Review

Crafting Genetic Diversity: Unlocking the Potential of Protein Evolution

Vamsi Krishna Gali [1] **, Kang Lan Tee** [1,2,*] **and Tuck Seng Wong** [1,2,*]

1. Department of Chemical & Biological Engineering, University of Sheffield, Sir Robert Hadfield Building, Mappin Street, Sheffield S1 3JD, UK; v.gali@sheffield.ac.uk
2. Evolutor Ltd., The Innovation Centre, 217 Portobello, Sheffield S1 4DP, UK
* Correspondence: k.tee@sheffield.ac.uk (K.L.T.); t.wong@sheffield.ac.uk (T.S.W.);
  Tel.: +44-114-222-7507 (K.L.T.); +44-114-222-7591 (T.S.W.); Fax: +44-114-222-7501 (K.L.T. & T.S.W.)

Abstract: Genetic diversity is the foundation of evolutionary resilience, adaptive potential, and the flourishing vitality of living organisms, serving as the cornerstone for robust ecosystems and the continuous evolution of life on Earth. The landscape of directed evolution, a powerful biotechnological tool inspired by natural evolutionary processes, has undergone a transformative shift propelled by innovative strategies for generating genetic diversity. This shift is fuelled by several factors, encompassing the utilization of advanced toolkits like CRISPR-Cas and base editors, the enhanced comprehension of biological mechanisms, cost-effective custom oligo pool synthesis, and the seamless integration of artificial intelligence and automation. This comprehensive review looks into the myriad of methodologies employed for constructing gene libraries, both *in vitro* and *in vivo*, categorized into three major classes: random mutagenesis, focused mutagenesis, and DNA recombination. The objectives of this review are threefold: firstly, to present a panoramic overview of recent advances in genetic diversity creation; secondly, to inspire novel ideas for further innovation in genetic diversity generation; and thirdly, to provide a valuable resource for individuals entering the field of directed evolution.

Keywords: protein engineering; directed evolution; genetic diversity; random mutagenesis; molecular cloning

1. Introduction

The year 2018 marked a significant milestone in the field of protein engineering, as it witnessed the well-deserved recognition of Prof. Frances H. Arnold with the Nobel Prize in Chemistry [1,2]. This prestigious accolade honoured her ground-breaking contributions to directed protein evolution, particularly through the utilization of a simple yet powerful algorithm. Directed evolution involves iterative cycles of genetic diversity creation, followed by screening and selection (Figure 1). The far-reaching applications and profound impacts of directed evolution are evident in various domains, including, but not limited to, the development of highly active or stable enzymes, the creation of novel enzymatic functions and chemistries, and the engineering of biopharmaceutical proteins.

At the core of protein evolution lies the essential prerequisite of genetic diversity. Indeed, without this diversity, the evolutionary process becomes untenable. In recognition of this pivotal aspect, this review article is dedicated to examining methodologies deliberately designed for the creation of genetic diversity, with a focus on publications from 2014 onward. This article serves as an extension to our prior reviews on the same subject in 2006 [3] and 2013 [4].
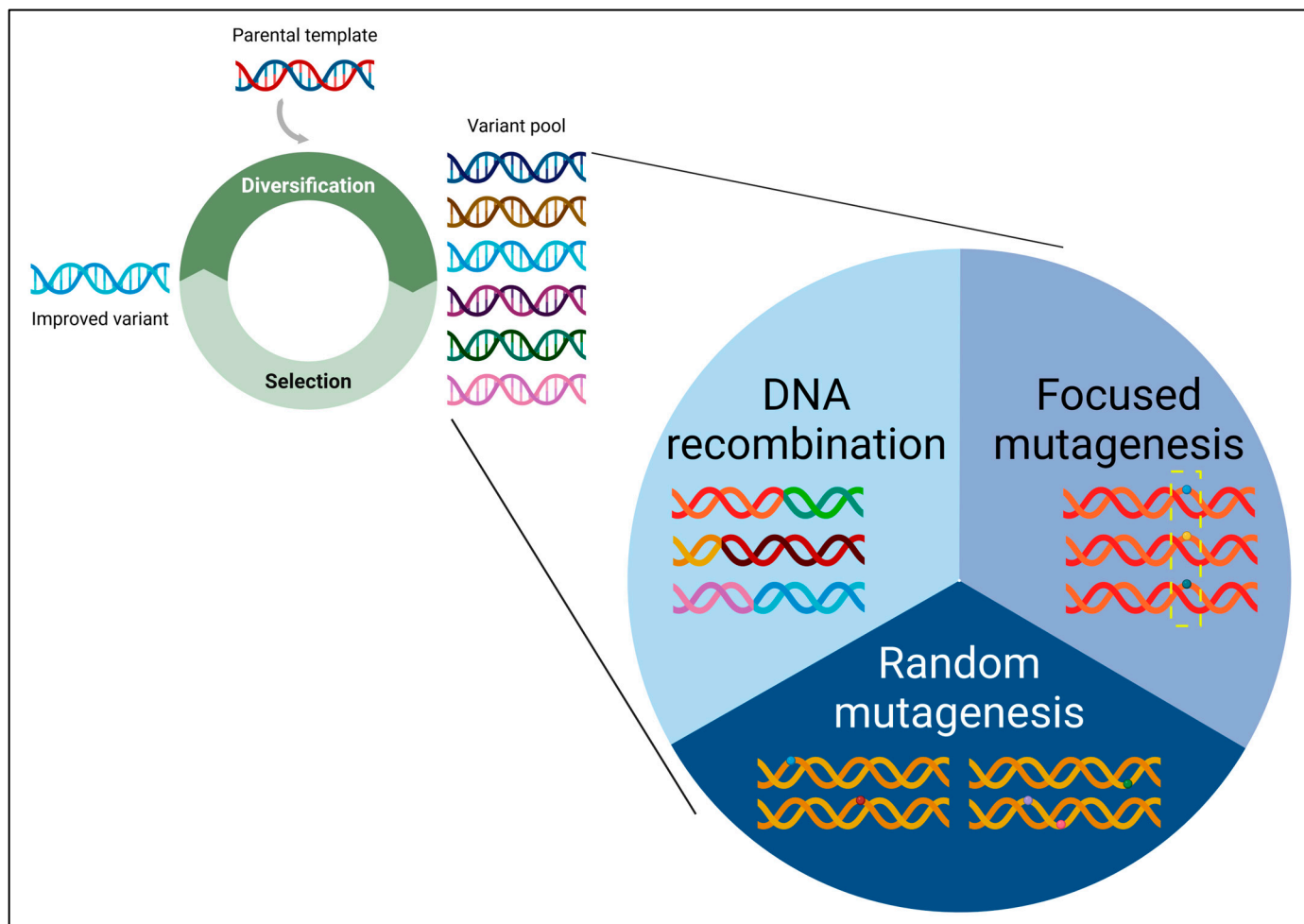
Figure 1. (**Top left**) The directed evolution cycle. The parental gene of interest (GOI) undergoes mutagenesis to generate a diverse pool of genetic variants. This pool is then subjected to a selection process targeting the desired phenotype, enabling the identification of improved variant(s). This iterative cycle is repeated until the desired trait is successfully achieved. (**Bottom right**) Classification of genetic diversity creation methods. The diverse methods for generating a genetically varied gene pool can be systematically categorized into three main classes: random mutagenesis, focused mutagenesis, and DNA recombination. Random mutagenesis involves the introduction of random mutations throughout the starting parental gene sequence. Focused mutagenesis targets mutations to specific pre-selected region(s) or amino acid residue(s) within the starting parental gene sequence. DNA recombination generates chimeric sequences by combining segments from a set of either homologous or non-homologous parental sequences.

The methodologies for generating genetic diversity can be categorized into three main classes: random mutagenesis, focused mutagenesis, and DNA recombination (Figure 1). There exist hybrid methods that integrate features derived from two or more of these classes, showcasing the versatility and adaptability of current approaches. In random mutagenesis, mutations are randomly introduced into the target gene of interest (GOI). Conversely, focused mutagenesis involves the randomization or semi-randomization of selected amino acid residues within the target gene. In DNA recombination, a collection of homologous or non-homologous genes undergo random fragmentation and subsequent reassembly to form chimeric constructs. The fundamentals of directed evolution and genetic diversity creation have been comprehensively documented in our recently published protein engineering textbook [5]. Therefore, this review primarily explores the significant advancements achieved in the past decade.

This review commences by discussing the latest advancements in molecular cloning, a crucial initial step in the process of generating genetic diversity. Subsequent to this, we delve into the assessment of gene library quality. Building upon this foundation, we offer succinct summaries of innovative methodologies within each category of genetic diversity creation, emphasizing their individual merits and limitations.

2. Molecular Cloning

DNA cloning serves as the cornerstone for mutagenesis experiments, involving the insertion of a target GOI or a gene library into a cloning/expression vector. Molecular cloning also plays an integral role in all synthetic biology projects at various stages. Conventionally, this process relies on the use of restriction endonucleases, which, while fundamental in molecular biology, present certain limitations. Drawbacks such as sequence specificity, time-consuming restriction digestion and ligation reactions, and the creation of undesirable 'scars' underscore the necessity for more versatile DNA cloning techniques. To address contemporary demands and challenges and facilitate rapid advancements, DNA assembly methods must be robust, cost-effective, swift, user-friendly, and highly scalable.

In this section, we present a comprehensive overview of the latest advancements in DNA cloning technology over the past decade. These innovations are categorized into *in vivo*, *in vitro*, modular DNA assembly, and automated DNA assembly methods. The cost per reaction for these methods varies significantly, ranging from a mere USD 0.0025 to a substantial USD 36.3 per reaction, as per 2019 pricing [6].

2.1. In Vivo Cloning

In vivo cloning techniques (Table 1) leverage the cellular machinery of a host organism, such as bacteria or yeast, to incorporate a DNA fragment into a vector. Recent descriptions of these techniques often rely on DNA recombination or repair systems, such as homologous recombination (HR) and non-homologous end joining (NHEJ). The appeal of *in vivo* cloning techniques lies in their distinct advantages: they obviate the need for *in vitro* ligation reactions and are not constrained by restriction endonuclease sites. Introducing a short homologous sequence (30–40 bp) to any vector or DNA fragment through PCR facilitates recombinational cloning, underscoring the simplicity and versatility of this approach.

An underutilized yet increasingly popular *in vivo* technique involves harnessing the RecA-independent recombination (RAIR) machinery [7] found in commonly employed laboratory bacterial strains, such as *Escherichia coli* DH5α. RAIR is thought to be dependent on XthA, a 3′-to-5′ exonuclease that resects the 3′-ends of linear DNA fragments introduced into *E. coli* cells, exposing the single-stranded 5′-overhangs [8]. Subsequently, the complementary single-stranded DNA ends hybridize with each other, and gaps are filled by DNA polymerase I and ligase LigA. Despite being originally described over three decades ago [9], the application of high-fidelity polymerases and low-cost oligonucleotides, coupled with insights into their molecular mechanisms [7,8,10–14], holds promise for expanding their widespread adoption.

Traditionally, the budding yeast *Saccharomyces cerevisiae* has served as a platform for HR-based cloning [15–17], employing a shuttle vector to propagate plasmids in both yeast and *E. coli*. However, this approach limited the utility of yeast-based cloning exclusively to these two organisms. Recent breakthroughs in yeast recombination cloning [18–22] have transcended these limitations, allowing the technique to be applied universally for cloning multiple fragments of interest into any vector. The key innovation involves incorporating an origin of replication and a selection marker for yeast as a cassette or DNA fragment, which is co-inserted with other fragments for cloning. Assembled plasmids can then be selected in and purified from yeast, ready for transformation or transfection into desired host organisms. HR-based cloning in yeast can also be adapted for high-throughput DNA assembly and is amenable to automation [23].

While *S. cerevisiae* exhibits efficient HR machinery, its NHEJ activity is limited. By leveraging previous findings that the thermotolerant yeast *Kluveromyces marxianus* pos-

sesses a highly efficient NHEJ pathway [24], an NHEJ-mediated cloning method was developed [25]. This method employs a functional marker selection system (*e.g.*, *ura3*) for the cloning of DNA fragments in *K. marxianus*. NHEJ-mediated cloning exploits the sequence-independent nature of NHEJ-based joining, allowing the cloning of DNA fragments without the need for homologous ends—a departure from the HR-based cloning approach.

CReasPy-cloning (Figure 2) harnesses the precision of Cas9 to cleave DNA at a user-specified locus, synergizing with the yeast's remarkably efficient homologous recombination [26]. This innovative approach enables the simultaneous cloning and engineering of a bacterial genome in yeast. The successful application of CReasPy-cloning was demonstrated through the cloning and engineering of the 0.816 Mbp genome of *Mycoplasma pneumonia*, showcasing its potential for bacterial genome manipulation.
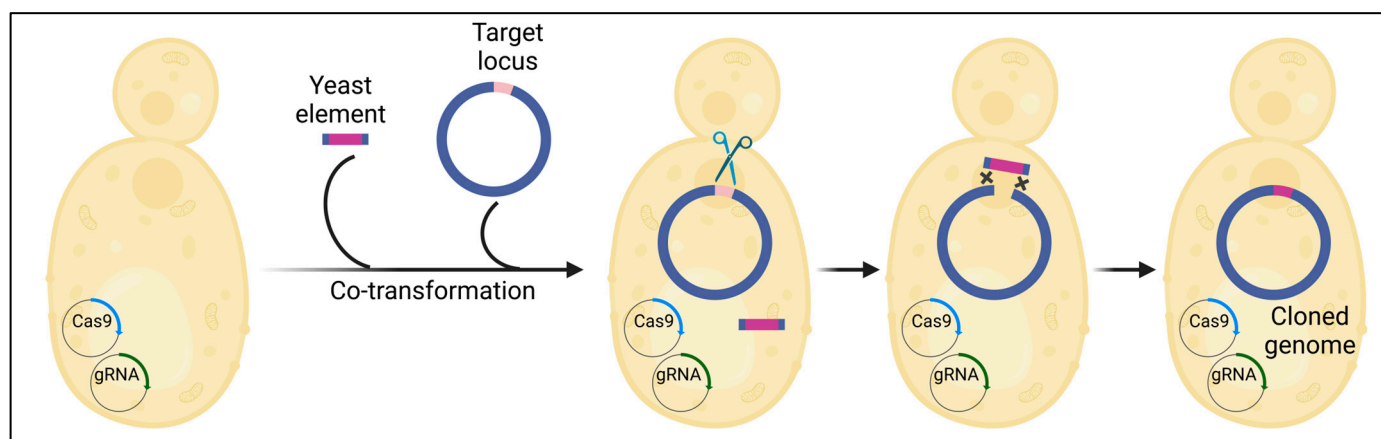

Figure 2. Illustration of the experimental procedure of CReasPy-cloning. Initially, a yeast is transformed with two plasmids—pCas9 and pgRNA—enabling the expression of the Cas9 nuclease and a guide RNA (gRNA). These plasmids carry the TRP1 and URA3 selection markers, respectively. Following this, the yeast undergoes simultaneous transformation with the target genome to be cloned and a linear DNA fragment containing yeast elements (CEN-HIS3, with or without ARS). The linear DNA fragment has recombination arms homologous to each side of the target locus. Upon entry into the cell, the Cas9/gRNA complex cleaves the target genome, and the yeast homologous recombination system repairs it using the provided linear DNA fragment as a template. Consequently, the bacterial genome incorporates the yeast elements precisely at the designated locus and is now carried by the yeast as an artificial chromosome.

Most *in vivo* cloning methods utilize bacteria or yeasts as hosts, necessitating transformation and plasmid extraction procedures for the introduction of extracted plasmids into other organisms. A recent advancement in this domain is the Phage Enzyme-Assisted In Vivo DNA Assembly (PEDA) method [27]. The simultaneous expression of phage-derived T5 DNA exonuclease and T4 DNA ligase facilitates *in vivo* DNA assembly in a diverse range of microorganisms, such as *Cupriavidus necator*, *Pseudomonas putida*, *Lactobacillus plantarum*, and *Yarrowia lipolytica*. Another cutting-edge technique, the Yeast Life Cycle (YLC) assembly method [28], capitalizes on CRISPR-Cas9 and yeast meiosis to iteratively assemble large DNA fragments. The advantage of the YLC method lies in bypassing challenging *in vitro* steps associated with handling and importing large DNA fragments into a host system.

Table 1. *In vivo* cloning methods developed in the past decade (**HR**, homologous recombination; **NHEJ**, non-homologous end joining).

In Vivo Cloning Method	Method Description	Cloning Accuracy *	Maximum Number of Fragments Assembled at Once *	Reference(s)
Bacterial *in vivo* cloning (HR-based)	Depends on RecA-independent recombination (RAIR) pathway. DNA fragments generated via PCR or restriction enzyme digestion with overlapping homologous sequences can be used to directly transform bacteria.	Close to 100% for a 2-fragment assembly. Efficiency decreases with increasing number of fragments assembled at once.	Up to 7 across the references cited	[7,8,10–17]
Yeast *in vivo* cloning (HR-based)	Highly efficient HR pathway in yeast (*S. cerevisiae*) can be used for assembling multiple DNA fragments with homologous sequences directing the order of assembly.	>95%, even for a 9-fragment assembly	4–9 across the references cited	[19–22]
Yeast *in vivo* cloning (NHEJ-based)	Highly efficient NHEJ pathway in thermotolerant yeast (*K. marxianus*) allows direct joining of two DNA fragments without the need for homologous ends.	97–100% for a 2-fragment assembly	2	[25]
CReasPy-cloning	Combines ability of CRISPR-Cas9 and HR pathway of yeast to clone and edit large genomes at multiple loci.	>50% for a 2-fragment assembly	4	[26]
Phage Enzyme-Assisted In Vivo DNA Assembly (PEDA) method	Simultaneous expression of exonuclease and ligase allows *in vivo* cloning in a wide range of microorganisms.	70–100% for a 2-fragment assembly, 71–100% for a 3-fragment assembly	3	[27]
Yeast Life Cycle (YLC) assembly method	Combines CRISPR-Cas9 and meiosis of yeast to iteratively assemble large DNA fragments.	67–100%	2 DNA fragments assembled per life cycle	[28]

* For clarity purposes, a gene cloned into a linearized vector is considered a 2-fragment assembly.

