CRISPR-GPT: An LLM Agent for Automated Design of Gene-Editing Experiments

Yuanhao Qu[1,2,†], Kaixuan Huang[3,†], Henry Cousins[4,5], William A. Johnson[1], Di Yin[1], Mihir Shah[4], Denny Zhou[6], Russ Altman[4,7], Mengdi Wang[3,*], Le Cong[1,*]

Affiliations:

[1]Department of Pathology, Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA

[2]Cancer Biology Program, Stanford University School of Medicine, Stanford, CA, USA

[3]Center for Statistics and Machine Learning and Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544, USA

[4]Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA

[5]Medical Scientist Training Program, Stanford University School of Medicine, Stanford, CA, USA
[6]Google DeepMind

[7]Department of Bioengineering, Department of Genetics, Stanford University, Stanford, CA, USA

*Correspondence to: mengdiw@princeton.edu (M.W.), congle@stanford.edu (L.C.)

† These authors contributed equally to this work.

Abstract

The introduction of genome engineering technology has transformed biomedical research, making it possible to make precise changes to genetic information. However, creating an efficient gene-editing system requires a deep understanding of CRISPR technology, and the complex experimental systems under investigation. While Large Language Models (LLMs) have shown promise in various tasks, they often lack specific knowledge and struggle to accurately solve biological design problems. In this work, we introduce CRISPR-GPT, an LLM agent augmented with domain knowledge and external tools to automate and enhance the design process of CRISPR-based gene-editing experiments. CRISPR-GPT leverages the reasoning ability of LLMs to facilitate the process of selecting CRISPR systems, designing guide RNAs, recommending cellular delivery methods, drafting protocols, and designing validation experiments to confirm editing outcomes. We showcase the potential of CRISPR-GPT for assisting non-expert researchers with gene-editing experiments from scratch and validate the agent's effectiveness in a real-world use case. Furthermore, we explore the ethical and regulatory considerations associated with automated gene-editing design, highlighting the need for responsible and transparent use of these tools. Our work aims to bridge the

gap between biological researchers across various fields with CRISPR genome engineering technology and demonstrate the potential of LLM agents in facilitating complex biological discovery tasks.

1. Introduction

Gene editing technology represents a groundbreaking scientific advancement that enables precise alterations to the genetic material of living organisms. This innovative technique has found extensive applications across various fields of biology and medicine, from correcting genetic defects responsible for disorders like cystic fibrosis, hemophilia, and sickle cell anemia, to offering new strategies in the battle against complex conditions such as cancer, cardiovascular diseases, neurodegenerative disorders, and infections. A most well-known gene editing system is called CRISPR-Cas9[1–9]. It was adapted from a naturally occurring genome editing system that bacteria use as an immune defense. Beyond CRISPR-Cas9, recent advancements have led to the development of CRISPR activation/interference, CRISPR-based prime editing and base editing technologies[10,11]. CRISPR activation/interference, also called CRISPRa/CRISPRi, are able to either enhance gene expression or silence the activities of specific genes via epigenetic regulation[12–16]. Prime editing, considered a "search and replace" method for DNA, allows for precise editing capabilities without introducing double-stranded breaks[17]. Base editing, on the other hand, enables the direct, irreversible conversion of one DNA base into another at targeted locations, further expanding the toolkit for precise genome modification[18]. All these technologies hold potential for broad applications in medicine, agriculture, and beyond, enhancing the scope of genome editing in pursuing treatments for genetic diseases and other applications.

Designing gene-editing experiments requires deep understanding of the suite of technologies and the associated biology of target organs. CRISPR Cas-based editing works by interacting with the RNA of a short "guide" sequence (guide RNA) that binds to a specific target sequence in a cell's DNA, much like the RNA segments bacteria produce from the CRISPR array. When introduced into cells, the guide RNA recognizes the intended DNA sequence, and the Cas enzyme (often Cas9 or others) cuts the DNA at the targeted location, mirroring the process in bacteria. There are numerous considerations when designing such experiments, including the selection of a well-suited gene-editing system, the development of optimal guide sequences, and validation approach. This typically requires significant domain expertise, understanding of biology of the target organ, as well as trial-and-error efforts. Developing AI-assisted computation tools to aid gene-editing holds great promise to make the technology more accessible and accelerate both scientific and therapeutic developments.

