UNDERGRADUATE RESEARCH IN NATURAL AND CLINICAL SCIENCE AND TECHNOLOGY (URNCST) JOURNAL
Read more URNCST Journal articles and submit your own today at: https://www.urncst.com

ENCYCLOPEDIA ENTRY **OPEN ACCESS**

Enhancing Genomic Medicine with AI-Integrated CRISPR-Cas9 Technologies

Alejandro Diaz Gonzalez, HBSc Student [1]*

[1] Faculty of Science, University of Western Ontario, London, Ontario, Canada N6A 3K7

Corresponding Author: adiazgon@uwo.ca

Abstract
Introduction: CRISPR-Cas9, a ground breaking gene editing tool, has transformed genetic engineering. With the advent of Artificial intelligence (AI) in recent years, this tool has been applied to the development of new CRISPR-Cas9 therapies. This integration with artificial intelligence (AI) “Helps genome editing achieve more precision, efficiency, and affordability in tackling various diseases”
Methods: A comprehensive literature review was conducted, with the assistance of AI to compile and assess the current state and future prospects of integrating AI with CRISPR-Cas9 technology in genomic medicine. Relevant articles were identified through database searches in PubMed, Web of Science, and Google Scholar, using keywords such as "CRISPR-Cas9," "artificial intelligence," "machine learning," "gene editing," and "genomic medicine." Studies published up to June 2024 were considered.
Current Research and Findings: Current research revolves around the development of CRISPR systems with improved efficiency and specificity, facilitated by advanced machine learning models. These focus specifically on sgRNA development and its consequences on genetic research.
New Research and Implications on Future Directions: New research avenues suggest exploring CRISPR’s use in complex genetic disorders involving multiple genes. AI's predictive capabilities are vital in designing multi-target strategies for such complex conditions. These include the diagnosis and treatment of cancer, the early identification of rare diseases, and the faster design of vaccines.
Summary: The convergence of these advanced technologies offers a pathway to more precise and personalized therapeutic interventions. By leveraging AI's capabilities in data analysis and pattern recognition, researchers can enhance the accuracy and efficiency of CRISPR-Cas9 gene editing. AI aids in predicting the most effective guide RNA (gRNA) designs, reducing off-target effects, and improving the specificity of gene edits. This synergy not only accelerates the gene editing process but also expands its applications across various medical fields. The ongoing refinement of AI algorithms and their integration with CRISPR technology could unlock new possibilities in medical science, potentially revolutionizing the diagnosis and treatment of genetic and multifactorial diseases.

Keywords: CRISPR-Cas9; artificial intelligence; genomic medicine; gene editing; machine learning

Introduction and Definition

The integration of Clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 complexes with artificial intelligence (AI) marks a pivotal advancement in genomic medicine, revolutionizing our approach to disease treatment and prevention. CRISPR-Cas9, a gene editing technology originally derived from a bacterial immune defense system [1], enables precise modifications at specific locations in the genome [2]. This capability is enhanced by AI, which aids in designing highly specific guide RNAs (gRNAs) and predicting potential off-target effects, thereby increasing the accuracy and safety of gene editing procedures [3].

AI technologies apply machine learning and deep learning models to handle and analyze large datasets, such as the Ensembl Genome Browser, the UCSC Genome Browser, and dbSNP. These resources provide detailed genomic sequences, annotations, and information on genetic variations across different populations. This data aids in identifying relevant unique target sites and avoiding regions with high similarity to other sequences across the genome, necessary for the design of optimal gRNA sequences. [4]. This not only improves the efficiency of the CRISPR-Cas9 system but also reduces the likelihood of unintended genetic modifications, which are crucial for clinical applications.

The synergy between CRISPR-Cas9 and AI is not merely a technical enhancement; it represents a transformative shift toward personalized medicine. By leveraging genetic data and AI's predictive capabilities, medical treatments can be tailored to individual genetic profiles. [5]. These profiles encompass information about an individual's DNA sequence, a curated subset of this data that highlights specific genetic markers of interest, variants, or mutations relevant to certain traits, diseases, or responses to treatments. This approach holds the potential to dramatically improve treatment outcomes by enabling therapies that are customized to the unique profile of each patient [6].