2.2. In Vitro Cloning

Gibson assembly, Seamless Ligation Cloning Extract (SLiCE), Sequence- and Ligation-Independent Cloning (SLIC), and In-Fusion cloning are all *in vitro* techniques. Their underlying principle relies on the presence of complementary overhangs at the ends of vector and insert DNA fragments. These methods have gained widespread popularity due to their ability to overcome limitations associated with restriction endonuclease-based cloning. Recent advancements have integrated the flexibility of Gibson assembly with the precision of CRISPR-Cas9, allowing for the targeting of double-strand breaks at any desired location. This is particularly advantageous in situations where the PCR amplification of vector fragments is challenging and restriction endonuclease sites are unavailable. Guide RNAs guide the Cas9 endonuclease to target any double-stranded DNA sequence, effectively linearizing the vector. The linearized vector can then be employed in Gibson assembly or other cloning techniques [29,30]. CRISPR-Cas9 has also been utilized to excise large DNA fragments (~100 kb) from bacterial chromosomes. Creating such lengthy fragments through PCR amplification could be challenging. The cut fragments are subsequently cloned using Gibson assembly [31]. Variations of the SLiCE technique, such as Zero-Background Redα (ZeBRα), leverage the *ccdB* gene to eliminate any background from uncut or re-ligated vectors during cloning [32].

Other recent advances, such as T5 Exonuclease-Dependent Assembly (TEDA), Single 3′-exonuclease-based multifragment DNA assembly (SENAX), and T5 exonuclease-mediated low-temperature DNA cloning (TLTC) [6,33,34], operate on the principle of exonuclease-generated overhangs within the overlapping regions of vectors and DNA fragments. These methods offer a cost-effective alternative to Gibson assembly, as they only require the exonuclease enzyme for the *in vitro* reaction to create the necessary complementary overhangs. The subsequent steps of gap repair and DNA fragment ligation take place *in vivo* post-transformation into bacterial cells.

Alternatives to generating compatible overhangs include techniques such as Uracil-Specific Excision Reagent (USER) Cloning. In USER Cloning, vector and insert fragments with short overlapping regions are produced through PCR. The primers used in this process introduce a single deoxyuracil (dU) residue, subsequently cleaved by the USER enzyme. This cleavage results in 3′ overhangs, facilitating the annealing of multiple DNA fragments. Although USER Cloning has a longstanding history [35], advancements in the past decade, such as optimizing the melting temperature of annealing DNA fragments [36] and employing *in silico* design tools like AMUSER [37], have enhanced its efficacy. A similar technique that introduces compatible overhangs through PCR is QuickStep-Cloning [38], where two parallel asymmetrical PCR reactions generate overhangs. An improvement on this method is PTO-QuickStep cloning [39,40], where phosphorothioate (PTO) bonds, introduced via primers, are processed by iodine cleavage to generate overhangs (Figure 3). Nicks in the plasmids resulting from both USER Cloning and QuickStep-Cloning are sealed following transformation into bacterial cells. For scarless and sequence-independent DNA assembly, the method using thermostable exonuclease and ligase (DATEL) [41,42] presents another alternative. This technique employs a combination of Taq and Pfu DNA polymerases to cleave single-stranded DNA flaps generated during the annealing of DNA fragments with overlaps. The resulting nicks are then ligated using a heat-stable DNA ligase.
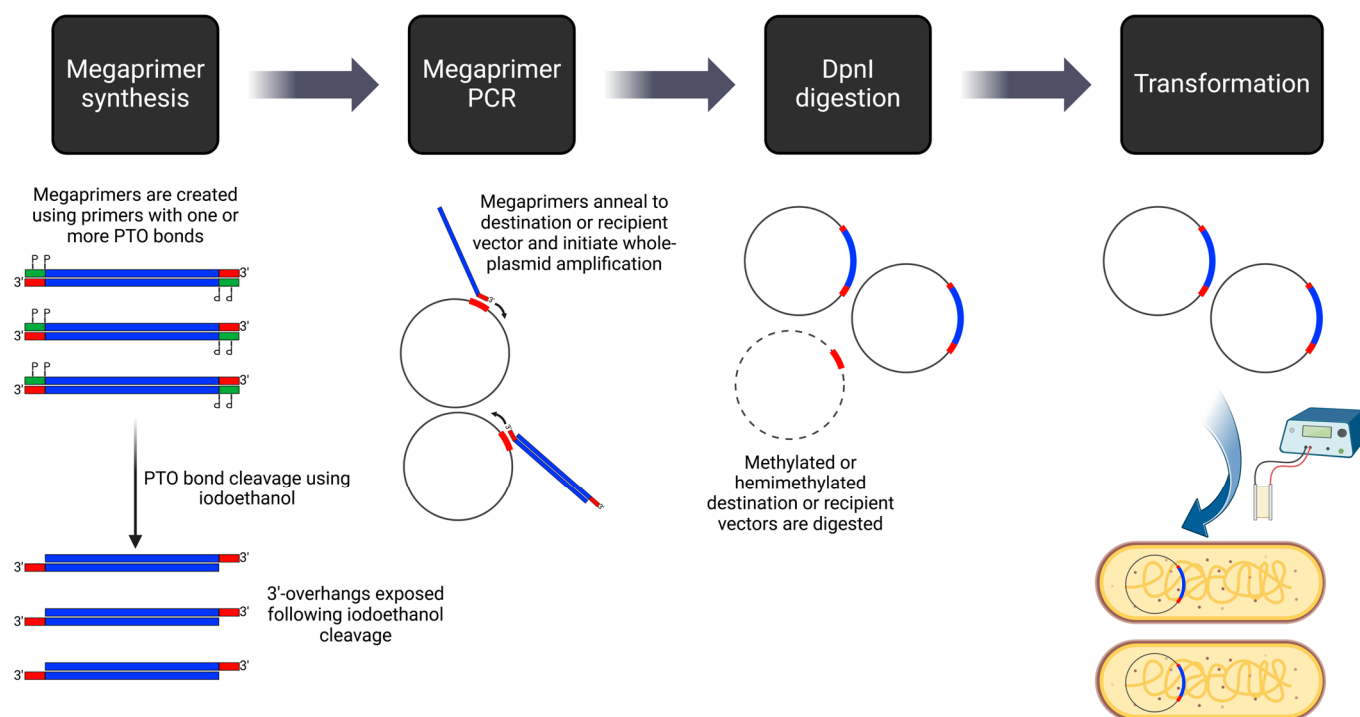

Figure 3. An overview of the PTO-QuickStep method. Initially, megaprimers (coloured in blue) are generated in a PCR using a set of PTO oligonucleotides containing phosphorothioate linkages (indicated with letters 'P'). Subsequently, iodoethanol treatment is applied to the megaprimers, breaking the phosphorothioate linkages and exposing 3′-overhangs. In the second step, these treated megaprimers anneal to the destination or recipient vector at the target locus, initiating the amplification of the entire

plasmid. Moving to the third step, DpnI is employed to remove the methylated or hemimethylated destination or recipient vector without the gene insert (shown as dotted circle). After DpnI digestion, the newly synthesized plasmids undergo transformation into *E. coli*, where any nicks are repaired *in vivo*.

In summary, the fundamental principle underlying the latest advancements in *in vitro* DNA assembly (Table 2) is to streamline the process, ensuring simplicity, cost-effectiveness, and versatility. Leveraging overlapping DNA sequences among fragments to be assembled allows for the insertion of any DNA fragment into any chosen vector without being confined by the limitations of restriction endonuclease sites. While each discussed technique presents its unique advantages and limitations, the optimal choice for a particular DNA assembly project will depend on specific requirements and contextual factors.

Table 2. *In vitro* cloning methods developed in the past decade.

In Vitro Cloning Method	Method Description	Cloning Accuracy *	Maximum Number of Fragments Assembled at Once *	Reference(s)
T5 Exonuclease-Dependent Assembly (TEDA)		70–99% for a 2-fragment assembly, depending on competent cell preparation	4	[6]
Single 3′-exonuclease-based multifragment DNA assembly (SENAX)	Exonuclease generates complementary overhangs, and gap repair and ligation are completed *in vivo*.	95% for a 3-fragment assembly	6	[33]
T5 exonuclease-mediated low-temperature DNA cloning (TLTC)		>95% for a 2-fragment assembly	4	[34]
Uracil-Specific Excision Reagent (USER)	Uracil-specific endonuclease generates complementary overhangs by digesting deoxyuracil introduced by primers.	~80% for a 4-fragment assembly	6	[36,37]
PTO-QuickStep cloning	Phosphorothioate bonds introduced by primers are processed by iodine cleavage to generate complementary overhangs.	95% for a 2-fragment assembly	2	[39,40]
Scarless and sequence-independent DNA assembly method using thermostable exonuclease and ligase (DATEL)	Thermostable exonuclease generates complementary overhangs, and thermostable ligase joins DNA fragments.	74–100% for assembly of 2–10 fragments	10	[41,42]

* For clarity purposes, a gene cloned into a linearized vector is considered a 2-fragment assembly.

2.3. Modular DNA Assembly

The *in vivo* and *in vitro* cloning techniques outlined above present numerous advantages compared to traditional methods relying on restriction endonucleases. However, they often involve custom-designed primers and constructs tailored for specific experiments or purposes. A paradigm shift occurs with modular DNA assembly in synthetic biology. This approach revolves around using standardized DNA 'parts' to construct complex assemblies, which, in turn, can serve as 'parts' for even more intricate constructions. Once these components are generated, sequence-verified, and validated, they become reusable and shareable among researchers. This fosters standardization, ultimately leading to significant time and cost savings in the field.

Initially, modular DNA assembly relied on type IIP restriction enzymes, leading to the development of the BioBrick system for standardization. More recent techniques have shifted to the Golden Gate cloning technology, leveraging type IIS restriction enzymes. Unlike type IIP enzymes, which recognize and cut within palindromic DNA sequences, type IIS enzymes identify non-palindromic sequences and cleave DNA outside of the recognition site. This innovation has emerged as a potent tool for seamless cloning, allowing user-defined DNA sequences to serve as cutting sites and facilitating the precise assembly of DNA fragments in a predetermined manner.

Golden Gate assembly has been embraced in the creation of two extensively utilized hierarchical modular DNA cloning methods: MoClo and Golden Braid. These methodologies employ DNA 'parts' organized in different hierarchical levels, with each level being more intricate than the one preceding it. Comprehensive reviews [43,44] provide invaluable resources for users seeking to initiate projects with these techniques. While these methods excel in seamlessly cloning multiple DNA fragments, their application has been constrained by limited vector choices—specifically, the requirement for destination vectors designed to contain one (or more) appropriately oriented pairs of recognition sites for a type IIS restriction enzyme. Recent advancements [45] have addressed this limitation by introducing type IIS restriction sites that are compatible with type IIP restriction sites, significantly broadening the range of vectors compatible with modular DNA assembly. A summary of other recently developed modular assembly techniques is provided in Table 3.

Table 3. Modular DNA assembly methods developed in the past 10 years.

Modular DNA Assembly Method	Method Description	Cloning Accuracy *	Maximum Number of Parts Assembled at Once *	Reference(s)
Biopart Assembly Standard for Idempotent Cloning (BASIC)	Makes use of reusable linkers and parts. Orthogonal oligonucleotide linkers with single-stranded overhangs are used to assemble DNA parts. Flexibility with the order of various DNA parts.	93% for a 4-part assembly with single antibiotic selection, 99.7% for a 4-part assembly with double antibiotic selection, 90% for a 7-part assembly with double antibiotic selection	7	[46]
Modular Idempotent DNA Assembly System (MIDAS)	Requires three type IIS restriction enzymes and is more complex than other modular DNA assemblies. Advantages include the ability to add new parts between existing parts rather than at the end.	86–98%	2	[47]
MetClo Assembly	Controlling methylation (which cuts or blocks the recognition site) of a single type IIS recognition enzyme allows for a simpler hierarchical DNA assembly system.	100% for a 3-part assembly	3	[48]
Start–Stop Assembly	3 bp overhangs corresponding to start and stop codons allow for scarless assembly of coding sequences.	90–100% for a 5-part assembly	5	[49,50]
PaperClip DNA Assembly	Unlike most other modular DNA assembly methods, it does not require restriction enzymes. Four oligos per DNA part allow flexible ordering of DNA parts and reuse of oligos.	100% for 5-part and 6-part assemblies	6	[51]

* For clarity purposes, a gene cloned into a linearized vector is considered a 2-part assembly.

2.4. Automated DNA Assembly

Amidst numerous breakthroughs, advancements, and refinements, a diverse array of *in vivo* and *in vitro* DNA cloning methods now cater to a spectrum of scenarios—from the straightforward insertion of a single DNA fragment into a vector to intricate assemblies involving large and/or multiple DNA fragments. As experimental designs grow increasingly complex, the norm is now the generation of hundreds, if not thousands, of library sequences. Consequently, there arises a demand for high-throughput methods capable of assembling a large number of DNA sequences to generate libraries.

The integration of automation into these processes (Table 4) offers several advantages, including the elimination of manual handling errors, enhanced reliability and reproducibility, and significant time and cost savings. This shift toward high-throughput methods represents a pivotal stride in accommodating the evolving demands of sophisticated experimental designs with large-scale library creation. Liquid-handling robots like Opentrons play a pivotal role in enabling this shift, offering benefits such as cost-effectiveness in hardware, open-source software accessibility, and user-friendly operation. These robotic systems prove invaluable for automating a diverse array of workflows, ranging from nucleic acid extraction to PCR and various DNA assembly techniques.

Table 4. Automated DNA assembly methods reported in the past decade.

Automated DNA Assembly Methods	Method Description	Cloning Accuracy *	Maximum Number of Fragments Assembled at Once *	Cost per Genetic Construct	Reference(s)
DNA assembly with BASIC on Opentrons (DNA-BOT)	Modular DNA assembly technique BASIC has been automated using robotic liquid-handling Opentrons	Not reported	10	USD1.50–USD5.50	[52]
PlasmidMaker	Combining the use of artificial restriction enzymes [*Pyrococcus furiosus* Argonaute (PfAgo)], custom software and robotic systems, an end-to-end system designed for automated DNA assembly.	18–69%, depending on the number of fragments assembled at once and fragment size	11	Not reported	[53]
AssemblyTron	Golden Gate and HR-dependent *in vivo* assemblies (IVA) or AQUA cloning are automated using Opentrons	98–100% for a 4-fragment assembly using the Golden Gate approach, 68–88% for a 4-fragment assembly using AQUA cloning, 33–50% for a 2-fragment assembly using an IVA approach.	4	Not reported	[54]

* For clarity purposes, a gene cloned into a linearized vector is considered a 2-fragment assembly.

3. Genetic Diversity Creation

The primary objective of creating genetic diversity is to generate various protein variants, with the hope that a subset of these variants, albeit often a small fraction, will exhibit favourable phenotypes compared to their parent or wildtype protein. The introduction of mutations into the GOI is a common approach in mutagenesis to achieve this diversity. The outcome of protein variants is influenced by several factors. Firstly, considering that there are 61 codons encoding 20 amino acids, altering the codon usage or the GC content of the DNA sequence can impact the mutagenesis outcome. This effect becomes particu-

larly prominent when the mutational spectrum of the employed mutagenesis method is highly biased. Secondly, the organization of the genetic code imposes constraints on the mutagenesis outcome.

While protein engineers predominantly work with DNA, it is worth noting that RNA mutagenesis methods also exist. Fukuda *et al.* developed an RNA mutagenesis method that leverages the intracellular RNA-editing mechanism [55]. In this approach, guide RNAs guide the editing enzyme, human adenosine deaminase acting on RNA (ADAR), inducing adenosine-to-inosine (A-to-I) mutations on RNA molecules. Other RNA-based mutagenesis strategies involve using Qβ replicase to generate complex mRNA libraries [56]. However, compared to DNA, RNA is a transient molecule, making it more challenging to track genotype–phenotype linkage. The conversion of RNA to cDNA is typically required to identify mutations introduced in RNA. In the subsequent sections, our focus will primarily be on DNA mutagenesis methods.

3.1. Quality and Size of a Gene Library

It is strongly advisable to assess the quality of a gene library, especially when employing a mutagenesis method for the first time on the target gene sequence. This evaluation is commonly achieved by sequencing a small subset, typically ranging from 10 to 20 randomly selected 'variants'. If the gene library's quality is found to be suboptimal, proceeding to the expression of the protein library and subsequent selection/screening may not be prudent. The probability of identifying enhanced protein variants would be too minimal to justify the resources invested in such an undertaking.

Several key indicators are commonly employed to evaluate the quality of a gene library. Taking random DNA mutagenesis as an example, a high-quality gene library should exhibit the following characteristics:

- Mutations are precisely targeted to the GOI (*i.e.*, no off-target mutations).
- Mutations are uniformly distributed along the entire GOI.
- All bases (A/T/G/C) experience mutations at the same frequency and are substituted with their three counterparts equally.
- The mutation frequency (number of errors per 1 kb of DNA) is not excessively high, preventing the predominance of non-functional protein variants.
- Duplicated sequences are avoided/eliminated.
- Wildtype sequences are absent in the gene library (*i.e.*, no template carry-over).