Large language models (LLMs) have demonstrated exceptional capabilities in language skills and encapsulate a tremendous amount of world knowledge, approximating aspects of artificial general intelligence[19–23]. Recent research has also explored enhancing LLMs with external tools, improving their problem-solving abilities and efficiency[24–26]. LLMs have also demonstrated potential as tool makers[27] and black-box optimizers[28]. Researchers have explored LLM-based specialized models for various

application domains[29,30], as well as for solving scientific and mathematical tasks. For example, ChemCrow[31] uses tool-augmented LM for solving a range of chemistry-related tasks such as paracetamol synthesis, whereas Coscientist[32], also driven by GPT-4 and integrated automated experimentation, achieved successful optimization of palladium-catalyzed cross-coupling reaction.

1.1 General-purpose LLMs do not know how to design biological experiments.

While leveraging large language models (LLMs) for aiding in the design of gene-editing experiments presents an enticing prospect, the current state-of-the-art general-purpose models exhibit significant shortcomings in this specialized domain. These models, despite their vast knowledge base, lack the precise, up-to-date domain-specific knowledge essential for the accurate design of biological experiments.

A critical limitation of general-purpose LLMs is their propensity for "hallucinations" or generating confident yet inaccurate responses when tasked with specialized biological queries. For instance, when tasked with designing a guide RNA (gRNA) sequence for targeting specific human genes, such as EMX1 or EGFR, general-purpose LLMs like ChatGPT-3/ChatGPT-4 can respond incorrect sequence with high confidence. However, the gRNA sequences they provide often do not correspond to any known genomic region. This discrepancy can be readily identified by comparing the LLM-generated sequences with reference sequences in databases such as the NCBI's BLAST tool, which aligns sequences to the human genome and transcriptome. Such hallucinated design sequences not only lack utility but can mislead researchers, potentially leading to wasted resources and time if not properly vetted.

In addition, responses generated from general-purpose LLMs usually lack essential details necessary for the experimental design, such as specific materials, protocols, considerations for off-target effects, gRNA efficiency, and specificity. Such gaps in information can leave researchers, especially those new to the field of gene editing, unprepared for the practical execution of experiments.

Furthermore, it is crucial to note that responses generated may contain overwhelming information that does not directly contribute to the gene-editing experimental designs. Such irrelevant text can lead to confusion and misdirection, complicating the researcher's task of identifying the most pertinent and practical information for their gene-editing objectives.

All of these limitations underscore the necessity for a new class of LLMs tailored specifically for the gene-editing experimental designs (We refer readers to **Supplementary Note 1** for more examples of failures). Such models would need to integrate deep, accurate domain knowledge with the ability to critically evaluate and generate experimentally viable solutions, thereby overcoming the current barriers faced by general-purpose LLMs in the design of CRISPR gene-editing experiments.

1.2 Overview of CRISPR-GPT

In the rapidly evolving field of genetic engineering, CRISPR technology has become a pivotal tool for precise gene editing. Despite its promise, the intricacy of designing CRISPR experiments—from guide RNA (gRNA) selection to predicting off-target effects—presents significant challenges, especially to those new to the field. To bridge this gap, we introduce CRISPR-GPT, a novel solution that combines the strengths of Large Language Models (LLMs) with domain-specific knowledge and computational tools, specifically tailored for CRISPR gene editing tasks.

CRISPR-GPT is centered around a tailor-made LLM-powered design and planning agent (**Figure 1**). This engine of the agent not only draws upon expert knowledge from leading practitioners in gene editing but also integrates a broad review of recent literature and a suite of computational toolkits including guideRNA design tool.

The innovation of CRISPR-GPT Agent allows automated designs of gene-editing experiments by simplifies the complex process into a series of manageable steps:

- Selection of CRISPR System: Tailoring the choice of CRISPR system to the experiment's needs.
- gRNA Design: Optimizing guide RNA sequences for efficiency and specificity based on the Broad Institute's gold-standard guideRNA library and CRISPRPick toolkit, including pre-designed gRNA libraries[33–36].
- Delivery Approach Selection: Advising on the most effective methods to introduce the CRISPR components into target cells.
- Prediction of Off-target Effects: Assessing potential unintended alterations alongside desired edits.
- Recommendation of Experimental Protocols: Outlining step-by-step procedures tailored to the experiment's objectives.
- Validation Approach Recommendation and Primer Design: Recommending gest ways to validate the edits and help design the associated primers.

This approach, leveraging a chain-of-thought reasoning model and state machines, ensures that even individuals new to gene editing can iteratively refine their experimental designs to achieve protocols that meet their specific research needs. In addition, CRISPR-GPT offers:

- A Freestyle Q&A Mode for addressing ad hoc queries with precision,
- An Off-target Prediction Mode for in-depth analysis of pre-designed gRNAs
-

These functions help the users when they meet additional issues during the experimental design process.