As research continues, the integration of CRISPR-Cas9 and AI is expected to yield further innovations in genomic medicine, offering new strategies for diagnosing and treating diseases at a molecular level. This convergence promises to enhance our understanding of genetic and multifactorial diseases, potentially leading to more effective interventions and a higher standard of care in medicine.

History

The CRISPR-Cas9 system, discovered as a component of a bacterial immune mechanism in *E. coli* by Ishino et al. in 1987 [1], has evolved significantly. In 2012, a group affiliated with the Howard Hughes Medical Institute demonstrated the potential transformation of CRISPR-Cas9 from a naturally occurring bacterial defence mechanism into a programmable genome editing tool. Led by Martin Jinek, a landmark study was published showcasing how the system could be adapted to function in eukaryotic cells, using guide RNA to direct Cas9 to specific DNA locations, thereby enabling precise genetic modifications across a range of organisms [2]

Due to this precise modification capacity in living organisms, the CRISPR-Cas9 system has been increasingly adopted in medical research and treatment. This tool is being harnessed in medicine to develop treatments for genetic disorders, where specific gene mutations can be corrected or altered to mitigate disease symptoms. . Today, the convergence of AI and CRISPR-Cas9 is not only streamlining the pathway for more precise genetic edits but is also shaping the future of personalized medicine. By integrating detailed genetic profiles with predictive AI models, it is possible to tailor genetic interventions to individual genetic profiles, thereby enhancing the efficacy and safety of treatments [6].

CRISPR-Cas9 Fundamental Concepts

The core components of the CRISPR-Cas9 system include the Cas-9 nuclease and a guide RNA (gRNA). The gRNA is a synthetic RNA composed of a scaffold sequence necessary for Cas9-binding and a user-defined ∼20 nucleotide spacer sequence that dictates target specificity [11].

Upon introduction into a cell, the Cas9-gRNA complex undergoes a conformational change enabling the gRNA to base-pair with its complementary DNA target, typically adjacent to a protospacer adjacent motif (PAM) - a short DNA sequence essential for Cas9 recognition and cleavage. The binding of the gRNA to its target sequence positions Cas9 to introduce a double-strand break (DSB) at a precise location within the genome [11].

The cell's innate repair mechanisms, primarily non-homologous end joining (NHEJ) or homology-directed repair (HDR), then address the DSB. NHEJ, an error-prone process, can lead to insertions or deletions (indels) at the repair site, potentially disrupting or altering gene function. Alternatively, the more precise HDR facilitates the introduction of specific mutations or gene inserts by using a donor DNA template with homology to the target site [11].

"Off-target" effects can occur when the CRISPR-Cas9 complex inadvertently binds and cuts DNA sequences similar but not identical to the intended target sequence. These unintended modifications can lead to genomic instability or unintended mutations, impacting safety, and efficacy [12] The likelihood of off-target effects is influenced by the target genomic sequence itself; sequences that have high similarity or overlap with bordering regions of the genome are prone to unintended interactions. Therefore, careful selection and design of target sequences are crucial to minimize overlap with similar genomic regions and reduce the risk of these effects.

Artificial Intelligence Fundamental Concepts

Although commonly vaguely defined, AI represents a significant domain within computer science, dedicated to developing computational mechanisms that emulate human cognitive functions such as learning, reasoning, and self-correction. These include an array of fields applicable to genomic research:

Machine Learning (ML), a core subset of AI, enables computational models to autonomously improve their performance on tasks through exposure to data. This methodology relies on statistical techniques that allow machines to "learn" from data, identifying patterns and making decisions with minimal human directive. ML models adapt through a process where they adjust internal parameters based on the input data they receive, continually improving their accuracy in task performance [14].

Deep Learning (DL), a specialized subset of ML, employs artificial neural networks with multiple layers of processing. Each layer uses a set of algorithms modeled on the human brain that are designed to interpret data features with varying abstraction levels [14]. Deep learning models perform exceptionally well on tasks that involve large amounts of data and complex patterns, such as speech and image recognition [14].

