RNA-recognition motif (RRM). Cas6 is an active endonuclease involved in the processing of crRNA [7-9]. Seven subtypes, I-A to I-F and I-U have been identified in Type I systems [7, 8].

### 2.2. Type II CRISPR-Cas systems

The signature gene for these systems is Cas9 which encodes a single multi-domain protein involved in the interference process, target DNA cleavage and adaptation process [7]. The Cas9 protein is composed of two lobes, REC (recognition) lobe and NUC (nuclease) lobe. The NUC lobe contains two nuclease domains, HNH and RuvC along with terminal PAM-interacting (PI) domain [2]. RNaseIII and tracrRNA are responsible for the processing of crRNA in these systems. The Type II systems are further divided into three subtypes, II-A, II-B, and II-C. The subtype II-A system comprises an additional gene, csn2 while subtype II-B comprises cas4 gene instead of csn2 [7, 8].

### 2.3. Type III CRISPR-Cas systems

Cas10 is the prominent gene encoding a multi-domain protein. It consists of two cyclase-like palm domains (RRM domain), helical domain comprising Zn-binding treble clef motif and a helical domain responsible for cleavage of target DNA [7, 8]. Cas6 in Type III systems are assumed to be associated with crRNA processing and are not a part of the multi-subunit crRNA-effector complex. Similar to Type I systems, these systems also consist of a larger subunit and a small subunit along with RAMP's Cas5 and Cas7. The small subunit is α-helical protein while a cyclase-related enzyme makes the large subunit. The Type III systems have four subtypes, III-A, III-B, III-C, and III-D. The effector complex in the subtypes III-A and III-D is known as csm while in case of subtypes III-B and III-C, it is known as cmr [8].

### 2.4. Type IV CRISPR-Cas systems

The Type IV system has been discovered in plasmids of several bacteria and has a unique minimalist effector complex architecture making it distinct from the other Class1 systems. The signature gene of this system is Csf1. The peculiar crRNA-effector complex consists of a highly reduced larger subunit, csf1, a presumed small subunit, RAMP'sCas5 and one Cas7 protein. This system is devoid of cas1 and cas2 genes responsible for spacer integration into CRISPR arrays [7, 8]. Type IV systems have two variants based on the presence of DinG family helicase [8].

Class 2 systems include a single multidomain crRNA-effector complex. Class 2 systems are classified into three different types, II, V, and VI on the criteria of different signature *cas* genes. All the Class 2 systems possess cas1 and cas2 proteins. The endonuclease activity of these proteins is requisite for the process of adaptation, that is, the integration of spacer into the CRISPR array [7, 8].

### 2.5. Type V CRISPR-Cas systems

Cas12 (formerly Cpf1) is the signature gene in Type V CRISPR-Cas systems. Cas12 is a single-RNA-guided nuclease which can function in absence of tracrRNA. It's a large protein composed of two RuvC-like nuclease domains along with TnpB protein. Some Type V systems also consist of Cas4 protein [2, 8]. The Type V systems are classified into three main subtypes, V-A, V-B, and V-U [2].

### 2.6. Type VI CRISPR-Cas systems

The preeminent gene in these systems is Cas13 encoding a multidomain effector complex consisting of two HEPN domains which may possess RNase activity. The three subtypes of Type VI systems are VI-A, VI-B, and VI-C [2, 9].

## 3. CONSTITUENTS OF THE CRISPR-CAS9 SYSTEM

Archaebacteria as well as the eubacteria have developed a specialized RNA guided adaptable immune system consisting of a two-component complex: CRISPR–Cas9 to fight against bacteriophages. Their genomes consist of unique CRISPR sequences or Clustered Regularly Interspaced Short Palindromic Repeats and spacer sequences [10]. Spacers are 17-84 nucleotide sequences that complement the invading foreign genome [11]. Successive transcription of the CRISPR system specifically the spacers produce many short mature CRISPR RNAs (crRNAs): 17-20 nucleotides long that complement the target DNA [12]. The CRISPR RNA consists of two prime ends; at the 5′ end it chains up the spacer sequences, whereas at the 3′ end it accomodates a segment of the CRISPR sequence. Small non-coding RNA, known as trans-activating crRNA (tracrRNA) associates with crRNA to construct a binary hybrid structure: Single guide RNA (sgRNA). The tracrRNA has two principal purposes: activation of pre-crRNA handling by RNase III and initiation of crRNA guide DNA cleavage by cas9 [13]. This sgRNA integrates the crRNA and tracrRNA into a single RNA code [14]. The single guide RNA is T-shaped which has one tetra-loop and two to three stem-loops [15].

### 3.1. The Cas9 Protein

Cas9 is a protein built up of 1,368 amino acids. It is a DNA endonuclease that spilts the double-stranded DNA (dsDNA) 3 bp above the PAM. The PAM or Protospacer Adjacent Motif is a 2-6 bp long nucleotide that contributes majorly to the ATP-independent cleavage of the viral DNA [16]. The cas9 protein is made up of six parts i.e., REC I, REC II, Bridge Helix, PAM Interacting, HNH and RuvC domains. REC I, II, or the recognition lobe works in the binding of the nucleic acids [17]. In the apostate, the structure of cas9 comprises of two lobes i.e., the alpha-helically shaped recognition (REC) lobe and the nuclease (NUC) lobe. The nuclease lobe has HNH nuclease (named due to the presence of characteristic histidine and asparagine residues) and spilt RuvC nuclease domains along with the more variable C-terminal domain (CTD). The HNH nuclease domain contributes to the cleavage of the DNA strand complementary to the guide RNA with the help of other HNH endonucleases by forming ββα-metal fold also called as the one-metal-ion mechanism. The RuvC domain belongs to the retroviral integrase superfamily identified by an RNase H fold. It cleaves the DNA opposite to the complementary strand by a two-metal-ion catalytic mechanism. These two lobes are interconnected to each other aided by two components: the former formed by a bridge helix which is arginine-rich and the latter by a disordered linker. The nucleic acids are bound by the two types of recognition lobes (nuclease and helical) to configure a four-way intersection that unzips the arginine-rich bridge helix [14]. If either HNH (H840A) or RuvC(D10A) domain is mutated, then cas9 is modified to a nickase. If both are mutated hindrance is not caused to the binding ability to the viral genome but it terminates the endonuclease activity. This cas9 is also called dead cas9 or dcas9 [13]. In type II CRISPR-Cas system, the cleavage of double-stranded DNA (dsDNA) of the viral genome by the cas9 protein allied with a single guide RNA which consists of a CRISPR RNA (crRNA) and a trans-activating CRISPR RNA (tracrRNA) by HNH and RuvC domains.

### 3.2. Cpf1 - an additional RNA guided endonuclease

Cpf1 is a RNA-guided nuclease observed in the CRISPR arrangement of *Prevotella* and *Francisella*. It is a 1,300 amino acid protein that does not require a tracrRNA hence, simplifying the cleavage process. Cpf1 associated CRISPR systems are transcribed to mature crRNAs. The cpf1-crRNA complex severs the

DNA by detecting a PAM that is T-rich and then instigates a double stranded break at a site away from the recognition site with a 4 or 5-nucleotide 5′ overhang and hence, forming a sticky end [18].

## 4. MECHANISM OF THE CRISPR-CAS9 COMPLEX

The defense mechanism of the bacteria against invading foreign genome is further categorized into three phases:

### 4.1. Adaptation phase

When the bacterial genome is exposed to the invading viral genome, small segments of the foreign DNA are incorporated into the CRISPR-spacer complex of the bacterial genome forming new spacers at the leading strand of the array. The leading region includes promoter, protein binding sites and elements required for spacer integration. Spacers build up the genetic memory of the cell by recording the infection and precluding future infection from the same virus. Hence, this phase is also known as 'Adaptation'. Palindromes aid the process of integration of spacers into the CRISPR system by providing direction as well as position [15].