Mindful of the ethical and safety considerations surrounding gene editing, especially in human applications, we have integrated safeguards into CRISPR-GPT. These include restrictions on its use in human subjects, measures to ensure the privacy of genetic information, and alerts for potential unintended consequences, reflecting our commitment to responsible use in alignment with the broader scientific and ethical discourse on gene editing technologies.

2. Methods and Algorithms

2.1 Large Language Model

The CRISPR-GPT agent consists of the following 4 core modules (**Figure 2**): LLM planner, Tool provider, Task executor, and the LLM Agent that serve as the interface with users for taking inputs and communicate outputs.

Task Executor operates as state machines, providing robust subgoal decomposition and progress control.

We implement 22 tasks, (summarized in **Table 1)**, in the form of state machines for CRISPR-GPT. The state machines are responsible for providing sufficient instruction for the current task and guiding the user to fulfill the decision-making through multiple rounds of textual interactions. Through these state machines, we manually decompose each task into sub-goals for the task executor. Specifically, each state is responsible for one particular sub-goal. The transition logic is well-defined so the task executor can properly transit to another sub-goal based on the current progress.

We have 4 predefined Meta-Tasks that support the full pipelines of 4 gene-editing-related experiments (see **Table 1)**. Besides, the LLM planner can generate a customized list of tasks depending on the user's meta-request. The state machines of the corresponding tasks are chained together as a bigger state machine to support the entire pipeline.

Tool Provider connects Task Executor with external APIs

To connect language models with external functionalities[37–41], the system needs to (1) analyze the current situation and judge whether it's suitable to call an external tool; (2) know what kinds of tools are available and choose the best from them. Instead of directly exposing the interfaces of the APIs to LLMs, in CRISPR-GPT, we wrap the usage of APIs inside the states and expose more user-friendly and LLM-friendly textual interfaces through hand-written instructions and responses. In plain words, we are teaching users (human agents & LLM agents) to use the tools. The tools include Google web search, running programs like Primer3[42], as well as retrieval from external guide RNA libraries, research papers, and experiment protocols.

LLM-planner automatically generates a list of tasks based on the user's request

Large Language Models (LLMs) such as GPT-4, Gemini, and Claude can serve as the reasoning core of the LLM-powered agent to solve real-world decision-making problems. We adopt the popular ReAct[43] prompting technique, where the LLM is prompted to output the chain-of-thought[44] reasoning path and the final action from the plausible action set (**Figure 3**). To let LLMs perform task decomposition[45], we provide a table of the descriptions and the dependencies of all the tasks as a prompt to the LLM. Based on LLM's internal knowledge as well as our manually written descriptions of tasks and instruction of task decomposition, LLM can intelligently analyze the user's request and decompose the user's request into a sequence of tasks, respecting the

dependencies of the tasks. After the decomposition, the corresponding state machines are chained together to complete all the tasks. The prompt format of the task decomposition can be found in **Supplementary Note 2**.

For robustness, we do not allow LLMs to dynamically add/delete new tasks (new state machines) during the automatic execution. However, we believe this is an important step toward a more intelligent CRISPR-GPT version and leave this as future work.

LLM-Agent automatically interacts with the Task Executor based on the user's meta request

In addressing the complex challenge of automating CRISPR gene editing tasks, we conceptualize the problem through the lens of sequential decision-making. This perspective frames the interaction between the user and the automated system as a series of steps, each requiring precise decisions to progress towards the ultimate goal of experiment design and execution. Central to our system is the LLM-agent, which acts as an intermediary between the user and a state machine. This state machine is derived from an initial task decomposition step, effectively breaking down the gene editing process into a structured sequence of actions and decisions. At each step in this sequence, the state machine presents a current state to the LLM-agent. This state encapsulates a description of the task at hand and specifies any input required from the user to move forward.

The LLM-agent's role is to interpret the current state and make informed decisions on behalf of the user. To do this effectively, the agent may draw upon a diverse set of information, including:

- The instruction inherent to the current state,
- The specific request made by the user,
- A history of past interactions within the current task session,
- Results from external computational tools that have been integrated into the system.

This information is synthesized into a prompt for the LLM-agent, which then uses its capabilities to determine the most appropriate next action. The format and structure of these prompts, designed to optimize the decision-making process, are detailed in the **Supplementary Note 2**.