Achieving an absolutely ideal gene library is practically impossible unless it exclusively comprises synthetic genes. The extent of deviation from this ideal serves as a practical measure of the gene library's quality. The Schwaneberg group conducted a comprehensive large-scale comparison among random mutagenesis libraries created using three approaches: error-prone polymerase chain reaction (epPCR) with low mutagenic conditions, epPCR with high mutagenic conditions, and Sequence Saturation Mutagenesis (SeSaM) [57–59]. After sequencing 1000 mutations for each library, the library quality was evaluated on both the DNA and protein levels [60,61]. The protein level assessment primarily employed the Mutagenesis Assistant Program (MAP) [62–64]. The SeSaM library exhibits a preference for transversion mutations, contrasting with epPCR libraries that display a transition bias. Additionally, the SeSaM library demonstrates a significantly higher number of consecutive nucleotide substitutions. These characteristics result in a greater number and more diverse amino acid substitutions in the SeSaM library.

In directed evolution, preserving genotype–phenotype linkage is crucial for tracing an improved phenotype back to its genetic origins. During the library transformation process, transformants incorporating more than one plasmid could account for over 20% of the constructed library, thereby compromising the library quality. To mitigate the occurrence of multiple-plasmid transformants, it is possible to reduce their frequency by optimizing the amount of plasmid DNA used for transformation [65].

The maximum library size of a mutagenesis method is determined by the transformation efficiency of the microbial host responsible for expressing the protein library. Table 5

provides a summary of the typical transformation efficiencies for commonly used microorganisms in biomanufacturing and protein evolution. Taking *E. coli* as an illustration, the achievable library size typically falls within the range of 10^9–10^{10}. For organisms with low transformation efficiencies, opting for an *in vivo* method for library creation is recommended. The ability to thoroughly explore this library of variants hinges on the capacity or throughput of the screening/selection method employed. The combination of a larger-sized and higher-quality library enables more efficient exploration of the protein space, particularly when the exploration is not constrained by screening/selection capacity. This review deliberately omits the discussion of screening or selection as it has been extensively covered in our textbook [5] and several recent, excellent reviews [66–71]. Readers are encouraged to consult these references for a comprehensive understanding of this aspect.

Table 5. Transformation efficiency of commonly used microorganisms in biomanufacturing and protein evolution.

Organism (Year Published)	Strain Used	Transformation Method	Transformation Efficiency (CFU/µg)	Ref.
E. coli (2022)	BW3KD	Chemical	$7.2 \pm 1.9 \times 10^9$	[72]
E. coli (2022)	BW25113	Chemical	$>10^9$	[73]
E. coli (2020)	TG1	Electroporation	$>8 \times 10^{10}$	[74]
E. coli (2018)	DH5α, JM109, TOP10	Chemical [1]	$3.1 \pm 0.3 \times 10^9$	[75]
S. cerevisiae (2023)	EBY100	Electroporation	10^8	[76]
S. cerevisiae (2016)	MaV203	Chemical [2]	1×10^7	[77]
K. pastoris (2021)	X-33	Magnetic nanoparticle-mediated gene delivery	1.3×10^3	[78]
K. pastoris (2020)	X-33	Chemical	6.52×10^3	[79]
K. pastoris (2019)	PPY12h	Electroporation	$1.3 \pm 0.1 \times 10^3$	[80]
K. marxianus (2021)	FIM-1ΔU	Protoplast	1.8×10^4	[81]

[1] Applied antimicrobial peptide LFcin-B to increase the permeability of the cell membrane and high concentration of Ca^{2+} and Mn^{2+} to suppress lethal antimicrobial properties. [2] Applied nutrient supplement to boost transformation efficiency.

3.2. Random Mutagenesis

Random mutagenesis continues to be a frequently employed method, particularly in cases where a structure–function relationship is lacking. In the following section, we will explore recent advancements in five key areas: (1) epPCR, (2) *in vivo* mutagenesis in *E. coli*, (3) random base editing, (4) virus-assisted mutagenesis, and (5) random insertion and deletion (InDel).

3.2.1. epPCR

Motivated by the technical simplicity of epPCR, several modified protocols have been developed. One adaptation is a modified epPCR protocol specifically optimized for small amplicons [82]. Another variation, known as Casting epPCR (cepPCR), directs random mutations to a specific region within the target GOI [82]. Furthermore, efforts are underway to extend epPCR methods beyond the conventional microbial hosts *E. coli* and *S. cerevisiae*, with attempts to optimize them for other microbial hosts. For example, in the transformation of *Bacillus subtilis* into a host for directed evolution, the epPCR product was fused with flanking regions and an antibiotic-resistant marker. This composite PCR product was then integrated into the chromosome through homologous recombination after transformation into the supercompetent cells of the *B. subtilis* strain SCK6 [83]. Additionally, a random mutagenesis protocol has been devised for the methylotrophic yeast *Pichia pastoris*

(Figure 4) [84]. This protocol involves the sequential amplification of plasmids using Phi29 DNA polymerase, encompassing error-prone rolling circle amplification (RCA) followed by multiple displacement amplification (MDA). Through these steps, it becomes feasible to obtain microgram amounts of plasmids for subsequent electroporation into *Pichia* cells, addressing a key challenge in employing *Pichia* for directed evolution.
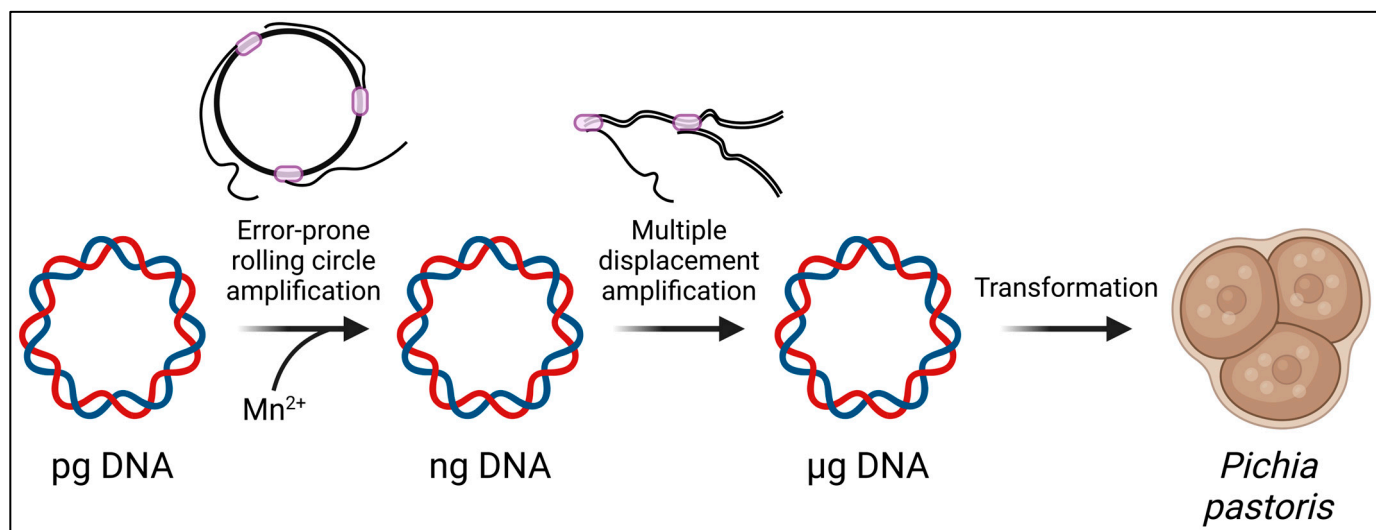

Figure 4. A schematic representation of the random mutagenesis method tailored for *Pichia pastoris*. Initially, the circular protein expression vector undergoes repeated amplification through strand displacement reactions facilitated by Phi29 DNA polymerase. Mutations are intentionally introduced by adding Mn^{2+} to lower the fidelity of the polymerase, a process known as error-prone rolling circle amplification (epRCA). Following this, subsequent amplification, achieved through Phi29 DNA polymerase (or multiple displacement amplification, MDA), yields microgram quantities of mutated DNA. This mutated DNA is then utilized for transformation into *P. pastoris* to enable enzyme production.

3.2.2. *In Vivo* Mutagenesis in *E. coli*

In vivo mutagenesis provides several distinct advantages over *in vitro* mutagenesis. This approach allows the integration of genetic diversity creation and selection, eliminating the need for a separate transformation step that often limits the size of gene libraries in *in vitro* mutagenesis. Additionally, it avoids the labour-intensive process associated with *in vitro* gene library creation.

Liu and colleague developed highly effective, inducible, broad-spectrum mutagenesis systems in *E. coli*, elevating mutation rates by over 320,000 times compared to basal levels [85]. These plasmid systems rely on the induced and combinatorial expression of proteins associated with proofreading (*dnaQ926*), translesion synthesis (*umuD'*, *umuC*, *reacA730*), mismatch repair (*dam*, *seqA*), base excision (*ugi*, *cda1*), and base selection (*emrR*). Although efficient, this global mutagenesis strategy could introduce extensive mutations throughout the genome of the host, leading to undesirable issues such as toxicity, a reduced library size, the silencing of mutagenic plasmids, or the introduction of parasite variants into DNA libraries (mutations outside the target GOI that allow the host to circumvent the selection scheme).

MutaT7 (Figure 5) is a targeted *in vivo* mutagenesis strategy that overcomes the challenges associated with global mutagenesis strategies [86]. It employs a DNA-damaging cytidine deaminase fused to a processive T7 RNA polymerase (T7RNAP), allowing continuous directed mutations to any DNA region downstream of a T7 promoter. To enhance mutation rates, eMutaT7 replaces the cytidine deaminase from rat APOBEC1 (rApo1) with *Petromyzon marinus* cytidine deaminase (PmCDA1), resulting in an increased mutation rate from 0.34 mutations/kb/day (MutaT7) to 4 mutations/kb/day [87]. The MutaT7

toolbox has recently been expanded further with the use of adenosine deaminase-T7 RNA polymerase fusion proteins [88], such as TadA8e fused to T7RNAP and TadA7.10 fused to T7RNAP. TadA8e and TadA7.10 are variants of *E. coli* tRNA adenosine deaminase, evolved to operate on DNA [89]. Despite its utility, the MutaT7 toolkit has limitations, including a restricted mutational spectrum, strand bias, and the necessity to prevent the repair of deoxyuridine (*e.g.*, deleting uracil DNA glycosylase in the host or using a uracil DNA glycosylase inhibitor) for significant mutagenesis when using cytidine deaminase.
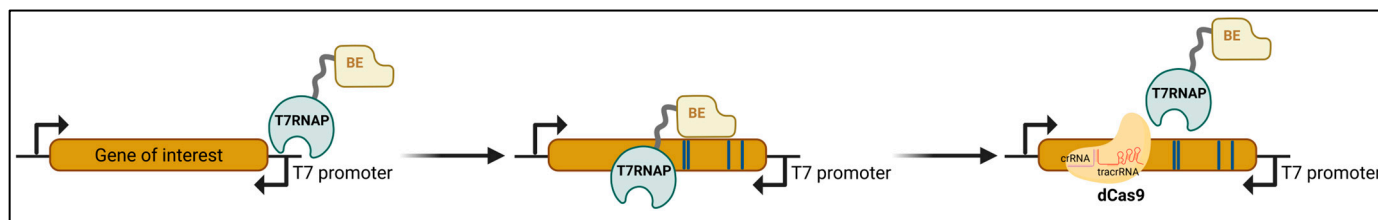
Figure 5. Graphic summary of the MutaT7 mutagenesis system and its derivatives. The T7 RNA polymerase fusion (T7RNAP) selectively binds to the T7 promoter, initiating transcription and traversing the gene of interest. As the fusion carries a base editor (BE), mutations are randomly introduced into the gene, represented by blue vertical stripes. The fusion halts and disengages from the DNA upon encountering a dCas9 molecule bound to a specific sequence dictated by the CRISPR RNA (crRNA). The termination process is also facilitated by the transactivating CRISPR RNA (tracrRNA). In the absence of dCas9, the movement of the fusion protein can be halted by incorporating one or multiple T7 terminators.

When employing a highly processive T7RNAP, mutations may extend beyond the intended target GOI. To confine the mutagenesis to a specific region, multiple copies of the T7 terminator are required, serving as a boundary to restrict mutagenesis [86]. Alternatively, in the T7-DIVA strategy [90], the catalytically dead Cas9 (dCas9), tethered to a custom-designed crRNA, acts as a 'roadblock' for the base deaminase–T7RNAP fusion proteins, effectively constraining the window of mutagenesis (Figure 5).

Instead of relying on base deaminase as a mutagenic agent and T7RNAP as a 'guide protein' (GP) to target mutations to specific loci, the EvolvR system offers continuous nucleotide diversification within a tunable window length at user-defined loci [91,92]. This is accomplished by directly generating mutations using engineered DNA polymerases (variant of *E. coli* PolI with reduced fidelity) targeted to specific loci through CRISPR-guided nickases (nCas9).

3.2.3. Random Base Editing

We introduced the use of base editors (BE, *e.g.*, cytidine deaminase and adenosine deaminase) for random mutagenesis in *E. coli* in Section 3.2.2. On first glance, this section on random base editing may seem redundant or repetitive. However, dedicating an additional section to random base editing is warranted for three reasons: (1) to discuss method variation, (2) to highlight the versatility of this approach, extending it to a wide range of organisms/cells beyond *E. coli*, and (3) to showcase its diverse applications. Indeed, the summary of methods in Table 6 emphasizes the popularity of this approach, further justifying a dedicated section on this topic.

Methods utilizing BEs for targeted random mutagenesis share a common chimeric protein design (Figure 6): a BE is tethered to a GP, with or without any accessory protein(s). The BE functions as the mutagenic agent by deaminating cytidine to uridine or adenosine to inosine. The GP directs the BE to specific gene loci for mutagenesis. The key variations among methods primarily reside in four areas: (1) the choice of BE, (2) the selection of GP, (3) the mechanisms of linkage between the GP and BE, and (4) the utilization or absence of accessory protein(s).
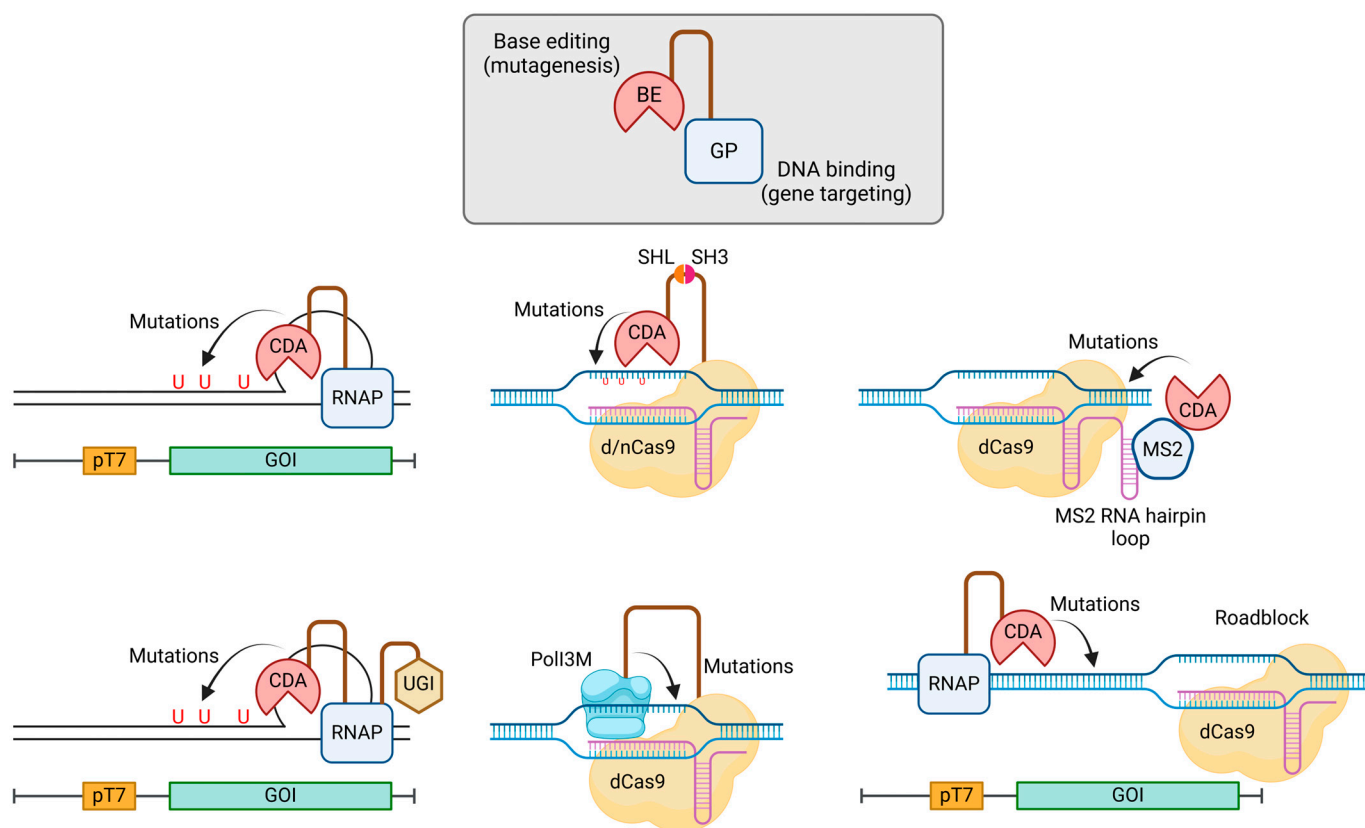
Figure 6. Targeted random mutagenesis using chimeric proteins comprising a base editor (BE) and a guide protein (GP), following the general BE-GP protein architecture. BE is the mutagenic agent, introducing random mutations through its base-editing activity (*e.g.*, cytidine and adenosine deamination). GP, with DNA-binding capability, guides or leads the BE to its target locus within the gene of interest (GOI) to effect mutagenesis. Typical BE choices include cytidine deaminase (CDA) or error-prone DNA polymerase (Pol). Frequently used GP candidates are T7 RNA polymerase (RNAP) or catalytically dead Cas9 (dCas9)/Cas9 nickase (nCas9). BE is tethered to GP via gene fusion or protein/protein or protein/RNA interactions through the utilization of SRC homology domain 3 (SH3) and the MS2 bacteriophage coat protein. In some methods, an accessory protein such as uracil DNA glycosylase inhibitor (UGI) is required.