### 4.2. Expression of crRNA and Cas proteins

Subsequent transcription of the spacers yields small pre-crRNA which is further processed to form smaller units of crRNA. Interaction between the CRISPR RNA and the complementary viral target DNA sequence also called protospacer during the second invasion, triggers disintegration of the viral DNA [19]. Sometimes, the mature crRNA binds to tracrRNA together called single guide RNA (sgRNA) and conducts cas9 mediated cleavage of viral DNA. In type I and III CRISPR systems, multiple Cas protein complexes are involved in the cleavage of foreign DNA. Whereas, in type II CRISPR system, one protein: Cas9 is required for the destruction of the viral genome [20].

### 4.3. CRISPR interference

The single guide RNA complex guides cas9 to cleave DNA carrying a complementary 20-nucleotide target sequence and adjacent PAM. It was experimentally found that seed sequences of RNA nucleotides within the crRNAs assist in target specification. This seed region is called PAM or Protospacer Adjacent Motif which is a 10-12 base pair sequence found at the 3′ end of the 20-nucleotide spacer in the type II CRISPR systems. PAM mutations severely impede target binding and cleavage [14]. The gRNA activates a significant conformational displacement in the Cas9 complex, which triggers a transformation from an inactive state to an active state of the protein [16]. The Cas protein consists of Recognition (REC) lobes and nuclease (NUC) lobes in addition to C-terminal domain (CTD), these simplify the identification and cleavage of viral DNA. The nuclease lobe comprises two domains namely HNH and RuvC domains. The complementary strand of DNA is cleaved by the HNH nuclease domain with a one-metal-ion mechanism signified by a conserved general base histidine. The RuvC domain slits the non-complementary DNA with the help of a two-metal-ion mechanism characterized by a conserved aspartate residue [21]. A positively charged furrow between the two lobes has the PAM duplex embedded in it. The C-terminal domain (CTD) accommodates the PAM-containing the non-target strand mainly through hydrogen-bonding interactivities with the phosphate spine of the PAM-holding the non-target DNA strand. NGG represents the PAM sequence with N being based twinned with its complement and does not interact with Cas9. The GG nucleotides are

processed in the major furrow by base-definite hydrogen-bonding interactions with two arginine residues (R1333 and R1335) located at a β-hairpin of the CTD [14] (Fig. 1).

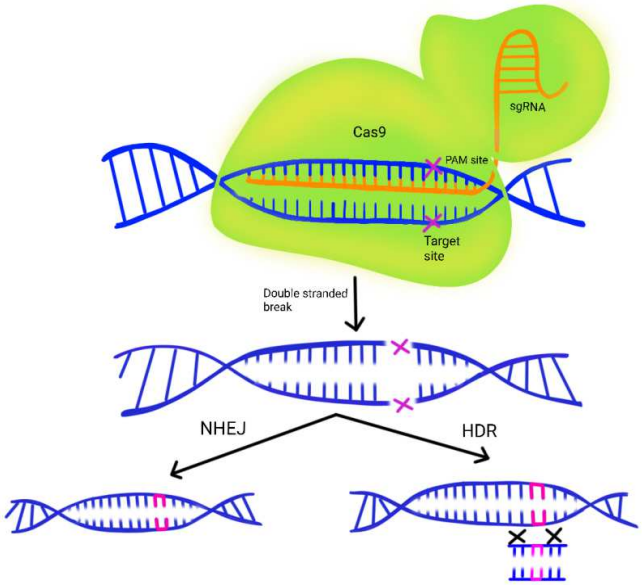


**Figure 1.** Mechanism of CRISPR-Cas9.

## 5. THE CONFORMATIONAL REARRANGEMENT

Considerable number of interactions of Cas9 with guide RNA's phosphate spine leads to the preordering of the 10-nucleotide RNA seed sequence essential for DNA recognition resulting in an A-form conformation. Additionally, the PAM-interacting sites R1333 and R1335, that are in charge of 5′-NGG-3′ PAM recognition is signaled prior to binding of the viral DNA so that formation of the sgRNA-Cas9 complex takes place. As soon as the sgRNA binds to the Cas9, an eminent conformational change is observed in the Hel-III domain of the REC lobe which moves by ~65A° toward the HNH domain. This suggests most of the prominent conformational changes take place before DNA binding [14]. The manifestation of the 5′ end of the guide RNA inside the cleft formed between the HNH and RuvC nuclease domains is observed from the electron microscopy (EM) structure of the Cas9 protein bound to a sgRNA. The conclusion drawn is that the 5′ end of the guide RNA is preserved from degradation, and a further conformational change is required so that the 5' end is released during target DNA binding [14]. Three-dimensional collisions take place and target recognition is triggered as Cas9 alienates from the DNA with inappropriate PAM sequence. As soon as the target DNA is recognized, Cas9 triggers a local DNA melting at the PAM binding site. The canonical 5′-NGG-3′ PAM nucleotide sequence is present on the non-target DNA strand resulting in a cleaved non target DNA strand and an untwined target DNA strand. Introduction of RNA strand builds an RNA-DNA hybrid. The local DNA melting is observed because of interactivities between the phosphate lock loop and +1 phosphate [16]. A noticeable sharp kink turn is observed in the target DNA strand (within the immediate vicinity of the PAM) to substantiate the binding of target DNA strand with the guide RNA instead of the non-target DNA strand. The +1-phosphate preserved by the phosphate lock loop ensures corroboration of the flipping and rotating of the first nucleobase of the target DNA towards the guide RNA. This explains how the presence of PAM on the non-target strand is essential in cleaving the single stranded DNA target strand [22]. It is the hydrophobic and van der Waals forces that hold the non-target strand. It laterally comes out through the positively charged tunnel betwixt the RuvC and HNH nucleases. This strand goes through a sharp kink turn at positions -2 and -3

(based on presence of PAM) where flipping occurs and then again at -4 position. These flipping of bases and kinks that occur in the untwined target strand and displaced non target strand (also called as R-loop) causes exposure of two seed nucleotides above the PAM to bulk solvents for initiating target DNA binding [22]. Thus, Cas9 cleaves the DNA strand preventing it from infecting the bacteria. The double stranded break (DSB) is repaired in two ways:

1. Non-homologous end joining (NHEJ)
2. Homology-directed repair (HDR)

### 5.1. The Non-Homologous End Joining (NHEJ)

Non-homologous end joining is a repair mechanism which is error-prone in many organisms because it leads to frameshift mutation that is deletion or insertion of a base [23]. It requires nucleases to cut DNA sequences, polymerases to introduce new DNA sequences, ligases to join the sequences along with a lot of enzymes for repairing the double stranded break [24]. It rapidly repairs the DNA and is a predominant repair mechanism in most organisms. Activation of NHEJ is observed throughout the cell cycle [25]. Figure 2 represents NHEJ.

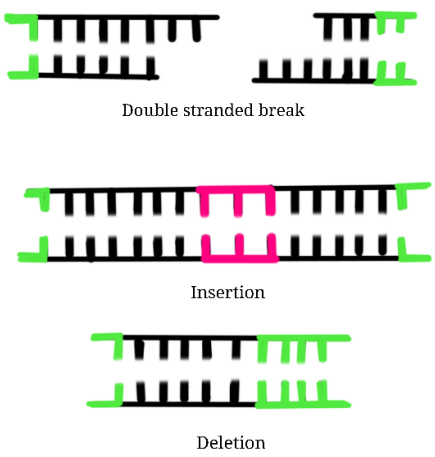


**Figure 2.** Non-homologous end joining.

### 5.2. Homology-Directed Repair (HDR)

Homology-directed repair mechanisms have high fidelity, but it cannot function in absence of a homologue close to the location of the double stranded break [24]. The requirement of homologous DNA nucleotides either from sister chromatids or foreign DNA assures homology-directed repair to occur [23]. It is confined to the S and G2 phase of the cell cycle as sister chromatids are only accessible during these phases [25]. The high fidelity of the homology-directed repair mechanism allows scientists for insertion, deletion and substitution of single nucleotide sequences or considerable amounts of genomic DNA sequences. HDR-based gene editing methods are recently being developed, opening more possibilities for genomic editing [26]. Figure 3 represents HDR.

