Splicing Speed: A Review of Breakneck Advances in Genetic Engineering and Biotech

John Cost

Introduction

The paper explores recent advancements in DNA sequencing, focusing on high-throughput technology's impact on science, medicine, and industry. It delves into CRISPR/Cas9 for genome editing, discussing challenges and innovative applications. Additionally, it highlights in-vitro transcription-modified mRNA therapeutics, addressing both potential and challenges. The intersection of data science, machine learning, and biological pathways is examined, emphasizing safety considerations. The revolutionary technique of optogenetics is discussed for precise cellular control, acknowledging challenges. Enzyme engineering challenges and the need for innovative strategies are also explored.

High Throughput and Next-Generation DNA Sequencing

The forefront of research in the sphere of biology and medicine is currently occupied chiefly by the study of the genetic code that underpins the natural world. There is a key technology that fuels this revolution— DNA sequencing— that in recent years, has developed in varied ways all with unique tradeoffs. The ability of researchers to accurately and economically determine this code represents significant value to an untold number of scientific and medical advances, but also to additional industries yet determined.

This latest generation of DNA sequencing technology is termed high-throughput (or Next-Generation) sequencing, after its technological contrast with initial chain-terminating (Sanger) methods[1]. This new generation has enhanced sequencing volume for lower cost (time and monetary) but with tradeoffs. Thus, the field of DNA sequencing boasts a suite of unique technologies all aiming to accomplish the same objective— accurate readout of genetic sequence information.

Challenges

All sequencing technologies face unique challenges. Chiefly of concern are economic constraints. While original processes cost exponentially more than the processes that fall under “high throughput,” the cost of sequencing infrastructure and upkeep is still a major factor in determining whether or not a particular technology will catch on.

Previous technologies have come and gone within the field on account of their economic viability. More recently, Helios filed for bankruptcy[2]. This group was a contender in the space but the infrastructure

required for their process was too costly and thus not adopted at scale. On the opposite end of the spectrum, nearly the very first iteration of sequencing technology, Sanger Sequencing, remains the most costly option per base (often measured now in dollars per gigabase, $ / Gb), but remains a foundational and widely used technology due to its read length (~1000bp) and superior accuracy[3]. Currently, most high throughput sequencing technologies lie in a similar range of cost per Gb, though they differ considerably in their equipment (startup) costs[1]. Ultimately, DNA sequencing aims to be widely accessible, and thus these costs should decrease.

The latest DNA sequencing technologies have improved significantly in their speed and throughput, enabling the rapid generation of vast amounts of genomic data. However, many still struggle with ensuring high levels of accuracy[1]. Accuracy must be a goal of any sequencing technology, as minute changes to genetic sequence can cause disproportionately large effects. Extreme phenotypes and changes are the result of a single differing nucleotide. Thus, accuracy is important. As of now, newer sequencing technologies do not yet match the accuracy of Sanger Sequencing, with some only capable of significantly lower read quality (but some with longer read length)[4].

DNA sequencing remains a physical process. The reactions and mechanisms that describe it take a highly variable amount of time from approach to approach. Some can take on the order of days and others take only hours[1]. Speed and volume are also a challenge to the field, as the ability to cheaply, accurately, but also rapidly assess a complete genetic makeup increases the technology's fitness for use in medicine.

Approaches

Sanger

Sanger Sequencing is often described as the “Gold Standard” of DNA sequencing[10], for its very high accuracy. Sanger sequencing also produces individual reads of ~1000 bases in length[1]. Although it has been largely supplanted by NGS technologies for large-scale projects with big volumetric demands, Sanger sequencing is still used for certain applications. Particularly, clinical diagnostics and validation of NGS results still rely on Sanger processes[1]. It still remains a foundational and actively utilized technology, known for its high accuracy in producing relatively long reads.

Reverse Terminator Sequencing & “Next Generation Sequencing”

Another more recent approach to DNA sequencing is referred to as Reverse Terminator Sequencing. As of now, this is one of the most widely used methods for DNA sequencing. The main proprietor of this method of sequencing is Illumina. Illumina relies on fluorescently labeled reversible terminators to read DNA sequences in parallel. Another representative of Next Generation Sequencing methods, PacBio, employs a single polymerase enzyme that synthesizes DNA in real-time as the template molecule passes through the enzyme's active site (SBS and SMRT respectively). As a whole, these platforms generate large amounts of data with relatively high accuracy[1]. Despite their youth as a process they are already currently in use for research, clinical diagnostics, and other applications.

Single Molecule Magnetic Sequencing

This method of DNA sequencing is highly economical, but also relatively highly inaccurate. This method requires very little in terms of infrastructure, leveraging common laboratory equipment such as

microscopes to perform sequencing. The tradeoff for this method is a highly involved process with a large time invested in library preparation. The method requires a cell-free library of DNA fragments to be created. Then these DNA fragments are amplified (with some error), followed by fixing the sequence via a system of charged beads and gel. These sequences are then determined via the cyclical addition of nucleotides, each of which emits a unique color of light to be observed via microscope. This method carries a higher error rate than other methods but for a much lower cost. This method is highly multiplexed in nature.

Nanopore

This technology, popularly Oxford Nanopore, works by passing DNA through a nanopore, measuring changes in electrical current as the nucleotides pass through. It can produce significantly longer reads than any other technique (4Mb individual reads) and has promising real-time sequencing capabilities and portability[6]. Nanopore sequencing currently struggles with achieving the accuracy of other methods (~95%). This method shows promise for improving accuracy as its cooperation with machine learning techniques has yet to be fully investigated. As it stands, this technology has some of the lowest costs in terms of startup investment[6]. Nanopore also is significantly more portable than other technologies.

