Systematic Review on the Usage of Artificial Intelligence in CRISPR-Cas9 Genome Editing Technology for Organoid Research
Ruoxi Zhu¹, Christine Yoon²

¹ The College Preparatory School, CA, USA (lexi.rxz@gmail.com)
² Albert Einstein College of Medicine, NY, USA (christine.yoon@einsteinmed.edu)

Keywords: Artificial Intelligence, Machine Learning, Deep Learning, CRISPR-Cas9, Genome Editing, Organoid, Stem Cell

Abstract
CRISPR-Cas9 engineered organoids represent the novel application of CRISPR-Cas9 genome editing technology in three-dimensional stem cell cultures, and have gained significant attention in recent years. These organoids, miniaturized 3D structures derived from stem cells, faithfully replicate the functional and structural characteristics of real organs. Their potential for disease modeling, providing insights into human development and disease mechanisms, and their ability to replicate complex organ structures and functions in vitro make organoids invaluable in medical research and drug discovery. The utilization of CRISPR-Cas9 technology enables precise genome editing of the stem cells, thereby enhancing the fidelity and accuracy of organoid models and enabling the study of mutations and diseases that were previously unable to be replicated in vitro. Recently, the integration of artificial intelligence (AI) has emerged as a promising approach to advance this technology. The convergence of CRISPR-Cas9 genome editing technology and AI can speed up and improve the CRISPR-Cas9 process for organoids by analyzing large and/or complicated datasets, predicting CRISPR-Cas9 off-target effects, improving CRISPR-Cas9 splicing accuracy, and even creating better CRISPR-Cas9 gene knockout plans. Though the technology is still at an early stage, this comprehensive review discusses the current and future utilization of AI in CRISPR-Cas9 genome editing technology for organoid creation, research, and improvement.

Introduction
Background
Organoids are a fairly new technology where researchers create individualized and often complex collections of cells in a lab that mimic organs in a patient’s body. This technology allows them to view the formation and growth process of organs in vitro under a controlled environment, which can help provide valuable information on human development and disease studies. Because organoids are much easier to manipulate and study compared to animal models, organoid cultures have been used for drug discovery, personalized medicine diagnostics, cell therapy, and more (2). Furthermore, as organoids are unique in that they provide insights into how organs work in vitro, they are a revolutionary stepping stone in simplifying the complicated processes of organ research.

Importantly, various AI models have recently emerged in the scientific field, combining with CRISPR-Cas9 and organoid technology to further organoid research. Specifically, computational algorithms based on machine learning and deep learning have been developed in order to aid in the facilitation of using CRISPR-Cas9 technology on organoids (13). As the specific link between AI in CRISPR-Cas9 technology for organoid research has yet not been
studied in depth, this literature review aims to effectively summarize how AI has been used in CRISPR-Cas9 genome editing technology for the advancement of organoid research, as well as provide background information on the intersection between organoids and CRISPR-Cas9 technology independent of artificial intelligence.

Methods
Overview and Scope
This literature review aims to present a holistic summary and analysis of how artificial intelligence (AI) is being used to improve CRISPR-Cas9 genome-editing technology for organoid creation, both now and in the future.

Inclusion and Exclusion Criteria
The inclusion criteria for this paper are set as full-text papers in the English language published from January 2020 - July 2023, with a global search of results including both research and review articles. Non-English articles, non-full text articles, books, interviews, and editorials were excluded from this review.

Information Sources and Search Strategy
A preliminary literature search on Google Scholar using the keywords mentioned below yielded 159 results, from which 4 were selected based on the inclusion and exclusion criteria outlined above. The relevant articles were acquired electronically. The articles retrieved were checked in-depth for the titles, abstracts, subject headings, and references, and only relevant papers were retained.

Keywords
Keywords used in Google Scholar: "artificial intelligence" AND "machine learning" AND "crispr" AND "organoid" AND "genome editing" OR "gene editing" AND "stem cell"

Data Extraction and Selection Process
Data was collected and extracted from Google Scholar between June 29, 2023 and July 29, 2023.