## 6. CRISPR-CAS9: A GENOME EDITING TOOL

Scientists manipulated the bacteria's defense mechanism: CRISPR CAS9 as a meticulous genome editing tool. A guide RNA is designed by the scientists to complement the gene they desire to edit. This designed gRNA is attached to Cas9 which leads Cas9 to the target gene. Cas9 is a dual-RNA-guided double stranded endonuclease. It possesses the ability to cause double stranded breaks in the DNA at sequences that

match the sequences in the gRNA. The molecular scissors cut the target DNA sequence by undergoing several rearrangements in the protein. Hence, scientists can edit any DNA from a viral immune mechanism and its manipulation [15, 18].

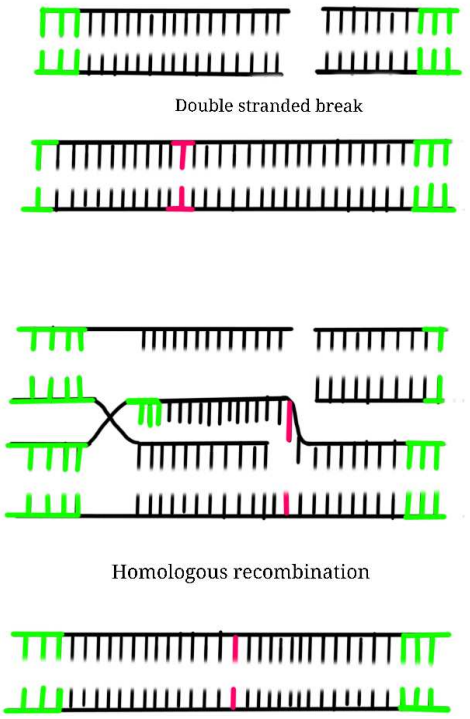


**Figure 3.** Homology-directed repair.

## 7. WHAT IS CANCER?

An extremely intricate nexus of 30 trillion cells make up the human body, with each cell following a communal mechanism. Normal cells interact with each other to regulate proliferation by restricting its growth according to the needs of the body by contact inhibition. Contrarily, cancer cells defy the specialized instruction system. They reproduce obstreperously and can migrate from the site of their origin [27]. Cancer arises from malignant tumors. Tumor is an abnormally growing mass of cells that divides uncontrollably and there are three types of tumors. Benign tumors do not possess the ability to develop into cancer. They either cannot spread or grow. Premalignant tumors are tumors that are not yet cancerous but may develop into malignant tumors. Lastly, malignant tumors are cancerous; they can grow and spread to other parts of the human body. These cancer cells have the property of metastasis: a multitudinous process in which a series of steps is followed. The cancer cells must encroach the surrounding stroma, intravasate, endure in the circulatory system, extravasate, invade the matrix and subsequently proliferate in the target organ [28]. Sometimes, cancer does not respond to medical treatment and hence is called refractory cancer. This becomes resistant either from the first treatment procedure or progresses during the treatment. Therefore, refractory cancer is also called resistant cancer. Cancer cells tend to exhibit distinguishing characters which are referred to as hallmarks and are represented in the Figure [29].

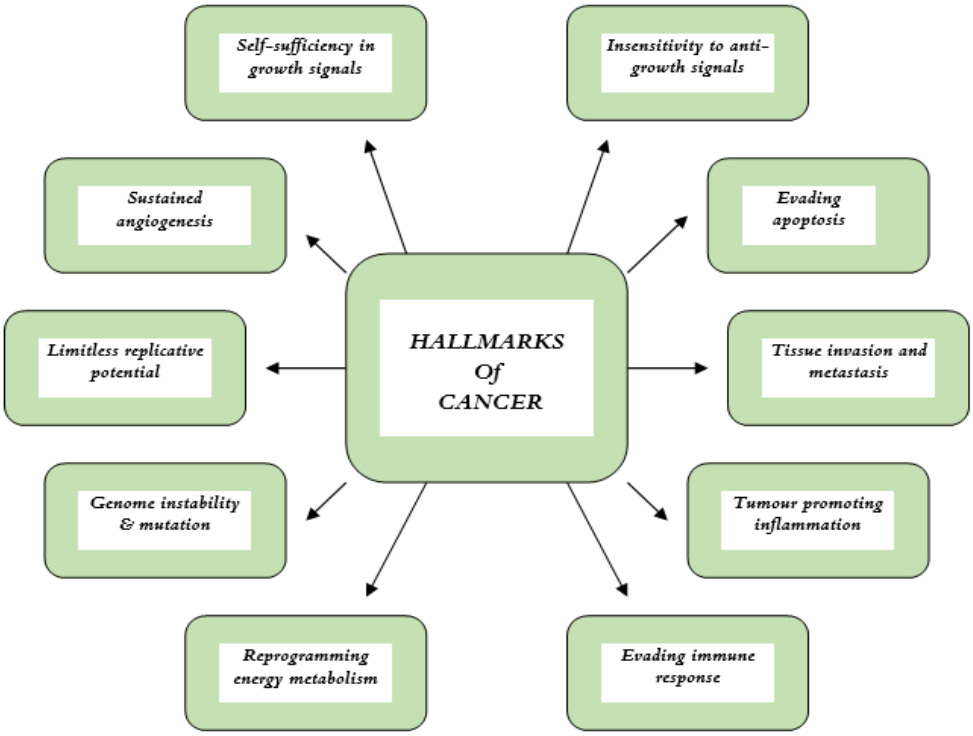


**Figure 4.** The 10 hallmarks of cancer (Adapted from [29]).

## 8. THE WARBURG EFFECT, DIFFERENCE BETWEEN NORMAL CELLS AND CANCER CELLS

Only in the absence of oxygen, normal cells metabolize glucose into lactate. In aerobic conditions, mitochondrial oxidative phosphorylation (OXPHOS) produces ATP which yields carbon dioxide ($CO_2$) and water ($H_2O$) as the products. On the other hand, cancer cells adopt anaerobic glycolysis even in aerobic conditions. Hence, cancer cells are highly glycolytic. Excessive expression of membrane glucose transporters (GLUTs) and its several isoforms aid the uptake of glucose by the cancer cells. This was experimentally proven with the help of Positron Emission Tomography (PET) imaging, a widely used imaging technique for diagnosing cancer. This observation of a peculiar mechanism adopted by cancer cells was made by Otto Warburg nearly a hundred years ago. He observed that even under aerobic conditions tumor slices exhibited higher rates of consumption of glucose and secretion of lactate than the normal cells. This effect was named after the scientist as the Warburg effect [30]. In normal cells, growth controlling messages from the outer surface of the cell is sent deep into the nucleus by signal pathways. These messages are collected by a molecular apparatus called the cell cycle clock which decides that cell division should take place or not. Due to genetic mutations in cancer cells, either stimulatory pathways send excessive signals for division or the inhibitory pathways fail to send inhibiting signals which lead to uncontrollable division of cancer cells. A mutation in a component like growth factor receptor triggers excessive functioning of the stimulatory pathway independently without any commands. If a mutation is caused in cytoplasmic relay, the inhibitory pathway is arrested, and this interrupts the signaling chain [27].

## 9. CANCER CELL CYCLE

The cell cycle clock is a molecular apparatus composed of a network of interacting proteins in the nucleus. It coalesces messages from the stimulatory and inhibitory pathways and decides whether cells should divide or not. It is mainly built up of two crucial components: cyclins and cyclin dependent kinases (CDK's).

These ally together and initiate the various stages of the cell cycle [27]. Multiple cyclin dependent kinases (CDK's) and cyclin complexes are present but only certain types of CDK-Cyclin complexes directly contribute to the cell cycle. These comprise of:

1. Three interphase CDK's (CDK2, CDK4, and CDK6)
2. CDK1-a mitotic CDK also called as a cell division control protein2 (CDC2)
3. Other ten cyclins that mainly belong to four distinct classes of cyclins namely the A-, B-, D- and E-type of cyclins [31].