The most commonly employed BEs include apolipoprotein B mRNA editing enzyme catalytic subunit 1 (rAPOBEC1) from rats (*Rattus norvegicus*), cytidine deaminase (PmCDA1) from sea lamprey (*Petromyzon marinus*), human activation-induced cytidine deaminase (hAID), *E. coli* tRNA adenosine deaminase (TadA), and their respective variants. CRISPR-Cas and T7RNAP represent the most popular choices for GPs. While gene fusion stands out as the most straightforward method for connecting a BE to its GP, alternative strategies have proven effective, including the utilization of the SRC homology domain 3 (SH3) and the MS2 bacteriophage coat protein. To augment base-editing efficiency, the inclusion of accessory proteins is frequently essential. Uracil DNA glycosylase inhibitor (UGI), an 83-residue protein derived from *Bacillus subtilis* bacteriophage PBS1, is extensively utilized for this purpose [93]. Additionally, the incorporation of mismatch and base excision proteins, Apn2p and Msh6p, has been shown to enhance editing efficiency.

With the BE-GP approach demonstrating growing maturity and proven success across diverse microbes and mammalian cells, we foresee ongoing developments and expect to witness adaptations of this method to additional microbial systems.

Table 6. Targeted and non-targeted (global) random base-editing methods reported in the past 10 years.

Method/First Author's Name (Year of Publication)	Guide Protein (GP)	Base Editor (BE)	Linkage between GP and BE	Organisms/Cells Validated	Ref.
Targeted mutagenesis					
Nishida *et al.* (2016)	dCas9 [1] or nCas9 [2]	PmCDA1	Gene fusion or interaction between SH3 (SRC homology domain 3) and SHL (SH3 interaction ligand)	*S. cerevisiae* and CHO	[94]
CRISPR-X (2016)	dCas9 [1]	hAID*Δ [4]	MS2 bacteriophage coat protein (MCP) binding to the MS2 RNA stem-loop	K-562 cell	[95]
TAM (2016)	dCas9 [1]	hAID, hAID CD [5], hAID P182X [6], or hAID R190X [7]	Gene fusion	K-562 cell and HEK293T cell	[96]
Komor *et al.* (2016)	dCas9 [1] or nCas9 [2]	hAID, rAPOBEC1, hAPOBEC3G, or PmCDA1	Gene fusion	U2OS cell, HEK293T cell, and HCC1954 cell	[97]
Gehrke *et al.* (2018)	nCas9 [2]	rAPOBEC1 or engineered hAPOBEC3A	Gene fusion	U2OS cell and HEK293T	[98]
TRACE (2020)	T7 RNA polymerase	rAPOBEC1 or hAID*Δ [4]	Gene fusion	HEK293T cell	[99]
TRIDENT (2021)	T7 RNA polymerase	PmCDA1 or yeTadA1.0 [8]	Gene fusion	*S. cerevisiae*	[100]
Volke *et al.* (2022)	nCas9 [2]	rAPOBEC1	Gene fusion	*P. putida* and *P. aeruginosa*	[101]
Skrekas *et al.* (2023)	dCas9 [1]	hAID*Δ [4], TadA8e [9], or TadA8e V106W [10]	Gene fusion	*S. cerevisiae*	[102]
CoMuTER (2023)	dCas3 [3]	PmCDA1 or rAPOBEC1	Gene fusion	*S. cerevisiae*	[103]
Global mutagenesis					
Pan *et al.* (2021)		rAPOBEC1	N/A	*S. cerevisiae*	[104]

[1] Catalytically dead Cas9, containing D10A and H840A mutations. [2] Cas9 nickase, containing D10A mutation. [3] Nuclease-deficient Cas3, containing H74A and D75A mutations. [4] Truncated AID, *i.e.*, AIDΔ196-198, containing K10E, T82I, and E156G mutations. [5] Truncated AID retaining only the catalytic domain, *i.e.*, AIDΔ94-198. [6] Truncated AID, *i.e.*, AIDΔ182-198. [7] Truncated AID, *i.e.*, AIDΔ190-198. [8] Variant of *E. coli* TadA, evolved in yeast. [9] Variant of *E. coli* TadA. [10] Variant of *E. coli* TadA.

3.2.4. Virus-Assisted Mutagenesis

Viruses provide a distinctive avenue for designing rapid laboratory evolution experiments, capitalizing on their inherent capacity to evolve at a much faster pace than many living organisms [105]. This accelerated evolution is facilitated by their smaller genome size, which tolerates a high frequency of mutations and a rapid rate of replication. These attributes present an excellent opportunity for the directed evolution of various biomolecules.

In the Viral Evolution of Genetically Actuating Sequences (VEGAS) method (Figure 7), the highly mutagenic RNA alphavirus Sindbis, belonging to the *Togaviridae* family and lacking known proof-reading capability, was employed to establish a mammalian-directed evolution system [106]. Estimates of RNA virus mutation frequencies range from 10^{-5} to 10^{-3} mutations per base replicated. The mutation rate of the Sindbis virus was quantified to be $1.0 \times 10^{-4} \pm 3.7 \times 10^{-5}$ mutations/base/hr. To create a robust directed evolution platform that capitalizes on the replicative and mutagenic potential of the Sindbis virus, artificial selective pressure must be applied. Each Sindbis viral particle requires 240 copies

of the structural proteins E1, E2, and capsid to form a functional viral particle capable of maturing and propagating; without this envelope, the virus is unable to mature and propagate. By engineering restrictions on structural genome transcription, it is possible to apply selective pressure to transgenic Sindbis virus carrying a GOI. The VEGAS system therefore allows for the simultaneous operation of viral mutagenesis, selection, and heredity. It is noteworthy that other reports on virus-assisted directed evolution which did not utilize viruses for creating genetic diversity are not included in this review.
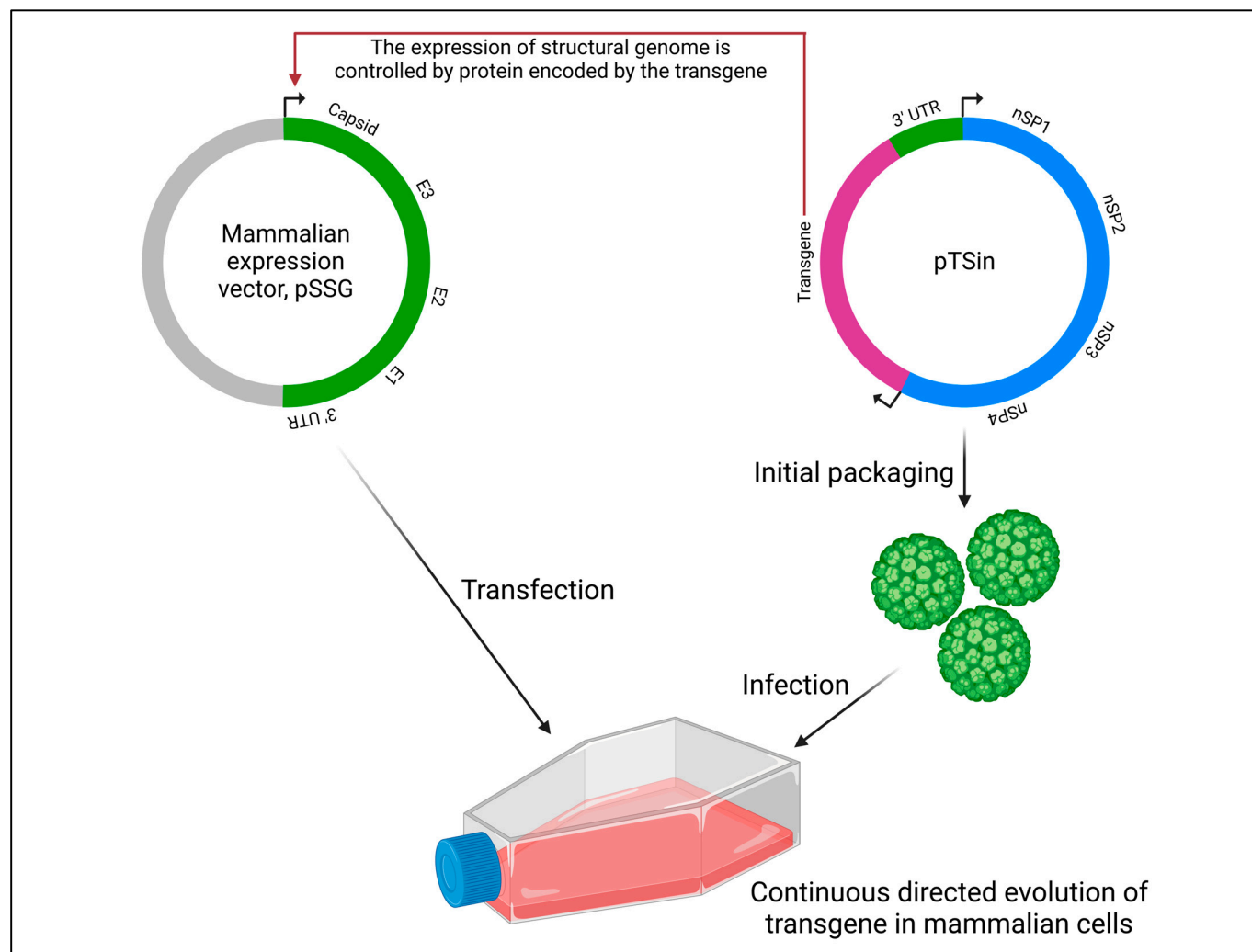
Figure 7. Schematic representation of the VEGAS platform for directed evolution, a technique for engineering DNA sequences in mammalian cells. This approach relies on the use of the Sindbis virus for efficient and mutagenic viral propagation in mammalian cell culture. To establish a robust directed evolution platform that harnesses the replicative and mutagenic potential of the Sindbis virus, artificial selective pressure must be applied. A crucial aspect involves the requirement for 240 copies of each structural protein (E1, E2, and capsid) in each Sindbis viral particle to form a functional unit capable of maturation and propagation. Without this envelope, the virus cannot mature or propagate. By strategically introducing limitations on the transcription of the structural genome, selective pressure can be applied to the transgenic Sindbis virus. In the VEGAS platform, the structural genome of the Sindbis virus is cloned into the mammalian expression vector pSSG, under the regulation of the tetracycline operator sequence. The structural genome elements of the Sindbis genome are then replaced with a transgene encoding a tetracycline transactivator. Propagation and selection can then be performed in mammalian cell culture, by infecting cells transfected with pSSG with the pTSin packaged virus.

3.2.5. Random Insertion and Deletion

Insertions and deletions in genomes occur naturally due to the replication slippage or error-prone NHEJ of double-stranded breaks. While their utilization in protein engineering is infrequent due to their tendency to be highly deleterious, often resulting in frame-shift mutations that significantly alter the protein sequence or prematurely terminate translation, there is evidence suggesting potential benefits. Instances of altered protein functionality through insertions and deletions have been reported, such as the broadening of substrate specificity in β-lactamase [107] and the modification of coenzyme specificity in Rossman fold enzymes [108]. These findings underscore the need for further exploration and investigation into the potential applications of this approach in protein engineering.

Methods for insertion and deletion in protein engineering have been developed over the last two decades, as comprehensively outlined in a recent review [109]. Some approaches leverage the higher slippage rates of certain polymerases to insert or delete one to two bases, inducing frame-shift mutations. Alternatively, other methods involve the fragmentation of GOI using DNase I or cleavage through exonucleases, endonucleases, or chemicals. Subsequently, nucleotides are added through processes like terminal deoxynucleotide transferase [110] or rolling circle amplification [111], incorporating random numbers or blocks of nucleotides. This results in libraries containing both in-frame and frame-shift variants.

Another set of methods relies on transposons to generate insertion and deletion libraries, and these approaches can prevent frame-shift occurrences through careful transposon sequence design [112–114]. Typically, transposons are designed with recognition sites for selected restriction enzymes, enabling cleavage and subsequent religation to generate the insertion or deletion of nucleotide triplets. A notable example is the recently developed TRIAD method (Figure 8) that randomly inserts or deletes one to three nucleotide triplets [115]. This method randomly inserts engineered mini-Mu transposons, defining the location of the insertion/deletion event. For deletion, the recognition site for the type II restriction enzyme MlyI is designed at both ends of the transposon. After MlyI digestion and religation, 3 bp are deleted. Repeating this process using MlyI custom cassettes enables longer deletions of up to 9 bp. For insertion, an asymmetric transposon with NotI and MlyI restriction sites is used. Following restriction digestion, a cassette carrying up to three NNN triplets can be inserted. The TRIAD method has been applied to evolve arylesterase activity in a phosphotriesterase [115] and enhance antibody affinity [116].

3.3. Focused Mutagenesis

Advancement in focused mutagenesis is marked by five noteworthy trends. These trends encompass (1) a transition from site-directed mutagenesis to multi-site-directed mutagenesis and even massive mutagenesis, (2) the incorporation of CRISPR/Cas9 in focused mutagenesis, mirroring trends seen in cloning and random mutagenesis, (3) the formulation of strategies to minimize library size, (4) the development of computational tools for automated oligo design, and (5) a shift from column-synthesized oligos to microarray-synthesized oligos for protein engineering. The subsequent sections will discuss each of these trends individually.
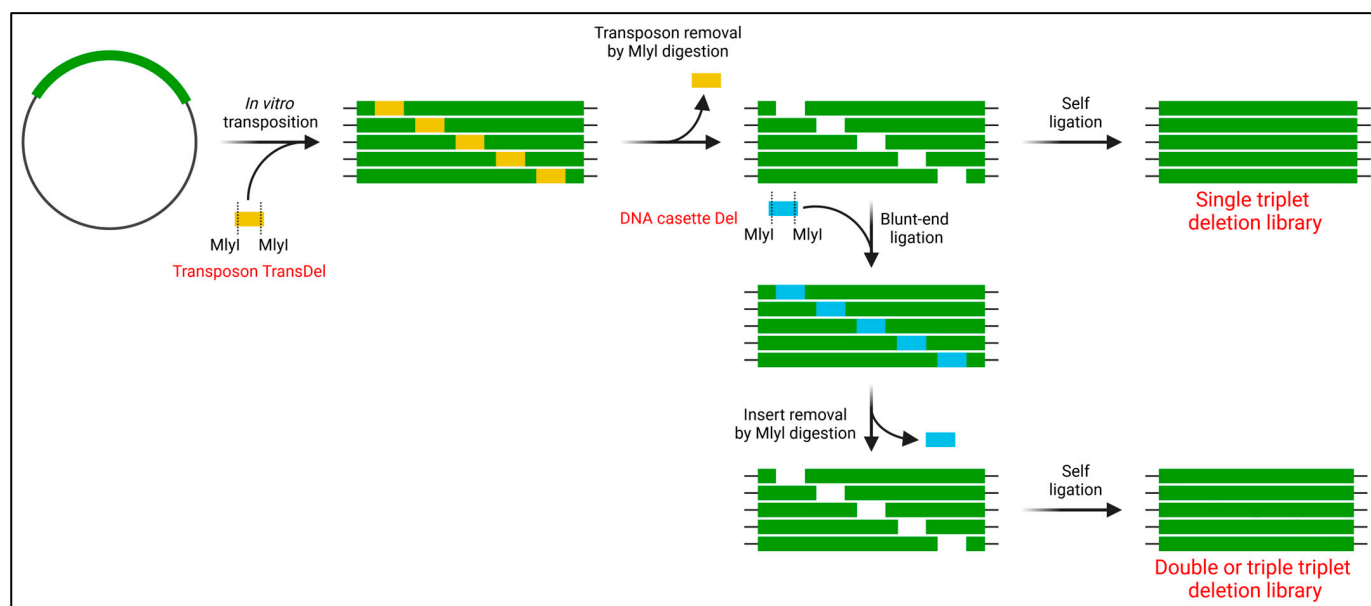
Figure 8. Schematic representation of the TRIAD process for generating deletion libraries. In the first step, the TransDel insertion library is formed through *in vitro* transposition of the engineered transposon TransDel into the target sequence on circular plasmid DNA. In the second step, MlyI digestion is applied to eliminate TransDel along with 3 base pairs of the original target sequence, creating a single break per variant. The third step involves self-ligation, leading to the reconstitution of the target sequence minus 3 base pairs. This results in a library of single variants, each featuring a deletion of one triplet. Alternatively, DNA cassette Dels can be inserted between the breaks in the target sequence to produce insertion libraries. MlyI digestion removes the DNA cassette Del along with 3 or 6 additional base pairs of the original target sequence, depending on the used DNA cassette. Subsequent self-ligation reestablishes the target sequence, now with a deletion of 2 or 3 triplets. This versatile approach also allows for the creation of insertion libraries.