Limitations
As only sources written completely in the English language were evaluated, information that is available in other languages but may have not been translated into English yet were not included in this review. Furthermore, this literature review only discusses the applications of AI in CRISPR-Cas9 genome editing optimization; a pinpoint analysis of other applications of AI in CRISPR-Cas9 for organoid creation were not discussed in detail.

Main Text
Organoids; An Overview
Organoids are self-organized cell-based three-dimensional tissue cultures that mimic many aspects of the structure and function of organs within a living being (2). They are usually derived from adult stem cells (ASCs) and pluripotent stem cells (PSCs), the latter of which includes both embryonic stem cells (ESCs) as well as induced pluripotent stem cells (iPSCs) (1, 3, 4). These organoids can be crafted to either replicate an entire in vitro organ or to only
express selected aspects of an in vitro organ, such as producing only certain types of cells. Using these stem cells, a variety of tissue-specific organoids have been created, including those mimicking the brain, kidney, lung, intestine, stomach, liver, and more (2).

To create PSC-derived organoids, researchers first attempted to reproduce the tissue development and homeostasis that occur naturally within a developing body. PSCs were then induced in vitro to differentiate down specific lineages to create specialized cell types, made possible due to their unique versatility even among stem cells. By exposing PSCs to specific biochemical factors and 3D scaffolding conditions, differentiated PSCs (including both iPSCs and ESCs) were able to self-organize into different tissue-specific organoids that mimicked their in vivo counterparts, including the optic cup, brain, intestine, liver, and kidney (5). Interestingly, given that PSCs are able to yield cells from all three cell germ layers (ectoderm, mesoderm, and endoderm) in vitro, PSC-derived organoids often consist of cells derived from more than just one germ layer. As such, PSC-derived organoids are useful for facilitating the study of interactions between various cell types, and also provide an accurate model that closely resembles endogenous organs.

Similarly, researchers have also worked with ASCs to successfully generate artificial organoids. Contrary to PSC-derived organoids, ASC-derived organoids do not require directed differentiation down specific lineages, as ASCs are simply directed to form organoids in vitro under an environment well-suited for optimal growth after being extracted from the in vivo organ through tissue dissociation (10). Under these specific culture conditions that support stem cell activity, researchers are able to reproduce the natural organ regeneration process in vitro to control self-renewal and differentiation, creating self-organized tissue organoids such as ones of the intestine, stomach, liver, and pancreas (5). However, though generating ASC-derived organoids requires significantly less time than generating PSC-derived organoids, ASC-generated cell types are more limited than PSC-generated cell types because ASC-derived organoids often only contain epithelial cells (10). Due to this limitation, they are extremely useful for studying the maintenance and regeneration of epithelial tissues, but not for studies pertaining to interactions between different cell types.

Since their discovery, 3D organoid systems have become popular tools used to model organ development, host-pathogen interactions, and diseases (10). Their popularity stems from the fact that they bridge the gap between in vivo animal models and in vitro 2D cell culture systems, the former of which is time-consuming and costly, the latter of which often contains cancer-associated genetic alterations while lacking 3D tissue organization. However, in order to accurately model monogenic diseases and cancer, the abilities to precisely introduce and repair specific genetic mutations in organoid systems were necessary. Here is where CRISPR-Cas9 genome editing technology steps in.

**What is CRISPR-Cas9 Technology?**
Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) refers to a natural antiviral defense mechanism found in bacteria and archaea (11). It consists of a pattern of repetitive palindromic DNA sequences interspersed with unique spacers derived from viral DNA fragments of bacteriophages that had previously infected the bacterial genome (6, 38). These spacers act as a sort of genetic memory of previous infections, and the CRISPR
mechanism uses them to efficiently recognize and destroy specific viral or plasmid DNA sequences in subsequent viral infections (7). On a separate but related note, the CRISPR-associated protein 9 (Cas9) is a bacterial enzyme coded by a CRISPR-associated gene. When used in conjunction with a guide RNA (gRNA) molecule, Cas9 can be directed to a particular location in the bacterial genome, where it recognizes and cuts the target DNA at that specific location to disable the virus as part of the CRISPR process of defending against additional virus attacks (38).