With an upsurge in the level of cyclins: the D type, followed by E, A and lastly B type, the advancement through the four stages of the cell cycle takes place. In the late G1 phase, the cell determines whether to proceed to another division or enter a resting phase- quiescent stage (G0) at the restriction point (R). A molecular switch needs to be turned on for the cell to continue the cell cycle by entering the S phase and passing through the restriction point (R). These checkpoints regulate the cell cycle by identifying the defects caused during DNA synthesis. Mutations in the DNA are maybe caused by endogenous and exogenous genotoxic agents such as chemicals, free radicals, ionizing radiation, side products of the intracellular metabolism or medical therapy which is inspected by the DNA damage checkpoint. The spindle assembly checkpoint (SAC) is responsible for chromosomal segregation. Unequal inheritance of the genetic instruction is induced under the absence of SAC which may lead to tumor succession by congregating numerical chromosomal aberrations (CIN) [31]. The consequent upsurge in levels of cyclin D and E triggers the switch for inhibition by combining to and activating enzymes titled cyclin dependent kinases. These seize the phosphate groups from the molecules of cyclin D and hence prevent further cell division process [27]. Deregulation of cyclin-dependent kinases (CDKs) mainly cause three cell cycle defects: unscheduled proliferation, GIN (genomic instability) and CIN (chromosomal inability). Genomic instability (GIN) is caused when the genome encounters mutations and chromosomal abnormalities leading to debilitated repair of the cell's genome. Chromosomal instability (CIN) is caused when numerical aberrations in the chromosomes are observed. These mutations lead to persistent proliferation or spontaneous recurrence in the cell cycle- two common properties of most tumor cells [31]. This further leads to excessive production of cyclins resulting in causing a lot of inhibitory pathways and, hence causing cancer. Some proteins such as pRB, p15, p16, p21, p53 also contribute majorly to the production of cancer cells [27].

## 10. CANCER METASTASIS

The process of metastasis is characterized by a malignant tumor cell migrating from a primary epithelial neoplastic lesion to a distant site through the circulatory system and establishing a secondary tumor which is no longer in contiguity with the primary tumor [32, 33]. Metastasis includes a course of discrete steps involving Epidermal-Mesenchymal transition (EMT), dissociation from bulk tumor, angiogenesis, invasion, intravasation, transport or cell migration and finally extravasation into a new distant site [34]. Most tumors originate from epithelial tissues which are composed of cells laid in sheets with lateral belts of cell-to-cell adhesion complexes and are cemented to a non-cellular basement membrane. The process of EMT helps the tumor cell migrate from the bulk tissue to the secondary site. During EMT, the tumor cells activate genes necessary for differentiation into mesenchymal cells. It is characterized by the upregulation of transcription factors like SNAI1/snail, SNAI2/slug, ZEB1, ZEB2, TWIST1, TWIST2 as well as E12/E47. These factors get activated through multiple pathways like Receptor Tyrosine Kinases (RTK); HIF1, HIF2, Notch signaling pathway in response to hypoxia and NF-κB, TGF-β for an anti-inflammatory response. TGF-β or

Transforming growth factor beta is responsible for activating snail and down regulating CDH16 gene encoding cadherin-16 (calcium dependent glycoprotein) and HNF-1β. Snail up-regulates Akt (protein kinase B) and Bcl-xL to restrict apoptosis induced by transcription growth factor. Along with this, Snail also down-regulates Cyclin D2 to restrict cell cycle progression [34, 35]. Expression of EMT inducing genes leads to the downregulation of epithelial proteins (like E-cadherin, occludin, claudins, cytokeratins or catenins) and upregulation of mesenchymal proteins (like N-cadherin, vimentin, tenascin C, laminin β1 or collagen type VI α) [34]. E-cadherin (epithelial cadherin) is a transmembrane protein linked with the actin cytoskeleton by α-catenin and β-catenin, involved in formation of adherens junctions and anchoring neighboring cells together. The crucial step in metastasis is the loss of E-cadherin. E-cadherin promoter genes are silenced through hypermethylation and histone deacetylation by E-cadherin repressors like snail, twist, zeb etc. Phosphorylation of β-catenin or proteolytic cleavage inactivates the existent E-cadherin proteins on the cell surface and transport of E-cadherin to the cell membrane is inhibited by the o-glycosylation of the protein post-translation process [35]. Integrins along with FAK (focal adhesion kinase) signaling and SRC signaling are essential for cell migration and inhibiting anoikis during EMT. PTK2 protein tyrosine kinase 2 (PTK2) or FAK is a type of protein kinase that phosphorylates β-catenin leading to detachment from E-cadherin and thus, dissociation from the bulk tumor. Actin–myosin 2-mediated cell contraction and adhesion to ECM (extracellular matrix) and release of adhesion mediated by integrin- and FAK-containing complexes are responsible for cell migration. The programmed cell death of anchorage-dependent cells upon detachment from ECM is called Anoikis. Thus, anoikis suppression plays an integral role in metastasis. Activation of integrins leads to its binding with ECM molecules and triggers an intracellular signaling cascade via FAK and SRC family kinases which suppresses anoikis [34, 35]. To invade the neighboring cells by breaking through the basement membrane, metastatic tumor cells secrete zinc dependent endopeptidases called MMP (Matrix metalloproteinases) which help in the degradation of components of ECM and cleavage of cell surface proteins [35, 36]. Cleavage of E-cadherin by MMP's results in the formation of a smaller fragment, sE-cad, which is responsible for maintaining EMT [36]. To produce distant metastases, malignant tumor cells invade tumor associated vasculature to reach distant sites which is facilitated by the process of angiogenesis or generation of new blood vessels. This particular feature termed as an angiogenic switch represents a crucial step in the process of producing secondary metastases [33, 34]. During angiogenesis, tumor cell promotes vascularisation and growth occurs past its diffusion limit as a result of the delicate balance between angiogenic activators [such as VEGF A (vascular endothelial growth factor A), FGFs (fibroblast growth factors), PDGFs (platelet derived growth factors or heparin binding growth factors) and EGFs (epidermal growth factors)] and angiogenic inhibitors (such as thrombospondin 1, angiostatin, endostatin, and tumstatin) tipping over to the pro-angiogenic side. A vital role in activating the angiogenic switch is performed by numerous tumor cell-intrinsic factors and stromal cells, particularly myeloid cells [34]. This process also involves the interactions of endothelial cells, tumor cells and extracellular matrix. Serine proteases and metalloproteinases mediate these interactions and are also affected by angiogenic factors [33]. Intratumor hypoxic conditions also support induced angiogenesis leading to invasive metastatic tumor and some hypoxia-inducible factors such as HIF1A, HIF2A play a vital part in the process [34]. Angiogenesis is followed by intravasation of malignant cells and then transported through the vessels to the secondary site. The process of intravasation involves local proteolysis of the extracellular matrix, followed by the pseudopodial extension, and cell migration. Certain factors affect the motility of tumor cells during intravasation. These factors include autocrine motility factors, matrix proteins, and host-secreted growth factors. The autocrine motility factors

such as hepatocyte growth factor/scatter factor (HGF/SF), insulin like growth factor II (IGF-II) and autotaxin (ATX) are secreted by the tumor cells while matrix proteins are extracellular matrix proteins (such as vitronectin, fibronectin, laminin type I collagen, type IV collagen and thrombospondin) that stimulate movement towards chemical gradient known as chemotaxis and motility stimulation towards a bound substrate known as haptotaxis. Host secreted growth factors or paracrine motility factors, such as insulin like growth factor I, interleukin 8, and histamine, are host secreted growth factors which are responsible for the motility of tumor cells towards the specific organ producing the factors. Cell shape, cytoskeletal rearrangements, and changes in cell adhesion and/or membrane fluidity are some of the changes that these motility factors cause through various mechanisms [33]. The mechanisms involved in tumor cell extravasation from the blood vessels into the organ may be similar to those contributing to invasion [32].