3.3.1. Multi-Site-Directed Mutagenesis

Significant progress in multi-site-directed mutagenesis has predominantly emerged in methodologies utilizing single-stranded DNA (ssDNA) templates, as summarized in Table 7. A key example is nicking mutagenesis (Figure 9A) [117], leveraging a nicking enzyme (Nt.BbvCI) and exonucleases (ExoIII, ExoI) for the preparation of ssDNA templates. This approach involves a pool of mutagenic primers defining targeted mutagenesis sites, which anneal to the ssDNA and undergo isothermal assembly. Following template strand removal, the complementary mutant strand is synthesized via PCR. Continuous protocol enhancements have been reported to bolster efficiency (Table 7). In contrast, the SLUPT method generates an ssDNA template by eliminating the phosphorylated strand in the linearly amplified GOI [118].

Another group of methods opts for the creation of gene fragments bearing the desired mutations, which are subsequently combined in a sequential assembly [*e.g.*, Combinatorial Codon Mutagenesis (Figure 9B)] using techniques like megaprimer PCR, overlap extension PCR, or Golden Gate assembly. Drawing inspiration from genome recombineering methods such as MAGE, Higgins *et al.* devised a plasmid recombineering method for *in vivo* multi-site-directed mutagenesis (Figure 9C) [119].

Another development involves leveraging solid-phase gene synthesis technologies to design libraries [120,121]. This approach minimizes the risk of stop codons and unintended mutations during PCR, thereby reducing screening efforts [120]. Notably, Öling *et al.* achieved an impressive 161 multi-site mutations using this methodology [122].
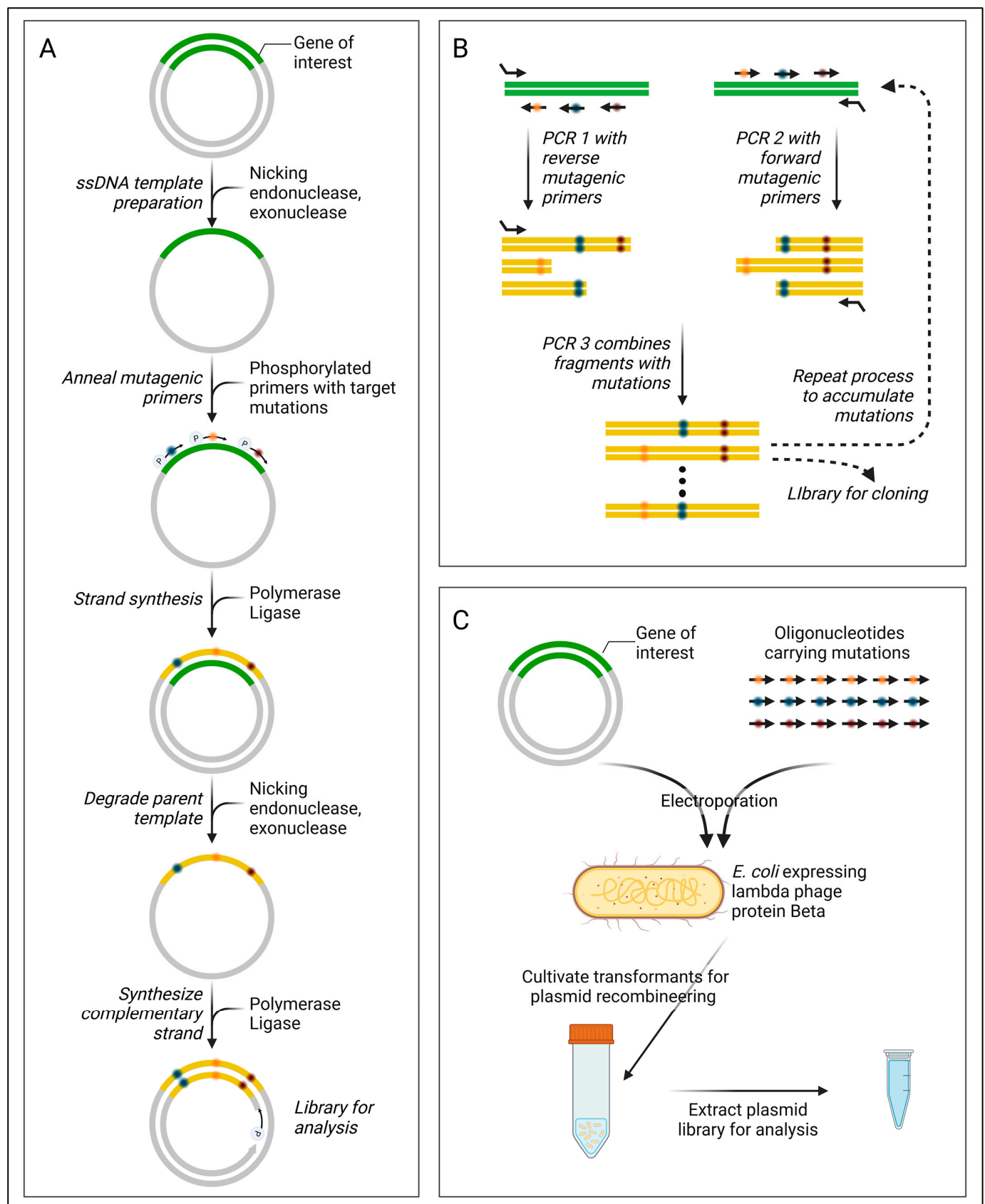
Figure 9. Multi-site-directed mutagenesis methods. (**A**) Nicking mutagenesis: This process begins with the wildtype (WT) plasmid dsDNA containing a 7-base-pair BbvCI recognition site, which is

selectively nicked by Nt.BbvCI. The resulting nicked strand undergoes degradation by Exonuclease III (ExoIII), creating a single-stranded DNA (ssDNA) template. To eliminate insufficiently digested DNA, Exonuclease I (ExoI) is employed. Next, phosphorylated mutagenic primers are annealed to the ssDNA parental template. The mutagenic strand is then synthesized through the collaborative action of a polymerase and ligase. Following this synthesis, the WT template strand is nicked by Nb.BbvCI and subsequently digested by ExoIII. The introduction of a second primer initiates the synthesis of the complementary mutant strand, resulting in the generation of mutagenized dsDNA. (**B**) Combinatorial codon mutagenesis: This process initiates with two parallel PCR reactions. In one, mutagenic reverse primers are employed, while in the other, mutagenic forward primers are utilized. In the subsequent step, a third PCR is employed to combine the fragments generated from the preceding PCR reactions. (**C**) Plasmid recombineering: This method involves the direct *in vivo* incorporation of synthetic oligonucleotides carrying desired mutations into a gene of interest. These oligonucleotides are introduced into *E. coli* cells via electroporation and can recombine with resident plasmids, facilitated by the lambda phage protein Beta.

Table 7. Multi-site-directed mutagenesis methods reported in the last decade.

Method/First Author (Year of Publication)	Method Description and Mutagenic Agent	Template	Maximum Mutated Sites per Gene	Ref.
ssDNA template				
Nicking mutagenesis (2016)	• ssDNA template generated with nicking enzyme (Nt.BbvCI) and exonucleases (ExoIII, ExoI). • Phosphorylated mutagenic primers annealed to template and extended by PCR using Phusion polymerase and Taq DNA ligase. • Template ssDNA removed with nicking enzyme (Nb.BbvCI) and exonucleases. • Complementary mutant strand synthesized by PCR.	Plasmid ssDNA	7	[117]
Darwin Assembly (2018)	• ssDNA template created with nicking enzyme (Nt.BbvCI) and exonuclease. • Phosphorylated mutagenic primers annealed to template and extended by PCR using Q5 polymerase and Taq DNA ligase. • Outnest primers used for PCR amplification of assembled fragment. • Template removal may or may not be required depending on complexity of construct. • Library cloned into vector using Golden Gate.	Plasmid ssDNA	19	[123]
Nicking mutagenesis (2021)	• Modified from nicking mutagenesis (2016). • Improved library coverage by using longer mutagenic primers and quick annealing conditions	Plasmid ssDNA	15	[124]

Table 7. *Cont.*

Method/First Author (Year of Publication)	Method Description and Mutagenic Agent	Template	Maximum Mutated Sites per Gene	Ref.
SLUPT (2021)	• GOI amplification with dUTP and one phosphorylated primer. • ssDNA template created by degradation of phosphorylated strand using lambda exonuclease. • Phosphorylated mutagenic primers annealed to template and extended by PCR using Phusion polymerase and Taq DNA ligase. • Template ssDNA removal using uracil DNA glycosylase. • Library cloned into vector by restriction digestion and ligation.	Linear GOI ssDNA	17	[118]
SUNi mutagenesis (2023)	• Derived from nicking mutagenesis. • Longer mutagenic primers with a 5′-G/C to improve efficiency. • Codon design at mutagenic site to minimize parent.	Plasmid ssDNA	n.r *	[125]
Assembly of mutated gene fragments				
Combinatorial codon mutagenesis (CCM) (2014, 2017)	• Megaprimers generated by one-pot amplification with mutagenic primers and an end primer. • Random combinations of mutations generated by PCR reassembly of GOI with mutant megaprimers and two end primers. • Template removed through gel extraction of assembled GOI variants. • Assembled GOI library used as template for further rounds of megaprimer generation and reassembly to accumulate mutations per gene. • Library cloned into vector by restriction digestion and ligation.	Plasmid dsDNA (2014) Linear GOI dsDNA (2017)	7	[126,127]
Chung *et al.* (2017)	• Megaprimers generated by one-pot or multi-pot amplification with mutagenic primers and an end primer. • Random combinations of mutations generated by PCR reassembly of GOI with mutant megaprimers and two end primers. • Library cloned by Gibson assembly.	Linear GOI dsDNA	14	[128]
Golden mutagenesis (2019)	• Fragments generated using mutagenic primers with Type IIS recognition sites. • Template removed by gel extraction of fragments. • Fragments assembled with vector via Golden Gate. • Web-tool for primer design and sequencing result analysis.	Plasmid dsDNA	5	[129]

Table 7. *Cont.*

Method/First Author (Year of Publication)	Method Description and Mutagenic Agent	Template	Maximum Mutated Sites per Gene	Ref.
Hejlesen *et al.* (2020)	• Fragments with overlapping regions generated by PCR amplification with mutagenic primers. • Prolonged overlap extension PCR created linear repetitive multimers of the template. • Linear multimers transformed and circularized in bacteria.	Plasmid dsDNA	3	[130]
In vivo method				
Plasmid recombineering (2017)	• *In vivo* method. • Plasmid and mutagenic primers electroporated into a recombineering strain. • Repeat recombineering using variant plasmids as template to increase sites of mutations.	Plasmid dsDNA	2	[119]

* n.r. not reported.

3.3.2. CRISPR/Cas9-Mediated Mutagenesis

She *et al.* introduced a PCR-free, two-step *In vitro* CRISPR/Cas9-mediated Mutagenic (ICM) system designed for both single-site- and multi-site-directed mutagenesis (Figure 10) [131]. In the first step, site-specific plasmid digestion is achieved by employing a complex of Cas9 with specific single guide RNA (sgRNA), followed by degradation with T5 exonuclease to create a 15-nucleotide complementary overhang. Subsequently, in step 2, primers containing the desired mutations are annealed to generate double-stranded DNA fragments, which are then ligated into the linearized plasmid. A distinct advantage of employing a PCR-free approach is the attainment of greater genetic diversity, attributed to the elimination of biases introduced by PCR, such as preferential primer binding to the template. Anticipating ongoing advancements, we foresee more methodologies integrating the use of CRISPR/Cas9 in the near future.
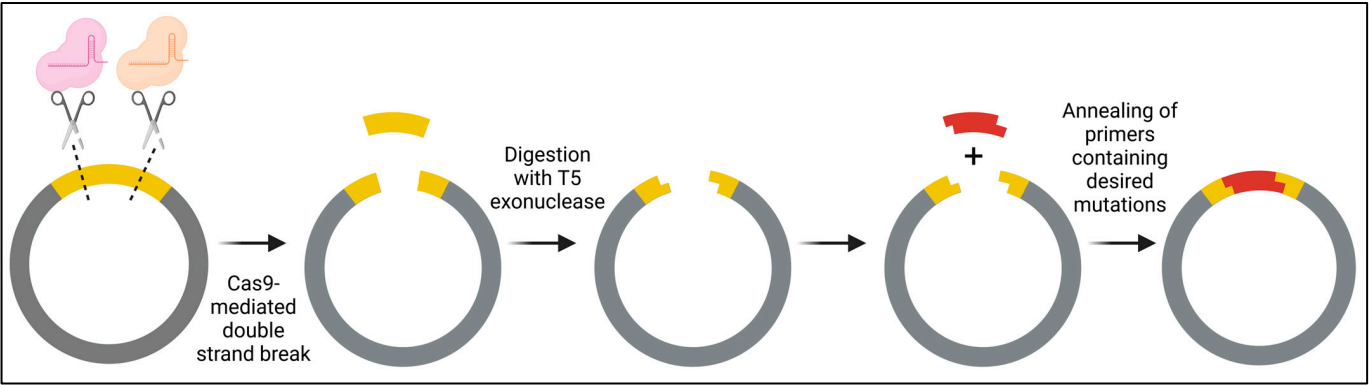

Figure 10. Schematic representation of the *In vitro* CRISPR/Cas9-mediated Mutagenic (ICM) system for site-directed mutagenesis: The target plasmid undergoes initial cleavage by the Cas9 protein and specific sgRNA complex at both sides of the mutational position, removing the wildtype sequence. Subsequent digestion with T5 exonuclease generates 15 nt sticky ends. Primers carrying the desired mutation are then annealed to create DNA fragments with 15 nt sticky ends that complement the digested plasmid. These fragments are inserted into the linearized vector through transformation into the host cell.

3.3.3. The Numbers Game

In focused mutagenesis, with the escalation of target sites, the library size (number of permutations) undergoes exponential expansion. When combined with oversampling to ensure a certain degree of diversity coverage, this exerts significant strain on library screening, making the navigation of the entire diversity space practically challenging—a scenario often denoted as permutation explosion. Commonly, three strategies are employed to tackle this issue: the divide-and-conquer approach, the codon randomization strategy, and the use of machine learning (ML).

GeneORator [132] employs a divide-and-conquer approach. In this method, target sites are partitioned into subsets, such as subsets A and B. A permutation library is established for each subset, generating libraries A and B. Both libraries A and B are then independently screened for the desired phenotype. Another notable divide-and-conquer strategy is iterative saturation mutagenesis (ISM) pioneered by Reetz [133], taking the approach one step further. Beneficial variants from each subset act as templates to randomize the other subset, leading to the creation of new libraries, AB or BA, for the subsequent round of screening. A potential limitation of the divide-and-conquer approach is that it may not capture some pairwise and higher-order effects, known as epistasis, between residues in the gene libraries.

M-ISM, a variant of the ISM method [134], employs a computational tool to design a 'small-intelligent' library for each subset. The term 'small-intelligent' denotes a minimal gene library size devoid of inherent amino acid biases, stop codons, or rare codons specific to the protein expression host [135]. Rather than sampling each target site with 19 amino acid alphabets, it is feasible to obtain improved protein variants using reduced amino acid alphabets, potentially even with just one or two alphabets [136].

Degenerate Codon Optimization for Informed Libraries (DeCOIL) is a recently introduced ML method that optimizes degenerate codon libraries to sample protein variants likely to exhibit both high fitness and high diversity within the sequence search space [137]. Like all ML approaches, DeCOIL necessitates a training library to develop its ML model.

3.3.4. Automated Oligo Design

Mutation Maker [138] and GeneGenie [139] serve as oligo design platforms for protein engineering. Taking Mutation Maker as an example, the platform designs oligos for various applications, including site-scanning saturation mutagenesis, multi-site-directed mutagenesis, and PCR-based accurate gene synthesis. The designed oligos can be assembled through overlap extension PCR, with the program optimizing melting temperature (T_m) and codon choice. The use of degenerate oligos remains popular in focused mutagenesis. Web-based applications like CodonGenie [140] facilitate the selection of degenerate codons. For high-throughput or massively parallel protein engineering, automated oligo design proves useful in minimizing errors and expediting experimental design.

3.3.5. Oligo Pool (oPool) for Cost-Effective Library Construction

Rather than relying on oligonucleotides with degenerate codons synthesized through solid-phase phosphoramidite chemistry in column-based synthesis, there is a growing preference for using microarray-synthesized oligonucleotides or oligo pools (oPools) [141]. oPools are mixes of thousands of individually designed polynucleotides, each up to 350 bases in length.

Individually designed oligos offer several distinct advantages, including the removal of redundancy from the gene library, the extensive coverage of mutational space, the precise selection of codons optimized for protein expression in a chosen host, and the facilitation of studies such as deep mutational scanning (DMS) [142] that necessitate large oligonucleotide pools. Programmed Allelic Series (PALS) is a technique devised for massively parallel single amino acid mutagenesis [143]. It achieves this by integrating a low-cost oPool with overlap extension mutagenesis. oPool has been demonstrated to yield higher-quality gene libraries compared to degenerate oligos [144].

While oPools show significant promise for protein engineering, it is essential to acknowledge certain challenges [141]. First, although the number of individual user-defined oligo sequences in a pool is large, their individual concentrations are low. Second, in focused mutagenesis applications, some oligos might preferentially bind to the template. Exacerbated by low concentrations, certain mutations may not be present in the final constructed gene library. Third, as the length of the oligos in the pool increases, the percentage of truncated molecules also rises, further diminishing the expected concentration of full-length molecules. Fourth, the error rates for oPools are typically higher than those for column-synthesized oligos.

Irrespective of the method chosen for creating a gene library, effective target site identification is crucial for focused mutagenesis. A recent comprehensive review by the Dalby group offers an excellent overview of methodologies for identifying target sites based on sequence or structural information [145]. Additionally, the review provides a valuable list of computational tools designed for target site identification.