Though three major types of CRISPR-Cas systems have already been identified to date, the type II CRISPR bacterial defense mechanism found in the Streptococcus pyogenes bacteria that uses the Cas9 enzyme has, in recent years, been repurposed into a powerful RNA-guided DNA targeting platform for genome editing, transcriptional perturbation, epigenetic modulation, and genome imaging called CRISPR-Cas9 (7, 8, 11). To utilize the genome editing tool, researchers first create a small piece of RNA, the gRNA, that is able to identify and guide the Cas9 enzyme to the desired location on the target DNA. Once the Cas9 enzyme cuts the DNA at the specified location, researchers are then able to exploit the cell’s own natural DNA repair mechanisms to introduce their desired changes, including inserting, deleting, or modifying specific pieces of genetic material on the target DNA (38). Furthermore, the system enables multiplex targeting, a revolutionary approach where multiple gRNAs or Cas enzymes are expressed at once, making it an efficient option for genetic engineering (9, 39). Thus, CRISPR-Cas9-mediated genome engineering is thought to hold immense promise to treat—or, perhaps, even cure–complicated genetic disorders (7).

**Use of CRISPR-Cas9 Genome Editing Technology in Organoid Development**

Conventional CRISPR-Cas9-mediated genome engineering is especially applicable in the field of organoid development and research. Currently, multiple methods of genetic engineering in organoids have already been established, opening up a new field of organoid research called organoid genetics (10). These methods enable the introduction of modifications of a genomic DNA sequence in a coding sequence, which can result in a specific change in the target protein, providing useful information about the biological role of a specific residue or even the protein itself. Though there are many different tools used for genetic engineering in organoids, including prominent ones such as RNA interference, retro/lentiviruses, and transposons, this paper will focus on the CRISPR-Cas9 method. The usage of the CRISPR-Cas9 system in organoid creation and research can be separated into two categories: CRISPR-Cas9 for ASC-derived organoid models, and CRISPR-Cas9 for PSC-derived organoid models.

**CRISPR-Cas9 Genome Editing in ASC-Derived Organoids**

Since ASCs are isolated directly from tissues within the body, one method of obtaining a genetically engineered organoid is to simply acquire ASCs from a genetically mutant animal line that already has the desired mutation (10, 4). However, as this method is not only difficult and time-consuming but also expensive, studies conclude that it is much more advantageous to directly edit gene expression in organoids using CRISPR-Cas9 technology (10). In fact, CRISPR-Cas9 technology has already been combined with organoid research in several different ways, many of which have been aptly summarized by Teriyapirom et. al. (10).
For example, in 2013, it was revealed that CRISPR-Cas9, facilitated by lipofection, could successfully carry out genetic knockout or repair tasks in intestinal organoids (25). Using the genome editing tool, multiple mutations in the cystic fibrosis transmembrane conductor receptor locus of patient-derived colonic organoids were repaired, demonstrating the potential of CRISPR-Cas9 in correcting monogenic disorders (10, 25).

CRISPR-Cas9 has also been used with organoid models in vitro to imitate the effects of oncogenes in tumor evolution (10). Through manipulating the genes of organoid models, researchers successfully confirmed the multi-hit oncogenesis model that had been previously hypothesized for different forms of cancer, which was then explored independently by two distinct studies (10, 40, 41). Both used the approach of introducing genetic mutations into colonic organoids using CRISPR-Cas9, targeting the APC, SMAD4, TP53, and KRAS genes, specifically. One of them also targeted PI3KCA, while the other chose not to. After inducing specific mutations in the aforementioned genes, the researchers obtained a loss-of-function mutation in all genes except for KRAS, which yielded a gain-of-function mutation. It was further discovered that successive oncogene alterations of APC and TP53 were enough for the organoids to gain autonomy from growth factors and acquire aneuploidy.