## 11. SEED-SOIL HYPOTHESIS

Stephen Paget, an English surgeon, published a seminal paper to explain the definite pattern of metastasis in 1889 called seed-soil hypothesis. He analyzed autopsy records of more than 900 patients with different primary tumors which revealed a peculiar nonrandom pattern of the process of metastasis to bones and visceral organs [37]. He observed low incidence of metastasis in the spleen as compared to the liver, ovary and specific bones and this disproportion was less pronounced in melanoma as compared to breast and uterine cancer. This led to the proposal of the 'seed and soil' principle which stated that seeds which fall on congenial soil will live and grow even though they are carried in all directions. The "seed" can be considered as a progenitor cell, cancer stem cell, or metastatic cell, while the "soil" as host factors, stroma, or the organ microenvironment which is suitable for the progenitor cell to grow [38, 39]. In 1929, James Ewing challenged Paget's theory of 'seed and soil' principle by theorizing that dissemination of metastasis is ensued by mechanical factors which are determined by the vascular system's anatomical structure but this was proved false in 1970's and thus the 'Seed soil hypothesis' remains the more widely accepted theory [39]. Figure 4 represents an example of the seed and soil hypothesis.

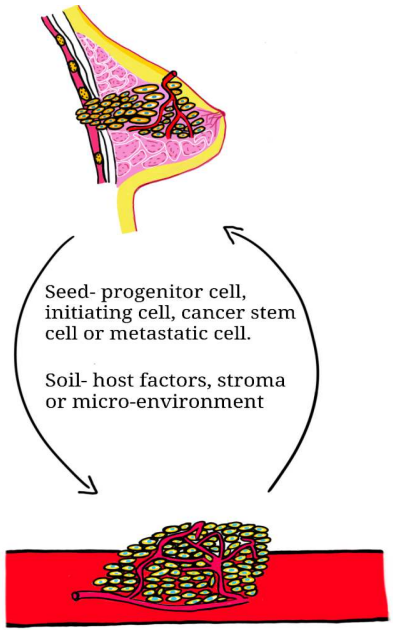


**Figure 4.** Seed and soil hypothesis.

## 12. ONCOGENES

Mutations in oncogenes, anti-oncogenes, and microRNAs cause cancer. The primary transformed cell undergoes secondary or tertiary genetic mutations and hence cytogenetically different clones of tumors are formed. Tumors can also encompass progenitor cancer cells besides the initial clone and subclones. These cells comprise a spectrum of cells with discrete genetic modifications and states of differentiation causing cancer [40]. Studies of Burkitt's lymphoma provide the first corroboration that cancer emerges from somatic genetic mutations. On chromosome 8q24, a MYC oncogene is translocated to one of the loci for immunoglobulin genes. Other translocating partners- chromosome 14q, 22q, and 2p bear enhancer elements in the immunoglobulin loci which activates the MYC oncogene [41]. Hence, every malignant lymphocyte undergoes MYC translocation. In chronic myelogenous leukemia, a reciprocal t (9;22) chromosomal translocation fuses the ABL protooncogene to the BCR gene. An oncogenic ABL fusion protein is produced from the fusion gene with excess tyrosine kinase activity. This chromosomal alteration is a characteristic of all the leukemic cells [42]. The function of oncogenes is to encode proteins that regulate cell proliferation, apoptosis, or both. Structural modifications caused by mutations, gene fusion, association with enhancer elements or amplification activate oncogenes. Translocations and mutations can transpire either as an initiating event or during tumor progression. On the other hand, amplification takes place during tumor progression. There are six broad groups of products yielded by oncogenes: transcription factors, chromatin remodelers, growth factors, growth factor receptors, signal transducers, and apoptosis regulators [40].

**Table 1.** Products of oncogenes in brief.

| Products of oncogenes | Description | Reference |
|---|---|---|
| Transcription factors | Transcription factors are constituents of multigene families that comprise of common structural domains. Interaction of these with other proteins is necessitated. For example, the AP1 transcription factor is formed by the dimerization of the Fos transcription protein and the Jun transcription factor which amplifies the expression of genes that control cell division. In lymphoid cancers, chromosomal translocation activates transcription factor genes and at times in solid tumors (e.g., prostate cancers). In Ewing's sarcoma, gene fusion of the EWS gene and the partner genes leads to anomalous transcriptional activity of the fused proteins | [40, 43] |
| Chromatin remodelers | The extent of condensation of chromatin is a factor in control of gene expression, replication, and repairing and of chromosome segregation. The chromatin is remodeled by two types of enzymes: ATP- dependent enzymes and enzymes that alter the N-terminal tails of histones. The structure of chromatin and its transcriptional activity is directed by the interactions between the nucleosomes and chromatin-associated proteins. These interactions are determined by an epigenetic code encrypted in the form of histone alteration | [44, 45] |
| Growth factors | When growth factor genes are activated, it leads to the malignancy of the cells. The phosphorylation of β-catenin is ceased by the WNT family of secreted glycoproteins. β-catenin adheres cells to cells, activates signal-transduction pathways, and is regulated by the APC protein. Mutations in the APC occlude the degradation of β-catenin by inhibiting the phosphorylation in familial adenomatous polyposis. The free β-catenin in the cytoplasm displaces to the nucleus, where genes involved in cell proliferation and invasion are activated. | [46] |
| Growth factor receptors | Epidermal growth factor receptor (EGFR), a transmembrane protein with tyrosine kinase activity is triggered by the deletion of the ligand-binding domain and hence absence of ligand binding in many tumors. Phosphorylation of tyrosine occurs in the intracellular domain of the receptor, providing interaction sites for cytoplasmic proteins containing the SRC homology domain and other binding domains. This leads to deregulation of signaling in several pathways. Hypoxia-dependent control of gene transcription is modulated by vascular endothelial growth factor (VEGF). It triggers angiogenesis in several cancers. | [40] |

| Products of oncogenes | Description | Reference |
|---|---|---|
| Signal transducers | Autophosphorylation of tyrosine residues occur in the intracellular part of the receptor upon its reorganization caused by binding of receptor tyrosine kinases to appropriate ligand. This intensifies the kinase activity of the receptor or stimulates the interaction of the receptor with domains of cytoplasmic proteins (e.g., the SRC homology 2 domain) that act as effectors and regulators of intracellular signaling. Upon mutations, oncogenes cipher members of signal transduction pathways, which are divided into two major groups: non-receptor protein kinases and guanosine-triphosphate–binding proteins | [47, 48] |
| Apoptosis regulators | The BCL2 gene, the initiator of almost all follicular lymphomas and some diffuse large B-cell lymphomas, chronic lymphocytic leukemia and lung cancer produces a cytoplasmic protein that localizes to mitochondria and escalates cell survival by inhibiting apoptosis. Apoptosis can be achieved by two different pathways: stress pathway and the death-receptor pathway. The former is activated proteins that contain the BCL2 homology 3 domain which disables inactivates BCL2 and BCL-XL and hence, stimulating apoptosis. The latter is stimulated by the binding of Fas ligand, TRAIL, and tumor necrosis factor α, to their corresponding (death) receptors on the cell surface. This causes cell death | [40, 49] |

## 13. ONCOGENE ACTIVATION

Chromosomal rearrangements, mutations, and gene amplification are three mechanisms that activate oncogenes by either causing a mutation in the oncogene structure or an upsurge in or demodulation of its expression. This highly increases the survival of cells with such modifications [50].

**Table 2.** Oncogene activation.

| Mechanisms of oncogene activation | Description | Reference |
|---|---|---|
| Chromosomal rearrangements | The characteristic cytogenetic abnormalities in cancerous cells are chromosome inversions and translocations. These amplify or demodulate transcription of the oncogene in hematopoietic cancers and solid tumors. Translocation of a gene that bears a promoter that is continually active in the target cells with a gene that has oncogenic activity (e.g., ERG1) occurs in prostate cancer. In cancers of B- and T-cells, activation of oncogenes is stimulated by MYC deregulation. On the other hand, gene fusion activates oncogenes in myeloid cancers and soft-tissue sarcomas. | [51] |
| Mutations | Different types of mutations take place in oncogenes. These modify the structure of the encoded protein thereby, strengthening its transforming activity. The RAS oncogenes (KRAS, HRAS, and NRAS) encrypt proteins with guanosine nucleotide–binding activity and intrinsic guanosine triphosphatase activity. Mutations in codon 12, 13, or 61 results in a protein, which constantly transports signals by linking tyrosine kinases to downstream serine and threonine kinases. This further leads to ceaseless cell growth. Carcinomas of the lung, colon, and pancreas commonly have mutated KRAS gene. In acute myelogenous leukemia and the myelodysplastic syndrome, the NRAS gene undergoes mutation. | [40] |
| Gene amplification | In methotrexate-resistant acute lymphoblastic leukemia, amplification of the dihydrofolate reductase gene (DHFR) is commonly observed during the progression of tumor. Along with amplification of DHFR, cytogenetic mutations take place that mimic amplification of oncogenes. Often mutations occur in four different oncogene families: MYC, cyclin D1 (or CCND1), EGFR, and RAS. In small-cell lung cancer, breast cancer, esophageal cancer, cervical cancer, ovarian cancer, and head and neck cancer, amplification of MYC gene is noticed. | [40] |