3.4. DNA Recombination

While considerable strides have been made in the random mutagenesis and focused mutagenesis categories, advancements in the DNA recombination category have been relatively modest. An article sought to tackle the challenge of recombining protein modules from distant parents with minimal disruption at crossover sites. To overcome this hurdle, an approach called key motif-directed recombination was introduced [146]. Members of the same protein superfamily often share common structural or sequence motifs. Validated through the creation of α/β-hydrolase chimeras, this method strategically conducted recombination at key sequence motif regions. These chimeras retained their biological functions and exhibited desirable properties.

4. Applications of Genetic Diversity Creation

Directed evolution has firmly established itself as a powerful tool in biotechnology, diversifying its applications from evolving single enzymes or proteins to fascinating areas such as metabolic pathway engineering, organismal engineering, viral engineering, and the engineering of molecular biology tools. In the section below, we provide recent examples of applications beyond engineering a single protein.

4.1. Biofuels and Biochemicals

Biofuels have emerged as a promising alternative to fossil fuels in sustainable energy solutions. Agricultural waste, a rich source of cellulose, provides a valuable feedstock for biofuel production [147]. Cellobiose, an intermediate in the conversion of cellulose into glucose monomers for fermentative processes, is crucial for efficient biofuel production. Directed evolution has played a key role in engineering two essential proteins in the cellobiose utilization pathway in *S. cerevisiae*: β-glucosidase and cellodextrin transporter [148,149]. These advancements have significantly increased ethanol production, contributing to the development of a more sustainable biofuel industry.

In Section 3.2.3 above, we extensively covered the use of the 'BE-GP' protein architecture for targeted random mutagenesis in a wide range of biological systems. By excluding the GP, this method transforms into a potent global mutagenesis approach suitable for adaptive laboratory evolution (ALE). Pan *et al.* showcased the application of BE for the ALE of *S. cerevisiae*, enhancing resistance to isobutanol and acetate, and boosting the production of β-carotene [104].

4.2. Bioremediation

Given the rising global population and the associated increase in demand for food, materials, chemicals, and energy to support contemporary lifestyles, there is an urgent need for bioremediation tools to address and alleviate the adverse environmental impacts of their production and manufacturing processes.

Polyethylene terephthalate (PET) stands as one of the most commonly used polyester plastics, finding extensive applications in fabrics and packaging. However, the thermo-mechanical recycling of PET encounters challenges, including the gradual decline in its mechanical properties over time, issues of contamination, and the energy-intensive nature inherent in these recycling methods [150]. Despite the identification of various PET hydrolases capable of converting PET into monomers, their widespread implementation on a large scale is hindered by limitations in efficiency.

A recent study aimed at enhancing the activity and thermal stability of leaf-branch compost cutinase (LCC) through enzyme engineering resulted in the creation of a mutant with high catalytic activity and increased thermostability, LCC^{ICCG} [151]. Achieved through the saturation mutagenesis of amino acid residues within the binding pocket and the introduction of a disulfide bridge, LCC^{ICCG} demonstrated the ability to convert 90% of PET into monomers on an industrial scale. In a separate study, Chen *et al.* explored the fusion of carbohydrate-binding modules (CBMs) to LCC^{ICCG}. This modification significantly enhanced the enzyme's binding affinity to PET [152]. These outputs exemplify the potential of protein engineering in addressing challenges linked to plastic utilization and contribute to the ongoing development of efficient and sustainable bioremediation solutions.

4.3. Agriculture and Food Production

Directed evolution is making contributions to the field of plant biology. The evolution of genes in crops has the potential to bolster their resistance against pests, diseases, or adverse environmental conditions, ultimately resulting in higher yields and enhanced food security. A recent study, leveraging the synergy of CRISPR/Cas and directed evolution, has successfully developed a herbicide-resistant spliceosomal protein, SF3B1 [153]. This plant-based directed evolution platform proves instrumental in investigating and evolving the molecular functions of essential biomolecules. Moreover, it facilitates the engineering of crop traits to enhance performance and adaptability under the changing conditions associated with climate change.

4.4. Diagnostics and Healthcare

In a recent study, directed evolution was leveraged to enhance the transduction capabilities of adeno-associated viruses (AAVs) [154]. Directed evolution of the AAV capsid protein resulted in variants exhibiting significantly improved transduction efficiencies. This breakthrough facilitated the genetic manipulation of microglia, specialized immune cells in the central nervous system, with unprecedented ease. Given the implication of microglial dysfunction in neurodegenerative disorders and brain cancers, this advancement opens avenues for a more profound understanding of their functions and potential therapeutic interventions.

Cholesterol oxidase finds extensive applications in diagnostics, as well as in the food and agriculture industries. It serves as a valuable diagnostic tool for detecting serum cholesterol levels and plays a pivotal role as a biocatalyst in steroid production. The thermal stability of this enzyme is crucial for its efficacy. In a recent study, epPCR was employed as a tool to engineer a mutant with three amino acid substitutions, resulting in enhanced thermal stability [155].

4.5. Advanced Molecular Biology and Protein Engineering Tools

In genome editing and targeting, the landscape has been revolutionized by the emergence of CRISPR-Cas9. The widely utilized Cas9 nuclease, originating from *Streptococcus pyogenes*, encounters challenges attributed to its substantial size, constraining its effectiveness in delivery into cells. A smaller alternative, the Cas9 orthologue from *Campylobacter Jejuni* (CjCas9), presents an appealing solution [156]. Nevertheless, the intricate PAM sequence recognized by CjCas9 imposes limitations on its versatility. Through successive rounds of directed evolution, a variant of CjCas9 named evoCjCas9 was found, boasting an altered PAM recognition sequence that is ten times more prevalent in the genome than

the canonical PAM sequence [157]. Beyond this, evoCjCas9 demonstrates increased nuclease activity compared to its wildtype counterpart, thereby broadening the horizons of CRISPR-Cas9 technology. Throughout this review, CRISPR-Cas9 has been consistently emphasized for its role in cloning and genetic diversity creation. It is truly inspiring to witness the 'ripple effect' where proteins crafted through protein engineering contribute to the continual expansion of the capabilities of protein engineering itself.

5. Emerging Trends and Prospective Trajectories

The convergence of automation and the emergence of artificial intelligence (AI) are poised to bring about a transformative shift in the domain of directed evolution, promising substantial improvements in the scale, efficiency, and precision of experiments. Automation stands to simplify labour-intensive processes like library construction, screening, and the identification of desired variants. AI tools, coupled with the capability of analysing extensive datasets, have the potential to steer experimental and variant design, expediting the discovery process.

Furthermore, we foresee that the effectiveness of directed evolution will be significantly amplified by synergistic integration with other well-established methodologies or tools. Specifically, NMR spectroscopy has been shown to enhance the capabilities of directed evolution [158]. Leveraging inhibitors and identifying chemical site perturbations through NMR spectroscopy allows for the identification of regions in a protein that are likely to yield beneficial mutations. This approach has proven successful in converting myoglobin into a highly efficient Kemp eliminase with just three amino acid substitutions.

Embarking on the era of transformation brought about by directed evolution necessitates the unlocking of untapped biological potential. At the heart of this unlocking process lies the art and science of creating genetic diversity. It is in the diverse tapestry of our genetic landscape that the blueprint for the future of the bioeconomy is etched. Through the strategic exploration and manipulation of genetic diversity, we forge a path into an era where the frontiers of bioengineering unfold boundlessly, promising novel discoveries and innovations that will shape the landscape of tomorrow.

References

1. Arnold:, F.H. Innovation by Evolution: Bringing New Chemistry to Life (Nobel Lecture). *Angew. Chem. Int. Ed. Engl.* **2019**, *58*, 14420–14426. [CrossRef] [PubMed]
2. Fasan, R.; Kan, S.B.J.; Zhao, H. A Continuing Career in Biocatalysis: Frances H. Arnold. *ACS Catal.* **2019**, *9*, 9775–9788. [CrossRef] [PubMed]
3. Wong, T.S.; Zhurina, D.; Schwaneberg, U. The Diversity Challenge in Directed Protein Evolution. *Comb. Chem. High Throughput Screen.* **2006**, *9*, 271–288. [CrossRef] [PubMed]
4. Tee, K.L.; Wong, T.S. Polishing the craft of genetic diversity creation in directed evolution. *Biotechnol. Adv.* **2013**, *31*, 1707–1721. [CrossRef] [PubMed]
5. Wong, T.S.; Tee, K.L. *A Practical Guide to Protein Engineering*; Springer: Cham, Switzerland, 2020.
6. Xia, Y.; Li, K.; Li, J.; Wang, T.; Gu, L.; Xun, L. T5 exonuclease-dependent assembly offers a low-cost method for efficient cloning and site-directed mutagenesis. *Nucleic Acids Res.* **2019**, *47*, e15. [CrossRef] [PubMed]
7. Watson, J.F.; García-Nafría, J. In vivo DNA assembly using common laboratory bacteria: A re-emerging tool to simplify molecular cloning. *J. Biol. Chem.* **2019**, *294*, 15271–15281. [CrossRef] [PubMed]

8. Nozaki, S.; Niki, H. Exonuclease III (XthA) Enforces In Vivo DNA Cloning of Escherichia coli To Create Cohesive Ends. *J. Bacteriol.* **2019**, *201*, e00660-18. [CrossRef] [PubMed]

9. Jones, D.H.; Howard, B.H. A rapid method for recombination and site-specific mutagenesis by placing homologous ends on DNA using polymerase chain reaction. *Biotechniques* **1991**, *10*, 62–66. [PubMed]

10. García-Nafría, J.; Watson, J.F.; Greger, I.H. IVA cloning: A single-tube universal cloning system exploiting bacterial In Vivo Assembly. *Sci. Rep.* **2016**, *6*, 27459. [CrossRef]

11. Huang, F.; Spangler, J.R.; Huang, A.Y. In vivo cloning of up to 16 kb plasmids in E. coli is as simple as PCR. *PLoS ONE* **2017**, *12*, e0183974. [CrossRef]

12. Jacobus, A.P.; Gross, J. Optimal Cloning of PCR Fragments by Homologous Recombination in Escherichia coli. *PLoS ONE* **2015**, *10*, e0119221. [CrossRef] [PubMed]

13. Kostylev, M.; Otwell, A.E.; Richardson, R.E.; Suzuki, Y. Cloning Should Be Simple: Escherichia coli DH5alpha-Mediated As-sembly of Multiple DNA Fragments with Short End Homologies. *PLoS ONE* **2015**, *10*, e0137466. [CrossRef] [PubMed]

14. Wang, Y.; Liu, Y.; Chen, J.; Tang, M.; Zhang, S.; Wei, L.; Li, C.; Wei, D. Restriction-ligation-free (RLF) cloning: A high-throughput cloning method by in vivo homologous recombination of PCR products. *Genet Mol. Res.* **2015**, *14*, 12306–12315. [CrossRef] [PubMed]

15. Chen, X.; Yuan, H.; He, W.; Hu, X.; Lu, H.; Li, Y. Construction of a novel kind of expression plasmid by homologous re-combination in Saccharomyces cerevisiae. *Sci. China C Life Sci.* **2005**, *48*, 330–336. [CrossRef] [PubMed]

16. Gibson, D.G. Synthesis of DNA fragments in yeast by one-step assembly of overlapping oligonucleotides. *Nucleic Acids Res.* **2009**, *37*, 6984–6990. [CrossRef] [PubMed]

17. Oldenburg, K.R.; Vo, K.T.; Michaelis, S.; Paddon, C. Recombination-mediated PCR-directed plasmid construction in vivo in yeast. *Nucleic Acids Res.* **1997**, *25*, 451–452. [CrossRef] [PubMed]

18. Ding, W.; Weng, H.; Jin, P.; Du, G.; Chen, J.; Kang, Z. Scarless assembly of unphosphorylated DNA fragments with a simplified DATEL method. *Bioengineered* **2017**, *8*, 296–301. [CrossRef] [PubMed]

19. Joska, T.M.; Mashruwala, A.; Boyd, J.M.; Belden, W.J. A universal cloning method based on yeast homologous recombination that is simple, efficient, and versatile. *J. Microbiol. Methods* **2014**, *100*, 46–51. [CrossRef]

20. Kuijpers, N.G.; Solis-Escalante, D.; Bosman, L.; van den Broek, M.; Pronk, J.T.; Daran, J.M.; Daran-Lapujade, P. A versatile, efficient strategy for assembly of multi-fragment expression vectors in Saccharomyces cerevisiae using 60 bp synthetic re-combination sequences. *Microb. Cell Fact.* **2013**, *12*, 47. [CrossRef]

21. Mashruwala, A.A.; Boyd, J.M. De Novo Assembly of Plasmids Using Yeast Recombinational Cloning. *Methods Mol. Biol.* **2016**, *1373*, 33–41.

22. van Leeuwen, J.; Andrews, B.; Boone, C.; Tan, G. Rapid and Efficient Plasmid Construction by Homologous Recombination in Yeast. *Cold Spring Harb. Protoc.* **2015**, *2015*, 085100. [CrossRef] [PubMed]

23. Ip, K.; Yadin, R.; George, K.W. High-Throughput DNA Assembly Using Yeast Homologous Recombination. *Methods Mol. Biol.* **2020**, *2205*, 79–89. [PubMed]

24. Abdel-Banat, B.M.A.; Nonklang, S.; Hoshida, H.; Akada, R. Random and targeted gene integrations through the control of non-homologous end joining in the yeast *Kluyveromyces marxianus*. *Yeast* **2009**, *27*, 29–39. [CrossRef] [PubMed]

25. Hoshida, H.; Murakami, N.; Suzuki, A.; Tamura, R.; Asakawa, J.; Abdel-Banat, B.M.A.; Nonklang, S.; Nakamura, M.; Akada, R. Non-homologous end joining-mediated functional marker selection for DNA cloning in the yeast *Kluyveromyces marxianus*. *Yeast* **2013**, *31*, 29–46. [CrossRef] [PubMed]

26. Ruiz, E.; Talenton, V.; Dubrana, M.P.; Guesdon, G.; Lluch-Senar, M.; Salin, F.; Sirand-Pugnet, P.; Arfi, Y.; Lartigue, C. CReasPy-Cloning: A Method for Simultaneous Cloning and Engineering of Megabase-Sized Genomes in Yeast Using the CRISPR-Cas9 System. *ACS Synth. Biol.* **2019**, *8*, 2547–2557. [CrossRef] [PubMed]

27. Pang, Q.; Ma, S.; Han, H.; Jin, X.; Liu, X.; Su, T.; Qi, Q. Phage Enzyme-Assisted Direct In Vivo DNA Assembly in Multiple Micro-organisms. *ACS Synth. Biol.* **2022**, *11*, 1477–1487. [CrossRef] [PubMed]

28. He, B.; Ma, Y.; Tian, F.; Zhao, G.-R.; Wu, Y.; Yuan, Y.-J. YLC-assembly: Large DNA assembly via yeast life cycle. *Nucleic Acids Res.* **2023**, *51*, 8283–8292. [CrossRef] [PubMed]

29. Jeong, Y.K.; Yu, J.; Bae, S. Construction of non-canonical PAM-targeting adenosine base editors by restriction enzyme-free DNA cloning using CRISPR-Cas9. *Sci. Rep.* **2019**, *9*, 4939. [CrossRef]

30. Wang, J.-W.; Wang, A.; Li, K.; Wang, B.; Jin, S.; Reiser, M.; Lockey, R.F. CRISPR/Cas9 nuclease cleavage combined with Gibson assembly for seamless cloning. *BioTechniques* **2015**, *58*, 161–170. [CrossRef]

31. Jiang, W.; Zhao, X.; Gabrieli, T.; Lou, C.; Ebenstein, Y.; Zhu, T.F. Cas9-Assisted Targeting of CHromosome segments CATCH enables one-step targeted cloning of large gene clusters. *Nat. Commun.* **2015**, *6*, 8101. [CrossRef]

32. Richter, D.; Bayer, K.; Toesko, T.; Schuster, S. ZeBRα a universal, multi-fragment DNA-assembly-system with minimal hands-on time requirement. *Sci. Rep.* **2019**, *9*, 2980. [CrossRef] [PubMed]

33. Dao, V.L.; Chan, S.; Zhang, J.; Ngo, R.K.J.; Poh, C.L. Single 3′-exonuclease-based multifragment DNA assembly method (SENAX). *Sci. Rep.* **2022**, *12*, 4004. [CrossRef] [PubMed]

34. Yu, F.; Li, X.; Wang, F.; Liu, Y.; Zhai, C.; Li, W.Q.; Ma, L.X.; Chen, W.P. TLTC, a T5 exonuclease-mediated low-temperature DNA cloning method. *Front. Bioeng. Biotechnol.* **2023**, *11*, 1167534. [CrossRef] [PubMed]

35. Nisson, P.E.; Rashtchian, A.; Watkins, P.C. Rapid and efficient cloning of Alu-PCR products using uracil DNA glycosylase. *PCR Methods Appl.* **1991**, *1*, 120–123. [CrossRef] [PubMed]

36. Cavaleiro, A.M.; Kim, S.H.; Seppälä, S.; Nielsen, M.T.; Nørholm, M.H.H. Accurate DNA Assembly and Genome Engineering with Optimized Uracil Excision Cloning. *ACS Synth. Biol.* **2015**, *4*, 1042–1046. [CrossRef] [PubMed]

37. Genee, H.J.; Bonde, M.T.; Bagger, F.O.; Jespersen, J.B.; Sommer, M.O.A.; Wernersson, R.; Olsen, L.R. Software-Supported USER Cloning Strategies for Site-Directed Mutagenesis and DNA Assembly. *ACS Synth. Biol.* **2014**, *4*, 342–349. [CrossRef] [PubMed]

38. Jajesniak, P.; Wong, T.S. QuickStep-Cloning: A sequence-independent, ligation-free method for rapid construction of re-combinant plasmids. *J. Biol. Eng.* **2015**, *9*, 15. [CrossRef]

39. Jajesniak, P.; Tee, K.L.; Wong, T.S. PTO-QuickStep: A Fast and Efficient Method for Cloning Random Mutagenesis Libraries. *Int. J. Mol. Sci.* **2019**, *20*, 3908. [CrossRef] [PubMed]

40. Jajesniak, P.; Tee, K.L.; Wong, T.S. Rapid Cloning of Random Mutagenesis Libraries Using PTO-QuickStep. *Methods Mol. Biol.* **2022**, *2461*, 123–135.