CRISPR-Cas9 genome editing in ASC-derived organoids have also proven to be capable of modeling diseases that previous technology found challenging to replicate in vitro. For example, a close examination of polyps from sessile serrated adenomas patients revealed an association between the activating BRAF proto-oncogene carrying the V600E mutation and SSA (10). After introducing the mutation into wild-type colorectal organoids using CRISPR-Cas9 technology, researchers found that organoids exhibited independent growth from transforming growth factor-β (TGF-β) signaling and displayed epithelial-mesenchymal tissue transition, the latter of which is a hallmark of the early stages of cancer development (10, 42).

Manipulating genomes within organoids can also serve as an excellent method for screening effective drug responses. Colorectal organoids with different CRISPR-induced RAS mutations have been tested for their impact on the responsiveness to EGFR and MEK inhibitors (10, 26). Results show that while normal organoids were sensitive to the inhibitors, those with oncogenic KRAS mutations exhibited diminished sensitivity due to cell-cycle arrest prompted by the drugs in place of cell death, highlighting the prospect of employing CRISPR-Cas9-altered organoid libraries for extensive screening purposes.

In 2016, a protocol delineating the techniques for using stably integrated retroviral transduction or transiently expressed liposomal transfection with CRISPR-Cas9 genome editing technology for genetically manipulating liver and pancreatic organoids was published (10, 28). Unfortunately, however, the methods presented in the protocol require for the organoids to be dissociated into single cells, severely reducing overall editing efficiency. As such, adeno-associated virus vectors have since been proposed as an alternative method of editing liver organoids (27).

CRISPR-Cas9 technology has also been used to knock in a BRAF mutation, overexpress the GREM1 gene, and generate long-range gene fusions in human colon-derived organoids in
order to create accurate models for traditional serrated adenoma (TSA), a rare subtype of colonic serrated adenomas (10, 29). After transplantation, the newly modified organoids displayed phenotypes that closely resembled those observed in TSA patients, demonstrating the potential of using CRISPR-Cas9 gene-edited organoids to accurately model and study specific subtypes of colonic serrated adenomas.

Finally, to better understand the clonal evolution found in breast cancer development, researchers explored the application of CRISPR-Cas9 genome editing technology in reduction-mammoplasty-patient-derived mammary epithelial organoids (10, 30). After knocking out the four tumor suppressor genes TP53, PTEN, RB1, and NF1 and transplanting the edited organoids into mice, the researchers were able to induce sustained culture viability and trigger tumor development, indicating that CRISPR-Cas9 technology holds great promise as a tool for modeling mammary epithelial-related diseases such as breast cancer.

**CRISPR-Cas9 Genome Editing in PSC-Derived Organoids**

Creating gene-edited PSC-derived organoids is comparatively easier than creating gene-edited ASC-derived organoids, as editing can be done directly in PSCs prior to differentiation, bypassing the initial requirement of breaking down organoids before transfecting the CRISPR-Cas9 gene-editing tools (10). As a result, the overall efficiency of organoid editing when using PSCs is enhanced compared to organoid editing using ASCs.

CRISPR-Cas9 genome editing technology is applicable in helping model diseases in brain organoids, which are exclusively derived from PSCs (10). Diseases such as Sandhoff disease, glioblastoma cancer, and Parkinson’s disease have all been successfully modeled by using CRISPR-Cas9 to edit and introduce mutations into PSC-derived cerebral organoids (10, 12, 32, 31). All of these studies indicate that CRISPR-Cas9 technology could be used to test the importance of disease-causing mutations, offering valuable insights into disease mechanisms and potential therapeutic strategies.

Gastrointestinal tract organoids (e.g. gastric, intestinal, and colorectal organoids) are another type of organoid that have successfully been derived from PSCs. The CRISPR-Cas9 system was used to introduce a *DKC1* mutation to iPSCs, modeling the mutation found in patients with dyskeratosis congenita (10, 33, 34, 35). Following differentiation, the mutant organoids exhibited shorter telomeres and lacked the ability to sustain budding crypts, mirroring characteristics observed in dyskeratosis congenita (10, 33). As such, PSC-derived gastrointestinal tract organoids edited by CRISPR-Cas9 prove to be great options for performing genetic studies of human diseases.