Current Applications

Whole Microbial Genomes

Already, high throughput technology has been demonstrated as effective at generating near-finished microbial genomes from cultures of single organisms or metagenomes. This is due in part to Nanopore's advantage in sequencing large repeat regions, where previous methods might struggle. Nanopore performs this feat also at a fraction of the cost of infrastructure and on a per-base basis. These reads also achieved in the case of these microbial genomes, a modal read accuracy of 99% with ~40-fold depth[9]. This would have been near impossible with previous technology, and magnitudes more costly in time and resources.

Whole Human Genomes with Illumina RTS

The method of reverse terminator sequencing has been proven as a cost-effective and comparatively simple method of sequencing in high throughput of an entire human genome. This approach specifically was used to sequence an entire human X-chromosome of a human male. The method was able to identify numerous SNPs accurately along with novel structural variants. Very importantly, this method was highly economical relative to previous processes[11].

Magnetic Bead Whole-Polony Sequencing

The method of bead sequencing via fluorescence has been used to determine the sequence of a bacterial colony. Researchers were able to highly automate the process— a process that was entirely available for a cost of ~$140,000, much cheaper than other methods. Researchers reported a cost close to 1/9 that of other methods, at a total accuracy of ~99.7%. This approach leverages readily available equipment and supplies but places most of the burden on library preparation at the expense of accuracy[12].

Beyond this, the widespread adoption of NGS processes supports their superiority over previous methods. These technologies boast the aforementioned advantages in the dimensions of useability and accuracy and are growingly accessible.

Future Development

Future development in the field may center on the application of ML methods to

sequencing. Especially with Nanopore, future gains with the computational models that are used to predict base seem to represent a large potential increase in accuracy. Other methods may also be developed that leverage these computational advances[7].

Beyond a direct increase in efficacy of individual methods, the field may be moving towards hybrid approaches to Sequencing, where different approaches are used on the same genetic material, where the sum is a read that does not suffer from inaccuracy as a result of long repeat regions via use of long-read techniques, but also leverages methods with higher accuracy for verification of critical intervals.

Future Applications

Advances in the field of high-throughput DNA sequencing have applications in the field of genomic medicines, and personalized medicine, where completely individualized treatments are provided in correspondence to the genome. Efficacy has already been demonstrated in the field of Cancer genomics[8], where treatments now are specialized on the individual level. Drug efficacy will also growingly be validated by analysis of an individual's genome, where by analyzing an individual's genetic makeup, healthcare providers can predict how a patient will respond to a specific drug.

As this technology improves, a greater population will have the opportunity to make genome-informed decisions about a variety of health decisions. Genome information can inform nutrition, and reproductive health decisions, and also provide diagnosis to individuals suffering from rare genetic diseases.

CRISPR Cas9 & Genome Editing

The last frontier for engineering—and the current state-of-art in Biology, BME, and Biochemistry— is the design and leveraging of those very systems that run the organic world.

The current best method of influencing biology directly is through DNA and the editing of the bases that govern the expression of all traits. DNA was previously editable, but at great cost. Now this is possible relatively swiftly and cheaply, all through the use of a technology stolen from prokaryotic cells— CRISPR/Cas9. This protein complex enables precise cutting at specific sequences within a genome.

There are, however, some drawbacks that represent areas of future research focus. These complexes cut based on sequence homology, and unintended shared homology represents unintended additional cutting (and associated InDel mutations). This, along with the structural realities of the current protein represents some of the challenges facing the use of Cas9.

Concurrently, researchers employ Cas9 editing in novel approaches to the problems of organ procurement, while also improving upon editing techniques.[13,14,15]

Challenges

Off-Target Mutagenesis

One of the primary challenges facing Cas9 technology derives from the intended functioning (and some unintended factors) of the protein complex. Cas9 cuts bases of DNA based on a guideRNA which is homologous to the region targeted for cutting.[13] If there are other regions of homology, then Cas9 may also induce DBSs at these sites. This represents a serious challenge. Similar issues may arise from another fact of the protein— the PAM sequence. This homology issue necessitates serious consideration when determining a target sequence, and also

when determining a specific Cas9 variant (Cas9s differ in PAM regions).[13]

DNA Repair Pathways

Another issue facing Cas9-based technology is the varying treatment of DNA repair within cells. More probabilistic repair pathways— NHEJ(Non-Homologous End Joining), especially in mammalian cells— prove troublesome for accurate genome editing.[13] These pathways often cause InDels, which result in faulty edits, and unintended mutations. This pathway can be effective for gene knockouts but is less so for the introduction of new genetic material. This risk of error is not present in an alternative pathway, HR(Homologous Recombination).[13] This pathway requires a number of temporally transient pieces of cell machinery to happen. This represents a challenge when these factors are present rarely in post-mitotic and somatic cells, and even further so in mammalian adult cells outside of the S and G2 phases of the cell cycle.[13]

Mosaicism

It remains particularly difficult to distribute accurate, and especially organism-wide changes to a genome. This is derived from difficulties with delivering Cas9 to the relevant cells at the correct time for accurate DNA editing.[13] Thus, if not all cells share a genetic lineage then genes may be expressed unevenly which represents potential complications for the organism.

Approaches

Base Editing

Base editing describes a suite of combined knowledge and techniques related to the CRISPR/Cas9 complex. This method strives for direct conversion of individual bases to other bases without the induction of double-stranded breaks.[14] This method effectively seeks to bypass DNA repair mechanisms and their associated errors. In base editing, a modified version of the Cas9 enzyme (or another base-editing enzyme even) is fused with an enzyme that can chemically alter specific bases. This technique is particularly potent for the correction of point mutations, or in the enaction of small but important and accuracy-dependant changes to DNA.