## 14. CRISPR-ENGINEERED T CELLS IN REFRACTORY CANCER PATIENTS

Myriad diseases have loomed over humans since time immemorial, but gene editing carries the budding potential to repair DNA alterations and hence, eliminate several genetic diseases. Demonstration of gene editing was first observed in mammalian cells when double-stranded DNA breaks were repaired by homologous and non-homologous recombination by an endonuclease. Some of the few edited nucleases

which have tremendous applications are homing endonucleases, zinc finger nucleases, CRISPR-Cas9 (clustered regularly interspaced short palindromic repeats associated with Cas9 endonuclease) and transcription activator-like effector nucleases. Recent progression in the CRISPR-Cas9 technology offers promising advancement in cancer therapy by causing DNA alterations in human T cells [55].

### 14.1. T-cell: killers of the immune system

T-cell, the specialized killer of the immune system is the heart of modern cancer immunotherapy. The T cell receptor (TCR) complex is superficially placed on the T cell. It is responsible for acknowledging foreign antigens/peptides attached to MHC-molecules causing anti-tumor responses. In adoptive cell therapy, T cells from patients are genetically edited and reinstalled in the body to produce a transgenic TCR that can destroy the tumor cells. Two genes, TCRα (TRAC) and TCRβ (TRBC) encode the endogenous T cell receptor (TCR) chains. Once the transgenic TCR is infused in the patient it exhibits mispairing and/or competition with alpha and beta chains of the endogenous TCR for expression. Mispairing of the therapeutic TCR alpha and beta chains with endogenous alpha and beta chains leads to excessive production of self-reactive TCRs and diminished expression of transgenic surface TCR. To eliminate this downfall, TCRα (TRAC) and TCRβ (TRBC) were knocked out in T cells which enhanced the expression of a synthetic, cancer-specific TCR transgene (NY-ESO-1). Another gene namely PDCD1 encoding PD-1 was deleted to enhance anti-tumor immunity. Mice with chronic lymphocytic choriomeningitis virus infection and PD-1 deficient T cells exhibited improved cytotoxicity and enhanced accumulation of terminally differentiated T cells. Antibody blockage of PD- 1 lead to improved chimeric antigen receptor (CAR) or TCR T cell-mediated killing of tumor cells in vitro and enhanced elimination of PD-L1+ tumor xenografts in vivo [55].

### 14.2. CAR (chimeric antigen receptor) T-cells

CARs or chimeric antigen receptors enable T cells to identify and produce an immune response against antigen that expresses cancer cells [56]. The antigen binding component and a spacer are the components that make up the extracellular domain of the CAR. Several antigen binding moieties present could be:

1. scFv (single-chain fragment variable) acquired from mouse monoclonal antibodies(mAbs), humanized Abs or fully human Abs is a variable monoclonal antibody fragment. It mainly identifies and binds to tumor-associated antigens (TAAs), expressed on the tumor cell surface.
2. A human Fab fragment, chosen from phage display libraries.
3. Nature ligands that engage their cognate receptor.

In contrast to TCRs, CARs identify the antigens, carbohydrate, and glycolipid structures (found on tumor cell surface) without the need of MHC. In CAR T-cell therapy, patient's T-cells are genetically edited to express a chimeric antigen receptor (CAR) for a tumor antigen, followed by ex vivo cell expansion and lastly re-infusion of the engineered T cells [57]. In B cell malignancies i.e., B cell acute lymphoblastic leukemia (B-ALL), B cell non-Hodgkin's lymphoma (BNHL), chronic lymphocytic leukemia (CLL), and Hodgkin's lymphoma (HL), CAR T- cell therapy has shown spectacular results. This is mainly achieved by targeting CD19, CD2O or CD30. In CD19 specific CAR T-cells for B-ALL 70~94% of high complete remission (CR) rates have been observed [58].

Six patients were enrolled at first out of which four were subjected to detailed release criteria testing as specified in the FDA accepted Investigational New Drug (IND) application. One patient among the four unique patient number (UPN) 27 experienced rapid clinical progression and was not eligible for infusion due

to mandated safety protocols. Two among these three, had refractory advanced myeloma and one had a refractory metastatic sarcoma unaffected by multiple other therapies. Lymphodepleting chemotherapy with cyclophosphamide and fludarabine on days -5 to -3 (i.e., prior to administration with CRISPR-Cas9 engineered T cells) and a single infusion of $1\times10^8$ manufactured CRISPR-Cas9 engineered T cells per kg on day 0 of the protocol was given to patients. There was no administration of cytokines given to them [55].

### 14.3. Manufacturing of the T-cell product

Incubation of protein with gRNA at a molar ratio of 1:1 at 25°C for 10 minutes done immediately prior to electroporation resulted in Cas9 Ribonucleoprotein complex (RNP). Manufacturing of the T cell product was done by electroporation of ribonucleoprotein complexes (RNP) consisting of recombinant Cas9 loaded with equimolar mixtures of sgRNA for TRAC, TRBC and PDCD1. Further, lentiviral transduction of the transgenic TCR was done by Adoptive cell therapy was conducted where patient's T cells were extracted, engineering and infused back. The engineered T cell product was named as "NYCE" (NY-ESO-1 transduced CRISPR 3X edited cells). Expansion of all products to $>1 \times 10^{10}$ T cells was done by the time of harvest [55].

## 15. WAYS TO GENETICALLY ENGINEER T-CELLS

To genetically engineer T-cells, viral and non-viral transfection methods with high transgene expression, less toxicity and less oncogenic adverse effects can be adopted.

### 15.1. Viral transduction

Due to ease of manufacturing, production, enhanced the ability of stable integration of genetic component into the host genome, viral vectors of the family Retroviridae (lentivirus and γ-retrovirus), adenovirus and adeno-associated viruses are used. According to clinical safety standards, viral vector platforms should exhibit replication incompetence, low genotoxicity, and low immunogenicity. The genes required to produce a CAR vector are *gag*, *pol*; and *env*, *rev* (for lentivirus). These are eliminated from the viral backbone. For viral production, they are provided in trans in helper plasmids. To create a stable virus-producing cell line for large-scale production, transfection of a packaging cell line is done with CAR transgene and the helper plasmids (with *gag, pol*, and *env* genes). Incubation of stimulated T-cells (with OKT3/CD28 beads) with retroviral particles is practiced for genomic incorporation. The virion core that is formed upon the unification of viral and host membrane is emancipated into the cytosol, followed by conduction along the microtubules to reach the nucleus. This method allows the production of T-cells with high amount of CAR [57].

### 15.2. Transposons

Transposons, the mobile genetic elements mainly consist of:

- One plasmid having the CAR (transposon)
- Another plasmid containing the transposase.

This dual component system enables stable integration of a transgene. The transposase acts on the inverted terminal repeats (ITRs) which leads to fringing of the CAR sequence. This further leads to excision and subsequent integration at a TA nucleotide sequence in the target cell genome. Electroporation of DNAs plasmid containing the CAR (transposon) and the transposase is done into the T-cells. After the transposition and stable genomic incorporation, the CAR is expressed on the T-cell surface [57].

### 15.3. CRISPR/Cas9

The short guide RNA (gRNA) that functions as an endonuclease can be transferred by liposome-mediated transfection, electroporation, chemical transductionor as part of a viral genome in the form of Cas9 protein/gRNA ribonucleoprotein (RNP), or in the form of a plasmid, driven by either U6 or H1 promoters for transcription after transfection of mammalian cells. A donor template in a plasmid form incorporates the desired transgene by homology-directed repair (HDR). An alternative non-viral method is adopted through nanomaterials. One of these approaches includes the biotin-streptavidin conjugate and the transport and binding of the templates from the donor to the Cas9 modified human cells. This increases the rates and efficiency up to 5 times more than traditional methods [57].