Similarly, CRISPR-Cas9 genome-editing technology was paired with organoid technology to create a polycystic kidney disease (PKD) organoid model in 2015 by inducing a biallelic mutation in the *polycystin-1* (*PDK1*) or *polycystin-2* (*PDK2*) genes of kidney organoids, both of which are known to be associated with PKD (10, 36). The functional loss of either *PDK1* or *PDK2* is usually enough to trigger autosomal dominant PDK, where those affected suffer from renal failure as a result of large renal cyst formation in their kidneys. As expected, similar renal cysts were observed after differentiating the mutant PSCs into kidney organoids in place of the normal structures that comprise standard kidney structure, further demonstrating
CRISPR-Cas9’s potential in disease modeling with organoids (10).

Usage of AI in CRISPR-Cas9 Genome Editing
In recent years, AI has emerged as a promising approach to further improve CRISPR-Cas9 genome editing technology in various fields, including vaccine design and cancer therapeutics. Overall, the integration of AI with CRISPR-Cas9 technology can be split into two main approaches: the knowledge discovery approach, in which AI analyzes data and knowledge to aid in the CRISPR-Cas9 process, and the modeling-based approach, in which AI helps with physical and digital modeling of complex animal systems using CRISPR. The resulting improvements with the CRISPR-Cas9 system can be applied to a variety of fields, including vaccine design, cancer therapeutics, and organoid research.

CRISPR/Cas9 genome editing technology is associated with the production of a plethora of information. With knowledge discovery approaches, AI facilitates and speeds up the treatment of cancer and other diseases in which CRISPR-Cas9 has been utilized by analyzing data from scientific papers, specialized databases, and clinical trials; it is able to discover patterns in genome editing while improving CRISPR-Cas9 efficiency (14). Furthermore, data-driven AI tools such as CRISPRIdentify are able to differentiate true CRISPR arrays from false ones, providing researchers with more accurate information and knowledge on which genomic regions are CRISPR arrays, from which they can then gather additional data about the natural CRISPR system to improve the accuracy of the CRISPR-Cas9 system (16).

Modeling-based approaches are also a popular form of AI/CRISPRCas9 integration. To current knowledge, there are several properties that affect the efficiency of gRNA cleavage and its off-target effects in the CRISPR-Cas9 genome editing system for physical models such as organoids. These properties include the specific gRNA sequence, the integration of site-specific nucleotides in the gRNA design, nucleotide composition surrounding the target site, protoscaler proximity, secondary structure of both the gRNA and the TracrRNA, epigenetic factors, and immune system barriers, all of which have a significant impact on gRNA efficiency (13). Currently, several machine learning (ML) and deep learning models are being leveraged to solve some of these issues with the gRNA design (43, 44, 45, 46, 47). These models are able to both predict the on/off-target impacts of gRNAs and score features for optimal CRISPR-Cas9 performance through extensive screenings of genome editing data from global reports, contributing to advancements in the success rate of CRISPR-Cas9 genome editing technology and improving the accuracy of CRISPR-generated organoid models. In contrast to other experimental detection tools like IDLV, GUIDE-seq, or HTGTS, artificial intelligence and machine learning-based methods offer superior cost-effectiveness, efficiency, and speed in making predictions due to their extensive training data.

Furthermore, machine learning algorithms have empirically been proven successful in guiding and regulating the spatial self-organization of multicellular patterns within CRISPR-engineered organoid models, circumventing the usage of any external patterning methods. Trained ML models have also been shown to be accurate in identifying mutations caused by the CRISPR-Cas9 processes, which could be taken into account for further refinement of disease models (17).
Both types of approaches play critical roles in facilitating the proper implementation and development of the CRISPR-Cas9 genome editing system. Unfortunately, many AI/ML models and algorithms currently face challenges such as data imbalance, heterogeneity, limited access to training datasets, and inefficiencies across different species, making it difficult to fully integrate the models into medical therapeutics (15). Nevertheless, increasing advancements in AI/ML algorithms are anticipated to improve the accuracy of the CRISPR-Cas9 system in the near future, which will be essential for its future clinical and therapeutic applications.