Traditional Genome Editing

This method is the culmination of much of the work surrounding the Cas9 protein family. This method seeks to ultimately induce a break in dsDNA for the purposes of introducing novel genetic information, or knock-out of already-present genes. This method cleaves DNA, either by “nicking”— where each strand of DNA is cleaved individually to leverage the less error-prone BER (Base Excision Repair) pathways— or by traditional double-stranded break and repair. This single-stranded break method requires that the Cas9 enzyme of choice has one of its two nuclease domains inactivated. “Nicking” also carries a lower probability of DNA being repaired by more accurate HDR (Homology Directed Repair) pathways.[13,14]

These methods are further subdivided into two categories: one which pursues the use of the inaccurate but statistically more prominent DNA repair pathway, NHEJ, and the other which pursues the rarer and more resource-intensive HDR pathway.[1]

Ultimately, the state-of-art today is a combination of so-called DNA “Nicking” and the pursuit of the HDR pathway to repair in the case of insertion.

Current Applications

Inactivation of PERVs in Pursuit of Porcine Xenotransplantation

A shortage of vital human organs for transplant represents a significant medical challenge. Access to vital organ replacements is extremely limited but extremely impactful to patients. Xenotransplantation represents a potential resolution to this challenge. The term refers to the transplantation of animal organs into humans. The highest confidence thus far has been as a result of Porcine Xenotransplantation — organs from pigs (who share many relevant proportions with humans). This method faces challenges, some of which are the presence of Retroviruses within the genome of Pigs (and thus, pig organ cells). These retroviruses or *PERVs*, if transplanted with the organs of their pig hosts, could potentially infect human organ-beneficiaries.

Recently, Cas9 was employed in an attempt to purge the porcine organs of such PERVs. Researchers achieved 100% PERV inactivation in embryos.[15] Researchers were able to implant these embryos, and the resulting offspring were also confirmed to be 100% PERV-free, with confirmation from additional mRNA sequencing of no presence of re-infection.[15]

Future Development

The Future development of Cas9 technology centers on improvements to the control of DNA repair pathways. Research and breakthroughs are also required on the frontier of the distribution of such permanent gene therapies to organisms in the avoidance of mosaicism. Development should also target methods of selecting Cas9 PAMs, where the discovery of additional Cas9s represents increased versatility. Methods in choosing guide RNA for Cas9 should also be an area of focus. These gRNA developments will likely leverage recent advancements in the field of machine learning. Research is also required to improve the stability of the Cas9 structure and its functioning.

Future Applications

Applications for the absolute ability to dictate and design genetic code are already seemingly endless, and this is only a growing area of knowledge. Particularly of interest are medical applications such as precision medicine which is offered to patients based on their specific genotype and condition, gene therapies for heritable conditions, cancer therapies that target oncogenic pathways or the immune system response, or infectious disease control. Applications are even broader, however, including such fields as agriculture and materials science, where command over the genome allows for the design of more productive crops, or biomaterials. Biological pathways may soon be leveraged in the production of industrial chemicals such as fuels, pharmaceuticals, or other anthropogenically relevant compounds.

mmRNA Technologies

In-vitro transcription-modified mRNA, or IVT mmRNA therapeutics, is emerging as one of the most promising technologies. mmRNA therapeutics seek to apply genetic and proteonomic knowledge to the realm of medicine and drug discovery, where engineered nucleases and proteins have the potential to fulfill demands in fields of vaccinology, cancer immunotherapy, enzyme, genetic and metabolic disorders, and immunology. This technology has innate advantages in its specificity on the scale of both individuals as well as cells, development throughput, lack of genomic integration (an advantage in safety — low risk of mutagenesis), and low manufacturing complexity.[16]

In contrast to the bulk of drug and medical technology of the past, this approach leverages biological pathways more than chemistry, and along with this come challenges, however, assessments currently state that this technology has “no platform-inherent major risks”.[16] mmRNA technology faces such challenges in immunogenicity, stability, industry size, and risk mitigation. Developers of mmRNA technology should be wary of haste which introduces unnecessary safety risks, and also of unrealistic expectations from investors and the public.[16] Recent developments in data science and machine learning represent significant growth opportunities for the field, which leverages the big-data nature of DNA, RNA, and protein approaches to medicine.

Challenges

Specificity

This technology leverages cellular pathways; some of which differ from cell to cell. For therapies to be effective, it is necessary that these differences be accounted for.[16] One of the largest areas of such diversity is Post-Translational Modification, where classes of proteins such as glycoproteins are composed in ways not governed by information encodable in the mmRNA.[16] Thus, some cells lack the ability to glycosylate proteins as intended. Additionally, proteolytic post-translational modification differs significantly from cell-to-cell, where essential classes of protein such as “growth factors, cytokines, receptors, neuropeptides, enzymes, hormones, and plasma proteins” are matured.[16] Research should aim to tackle challenges of this category via more specific drug delivery, such as organ-specific delivery, but also in approaching the design of mmRNAs.