User oversight is a critical component of this system. While the LLM-agent operates autonomously, the user is not removed from the process. Instead, they are encouraged to monitor the progression of tasks and interact with the agent. This setup ensures that any errors or misinterpretations by the LLM-agent can be quickly identified and corrected by the user, maintaining the accuracy and integrity of the gene editing experiment design. This approach to automation emphasizes a collaborative synergy between human expertise and artificial intelligence. By leveraging the LLM-agent's ability to process and act on complex information, we facilitate a more efficient and user-friendly experience in designing CRISPR gene editing experiments. The sequential decision-making framework not only streamlines the task execution process but also ensures that user input remains a cornerstone of experiment planning and design.

2.2 Human evaluation

To evaluate the CRISPR-GPT agent's effectiveness in aiding gene editing and experimental design, we assembled a diverse group of 12 experts in the field of CRISPR and gene editing research. Each of the 12 experts rated the responses to the experimental design tasks from three modes on a scale from 1 (Poor) to 5 (Excellent) according to the established criteria (all rubrics for the human evaluation are detailed in Supplementary Note 3). To offer a comparative perspective, outputs from both ChatGPT 3.5 and ChatGPT 4.0 (model version gpt-4-0613) were generated using similar prompts and evaluated using the same criteria.

2.3 Biological experiment and wet-lab validation

We performed biological experiments through human-agent collaboration using CRISPR-GPT with ChatGPTv4 API, as wet-lab based real world validation of our approach. Specifically, we have independent scientist who are not familiar with gene-editing experiment to use CRISPR-GPT to facilitate their knock-out (KO) gene-editing experiments in a cancer research project. Detailed methods are provided below.

Cell line and cell culture

A375 cell line was cultured in DMEM, high glucose, GlutaMAX (Gibco) supplemented with 10% fetal bovine serum (FBS, Gemini Bio), 100 U/ml penicillin and 100ug/ml streptomycin (Gibco) at 37 °C with 5% CO2.

crRNA cloning

Cloning of 4 crRNAs (TGFBR1/SNAI1/BAX/BCL2L1) was performed with BbsI or Esp3I (NEB) through a Golden Gate assembly approach into a Cas12a expressing backbone backbone. Constructs were sequence verified by Sanger sequencing using a U6 sequencing primer: 5’-GACTATCATATGCTTACCGT-3’.

Lentivirus packaging and transduction

Lentivirus was produced by co-transfecting the assembled lentiviral vector with VSV-G envelope and Delta-Vpr packaging plasmids into HEK-293T cells using PEI transfection reagent (Sigma-Aldrich). Supernatant was harvested 48 hr after transfection. A375 cells were transduced at low MOI with 8 µg/mL polybrene using a spin-infection at 1000∗g for 45 minutes. After 24 hours, cells were selected with 1 µg / mL puromycin to establish stably expressing cell lines.

gDNA extraction, PCR and sequencing

Genomic DNA was extracted from selected cells 7 days later using QuickExtract (Lucigen). The targeted loci were then amplified using Phusion Flash High-Fidelity PCR

Master Mix (ThermoFisher Scientific) according to the manufacturer's instructions with primers containing Illumina sequencing adapters. Paired-end reads (150 bp) were generated on an Illumina MiSeq platform.

3. Results

CRISPR-GPT leverages the reasoning abilities of LLM, domain knowledge, retrieval techniques and external tools to provide a comprehensive solution to gene-editing experimental design tasks. It supports a wide arrange of gene editing scenarios, including single gene knockout, base editing without double strand breaks, insertions/deletions/replacement via prime editing, epigenetic editing for activation or repression of genes (CRISPRa and CRISPRi).

3.1 CRISPR-GPT assists researchers with gene-editing experimental design through three modules.

The CRISPR-GPT agent aids researchers in designing gene-editing experiments via three distinct modules. "Meta Mode" equips users, especially newcomers to the gene-editing field, with expertly defined pipelines for general gene editing scenarios (termed Meta-tasks). "Auto Mode" automatically generates a tailored list of necessary design tasks based on user input, facilitating goal achievement for users of all experience levels. "Q&A Mode" functions as an advanced GPT-4 chatbot, addressing users' CRISPR and gene-editing related queries throughout the design process (**Figure 4**).

3.1.1 Meta Mode

"Meta Mode" involves the planning and implementation of 22 unique gene-editing experimental design tasks utilizing four types of CRISPR-based gene editing systems (Meta-tasks) (**Table 1**). It leverages predefined pipelines to help users complete a Meta-task thoroughly. In this mode, the CRISPR-GPT agent guides users through each task necessary for designing gene editing experiments. This includes selecting the appropriate CRISPR system, recommending delivery methods, designing the sgRNA, predicting sgRNA off-target efficiency, selecting experimental protocols, and planning validation experiments.