At the heart of many AI applications, neural networks consist of interconnected nodes, akin to biological neurons. These networks use activation functions that determine the output at each node based on inputs [14]. Learning occurs as the network outputs are adjusted in response to the discrepancy between actual and desired outputs, a process known as backpropagation [14].

This concept, coupling ML and DL with user interactions can train an AI model to provide desired responses to mass amounts of data and can streamline the workflow of processing CRISPR experiment results. As seen in the next section, these include finding the most optimal PAM sites, or the most effective sgRNA structures.

Current Research and Findings

The effectiveness of CRISPR-Cas9 gene editing is significantly influenced by the design of the gRNA, which determines the specificity and efficiency of the target DNA cleavage [3]. The accuracy of gRNA design is thus crucial for minimizing off-target effects and enhancing on-target activity. Current studies applying AI to CRISPR-Cas9 systems have focused on this barrier.

In a 2022 study, researchers analyzed and compared the performance of existing gRNA efficiency prediction tools, categorizing them based on their methodological approaches. These include a rules-based approach, which considers factors such as the number, position, and type of mismatches between the gRNA and potential off-target DNA sequences. They also include statistical models that quantify the relationship between sequence features and the likelihood of off-target activity. Finally, they assessed a machine learning approach by training algorithms on high-dimensional data, capturing nonlinear relationships that simpler models might miss. In this study, the machine-learning approach integrated a wide array of sequence features, including nucleotide composition, secondary structures, chromatin accessibility, and epigenetic markers to predict Cas9 activity more accurately. The study highlights that while traditional models rely on general features of gRNA design, such as guanine and cytosine (GC) content and target sequence homology, modern machine learning approaches, particularly deep learning, can integrate complex patterns and interactions that are less obvious to human experts [4].

This study utilized a recently developed machine learning tool named Azimuth 2.0 as a model to predict both the on-site and off-site effects of specific sgRNAs and rank them on a scale to find the most adequate one [4]. It not only utilizes the commonly used models mentioned before, but it also processes an online library of genomic data. This data includes the cutting frequency of sgRNAs (CFD) score, quantifying the efficiency and specificity of sgRNAs in guiding the CRISPR-Cas9 system to cut DNA at intended target sites [4]. This score evaluates how frequently a sgRNA leads to a successful cut at the designated genomic location, a crucial factor in CRISPR experiments, as higher cutting frequencies increase the likelihood of effective gene editing while minimizing off-target effects [10].

The risk of using a chosen sgRNA is assessed using a CRISPRko scale, which identifies possible off-target sites where the CRISPR-Cas9 complex could bind, and ranks them from most to least likely to result in off-target effects. This ranks from known coding regions to non-coding regions without any known genetic activity [4]. Once these parameters are assessed, after obtaining them from the relevant databases, the model will produce a list of potential sgRNAs that are deemed to be most efficient and produce the least non-desirable effects.

As a constantly learning model, it can learn from past outputs. For example, if a previously designed sgRNA is deemed to be inefficient, or it has produced unexpected off-site effects, this data will be automatically uploaded to the cloud and will condition its future reasoning and the reasoning of all Azimuth models used worldwide by other research teams [14].

Researchers have also acknowledged that in-vitro sgRNA efficiency is also influenced by factors such as a cell’s microenvironment, or experimental conditions, also currently developing ML models to account for these variables and their potential effects on sgRNA outcomes [15]. This growth is encouraged by the open-source nature of the research, as models worldwide contribute to an ever-growing dataset through an interconnected cloud that assesses the effectiveness of different experimental conditions.

New Research and Implications on Future Directions

The advent of AI, coupled with numerous other technological advances in medicine has the potential to revolutionize the next generation of genomic medicine. This includes the convergence of AI-driven gene machinery with personalized medicine, which aims to tailor healthcare to the individual scale. AI assessment of genetic information, driven by the ample data obtained from CRISPR-Cas9 systems research can streamline the process of obtaining genetic information from patients, assessing phenotypic changes that their environment could have caused, and overall streamline and cheapen the cost of such care [16].