Immune response to mRNA

IVT mmRNA technologies must be based on the inoculation of cells with foreign synthetic mRNAs. There is an inherent obstacle to this direct strategy— the cellular immune response to foreign material, especially RNA. This foreign mmRNA, if not synthesized strategically, will be recognized by the immune system as a threat and destroyed before synthesis of relevant polypeptides is possible.[16] The introduction of such foreign material may trigger the destruction of cells expressing the desired trait. Furthermore, in time, and especially with the treatment of immune response dysfunctions, individuals may develop an “immune memory” over time to the same or related mmRNA segments which would require even further strategy.[16] Some mitigatory solutions have been developed, such as the synthesis of mmRNA sequences that resemble natural sequences, or methods of mmRNA transport into targets with lower rates of immunogenicity, but this remains a relevant dimension of challenge to the technology, especially considering any potential for learned immune response.[16,17]

mRNA Instability

The short half-life, owing to the relatively unstable structure of RNA, seemed to be the original roadblock to mRNA technology. This has since largely been addressed by developments in formulations of nanoparticle carriers.[16]

Dosage

Determining dosage for mmRNA-based treatments is complex. Expression levels may be influenced by the careful design of caps, untranslated regions (UTRs), and poly-A tails, but variation between individuals accounts for much of the challenge. Part of the challenge is also inherent in the fact that IVT mmRNA is not

the active agent in treatment— this is instead the encoded protein. Additionally, dosage and treatment are confounded by the presence of other drugs that influence expression of mmRNAs.[16]

Approaches

siRNA

IV siRNA approaches seek to knock out the expression of certain RNAs. siRNAs are RNA molecules that have the specific ability to destroy a target sequence of RNA.[16,17] This allows for the selective elimination of certain expressions through the breakdown of their progenitor mRNA sequences pre-translation. This approach, by disrupting the expression of oncogenic factors, may be used in the treatment of cancers. Additionally, siRNA may be prescribed and delivered to target viral, metabolic, neurological, cardiovascular, or other factors.[17]

mRNA

Treatments can leverage IVT mRNA to induce the expression of a particular polypeptide/trait. This approach includes mRNA molecules which are generated and accumulated, then distributed to increase the expression of a target polypeptide. Here, mRNAs are duplicated, but not modified to increase expression, stability, or reduce immunogenicity.

mmRNA

mmRNA, or modified messenger RNA, includes the strategic design of mRNA for increased expression, decreased immunogenicity, and increased stability. This approach is the current focal point of research into mRNA treatment and is magnitudes more effective.

mRNA delivery vectors

Viral

This method leverages viral vectors to deliver and express mRNA sequences. These factors are efficient at delivering nucleic acids but have as-of-yet incompletely determined efficacy due to immune and general safety concerns.

Non-viral

Non-viral methods of delivery include lipid nanoparticles (LNPs) and polymeric nanoparticles. These methods are considered generally safer and more versatile than approaches that leverage viruses, but this comes at the cost of efficiency in delivery. This method struggles with cell-specific delivery.

Other

Other approaches to mRNA delivery include electroporation and direct injection. These two methods are far less efficient than the aforementioned methods but can be more specific in delivery to exact positions.

Current Applications

Treatment of Inflammatory Leukeocytes in IBD-induced mice

IVT mmRNAs have been used to treat Inflammatory Bowel Disease (IBD) in IBD-induced mice models. In treatment, researchers induced mice to express IBD, then through precision targeting of expression of one gene responsible for the condition, interleukin 10 expressed in inflammatory leukocyte cells, disease activity was reduced. Researchers developed for this purpose a targeting platform, ASSET (Anchored Secondary scFv Enabling Targeting) to achieve the required level of specificity of treatment, This platform involves the use of mmRNA-filled lipid nanoparticles coated in antibodies (mAbs).[17] This suggests that the approach has significant therapeutic potential in IBD conditions and that the modality of treatment— targeted mmRNA— could be used to tackle other conditions as well.[17]

Regenerative Cardiovascular Therapeutics

IVT mRNAs were applied to the field of regenerative cardiovascular therapeutics, and more specifically the regeneration of a specific class of cells labeled “heart progenitor cells.” This required immense specificity in delivery. In this research, mmRNAs were able to induce vascular regeneration following a myocardial infarction (tissue death due to lack of blood supply, common following heart attack)[17]. These heart progenitor cells have the ability to differentiate into any class of cardiovascular cell, and thus regenerate any region of the heart. Researchers applied delivery of mmRNA which encoded for an endothelial growth factor (VEGF-A)[17], to induce regeneration. mmRNA was delivered via direct injection into mice models. This therapy was successful and led to the desired vascular regeneration. This research suggests future use of mmRNA technology to induce similar progenitor-cell activity.[17]

Future Development

Future research on IVT mmRNA technology should center around improving (lessening) immune response to treatments. Additionally, the long-term consequences (and if any exist despite the short half-life of mmRNA molecules) should be investigated further. Researchers should be wary of “moving too fast”[16], for the risk of loss of public trust in the vast potential benefits of the technology. Additionally, recent developments in the field of Machine Learning, and the growing availability of computing power should be leveraged for maximum efficiency of the design of mmRNA sequences and individual specificity.

Future Applications

Future applications for IVT mmRNA technology are vast. They include the addressing of cancers, neurological disorders, rare diseases, genetic disorders, and immune disorders among many others. The future should seek to leverage this technology for the purposes of addressing cancer and tumors at the individual level, with drugs curated to genotype and specific cancer. Additionally, protein, viral, and metabolic needs could be addressed via the intermittent use of such therapies. Overall, the future applications of this technology are seemingly limitless, but the potential risks and drawbacks are also tangible.