41. Jin, P.; Ding, W.; Du, G.; Chen, J.; Kang, Z. DATEL: A Scarless and Sequence-Independent DNA Assembly Method Using Thermostable Exonucleases and Ligase. *ACS Synth. Biol.* **2016**, *5*, 1028–1032. [CrossRef]

42. Kang, Z.; Ding, W.; Jin, P.; Du, G.; Chen, J. DNA Assembly with the DATEL Method. *Methods Mol. Biol.* **2018**, *1772*, 421–428. [PubMed]

43. Bird, J.E.; Marles-Wright, J.; Giachino, A. A User's Guide to Golden Gate Cloning Methods and Standards. *ACS Synth. Biol.* **2022**, *11*, 3551–3563. [CrossRef] [PubMed]

44. Marillonnet, S.; Grützner, R. Synthetic DNA Assembly Using Golden Gate Cloning and the Hierarchical Modular Cloning Pipeline. *Curr. Protoc. Mol. Biol.* **2020**, *130*, e115. [CrossRef] [PubMed]

45. Sorida, M.; Bonasio, R. An efficient cloning method to expand vector and restriction site compatibility of Golden Gate As-sembly. *Cell Rep. Methods* **2023**, *3*, 100564. [CrossRef] [PubMed]

46. Storch, M.; Casini, A.; Mackrow, B.; Fleming, T.; Trewhitt, H.; Ellis, T.; Baldwin, G.S. BASIC: A New Biopart Assembly Standard for Idempotent Cloning Provides Accurate, Single-Tier DNA Assembly for Synthetic Biology. *ACS Synth. Biol.* **2015**, *4*, 781–787. [CrossRef] [PubMed]

47. van Dolleweerd, C.J.; Kessans, S.A.; Van de Bittner, K.C.; Bustamante, L.Y.; Bundela, R.; Scott, B.; Nicholson, M.J.; Parker, E.J. MIDAS: A Modular DNA Assembly System for Synthetic Biology. *ACS Synth. Biol.* **2018**, *7*, 1018–1029. [CrossRef] [PubMed]

48. Lin, D.; O'callaghan, C.A. MetClo: Methylase-assisted hierarchical DNA assembly using a single type IIS restriction enzyme. *Nucleic Acids Res.* **2018**, *46*, e113. [CrossRef] [PubMed]

49. Taylor, G.M.; Heap, J.T. Design and Implementation of Multi-protein Expression Constructs and Combinatorial Libraries using Start-Stop Assembly. *Methods Mol. Biol.* **2020**, *2205*, 219–237. [PubMed]

50. Taylor, G.M.; Mordaka, P.M.; Heap, J.T. Start-Stop Assembly: A functionally scarless DNA assembly system optimized for metabolic engineering. *Nucleic Acids Res.* **2018**, *47*, e17. [CrossRef] [PubMed]

51. Trubitsyna, M.; Honsbein, A.; Jayachandran, U.; Elfick, A.; French, C.E. PaperClip DNA Assembly: Reduce, Reuse, Recycle. *Methods Mol. Biol.* **2020**, *2205*, 161–177.

52. Storch, M.; Haines, M.C.; Baldwin, G.S. DNA-BOT: A low-cost, automated DNA assembly platform for synthetic biology. *Synth. Biol.* **2020**, *5*, ysaa010. [CrossRef] [PubMed]

53. Enghiad, B.; Xue, P.; Singh, N.; Boob, A.G.; Shi, C.; Petrov, V.A.; Liu, R.; Peri, S.S.; Lane, S.T.; Gaither, E.D.; et al. PlasmidMaker is a versatile, automated, and high throughput end-to-end platform for plasmid construction. *Nat. Commun.* **2022**, *13*, 2697. [CrossRef] [PubMed]

54. Bryant, J.A.; Kellinger, M.; Longmire, C.; Miller, R.; Wright, R.C. AssemblyTron: Flexible automation of DNA assembly with Opentrons OT-2 lab robots. *Synth. Biol.* **2022**, *8*, ysac032. [CrossRef]

55. Fukuda, M.; Umeno, H.; Nose, K.; Nishitarumizu, A.; Noguchi, R.; Nakagawa, H. Construction of a guide-RNA for site-directed RNA mutagenesis utilising intracellular A-to-I RNA editing. *Sci. Rep.* **2017**, *7*, 41478. [CrossRef] [PubMed]

56. Kopsidas, G.; Carman, R.K.; Stutt, E.L.; Raicevic, A.; Roberts, A.S.; Siomos, M.-A.V.; Dobric, N.; Pontes-Braz, L.; Coia, G. RNA mutagenesis yields highly diverse mRNA libraries for in vitro protein evolution. *BMC Biotechnol.* **2007**, *7*, 18. [CrossRef] [PubMed]

57. Wong, T.S.; Roccatano, D.; Loakes, D.; Tee, K.L.; Schenk, A.; Hauer, B.; Schwaneberg, U. Transversion-enriched sequence saturation mutagenesis (SeSaM-Tv+): A random mutagenesis method with consecutive nucleotide exchanges that complements the bias of error-prone PCR. *Biotechnol. J.* **2008**, *3*, 74–82. [CrossRef] [PubMed]

58. Wong, T.S.; Tee, K.L.; Hauer, B.; Schwaneberg, U. Sequence saturation mutagenesis (SeSaM): A novel method for directed evolution. *Nucleic Acids Res.* **2004**, *32*, e26. [CrossRef] [PubMed]

59. Wong, T.S.; Tee, K.L.; Hauer, B.; Schwaneberg, U. Sequence saturation mutagenesis with tunable mutation frequencies. *Anal. Biochem.* **2005**, *341*, 187–189. [CrossRef] [PubMed]

60. Zhao, J.; Frauenkron-Machedjou, V.J.; Kardashliev, T.; Ruff, A.J.; Zhu, L.; Bocola, M.; Schwaneberg, U. Amino acid substitutions in random mutagenesis libraries: Lessons from analyzing 3000 mutations. *Appl. Microbiol. Biotechnol.* **2017**, *101*, 3177–3187. [CrossRef] [PubMed]

61. Zhao, J.; Kardashliev, T.; Ruff, A.J.; Bocola, M.; Schwaneberg, U. Lessons from diversity of directed evolution experiments by an analysis of 3,000 mutations. *Biotechnol. Bioeng.* **2014**, *111*, 2380–2389. [CrossRef]

62. Verma, R.; Wong, T.S.; Schwaneberg, U.; Roccatano, D. The Mutagenesis Assistant Program. *Methods Mol. Biol.* **2014**, *1179*, 279–290. [PubMed]

63. Wong, T.S.; Roccatano, D.; Schwaneberg, U. Are transversion mutations better? A Mutagenesis Assistant Program analysis on P450 BM-3 heme domain. *Biotechnol. J.* **2007**, *2*, 133–142. [CrossRef] [PubMed]

64. Wong, T.S.; Roccatano, D.; Zacharias, M.; Schwaneberg, U. A Statistical Analysis of Random Mutagenesis Methods Used for Directed Protein Evolution. *J. Mol. Biol.* **2006**, *355*, 858–871. [CrossRef] [PubMed]

65. Li, H.; Li, J.; Jin, R.; Chen, W.; Liang, C.; Wu, J.; Jin, J.-M.; Tang, S.-Y. Towards the construction of high-quality mutagenesis libraries. *Biotechnol. Lett.* **2018**, *40*, 1101–1107. [CrossRef] [PubMed]

66. Li, L.; Liu, X.; Bai, Y.; Yao, B.; Luo, H.; Tu, T. High-Throughput Screening Techniques for the Selection of Thermostable Enzymes. *J. Agric. Food Chem.* **2024**, *72*, 3833–3845. [CrossRef] [PubMed]

67. Li, Z.; Deng, Y.; Yang, G.-Y. Growth-coupled high throughput selection for directed enzyme evolution. *Biotechnol. Adv.* **2023**, *68*, 108238. [CrossRef] [PubMed]

68. Manteca, A.; Gadea, A.; Van Assche, D.; Cossard, P.; Gillard-Bocquet, M.; Beneyton, T.; Innis, C.A.; Baret, J.-C. Directed Evolution in Drops: Molecular Aspects and Applications. *ACS Synth. Biol.* **2021**, *10*, 2772–2783. [CrossRef]

69. Wang, Y.; Xue, P.; Cao, M.; Yu, T.; Lane, S.T.; Zhao, H. Directed Evolution: Methodologies and Applications. *Chem. Rev.* **2021**, *121*, 12384–12444. [CrossRef]

70. Stucki, A.; Vallapurackal, J.; Ward, T.R.; Dittrich, P.S. Droplet Microfluidics and Directed Evolution of Enzymes: An Intertwined Journey. *Angew. Chem. Int. Ed. Engl.* **2021**, *60*, 24368–24387. [CrossRef]

71. Markel, U.; Essani, K.D.; Besirlioglu, V.; Schiffels, J.; Streit, W.R.; Schwaneberg, U. Advances in ultrahigh-throughput screening for directed enzyme evolution. *Chem. Soc. Rev.* **2019**, *49*, 233–262. [CrossRef]

72. Yang, Y.; Liu, M.; Wang, T.; Wang, Q.; Liu, H.; Xun, L.; Xia, Y. An Optimized Transformation Protocol for Escherichia coli BW3KD with Supreme DNA Assembly Efficiency. *Microbiol. Spectr.* **2022**, *10*, e0249722. [CrossRef] [PubMed]

73. Yang, Y.; Yu, Q.; Wang, M.; Zhao, R.; Liu, H.; Xun, L.; Xia, Y. Escherichia coli BW25113 Competent Cells Prepared Using a Simple Chemical Method Have Unmatched Transformation and Cloning Efficiencies. *Front. Microbiol.* **2022**, *13*, 838698. [CrossRef] [PubMed]

74. Chai, D.; Wang, G.; Fang, L.; Li, H.; Liu, S.; Zhu, H.; Zheng, J. The optimization system for preparation of TG1 competent cells and electrotransformation. *Microbiologyopen* **2020**, *9*, e1043. [CrossRef] [PubMed]

75. Liu, J.; Chang, W.; Pan, L.; Liu, X.; Su, L.; Zhang, W.; Li, Q.; Zheng, Y. An Improved Method of Preparing High Efficiency Transformation Escherichia coli with Both Plasmids and Larger DNA Fragments. *Indian J. Microbiol.* **2018**, *58*, 448–456. [CrossRef] [PubMed]

76. Loock, M.; Antunes, L.B.; Heslop, R.T.; De Lauri, A.A.; Lira, A.B.; Cestari, I. High-Efficiency Transformation and Expression of Genomic Libraries in Yeast. *Methods Protoc.* **2023**, *6*, 89. [CrossRef] [PubMed]

77. Yu, S.-C.; Dawson, A.; Henderson, A.C.; Lockyer, E.J.; Read, E.; Sritharan, G.; Ryan, M.; Sgroi, M.; Ngou, P.M.; Woodruff, R.; et al. Nutrient supplements boost yeast transformation efficiency. *Sci. Rep.* **2016**, *6*, 35738. [CrossRef] [PubMed]

78. Yildiz, S.; Solak, K.; Acar, M.; Mavi, A.; Unver, Y. Magnetic nanoparticle mediated-gene delivery for simpler and more effective transformation of *Pichia pastoris*. *Nanoscale Adv.* **2021**, *3*, 4482–4491. [CrossRef] [PubMed]

79. Kumar, S.; Mannil, A.; Mutturi, S. Modified chemical method for efficient transformation and diagnosis in Pichia pastoris. *Protein Expr. Purif.* **2020**, *174*, 105685. [CrossRef]

80. Kumar, R. Simplified protocol for faster transformation of (a large number of) *Pichia pastoris* strains. *Yeast* **2019**, *36*, 399–410. [CrossRef]

81. Lyu, Y.; Wu, P.; Zhou, J.; Yu, Y.; Lu, H. Protoplast transformation of *Kluyveromyces marxianus*. *Biotechnol. J.* **2021**, *16*, 2100122. [CrossRef]

82. Yang, J.; Ruff, A.J.; Arlt, M.; Schwaneberg, U. Casting epPCR (cepPCR): A simple random mutagenesis method to generate high quality mutant libraries. *Biotechnol. Bioeng.* **2017**, *114*, 1921–1927. [CrossRef]

83. Ye, B.; Li, Y.; Tao, Q.; Yao, X.; Cheng, M.; Yan, X. Random Mutagenesis by Insertion of Error-Prone PCR Products to the Chromosome of Bacillus subtilis. *Front. Microbiol.* **2020**, *11*, 570280. [CrossRef] [PubMed]

84. Tachioka, M.; Sugimoto, N.; Nakamura, A.; Sunagawa, N.; Ishida, T.; Uchiyama, T.; Igarashi, K.; Samejima, M. Development of simple random mutagenesis protocol for the protein expression system in Pichia pastoris. *Biotechnol. Biofuels* **2016**, *9*, 8425. [CrossRef] [PubMed]

85. Badran, A.H.; Liu, D.R. Development of potent in vivo mutagenesis plasmids with broad mutational spectra. *Nat. Commun.* **2015**, *6*, 8425. [CrossRef] [PubMed]

86. Moore, C.L.; Papa, L.J.; Shoulders, M.D. A Processive Protein Chimera Introduces Mutations across Defined DNA Regions *In Vivo*. *J. Am. Chem. Soc.* **2018**, *140*, 11560–11564. [CrossRef]

87. Park, H.; Kim, S. Gene-specific mutagenesis enables rapid continuous evolution of enzymes in vivo. *Nucleic Acids Res.* **2021**, *49*, e32. [CrossRef]

88. Mengiste, A.A.; Wilson, R.H.; Weissman, R.F.; Iii, L.J.P.; Hendel, S.J.; Moore, C.L.; Butty, V.L.; Shoulders, M.D. Expanded MutaT7 toolkit efficiently and simultaneously accesses all possible transition mutations in bacteria. *Nucleic Acids Res.* **2023**, *51*, e31. [CrossRef]

89. Gaudelli, N.M.; Komor, A.C.; Rees, H.A.; Packer, M.S.; Badran, A.H.; Bryson, D.I.; Liu, D.R. Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* **2017**, *551*, 464–471. [CrossRef]

90. Álvarez, B.; Mencía, M.; de Lorenzo, V.; Fernández, L. In vivo diversification of target genomic sites using processive base deaminase fusions blocked by dCas9. *Nat. Commun.* **2020**, *11*, 6436. [CrossRef]

91. Halperin, S.O.; Tou, C.J.; Wong, E.B.; Modavi, C.; Schaffer, D.V.; Dueber, J.E. CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. *Nature* **2018**, *560*, 248–252. [CrossRef]

92. Sadanand, S. EvolvR-ing to targeted mutagenesis. *Nat. Biotechnol.* **2018**, *36*, 819. [CrossRef] [PubMed]

93. Wang, L.; Xue, W.; Yan, L.; Li, X.; Wei, J.; Chen, M.; Wu, J.; Yang, B.; Yang, L.; Chen, J. Enhanced base editing by co-expression of free uracil DNA glycosylase inhibitor. *Cell Res.* **2017**, *27*, 1289–1292. [CrossRef] [PubMed]

94. Nishida, K.; Arazoe, T.; Yachie, N.; Banno, S.; Kakimoto, M.; Tabata, M.; Mochizuki, M.; Miyabe, A.; Araki, M.; Hara, K.Y.; et al. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* **2016**, *353*, aaf8729. [CrossRef] [PubMed]

95. Hess, G.T.; Frésard, L.; Han, K.; Lee, C.H.; Li, A.; Cimprich, K.A.; Montgomery, S.B.; Bassik, M.C. Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. *Nat. Methods* **2016**, *13*, 1036–1042. [CrossRef] [PubMed]

96. Ma, Y.; Zhang, J.; Yin, W.; Zhang, Z.; Song, Y.; Chang, X. Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. *Nat. Methods* **2016**, *13*, 1029–1035. [CrossRef] [PubMed]

97. Komor, A.C.; Kim, Y.B.; Packer, M.S.; Zuris, J.A.; Liu, D.R. Programmable editing of a target base in genomic DNA without dou-ble-stranded DNA cleavage. *Nature* **2016**, *533*, 420–424. [CrossRef] [PubMed]

98. Gehrke, J.M.; Cervantes, O.; Clement, M.K.; Wu, Y.; Zeng, J.; Bauer, D.E.; Pinello, L.; Joung, J.K. An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities. *Nat. Biotechnol.* **2018**, *36*, 977–982. [CrossRef] [PubMed]

99. Chen, H.; Liu, S.; Padula, S.; Lesman, D.; Griswold, K.; Lin, A.; Zhao, T.; Marshall, J.L.; Chen, F. Efficient, continuous mutagenesis in human cells using a pseudo-random DNA editor. *Nat. Biotechnol.* **2019**, *38*, 165–168. [CrossRef]