Clinical trials using CRISPR-Cas9 are underway for conditions such as sickle cell disease and beta-thalassemia. These trials focus on correcting the genetic defects causing these hematologic diseases, by disrupting the enhancer region of the BCL11A gene, a key regulator that suppresses fetal hemoglobin (HbF) production in adults. By editing this enhancer, the repression of HbF is lifted, leading to increased production of HbF in red blood cells, which can function effectively in place of defective adult hemoglobin [7, 8].

Avenues have also been identified to implement AI CRISPR-Cas9 systems into oncological therapy and diagnostics. CRISPR can be utilized to analyze cancer mechanisms through detailed individual genomic screenings, and the creation of sophisticated cancer models that predict the success of potential therapies on individual patients, depending on their type of cancer and genetic profile [5]. On the therapeutic front, CRISPR shows promise in directly modifying cancerous genomes, as well as engineering T-cells to improve immune responses against tumors [5]. CRISPR-based diagnostic tools can detect cancer-specific genetic alterations, thus facilitating early diagnosis and tailored treatment strategies [5]. If this is combined with the efficiency and speed at which ML models can process and interpret data, this type of care could be democratized much faster than current projections suggest.

Applications have also been proposed for identifying potential rare diseases before the mutations manifest, with researchers training a DL network to identify and predict rare diseases in primates. This model used the predicted changes in gene sequences to predict the pathological manifestation of said mutations, as well as how likely the mutation would be, with an 88% accuracy rate [17]. If this was applied to humans at a mass scale with the help of AI, rare diseases could be investigated, and possible treatment developed, before they even manifested in populations.

Finally, the ability of AI CRISPR-Cas9 systems to create accurate models can also be applied to the design of vaccines [18], with pharmacological compounds being assessed for their effectiveness without the need to use live models. This could streamline the process of choosing which are ready to be tested in animal models and decrease the time to market, as well as potentially lead to a future where vaccines are designed at an individual scale, depending on the pathological predispositions of each patient in a population.

Conclusion

The convergence of CRISPR-Cas9 technology with AIcould represent a forward leap in genomic medicine, paving the way for more precise and personalized therapeutic interventions. CRISPR-Cas9 can be refined by AI to improve its accuracy and efficiency, significantly enhancing its clinical applicability. By leveraging AI's capabilities in data analysis and pattern recognition, researchers can already predict the most effective gRNA designs, reducing off-target effects and improving the specificity of gene edits, and consequently the success of future genetic research.

This synergy between CRISPR-Cas9 and AI not only accelerates the process of gene editing but also broadens its potential applications across various medical fields. AI-driven CRISPR systems could open the door to the development of personalized medicine strategies, maximizing therapeutic efficacy and minimizing side effects.

As this field evolves, the ongoing refinement of AI algorithms and the continuous integration with CRISPR technology will likely unlock new possibilities in medical science. This advancement holds the promise to revolutionize how we understand, diagnose, and treat genetic and multifactorial diseases, making a significant impact on global health outcomes. The journey of CRISPR-Cas9 from a basic bacterial defense mechanism to a cornerstone of genomic medicine could represent the potential AI technologies hold to reshape the future of healthcare.

List of Abbreviations

AI: artificial intelligence
CFD: cutting frequency determination
DL: deep learning
DSB: double stranded break
GC: guanine and cytosine
HDR: homology directed repair
ML: machine learning
NHEJ: non-homologous end joining
PAM: photospacer adjacent motif
sgRNA: single guide ribonucleic acid

Conflicts of Interest

The author(s) declare that they have no conflict of interests

Authors' Contributions

ADG: Made contributions to the design of the entry, collected and analysed data, drafted the manuscript, and gave final approval of the version to be published.

Acknowledgements

ChatGPT 4.0 was used in the assistance of writing this article.

Funding

The development of this encyclopedia entry was not funded.