Optogenetics

Optogenetics, a field that combines genetics and optics to enable precise control of cellular-level activities of enzymes and other proteins using light, is a revolutionary technique which has the potential to transform our understanding of neural circuits, cellular signaling, and various biological processes. By incorporating light-sensitive proteins such as opsins into target cells, researchers modulate cellular behavior with temporal and spatial precision magnitudes higher than with any other method. Optogenetic systems have potential applications in fields such as neuroscience, and fields like cardiology, endocrinology, and regenerative medicine. The versatility of optogenetics, coupled with its ability to unravel cellular dynamics and prod at specific proteins in an extremely precise manner positions it as a transformative tool with serious implications for the understanding and manipulating of cellular behavior.[19]

However, optogenetics is not without its challenges. Achieving precise and specific expression of light-sensitive proteins within target cells remains the

primary concern, which is only further confounded by the lack of availability and consistency of hardware in the field, as well as facts concerning phototoxicity. Additionally, striking a balance between high expression of key factor levels and maintaining cellular function is crucial for reliable results. Currently, effective light delivery can be limited as well, necessitating the further development of more advanced optical techniques for expansion into in-vivo applications.[19]

Challenges

Light Delivery and Penetration

Effective light delivery to the target cells is crucial for successful optogenetic experiments. The choice of light source, wavelength, and delivery method must be carefully considered to achieve the desired spatial and temporal control. In complex biological tissues, especially in deep brain regions, light penetration can be limited, necessitating the development of advanced optical techniques for optimal activation, especially for future in-vivo applications where tissues are considerably more complex.[19]

Cost/Scalability

The initial investment required for establishing optogenetics laboratories, including the procurement of specialized equipment, optoelectronic devices, and the development of transgenic animal models or viral vectors, can be substantial. High upfront costs currently pose a barrier to entry and limit access. Additionally, there is a significant challenge to the use of optogenetic models in experimentation currently due to a lack of standardization across the field. Addressing such economic concerns and promoting standardization, along with cost-effective solutions will be crucial to the success of optogenetics as a method of inquiry, etc.[19]

Phototoxicity

Phototoxicity poses a significant challenge to optogenetics due to the inherent nature of the technique, which relies on the introduction of exogenous light. However, the intense illumination can inadvertently lead to photodamage, cellular stress, altered physiological responses, and cell death. These adverse effects on the cell can confound data, not only compromising the accuracy and reproducibility of experiments but also raising concerns about the long-term effects on biological tissues if it were to ever be platformed in-vivo.

Hence, mitigating phototoxicity remains a chief objective for optogenetic advancement, requiring innovative strategies such as the development of less invasive light delivery methods, optimization of illumination parameters, and the engineering of photostable light-sensitive molecules. [19]

Low-Light Sensitivity

Going hand-in-hand with previous challenges, sensitivity to low levels of activation light is a critical hurdle to optogenetics. Achieving precise control with minimal light exposure proves difficult. Some molecules only respond at high levels of light or only take effect at high concentrations. Such demands place strain on experiments and potentially cause damage to cells which may be relevant for in-vivo applications. Lowering the intensity of light can demand longer exposure times, potentially causing unintended consequences. This challenge necessitates the development of molecules with higher sensitivity, as well as technology for delivering light with higher precision.[19]

Approaches

Selective Activation of Thalamocortical/Corticothalamic Axons in Treatment of Epilepsy

Researchers were able to identify neural circuits involved in injury-induced epilepsy, specifically focusing on the role of long-range connections in maintaining normal and pathologic oscillations. They found that selective activation of thalamocortical or corticothalamic axons in rats with induced stroke or Rose Bengal treatment led to seizure suppression, suggesting previously unknown roles for long-range connections in maintaining normal and pathologic oscillations. Upon identifying the thalamocortical or corticothalamic axons, researchers further found that targeting a specific spatial region of the thalamus might be an effective therapeutic target for neural activity modulation. It was identified then that selectively (temporally and spatially) silencing these specific structures could serve as effective treatment.[19,20]

As proof, researchers attempted to selectively activate thalamocortical or corticothalamic axons in rats with induced stroke or Rose Bengal treatment. Researchers decided to use a viral vector (stereotactic viral injections) to get rats to express a suite of genes for fluorescent proteins as well as eNpHR, a protein which inhibits neural activity in response to light. Against experimental control, the optogenetic treatment was able to abort seizures within a second of beginning. This suggests efficacy of such an approach in treating similar conditions in humans.[19,20]

Millisecond-Timescale Optical Control of Neural Activity

Researchers induced expression of a fungal protein, Channelrhodopsin-2, a “rapid gated” cation channel, in mammalian neurons, allowing for extremely precise influencing of neural activity via light. Researchers demonstrated the effectiveness of this technique in both cultured hippocampal neurons and in vivo in the brain of living mice. Researchers performed lentiviral (retrovirus) delivery of the gene, then stimulated the cation channel with millisecond-timescale optical stimulus. The method was relatively non-invasive and successful. This method holds immense potential for diverse experimental setups. It has the potential to offer insights into neural circuits and brain function, but also represents an entirely new way of treating and interacting with the human brain and neural circuits. Such a technique has potential applications in the treatment of most immediately, neurological disorders.[19,21]

Compare/Contrast Approaches

The two approaches offer distinct methods of modulating neural activity. While both approaches have their merits, the choice between them depends on the specific experimental and treatment requirements. Viral-mediated delivery, as employed in both studies, offers flexibility and speed in introducing optogenetic tools, making it suitable for short-term research applications. In the future, this technology may demand more exhaustive investigation of the safety of retroviruses as a vector for distribution of the therapy. The primary difference between these two approaches seems to be the selection of light-responsive molecule. One seeks to “down-regulate” neural activity, while the other seeks to do the opposite. Both approaches represent a revolutionary vector for treatment of neural dysfunction and investigation.w

Future Development

The field of optogenetics is poised for significant advancements. The

development of more efficient and potent Opsin proteins with improved properties such as increased sensitivity to light or stronger kinetics are actively being developed. Furthermore, this technology is capable of distributing therapies to many disparate sites and systems. This multiplexing of optogenetic treatments, which would involve the simultaneous control of potentially many different cell types or signaling pathways, holds great promise for revealing truths concerning some of the more currently-intractable problems in neurology, as well as represents a revolutionary means of treatment.