#### *15.3.1. Observed responses after infusion of the NYCE*

The three patients were given $1 \times 10^8$ cells/kg but due to variation in TCR transduction efficiency the number of infused engineered T cells ranged from $6.0 \times 10^7$ to $7.1 \times 10^8$ cells. No development of humoral response to Cas9 was observed in the three patients tested at different time points after the infusion. The infusions of engineered T cells had no adverse effects such as cytokine release syndrome which is quite common in cancer immunotherapies. Endurance of the infused cells, low content of Cas9 in the infused product and/or immunodeficiency in the patients due to substantial previous treatments can be responsible for the lack of immunization to Cas9. Chip-based digital PCR was used to ascertain the engraftment frequency of the CRISPR-Cas9 gene-edited cells. Patient UPN35 with the lowest transduction efficiency, had the lowest amount of steady state engineered T cells. The stability period of the transduced T cells varies from three to nine months after infusion i.e.,5 to 50 cells per µl of blood. Biopsy of the bone marrow in myeloma patients and tumor in the sarcoma suggested crowding of the engineered T cells to the tumor. One patient had a 50% decrease in a huge abdominal mass that was persistent for four months, along with lesions. Until December 2019, two are receiving additional therapies and one (UPN07) deceased due to progressive myeloma. Genetics modifications at the TRAC and PDCD1 locus were observed in all the patients. In patients UPN39 and UPN07, genetics edits were persistent at the TRAC and PDCD1 locus at the frequency of 5 to 10% of circulating peripheral blood mononuclear cells (PBMC). Edits at the TRBC locus were the lowest in frequency and barely detectable. Hence, the TRBC locus contains the lowest level of editing efficiency [55].

#### *15.3.2. Efficiency of the CRISPR-Cas9 genome editing*

To inspect the Cas9-mediated cleavage specificity, the iGUIDE method was adopted- a moderation of the GUIDE-seq method. On and off- target editing efficiency was evaluated in the NYCE cells at the end of product manufacturing. An impediment caused during assays is that DNA double-strand breaks are formed incessantly during cell division at high rates in the absence of added nucleases, which surges the background in assays of off-target cleavage. Majority of the off-target mutations were detected for TRBC than for other loci out of the three sgRNA. Less number of off-target edits were detected in over 7000 sites of cleavage in the sgRNA of PDCD1. Fewer off-target edits were observed at the TRAC1 and TRAC2 loci. Relatively lower mutations were identified within the transcriptional unit of CLIC2 (chloride intracellular channel 2) for the TRAC sgRNA. Whereas, for the TRBC sgRNA off-target reads were detected in genes encoding a transcriptional regulator (ZNF609) and a long intergenic non-protein coding RNA (LINC00377) [55].

### *15.3.3. Identification of chromosomal translocations in the engineered T cells*

Detection of translocations at frequencies $10^{-4}$ to $10^{-2}$ was observed in gene editing of TRAC and CD52 using transcription activator-like effector nucleases (TALENs) in the preclinical studies. In successive clinical report using TALENs, chromosomal rearrangements have been determined in 4% of infused cells. To look at the protection and genotoxicity of multiplex CRISPR-Cas9 genome editing on three chromosomes, stringent launch standards of the synthetic cells and assays to identify translocations were adopted. Certified qPCR assays were used to quantify the twelve likely translocations that would arise with the simultaneous modifications of four loci: TRAC, TRBC1, TRBC2, PDCD1. Translocations were identified in all manufactured products, but the translocations have been on the restriction of detection for the assay in patient UPN39.

TRBC1:TRBC2 was the most plentiful rearrangement, ensuing in a 9.3 kB deletion. The deletion and translocations peaked on days five to seven of producing after which declined in frequency until cell harvest. The translocations and the TRBC1:TRBC2 deletion were perceptible in the three patients among 10 days after infusion and 30 to 170 days after infusion. Frequency of rearrangements diminished in vivo implying that no proof of a growth advantage over many generations of expansion in the patients on this trial. At day 30, 150 and 170 in UPN07, UPN35 and UPN39, chromosomal translocations have been on the limits of detection are now no longer detected for all rearrangements besides for the 9.3 kB deletion for TRBC1:TRBC2 [55].

## 16. EVOLUTION OF ENGINEERED T-CELLS OVER TIME IN PATIENTS

To analyze the transcriptomic phenotype and the evolution of the engineered T-cells over time single cell RNA sequencing (scRNAseq) was employed in patient UPN39. UPN34 was selected as there was evidently the highest level of cell engraftment in the patient. Hence, tumor regression was clearly perceptible. The patient UPN39 was infused with CRISPR-Cas9 engineered T cells, after which recovery was achieved from the blood on day 10 (D10) and at ~4 months (day 113). For each sample (infusion product, D10 and D113), T cells had been organized primarily based upon the expression of CD4 or CD8. Further, processing was done using the droplet-based 5' scRNAseq. To genotype single cells as wild-type or mutant, PCR was utilized to amplify the cellular cDNA alike to the NY-ESO-1 TCR transgene from the gene expression libraries. Cells with mutations in all three target sequences were detected in the infusion product. TRAC was accepted as the most mutated gene. Around 10% of the T cells were triple-mutated at the target sequences, 20%: double-mutated at the target sequences, 30% of cells had no detectable mutations whereas, ~40% had one mutation [55].

## 17. CONCLUSION

Since the advancement in technology and developments in the field of diagnostics and medicine, there have been various studies to analyze and fight the deadly disease, cancer. Scientists have still been exploring every nook and corner to fight and ultimately cure cancer. From peptides extracted from bacteria [59] to the use of micro-RNAs as therapeutics [60], apart from chemotherapy and radiation, scientists have been working hard to contract cancer. Since the past few years, scientists have started using CRISPR-Cas9 to improve the genetic aspects in various fields, one of them being cancer therapeutics. Using this engineered T-cell to fight cancer is a novel idea and further improvement and perfection in this technique can lead to the production of T-cells that can be administered to most of the cancer types and ultimately help combat this ailment in a better way.

**Authors' Contributions:** RB, VM, EJ and SS planned, conceptualized, designed and structured the review article under the supervision of AB and KS. RB, VM, EJ and SS wrote the article with the inputs from all the authors. The manuscript was amalgamated and formatted by EJ. The figures were designed and made by SS. The article was scrutinized and finalized by SS, EJ, AB and KS. The article was written under the supervision of KS and AB. RB, VM, EJ and SS share the first authorship and have the right to list their names first in their CVs. AB and KS share the second authorship for this review paper. All authors are read and approved the final manuscript.