References

[1] Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. J Bacteriol. 1987 Dec;169(12):5429-33. https://doi.org/10.1128/jb.169.12.5429-5433.1987.
[2] Jinek M, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012;337(6096):816-21. https://doi.org/10.1126/science.1225829
[3] Konstantakos V, Nentidis A, Krithara A, Paliouras G. CRISPR–Cas9 gRNA efficiency prediction: an overview of predictive tools and the role of deep learning. Nucleic Acids Res. 2022 Apr 22;50(7):3616-37. https://doi.org/10.1093/nar/gkac192
[4] Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat Biotechnol. 2016;34(2):184-91. https://doi.org/10.1038/nbt.3437
[5] Katti A, Diaz BJ, Caragine CM, Sanjana NE, Dow LE. CRISPR in cancer biology and therapy. Nat Rev Cancer. 2022;22:259-79. https://doi.org/10.1038/s41568-022-00441-w

[6] Bhat AA, Nisar S, Mukherjee S, Saha N, Yarravarapu N, Lone SN, et al. Integration of CRISPR/Cas9 with artificial intelligence for improved cancer therapeutics. J Transl Med. 2022;20:534. https://doi.org/10.1186/s12967-022-03765-1
[7] Frangoul H, Altshuler D, Cappellini MD, Chen YS, Domm J, Eustace BK, et al. CRISPR-Cas9 gene editing for sickle cell disease and beta-thalassemia. N Engl J Med. 2021;384:252-60. https://doi.org/10.1056/nejmoa2031054
[8] Thompson AA, Walters MC, Kwiatkowski J, Rasko JE, Ribeil JA, Hongeng S, Magrin E, et al. Gene therapy in patients with transfusion-dependent beta-thalassemia. N Engl J Med. 2021;376:848-60. https://doi.org/10.1056/nejmoa1705342
[9] Stadtmauer EA, Fraietta JA, Davis MM, Cohen AD, Weber KL, Lancaster E, et al. CRISPR-engineered T cells in patients with refractory cancer. Science. 2020;367. https://doi.org/10.1126/science.aba7365
[10] iu Q, He D, Xie L. Prediction of off-target specificity and cell-specific fitness of CRISPR-Cas System using attention boosted deep learning and network-based gene feature. PLoS Comput Biol. 2019;15(10). https://doi.org/10.1371/journal.pcbi.1007480
[11] Adli M. The CRISPR tool kit for genome editing and beyond. Nat Commun. 2018;9:1911. https://doi.org/10.1038/s41467-018-04252-2
[12] Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S, et al. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome research. 2014 Jan 1;24(1):132-41. https://doi.org/10.1101/gr.162339.113
[13] Alpaydin E. Introduction to machine learning. MIT press; 2020 Mar 24.
[14] Chuai G, Ma H, Yan J, Chen M, Hong N, Xue D, et al. DeepCRISPR: optimized CRISPR guide RNA design by deep learning. Genome Biol. 2018;19:80. https://doi.org/10.1186/s13059-018-1459-4
[15] Störtz F, Mak JK, Minary P. piCRISPR: physically informed deep learning models for CRISPR/Cas9 off-target cleavage prediction. Artificial Intelligence in the Life Sciences. 2023 Dec 1;3:100075. https://doi.org/10.1016/j.ailsci.2023.100075
[16] Dixit S, Kumar A, Srinivasan K, Vincent PMDR, Ramu Krishnan N. Advancing genome editing with artificial intelligence: opportunities, challenges, and future directions. Front Bioeng Biotechnol. 2024;11:1335901. https://doi.org/10.3389/fbioe.2023.1335901
[17] Sundaram L, Gao H, Padigepati SR, McRae JF, Li Y, Kosmicki JA, et al. Predicting the clinical impact of human mutation with deep neural networks. Nat Genet. 2018;50(8):1161-70. https://doi.org/10.1038/s41588-018-0167-z
[18] Maserat E. Integration of Artificial Intelligence and CRISPR/Cas9 System for Vaccine Design. Cancer Informatics. 2022;21. https://doi.org/10.1177/11769351221140102

Article Information
Managing Editor: Jeremy Y. Ng
Peer Reviewers: Chun Ju Liang, Alita Gideon
Article Dates: Received Jul 08 24; Accepted Nov 26 24; Published Feb 18 25

Citation
Please cite this article as follows:
Diaz Gonzalez A. enhancing genomic medicine with AI-Integrated CRISPR-Cas9 technologies URNCST Journal. 2025 Feb 18: 9(2). https://urncst.com/index.php/urncst/article/view/685
DOI Link: https://doi.org/10.26685/urncst.685