Additionally, advancements in the technology of optogenetics continues. Future devices should allow for increasingly precise—and less invasive— light delivery, enabling optogenetic experimentation in regions of the body which were previously inaccessible due to tissue complexity.

Overall, future advances in optogenetics should focus on shoring up current weaknesses of the technology, and also economization which allows for a global base of research and development, but also on the creation of novel applications and treatments.

Future Applications

Optogenetics has the potential to revolutionize multiple fields of current study. In neuroscience, optogenetics may prove invaluable in the uncovering of underlying mechanisms of complex neural circuitry and neurodegenerative diseases. Beyond the field of neuroscience, optogenetic systems have applications in fields such as cardiology or endocrinology, where the techniques may be used to precisely control any system from the release of hormones to beating of the heart. Furthermore, optogenetics may be employed in the development of therapies which seek to provide precise tissue repair and regeneration.

In conclusion, optogenetic systems represent a new, powerful way of influencing cellular mechanisms. While challenges to the field are rapidly being investigated, continuous advancements in the suite of proteins available along with optical technology promise the potential for groundbreaking discovery and transformative therapy.

Enzyme Engineering

Enzymes, the hyper-specific biological catalysts of reactions, are now the subject of extensive investigation. Today, researchers and industry seek to apply the power and specificity of enzymes to outstanding problems faced by humanity in the spheres of medicine and manufacturing. Unlike most mechanical constructs, enzymes cannot be directly shaped or constructed— they exist on a subcellular scale. Thus, the production and design of enzymes leverage those protein production factories which permeate the natural world— cells. This natural reality of enzymes as proteins also implies a more ecologically sound functioning— one of the main advantages to protein-based biotechnology.[22]

This natural origin of proteins also demands a unique process for development that mirrors the biological reality of the organisms that employ them— evolution. Currently, most approaches for enzyme engineering aim to direct some process of automated selection. This approach demands the highest throughput in facets of alteration, screening, and selection— all areas of current focus for researchers. Currently, these approaches demand significant resources and struggle to produce molecules capable of performing

novel processes not discovered in nature. Some new approaches aim to tackle these challenges.[22]

Challenges

Screening/Selection

One of the more intractable challenges faced by the field seems to be the efficient screening and selection of enzyme variants. This process seems to be necessarily time and resource-intensive. Much of the current research focuses on "scaling up"— increasing efficiency in operation. This currently chiefly includes microfluidics and computational approaches to determining the conformation and sequence of potentially relevant enzymes.[22]

Novel Catalyses

Current approaches to enzyme design share a commonality — the necessity for a progenitor enzyme sequence. Often this is an enzyme that performs the desired function but with some weakness. Here the process of design is comparatively straightforward— multiply with variation and select. These approaches become less straightforward as solutions to problems that henceforth have no solution yet discovered in nature; that is, no enzyme exists with a function related to the relevant process. Currently, research aiming to overcome this drawback centers on evolving enzymes from non-enzymatic protein sequences, or from the use of artificial metalloenzymes.[22]

Cost/Scalability

The governor of any ambition— cost— is a primary dimension of concern currently for techniques that aim to engineer enzymes. The extremely high-throughput nature of methods represents one cost, whereas the purification of selected enzymes is another area of note. Both these components of the enzyme design pipeline are currently resource-intensive and necessitate more cost-effective processes.[22]

Approaches

Dual-Channel Microfluidic Droplet Screening (DMDS)

This method employs lab-on-a-chip microfluidics architecture, which enables a multi-fold higher throughput than other methods. This system employs a dual-fluorescence detection/sorting mechanism within the chip, which allows for the simultaneous evaluation of two reaction channels.[23] This system varies the ratio of two-color fluorogenic substrates allowing for simultaneous selection along two axes of interest, be it improved catalytic activity, regioselectivity, enantioselectivity, chemoselectivity, etc. This method additionally carries advantages in its cost-effectiveness. This method is, however, limited in that it can only select for properties detectable via fluorescence.[23,24] This method is also not a direct model for enzyme design. This method leverages selectivity and high-throughput statistical approaches to develop relevant proteins, and it is thus capable of translating direct dictation of specifications into an enzymatic format.

Absorbance-Activated Droplet Sortation (AADS)

This approach seeks to expand the range of protein design into categories formerly not available to high-throughput methods by allowing for the selection of non-fluorescence-based traits. Previously, fluorescence-based readouts and selectivity were the only option. This method is in contrast, capable of developing enzymes for which no fluorescent assays are available.[24] Additionally, this method differs in that it carries advantages to previous fluorescence-based techniques in that fluorogenic compounds are often bulky and obscure truly accurate assessment of

enzyme characteristics. This method, in contrast, can provide more accurate kinetic data, which is of significant value to the effect of creating more efficiently functioning enzymes.[24] However, This method currently carries the limitation of only being able to screen enzymes that are available in water-in-oil emulsion droplets (this limitation is shared by many microfluidics approaches).[23,24]

Compare/Contrast Methods

Dual-Channel Microfluidic Droplet Screening, employing lab-on-a-chip microfluidics, enables high throughput and simultaneous evaluation of two reaction channels. It is cost-effective but limited to properties detectable via fluorescence. Absorbance-Activated Droplet Sortation expands trait selection beyond fluorescence-based methods, offering more precise kinetic data, but is less mature. Both are lab-on-a-chip in architecture, and both thus require water-in-oil emulsion droplets for screening, restricting the range of assessable enzymes. While the former excels in throughput and cost-effectiveness, the latter stands out for its capacity to assess non-fluorescent traits and provide accurate kinetic data.[23,24]