**Conflict of Interest:** The author has no conflict of interest to declare.

**REFERENCES**

1. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. J Bacteriol. 1987; 169(12): 5429-5433.
2. Ishino Y, Krupovic M, Forterre P. History of CRISPR-Cas from encounter with a mysterious repeated sequence to genome editing technology. J Bacteriol. 2018; 200(7): e00580-617.
3. Han W, She Q. CRISPR history: Discovery, characterization, and prosperity. Prog Mol Biol Transl. 2017; 152: 1-21.
4. Barrangou R. CRISPR provides acquired resistance against viruses in prokaryotes. Science. 2007; 315(5819): 1709-1712.
5. Liang P, Zhang X, Chen Y, Huang J. Developmental history and application of CRISPR in human disease. J Gene Med. 2017; 19: 6-7.
6. Javed MR. CRISPR-Cas system: History and prospects as a genome editing tool in microorganisms. Curr Microbiol. 2018; 75(12): 1675-1683.
7. Makarova KS, Koonin EV. Annotation and classification of CRISPR-Cas systems. Methods Mol Biol. 2015; 1311: 47-75.
8. Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, Koonin EV. An updated evolutionary classification of CRISPR-Cas systems. Nat Rev Microbiol. 2015: 13(11): 722-736.
9. Koonin EV, Makarova KS, Zhang F. Diversity, classification and evolution of CRISPR-Cas systems. Curr Opin Microbiol. 2017; 37: 67-78.
10. van Belkum A, Scherer S, van Alphen L, Verbrugh H. Short-sequence DNA repeats in prokaryotic genomes. Microbiol Mol Biol Rev. 1998; 62(2): 275-293.
11. Marraffini LA. CRISPR-Cas immunity in prokaryotes. Nature. 2015; 526(7571): 55-61.
12. Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJH, Snijders APL, van der Oost J. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science. 2008; 321(5891): 960-964.
13. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012; 337(6096): 816-821.
14. Jiang F, Doudna JA. CRISPR–Cas9 Structures and Mechanisms. Annu Rev Biophys. 2017; 46(1): 505-529.
15. Gupta D, Bhattacharjee O, Mandal D, Sen MK, Dey D, Dasgupta A, Ghosh D. CRISPR-Cas9 system: A new-fangled dawn in gene editing. Life Sci. 2019; 232: 116636.
16. Anders C, Niewoehner O, Duerst A, Jinek M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. Nature. 2014; 513(7519): 569-573.
17. Palermo G, Chen JS, Ricci CG, Rivalta I, Jinek M, Batista VS, McCammon JA. Key role of the REC lobe during CRISPR-Cas9 activation by "sensing", "regulating", and "locking" the catalytic HNH domain. Q Rev Biophys. 2018; 51: e9.

18. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Zhang F. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell. 2015; 163(3): 759-771.
19. Hale CR, Zhao P, Olson S, Duff MO, Graveley BR, Wells L, Terns MP. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. Cell. 2009; 139(5): 945-956.
20. van der Oost J, Westra ER, Jackson RN, Wiedenheft B. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. Nat Rev Microbiol. 2014; 12(7): 479-492.
21. Yang W. An equivalent metal ion in one- and two-metal-ion catalysis. Nat Struct Mol Biol. 2008; 15(11): 1228-1231.
22. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature. 2014; 507(7490): 62-67.
23. Tang X-D, Gao F, Liu M-J, Fan Q-L, Chen D-K, Ma W-T. Methods for enhancing clustered regularly interspaced short palindromic repeats/Cas9-mediated homology-directed repair efficiency. Front Genet. 2019; 10: 551.
24. Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem. 2010; 79(1): 181-211.
25. Ray, U., & Raghavan, S. C. Modulation of DNA double-strand break repair as a strategy to improve precise genome editing. Oncogene. 2020; 39(41):6393–6405.
26. Salsman J, Dellaire G. Precision genome editing in the CRISPR era. Biochem Cell Biol. 2017; 95(2): 187-201.
27. Weinberg RA. How cancer arises. Sci Am. 1996; 275(3): 62-70.
28. Duda DG, Duyverman AMMJ, Kohno M, Snuderl M, Steller EJA, Fukumura D, Jain RK. Malignant cells facilitate lung metastasis by bringing their own soil. PNAS. 2010; 107(50): 21677-21682.
29. Joseph E, Praveena EA, Fatma S. Non-small cell lung cancer and necroptosis as a Potential Biomarker. Recent Advances and Current Trends in Chemical and Life Sciences, Immortal publications, 2021.
30. Kalyanaraman B. Teaching the basics of cancer metabolism: Developing antitumor strategies by exploiting the differences between normal and cancer cell metabolism. Redox Biol. 2017; 12: 833-842.
31. Malumbres M, Barbacid M. Cell cycle, CDKs and cancer: a changing paradigm. Nat Rev Cancer. 2009; 9(3): 153-166.
32. Gupta GP, Massagué J. Cancer metastasis: building a framework. Cell. 2006; 127(4): 679-695.
33. Woodhouse EC, Chuaqui RF, Liotta LA. General mechanisms of metastasis. Cancer. 1997; 80(S8): 1529-1537.
34. Geiger TR, Peeper DS. Metastasis mechanisms. Biochim Biophys Acta. 2009; 1796(2): 293-308.
35. Heerboth S, Housman G, Leary M, Longacre M, Byler S, Lapinska K, Sarkar S. EMT and tumor metastasis. Clin Transl Med. 2015; 4(1): 6.
36. Kessenbrock K, Plaks V, Werb Z. Matrix metalloproteinases: regulators of the tumor microenvironment. Cell. 2010; 141(1): 52-67.
37. Fidler IJ. The pathogenesis of cancer metastasis: the “seed and soil” hypothesis revisited. Nat Rev Cancer. 2003; 3(6): 453-458.
38. Fidler IJ, Poste G. The “seed and soil” hypothesis revisited. Lancet Oncol. 2008; 9(8): 808.
39. Mathot L, Stenninger J. Behavior of seeds and soil in the mechanism of metastasis: a deeper understanding. Cancer Sci. 2012; 103(4): 626-631.
40. Croce CM. Oncogenes and cancer. N Engl J Med. 2008; 358(5): 502-511.
41. Dalla-Favera R, Bregni M, Erikson J, Patterson D, Gallo RC, Croce CM. Human c-myconc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. PNAS. 1982; 79(24): 7824-7827.

42. Shtivelman E, Lifshitz B, Gale RP, Canaani E. Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. Nature. 1985; 315(6020): 550-554.
43. Shaulian E, Karin M. AP-1 in cell proliferation and survival. Oncogene. 2001; 20(19): 2390-2400.
44. Peterson CL, Workman JL. Promoter targeting and chromatin remodeling by the SWI/SNF complex. Curr Opin Genet Dev. 2000; 10(2): 187-192.
45. Jenuwein T, Allis CD. Translating the histone code. Science. 2001; 293(5532): 1074-1080.
46. Giles RH, van Es JH, Clevers H. Caught up in a Wnt storm: Wnt signaling in cancer. Biochim Biophys Acta Rev Cancer. 2003; 1653(1): 1-24.
47. Pawson T, Warner N. Oncogenic re-wiring of cellular signaling pathways. Oncogene. 2007; 26(9): 1268-1275.
48. Kaziro Y, Itoh H, Kozasa T, Nakafuku M, Satoh T. Structure and function of signal-transducing GTP-binding proteins. Annu Rev Biochem. 1991; 60(1): 349-400.
49. Tsujimoto Y, Cossman J, Jaffe E, Croce CM. Involvement of the bcl-2 gene in human follicular lymphoma. Science. 1985; 228(4706): 1440-1443.
50. Bishop JM. Molecular themes in oncogenesis. Cell. 1991; 64(2): 235-248.
51. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun X-W, Chinnaiyan AM. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science. 2005; 310(5748): 644-648.
52. Borst J, Ahrends T, Bąbała N, Melief CJM, Kastenmüller W. CD4+ T cell help in cancer immunology and immunotherapy. Nat Rev Immunol. 2018; 18(10): 635-647.
53. Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. Immunity. 2013; 39(1): 1-10.
54. Finn OJ. Cancer immunology. N Engl J Med. 2008; 358(25): 2704-2715.
55. Stadtmauer EA, Fraietta JA, Davis MM, Cohen AD, Weber KL, Lancaster E, June C. H.CRISPR-engineered T cells in patients with refractory cancer. Science. 2020; 367(6481): 7365.
56. Kaushik I, Ramachandran S, Srivastava SK. CRISPR-Cas9: A multifaceted therapeutic strategy for cancer treatment. Semin Cell Dev Biol. 2019; 96: 4-12.
57. Miliotou AN, Papadopoulou LC. CAR T-cell therapy: A New Era in cancer immunotherapy. Curr Pharm Biotechnol. 2018; 19(1): 5-18.
58. Wang Z, Wu Z, Liu Y, Han W. New development in CAR-T cell therapy. J Hematol Oncol. 2017; 10(1): 53.
59. Karpiński TM, Szkaradkiewicz AK. Anti-cancer peptides from bacteria. Bangladesh J Pharmacol. 2013; 8(3): 343-348.
60. Kasinski AL, Slack FJ. MicroRNAs en route to the clinic: progress in validating and targeting microRNAs for cancer therapy. Nat Rev Cancer. 2011; 11: 849-864.