**AI + CRISPR-Cas9 for Organoid Research**

As mentioned above, the integration of AI into CRISPR-Cas9 technology has helped improve both the efficiency and accuracy of the system. Due to the importance of CRISPR-Cas9 technology in organoid creation, this combination of AI and CRISPR-Cas9 has also led to significant advancements in organoid creation, research, and improvement.

For example, AI and CRISPR-Cas9 genome editing technology have been combined in organoids in the field of cancer therapeutics. In recent years, researchers have applied the CRISPR-Cas9 system to organoids in order to recapitulate tumor heterogeneity through creating cancer models that accurately portray tumors in diseases like ovarian cancer (15, 18). By leveraging globally available genome editing datasets to help predict gRNA activity and specificity scores, researchers discovered that AI not only sped up the efficiency of the genome editing process for cancer organoid modeling, but was also much more cost-effective than other tools used to predict gRNA activity and specificity scores (15).

Another area of research in which the integration of AI and CRISPR-Cas9 genome editing technology has been applied is brain organoids (19). Since there is currently extensive data available for genome editing in brain organoids, AI has been utilized to help manage off-target editing and generate blunt ends or staggered breaks, improving the overall accuracy of CRISPR-Cas9 editing in organoids. For instance, Allen et. al. (2018) successfully used a ML algorithm to predict specific factors of CRISPR-Cas9 critical to in vivo applications, ensuring accurate genome splicing in brain organoids (20). As brain organoid research and technology improves, they could eventually grow to resemble the full human brain, potentially becoming a vital part of studies on neurological diseases in the future (23). With the aid of AI enhancements on CRISPR-Cas9 splicing and editing, the treatment of neurological disorders (leading cause of disability and second leading cause of death worldwide) could be expedited immensely (24).

On a similar note, AI/ML and CRISPR-Cas9 genome editing technology have both been integrated into drug discovery efforts for the treatment of neurological disorders (21). In order to improve the utility of neural organoids in drug discovery, CRISPR-Cas9 has helped generate robust isogenic models, improving the accuracy of disease modeling, while ML was used to speed up dataset processing and analysis for larger or more complicated datasets.

Furthermore, CRISPR-Cas9 genome editing technology has also been combined with AI for the developmental phase of organoids, leveraging CRISPR-Cas9’s precise control of gene expression and mutations to carefully regulate organoid creation (22). As it is difficult to quickly determine the optimal gene knockout groups due to the many variables present during the genome editing and knockout process, AI has been used to build a ML model capable of swiftly
identifying optimal gene knockout groups and protocols to develop a better gene knockout plan. In the near future, it is predicted that researchers will be able to better control the spatial self-organization of stem cells through the combination of AI/ML and CRISPR-Cas9 with organoids.

Conclusion
The integration of AI and CRISPR-Cas9 genome editing technology represents an emerging idea that will radically improve the field of organoid creation, research, and improvement. AI is able to improve CRISPR-Cas9 technology through both knowledge discovery approaches where it analyzes complex datasets and modeling-based approaches where it helps with improving CRISPR-Cas9 models. As CRISPR-Cas9 genome editing has become a leading factor in organoid research and advancement in recent years, these AI-led improvements in the technology naturally reflect onto the same field.

Since both AI and CRISPR-Cas9 genome editing technology are still considered fairly novel applications in the study of organoids, more research is necessary regarding AI in CRISPR-Cas9 genome editing technology for organoid creation, research, and improvement. Current AI/ML models for CRISPR-Cas9 in organoids are trained from a limited number of available training data or suffer from data imbalances or inefficiency; as such, there is much room for improvement for the accuracy of these AI/ML models for CRISPR-Cas9 and organoids in the future.

References