Current Applications

700-fold Increase in Enantioselectivity of Drug Enzyme with DMDS

Researchers employed a DMDS platform to enhance the enantioselectivity of an esterase for synthesizing specific enantiomers of anti-inflammatory drugs. Profens, a chiral molecule widely distributed for the treatment of inflammation, have curative or aggravative effects depending on their enantiomerization. (S)-enantiomers are desired, but drugs usually include unfortunately racemic mixtures of both. Thus, it would be highly advantageous to develop a highly specific enzymatic method of producing the desired molecular confirmation. Here, in only five rounds of directed evolution using the DMDS platform, researchers successfully identified an enzyme variant with a remarkable 700-fold improvement in enantioselectivity for the desired (S)-profens. This evidences the DMDS platforms' potential as a huge breakthrough in the development of highly specific enzymes.[23]

Engineered Evolution of Dehydrogenaze Enzyme With AADS

Researchers employed AADS technology to deliberately evolve a dehydrogenase enzyme, wtPheDH. First, researchers generated and screened a library of half a million variants of wtPheDH. The AADS device sorted the droplets based on absorbance readout, allowing researchers to identify variants with significantly improved activity towards the substrate, L-Phe. Based on specific readouts generated by the platform, researchers managed to identify variants with improved activity. This experiment served as a proof-of-concept for AADS' fluorescence-independent effectiveness in the directed evolution of enzymes. Additionally, researchers compared the performance of the AADS method to that of the standard-industrial practice DMDS system and found that the AADS method was more efficient, boasting a lower rate of false positives.[24]

Future Development

Research should focus on increasing the throughput of existing methods, but also on the development of completely novel methods that aim to fulfill the greater needs that are inherent to the growing field of engineered enzymes. This includes factors such as cost, useability, and versatility. Future methods should seek to increase the rapidity with which enzyme

variants of interest are recognized, or these methods should center on complete dictation of the features and specificities of completely novel enzymes.

Much of future research should center on the application of breakthroughs in adjacent fields of machine learning and artificial intelligence, where large datasets may be leveraged to create more specific templates or starting enzymes for novel purposes.

Additionally, significant opportunity exists in the development of scaling enzyme production following discovery.

Future Applications

Many of the most obvious applications of engineered enzymes lie in the field of medicine, where evolved proteins have the potential to revolutionize drug synthesis, enhance the specificity of treatments, and reduce side effects. In the future, tailored enzymes could prove pivotal in the synthesis of complex pharmaceutical compounds with greater efficiency and precision, which could significantly contribute to the development of more effective, environmentally aware, and personalized therapeutic interventions. Furthermore, such engineered enzymes could prove instrumental in the creation of highly specific drug delivery systems that aim to ensure that medications reach their intended destinations in the body with heightened accuracy. This field represents a new era of medicine, where highly specific and potentially personal enzyme technology could significantly improve many medical outcomes

Conclusion

In conclusion, the paper aims to shine a light on the current dynamic landscape of genetic engineering research, and some recent compelling applications, underscored by the power of DNA sequencing, CRISPR/Cas9 genome editing, IVT mmRNA therapeutics, and the convergence of data science with biological pathways. The potential applications and challenges of these technologies are unlimited. It is evident that the building wealth of knowledge surrounding the code which underwrites all of biology is beginning to reap tangible reward to medicine and technology. In this paper, optogenetics offers a glimpse into a future of precise cell-by-cell control over gene expression, while enzyme engineering also represents an entirely new field of medicine and technology. This overview serves as a introduction to the current state of such technologies, for researchers, investors, and the public alike.