100. Cravens, A.; Jamil, O.K.; Kong, D.; Sockolosky, J.T.; Smolke, C.D. Polymerase-guided base editing enables in vivo mutagenesis and rapid protein engineering. *Nat. Commun.* **2021**, *12*, 1579. [CrossRef]

101. Volke, D.C.; Martino, R.A.; Kozaeva, E.; Smania, A.M.; Nikel, P.I. Modular (de)construction of complex bacterial phenotypes by CRISPR/nCas9-assisted, multiplex cytidine base-editing. *Nat. Commun.* **2022**, *13*, 3026. [CrossRef]

102. Skrekas, C.; Limeta, A.; Siewers, V.; David, F. Targeted In Vivo Mutagenesis in Yeast Using CRISPR/Cas9 and Hyperactive Cytidine and Adenine Deaminases. *ACS Synth. Biol.* **2023**, *12*, 2278–2289. [CrossRef] [PubMed]

103. Zimmermann, A.; Prieto-Vivas, J.E.; Cautereels, C.; Gorkovskiy, A.; Steensels, J.; Van de Peer, Y.; Verstrepen, K.J. A Cas3-base editing tool for targetable in vivo mutagenesis. *Nat. Commun.* **2023**, *14*, 3389. [CrossRef] [PubMed]

104. Pan, Y.; Xia, S.; Dong, C.; Pan, H.; Cai, J.; Huang, L.; Xu, Z.; Lian, J. Random Base Editing for Genome Evolution in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* **2021**, *10*, 2440–2446. [CrossRef] [PubMed]

105. Jewel, D.; Pham, Q.; Chatterjee, A. Virus-assisted directed evolution of biomolecules. *Curr. Opin. Chem. Biol.* **2023**, *76*, 102375. [CrossRef] [PubMed]

106. English, J.G.; Olsen, R.H.J.; Lansu, K.; Patel, M.; White, K.; Cockrell, A.S.; Singh, D.; Strachan, R.T.; Wacker, D.; Roth, B.L. VEGAS as a Platform for Facile Directed Evolution in Mammalian Cells. *Cell* **2019**, *178*, 748–761.e17. [CrossRef] [PubMed]

107. Hwang, J.; Cho, K.H.; Song, H.; Yi, H.; Kim, H.S. Deletion mutations conferring substrate spectrum extension in the class A be-ta-lactamase. *Antimicrob. Agents Chemother.* **2014**, *58*, 6265–6269. [CrossRef] [PubMed]

108. Toledo-Patiño, S.; Pascarelli, S.; Uechi, G.-I.; Laurino, P. Insertions and deletions mediated functional divergence of Rossmann fold enzymes. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2207965119. [CrossRef] [PubMed]

109. Savino, S.; Desmet, T.; Franceus, J. Insertions and deletions in protein evolution and engineering. *Biotechnol. Adv.* **2022**, *60*, 108010. [CrossRef] [PubMed]

110. Fujii, R.; Kitaoka, M.; Hayashi, K. RAISE: A simple and novel method of generating random insertion and deletion mutations. *Nucleic Acids Res.* **2006**, *34*, e30. [CrossRef]

111. Kipnis, Y.; Dellus-Gur, E.; Tawfik, D.S. TRINS: A method for gene modification by randomized tandem repeat insertions. *Protein Eng. Des. Sel.* **2012**, *25*, 437–444. [CrossRef]

112. Liu, S.S.; Wei, X.; Ji, Q.; Xin, X.; Jiang, B.; Liu, J. A facile and efficient transposon mutagenesis method for generation of mul-ti-codon deletions in protein sequences. *J. Biotechnol.* **2016**, *227*, 27–34. [CrossRef] [PubMed]

113. Hayes, F.; Hallet, B. Pentapeptide scanning mutagenesis: Encouraging old proteins to execute unusual tricks. *Trends Microbiol.* **2000**, *8*, 571–577. [CrossRef]

114. Jones, D.D. Triplet nucleotide removal at random positions in a target gene: The tolerance of TEM-1 beta-lactamase to an amino acid deletion. *Nucleic Acids Res.* **2005**, *33*, e80. [CrossRef] [PubMed]

115. Emond, S.; Petek, M.; Kay, E.J.; Heames, B.; Devenish, S.R.A.; Tokuriki, N.; Hollfelder, F. Accessing unexplored regions of sequence space in directed enzyme evolution via insertion/deletion mutagenesis. *Nat. Commun.* **2020**, *11*, 3469. [CrossRef] [PubMed]

116. Skamaki, K.; Emond, S.; Chodorge, M.; Andrews, J.; Rees, D.G.; Cannon, D.; Popovic, B.; Buchanan, A.; Minter, R.R.; Hollfelder, F. In vitro evolution of antibody affinity via insertional scanning mutagenesis of an entire antibody variable region. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 27307–27318. [CrossRef] [PubMed]

117. Wrenbeck, E.E.; Klesmith, J.R.; Stapleton, J.A.; Adeniran, A.; Tyo, K.E.; Whitehead, T.A. Plasmid-based one-pot saturation mu-tagenesis. *Nat. Methods* **2016**, *13*, 928–930. [CrossRef] [PubMed]

118. Meinke, G.; Dalda, N.; Brigham, B.S.; Bohm, A. Synthesis of libraries and multi-site mutagenesis using a PCR-derived, dU-containing template. *Synth. Biol.* **2021**, *6*, ysaa030. [CrossRef] [PubMed]

119. Higgins, S.A.; Ouonkap, S.V.Y.; Savage, D.F. Rapid and Programmable Protein Mutagenesis Using Plasmid Recombineering. *ACS Synth. Biol.* **2017**, *6*, 1825–1833. [CrossRef] [PubMed]

120. Hoebenreich, S.; Zilly, F.E.; Acevedo-Rocha, C.G.; Zilly, M.; Reetz, M.T. Speeding up Directed Evolution: Combining the Advantages of Solid-Phase Combinatorial Gene Synthesis with Statistically Guided Reduction of Screening Effort. *ACS Synth. Biol.* **2014**, *4*, 317–331. [CrossRef]

121. Li, A.; Sun, Z.; Reetz, M.T. Solid-Phase Gene Synthesis for Mutant Library Construction: The Future of Directed Evolution? *ChemBioChem* **2018**, *19*, 2023–2032. [CrossRef]

122. Öling, D.; Lawenius, L.; Shaw, W.; Clark, S.; Kettleborough, R.; Ellis, T.; Larsson, N.; Wigglesworth, M. Large Scale Synthetic Site Saturation GPCR Libraries Reveal Novel Mutations That Alter Glucose Signaling. *ACS Synth. Biol.* **2018**, *7*, 2317–2321. [CrossRef] [PubMed]

123. Cozens, C.; Pinheiro, V.B. Darwin Assembly: Fast, efficient, multi-site bespoke mutagenesis. *Nucleic Acids Res.* **2018**, *46*, e51. [CrossRef] [PubMed]

124. Kirby, M.B.; Medina-Cucurella, A.V.; Baumer, Z.T.; Whitehead, T.A. Optimization of multi-site nicking mutagenesis for generation of large, user-defined combinatorial libraries. *Protein Eng. Des. Sel.* **2021**, *34*, gzab017. [CrossRef] [PubMed]

125. Mighell, T.L.; Toledano, I.; Lehner, B. SUNi mutagenesis: Scalable and uniform nicking for efficient generation of variant libraries. *PLoS ONE* **2023**, *18*, e0288158. [CrossRef] [PubMed]

126. Belsare, K.D.; Andorfer, M.C.; Cardenas, F.S.; Chael, J.R.; Park, H.J.; Lewis, J.C. A Simple Combinatorial Codon Mutagenesis Method for Targeted Protein Engineering. *ACS Synth. Biol.* **2017**, *6*, 416–420. [CrossRef] [PubMed]

127. Bloom, J.D. An Experimentally Informed Evolutionary Model Improves Phylogenetic Fit to Divergent Lactamase Homologs. *Mol. Biol. Evol.* **2014**, *31*, 2753–2769. [CrossRef]

128. Chung, D.H.; Potter, S.C.; Tanomrat, A.C.; Ravikumar, K.M.; Toney, M.D. Site-directed mutant libraries for isolating minimal mutations yielding functional changes. *Protein Eng. Des. Sel.* **2017**, *30*, 347–357. [CrossRef]

129. Püllmann, P.; Ulpinnis, C.; Marillonnet, S.; Gruetzner, R.; Neumann, S.; Weissenborn, M.J. Golden Mutagenesis: An efficient multi-site-saturation mutagenesis approach by Golden Gate cloning with automated primer design. *Sci. Rep.* **2019**, *9*, 10932. [CrossRef]

130. Hejlesen, R.; Füchtbauer, E.-M. Multiple site-directed mutagenesis via simple cloning by prolonged overlap extension. *BioTechniques* **2020**, *68*, 345–348. [CrossRef]

131. She, W.; Ni, J.; Shui, K.; Wang, F.; He, R.; Xue, J.; Reetz, M.T.; Li, A.; Ma, L. Rapid and Error-Free Site-Directed Mutagenesis by a PCR-Free In Vitro CRISPR/Cas9-Mediated Mutagenic System. *ACS Synth. Biol.* **2018**, *7*, 2236–2244. [CrossRef]

132. Currin, A.; Kwok, J.; Sadler, J.C.; Bell, E.L.; Swainston, N.; Ababi, M.; Day, P.J.; Turner, N.J.; Kell, D.B. GeneORator: An Effective Strategy for Navigating Protein Sequence Space More Efficiently through Boolean OR-Type DNA Libraries. *ACS Synth. Biol.* **2019**, *8*, 1371–1378. [CrossRef]

133. Reetz, M.T.; Carballeira, J.D. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat. Protoc.* **2007**, *2*, 891–903. [CrossRef] [PubMed]

134. Wang, X.; Zheng, K.; Zheng, H.; Nie, H.; Yang, Z.; Tang, L. DC-Analyzer-facilitated combinatorial strategy for rapid directed evolution of functional enzymes with multiple mutagenesis sites. *J. Biotechnol.* **2014**, *192*, 102–107. [CrossRef] [PubMed]

135. Tang, L.; Gao, H.; Zhu, X.; Wang, X.; Zhou, M.; Jiang, R. Construction of "small-intelligent" focused mutagenesis libraries using well-designed combinatorial degenerate primers. *BioTechniques* **2012**, *52*, 149–158. [CrossRef] [PubMed]

136. Sun, Z.; Lonsdale, R.; Li, G.; Reetz, M.T. Comparing Different Strategies in Directed Evolution of Enzyme Stereoselectivity: Single-versus Double-Code Saturation Mutagenesis. *ChemBioChem* **2016**, *17*, 1865–1872. [CrossRef] [PubMed]

137. Yang, J.; Ducharme, J.; Johnston, K.E.; Li, F.-Z.; Yue, Y.; Arnold, F.H. DeCOIL: Optimization of Degenerate Codon Libraries for Machine Learning-Assisted Protein Engineering. *ACS Synth. Biol.* **2023**, *12*, 2444–2454. [CrossRef] [PubMed]

138. Hiraga, K.; Mejzlik, P.; Marcisin, M.; Vostrosablin, N.; Gromek, A.; Arnold, J.; Wiewiora, S.; Svarba, R.; Prihoda, D.; Clarova, K.; et al. Mutation Maker, An Open Source Oligo Design Platform for Protein Engineering. *ACS Synth. Biol.* **2021**, *10*, 357–370. [CrossRef] [PubMed]

139. Swainston, N.; Currin, A.; Day, P.J.; Kell, D.B. GeneGenie: Optimized oligomer design for directed evolution. *Nucleic Acids Res.* **2014**, *42*, W395–W400. [CrossRef] [PubMed]

140. Swainston, N.; Currin, A.; Green, L.; Breitling, R.; Day, P.J.; Kell, D.B. CodonGenie: Optimised ambiguous codon design tools. *PeerJ Comput. Sci.* **2017**, *3*, e120. [CrossRef]

141. Kuiper, B.P.; Prins, R.C.; Billerbeck, S. Oligo Pools as an Affordable Source of Synthetic DNA for Cost-Effective Library Construction in Protein- and Metabolic Pathway Engineering. *ChemBioChem* **2021**, *23*, e202100507. [CrossRef]

142. Fowler, D.M.; Fields, S. Deep mutational scanning: A new style of protein science. *Nat. Methods* **2014**, *11*, 801–807. [CrossRef] [PubMed]

143. Kitzman, J.O.; Starita, L.M.; Lo, R.S.; Fields, S.; Shendure, J. Massively parallel single-amino-acid mutagenesis. *Nat. Methods* **2015**, *12*, 203–206. [CrossRef] [PubMed]

144. Medina-Cucurella, A.V.; Steiner, P.J.; Faber, M.S.; Beltrán, J.; Borelli, A.N.; Kirby, M.B.; Cutler, S.R.; Whitehead, T.A. User-defined single pot mutagenesis using unamplified oligo pools. *Protein Eng. Des. Sel.* **2019**, *32*, 41–45. [CrossRef] [PubMed]

145. Yu, H.; Ma, S.; Li, Y.; Dalby, P.A. Hot spots-making directed evolution easier. *Biotechnol. Adv.* **2022**, *56*, 107926. [CrossRef] [PubMed]

146. Zhou, X.; Gao, L.; Yang, G.; Liu, D.; Bai, A.; Li, B.; Deng, Z.; Feng, Y. Design of Hyperthermophilic Lipase Chimeras by Key Motif-Directed Recombination. *ChemBioChem* **2014**, *16*, 455–462. [CrossRef] [PubMed]

147. Kumar, J.A.; Sathish, S.; Prabu, D.; Renita, A.A.; Saravanan, A.; Deivayanai, V.; Anish, M.; Jayaprabakar, J.; Baigenzhenov, O.; Hosseini-Bandegharaei, A. Agricultural waste biomass for sustainable bioenergy production: Feedstock, characterization and pre-treatment methodologies. *Chemosphere* **2023**, *331*, 138680. [CrossRef] [PubMed]

148. Hu, M.-L.; Zha, J.; He, L.-W.; Lv, Y.-J.; Shen, M.-H.; Zhong, C.; Li, B.-Z.; Yuan, Y.-J. Enhanced Bioconversion of Cellobiose by Industrial Saccharomyces cerevisiae Used for Cellulose Utilization. *Front. Microbiol.* **2016**, *7*, 241. [CrossRef] [PubMed]

149. Kim, H.; Oh, E.J.; Lane, S.T.; Lee, W.-H.; Cate, J.H.; Jin, Y.-S. Enhanced cellobiose fermentation by engineered Saccharomyces cerevisiae expressing a mutant cellodextrin facilitator and cellobiose phosphorylase. *J. Biotechnol.* **2018**, *275*, 53–59. [CrossRef] [PubMed]

150. Magalhães, R.P.; Cunha, J.M.; Sousa, S.F. Perspectives on the Role of Enzymatic Biocatalysis for the Degradation of Plastic PET. *Int. J. Mol. Sci.* **2021**, *22*, 11257. [CrossRef]

151. Tournier, V.; Topham, C.M.; Gilles, A.; David, B.; Folgoas, C.; Moya-Leclair, E.; Kamionka, E.; Desrousseaux, M.-L.; Texier, H.; Gavalda, S.; et al. An engineered PET depolymerase to break down and recycle plastic bottles. *Nature* **2020**, *580*, 216–219. [CrossRef]

152. Chen, Y.; Zhang, S.; Zhai, Z.; Zhang, S.; Ma, J.; Liang, X.; Li, Q. Construction of Fusion Protein with Carbohydrate-Binding Module and Leaf-Branch Compost Cutinase to Enhance the Degradation Efficiency of Polyethylene Terephthalate. *Int. J. Mol. Sci.* **2023**, *24*, 2780. [CrossRef] [PubMed]

153. Butt, H.; Eid, A.; Momin, A.A.; Bazin, J.; Crespi, M.; Arold, S.T.; Mahfouz, M.M. CRISPR directed evolution of the spliceosome for resistance to splicing inhibitors. *Genome Biol.* **2019**, *20*, 73. [CrossRef] [PubMed]

154. Lin, R.; Zhou, Y.; Yan, T.; Wang, R.; Li, H.; Wu, Z.; Zhang, X.; Zhou, X.; Zhao, F.; Zhang, L.; et al. Directed evolution of adeno-associated virus for efficient gene delivery to microglia. *Nat. Methods* **2022**, *19*, 976–985. [CrossRef] [PubMed]

155. Fana, S.E.; Fazaeli, A.; Aminian, M. Directed evolution of cholesterol oxidase with improved thermostability using error-prone PCR. *Biotechnol. Lett.* **2023**, *45*, 1159–1167. [CrossRef] [PubMed]

156. Kim, E.; Koo, T.; Park, S.W.; Kim, D.; Kim, K.; Cho, H.-Y.; Song, D.W.; Lee, K.J.; Jung, M.H.; Kim, S.; et al. In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni. *Nat. Commun.* **2017**, *8*, 14500. [CrossRef] [PubMed]

157. Schmidheini, L.; Mathis, N.; Marquart, K.F.; Rothgangl, T.; Kissling, L.; Böck, D.; Chanez, C.; Wang, J.P.; Jinek, M.; Schwank, G. Continuous directed evolution of a compact CjCas9 variant with broad PAM compatibility. *Nat. Chem. Biol.* **2023**, *20*, 333–343. [CrossRef]

158. Bhattacharya, S.; Margheritis, E.G.; Takahashi, K.; Kulesha, A.; D'souza, A.; Kim, I.; Yoon, J.H.; Tame, J.R.H.; Volkov, A.N.; Makhlynets, O.V.; et al. NMR-guided directed evolution. *Nature* **2022**, *610*, 389–393. [CrossRef]