References

1. Kircher, M., & Kelso, J. (2010). High-throughput DNA sequencing - concepts and limitations. *BioEssays*, *32*(6), 524–536. https://doi.org/10.1002/bies.200900181
2. Helicos Biosciences Files for Chapter 11 bankruptcy protection. (n.d.). https://www.genomeweb.com/sequencing/helicos-biosciences-files-chapter-11-bankruptcy-protection
3. Cheng, C., Fei, Z., & Xiao, P. (2023). Methods to improve the accuracy of next-generation sequencing. *Frontiers in Bioengineering and Biotechnology*, *11*. https://doi.org/10.3389/fbioe.2023.982111
4. Amarasinghe, S. L., Su, S., Dong, X., Zappia, L., Ritchie, M. E., & Gouil, Q. (2020b, February 7). *Opportunities and challenges in long-read sequencing data analysis - genome biology*. BioMed Central. https://genomebiology.biomedcentral.com/articles/10.1186/s13059-020-1935-5
5. *Overview of pacbio SMRT sequencing: Principles, workflow, and applications*. Overview of PacBio SMRT sequencing: principles, workflow, and applications - CD Genomics. (n.d.). https://www.cd-genomics.com/pacbio-smrt-system-single-molecule-real-time-sequencing.html#:~:text=PacBio%20SMRT%20Sequencing%20uses%20the,glass%20surface%20of%20a%20ZMW.
6. *Nanopore sequencing: Principles, platforms and advantages*. Nanopore Sequencing: Principles, Platforms and Advantages - CD Genomics. (n.d.). https://www.cd-genomics.com/nanopore-sequencing-principles-platforms-and-advantages.html
7. Amarasinghe, S. L., Su, S., Dong, X., Zappia, L., Ritchie, M. E., & Gouil, Q. (2020a, February 7). *Opportunities and challenges in long-read sequencing data analysis - genome biology*. BioMed Central. https://genomebiology.biomedcentral.com/articles/10.1186/s13059-020-1935-5
8. Sebastião, M., Ho, R., de Carvalho, J. P., & Nussbaum, M. (2020). Diagnostic accuracy of next generation sequencing panel using circulating tumor DNA in patients with advanced non-small cell lung cancer: A systematic review and meta-analysis. *Journal of Health Economics and Outcomes Research*, 158–163. https://doi.org/10.36469/jheor.2020.17088
9. Sereika, M., Kirkegaard, R. H., Karst, S. M., Michaelsen, T. Y., Sørensen, E. A., Wollenberg, R. D., & Albertsen, M. (2021). Oxford nanopore R10.4 long-read sequencing enables near-perfect bacterial genomes from pure cultures and metagenomes without short-read or reference polishing. *Nature*. https://doi.org/10.1101/2021.10.27.466057
10. Crossley, B. M., Bai, J., Glaser, A., Maes, R., Porter, E., Killian, M. L., Clement, T., & Toohey-Kurth, K. (2020). Guidelines for Sanger sequencing and Molecular Assay Monitoring. *Journal of Veterinary Diagnostic Investigation*, *32*(6), 767–775. https://doi.org/10.1177/1040638720905833
11. Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., Hall, K. P., Evers, D. J., Barnes, C. L., Bignell, H. R., Boutell, J. M., Bryant, J., Carter, R. J., Keira Cheetham, R., Cox, A. J., Ellis, D. J., Flatbush, M. R., Gormley, N. A., Humphray, S. J., … Smith, A. J. (n.d.). *Accurate whole human genome sequencing using reversible terminator chemistry*. Nature News. https://www.nature.com/articles/nature07517
12. Shendure, J., Porreca, G. J., Reppas, N. B., Lin, X., McCutcheon, J. P., Rosenbaum, A. M., Wang, M. D., Zhang, K., Mitra, R. D., & Church, G. M. (2005). Accurate multiplex polony sequencing of an evolved bacterial genome. *Science*, *309*(5741), 1728–1732. https://doi.org/10.1126/science.1117389
13. Hsu, P. D., Lander, E. S., & Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, *157*(6), 1262–1278. https://doi.org/10.1016/j.cell.2014.05.010
14. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., & Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*, *533*(7603), 420–424. https://doi.org/10.1038/nature17946
15. Niu, D., Wei, H.-J., Lin, L., George, H., Wang, T., Lee, I.-H., Zhao, H.-Y., Wang, Y., Kan, Y., Shrock, E., Lesha, E., Wang, G., Luo, Y., Qing, Y., Jiao, D., Zhao, H., Zhou, X., Wang, S., Wei, H., … Yang, L. (2017). Inactivation of porcine endogenous retrovirus in pigs using CRISPR-Cas9. *Science*, *357*(6357), 1303–1307. https://doi.org/10.1126/science.aan4187
16. Sahin, U., Karikó, K., & Türeci, Ö. (2014). MRNA-based therapeutics — developing a new class of drugs. *Nature Reviews Drug Discovery*, *13*(10), 759–780. https://doi.org/10.1038/nrd4278
17. Veiga, N., Goldsmith, M., Granot, Y., Rosenblum, D., Dammes, N., Kedmi, R., Ramishetti, S., & Peer, D. (2018). Cell specific delivery of modified mrna expressing therapeutic proteins to leukocytes. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-06936-1
18. Zangi, L., Lui, K. O., von Gise, A., Ma, Q., Ebina, W., Ptaszek, L. M., Später, D., Xu, H., Tabebordbar, M., Gorbatov, R., Sena, B., Nahrendorf, M., Briscoe, D. M., Li, R. A., Wagers, A. J., Rossi, D. J., Pu, W. T., & Chien, K. R. (2013). Modified mrna directs the fate of heart progenitor cells and induces vascular regeneration after myocardial infarction. *Nature Biotechnology*, *31*(10), 898–907. https://doi.org/10.1038/nbt.2682
19. Repina, N. A., Rosenbloom, A., Mukherjee, A., Schaffer, D. V., & Kane, R. S. (2017). At light speed: Advances in optogenetic systems for regulating cell signaling and behavior. *Annual Review of Chemical and Biomolecular Engineering*, *8*(1), 13–39. https://doi.org/10.1146/annurev-chembioeng-060816-101254
20. Paz, J. T., Davidson, T. J., Frechette, E. S., Delord, B., Parada, I., Peng, K., Deisseroth, K., & Huguenard, J. R. (2012). Closed-loop optogenetic control of thalamus as a tool for interrupting seizures after cortical injury. *Nature Neuroscience*, *16*(1), 64–70. https://doi.org/10.1038/nn.3269
21. Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G., & Deisseroth, K. (2005). Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neuroscience*, *8*(9), 1263–1268. https://doi.org/10.1038/nn1525
22. Zeymer, C., & Hilvert, D. (2018). Directed evolution of protein catalysts. *Annual Review of Biochemistry*, *87*(1), 131–157. https://doi.org/10.1146/annurev-biochem-062917-012034
23. Ma, F., Chung, M. T., Yao, Y., Nidetz, R., Lee, L. M., Liu, A. P., Feng, Y., Kurabayashi, K., & Yang, G.-Y. (2018). Efficient molecular evolution to generate enantioselective enzymes using a dual-channel microfluidic droplet screening platform. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-03492-6
24. Gielen, F., Hours, R., Emond, S., Fischlechner, M., Schell, U., & Hollfelder, F. (2016). Ultrahigh-throughput–directed enzyme evolution by absorbance-activated droplet sorting (AADS). *Proceedings of the National Academy of Sciences*, *113*(47). https://doi.org/10.1073/pnas.160692