For every design task, the CRISPR-GPT agent interacts with users, applying various techniques and external tools to deliver the optimal solution. For example, in choosing CRISPR systems, CRISPR-GPT continuously interacts with users, providing instructions and collecting information to suggest options based on published protocols (See example in **Figure 5** General Task 1). For context-sensitive tasks like delivery method recommendations, CRISPR-GPT not only suggests common methods but also offers customized solutions based on the user's requests through web search (See example in **Figure 5** General Task 2). For sgRNA/pegRNA designs, a multi-species database derived from existing designs and publications enables CRISPR-GPT to swiftly suggest pre-designed sgRNA based on user information (See example in **Figure 5** General Task 3). Following sgRNA/pegRNA design, users can evaluate designed guides' potential off-target effects with detailed instructions and code from CRISPR-GPT

(See example in **Figure 5** General Task 4). After completing design tasks, CRISPR-GPT offers selected protocols based on the interaction history, including CRISPR system selection and delivery methods (See example in **Figure 5** General Task 5). Finally, for the validation task, CRISPR-GPT utilizes external APIs, like Primer3, to assist users in designing primers for validation experiments (See example in **Figure 5** General Task 6).

3.1.2 Auto Mode

"Auto Mode" also facilitates the planning and execution of 13 unique gene-editing experimental design tasks. Unlike "Meta Mode," it does not rely on predefined meta-tasks and pipelines; instead, it uses an LLM-planner to break down a user's request into a sequence of dependent tasks. For instance, if a user requests to “design sgRNA to knockout human EGFR,” the CRISPR-GPT agent identifies the keywords from the request and lists the necessary design tasks, like “CRISPR/Cas system selection” and “sgRNA design for knockout.” Additionally, it uses information from the initial request (e.g., target gene “EGFR” and species “human”) to autofill relevant fields and generate sgRNA designs without needing repeated inputs from the user. Simultaneously, CRISPR-GPT elucidates the rationale behind its choices, allowing users to track the process and make corrections if necessary.

3.1.3 Q&A Mode

During the design tasks in "Meta Mode" and "Auto Mode," the CRISPR-GPT agent offers immediate responses or advice for CRISPR and gene editing related inquiries through "Q&A Mode." For example, after selecting a CRISPR system, users seeking more information about the chosen system (e.g., Cas12a) can quickly obtain answers by asking, “Q: What is Cas12a?”. The CRISPR-GPT uses its knowledge base and document retrieval from expert-selected databases in the field to provide accurate and relevant information swiftly.

3.2 CRISPR-GPT outperforms general LLMs in gene-editing design tasks through human expert evaluations.

To evaluate the performance of CRISPR-GPT agent, we invited 12 researchers with expertise in CRISPR and gene editing to design sets of tasks to test the ability of CRISPR-GPT in assisting researchers with experimental design. The results are evaluated in four different aspects: Accuracy, Reasoning, Completeness and Conciseness (**Supplementary Note 3**). Accuracy reflects whether CRISPR-GPT could provide accurate information of current state of CRISPR research and methodologies. Reasoning evaluates if CRISPR-GPT could provide insightful, well-supported explanation of the suggested designs. Completeness makes sure users receive all required information needed for CRISPR experimental design. Lastly, conciseness ensures that CRISPR-GPT provides users with directly relevant information for the design tasks with minimal unnecessary information. All evaluators were asked to score sets of tasks with these four aspects from 1 (Poor) to 5 (Excellent) for all three modes.

Responses from ChatGPT 3.5 and ChatGPT 4.0 were generated and scored alongside those of CRISPR-GPT, using equivalent prompts in all cases.

We observed that CRISPR-GPT achieved significantly higher accuracy in our designed sets of tasks over general LLM-agents across all three modes, as we employed vast domain knowledge in the CRISPR and gene editing field to ensure the robustness of the CRISPR-GPT agent (**Figure 6**). While, responses generated by general LLM agents including ChatGPT 3.5 and ChatGPT 4.0 contain more minor factual errors due to known issues including inadequate domain knowledge and hallucination. At the same time, we found both CRISPR-GPT and general LLM agents show good reasoning ability over different sets of tasks. For “Auto Mode” related tasks, CRISPR-GPT showed even better reasoning, potentially due to better prompting techniques encoded in the agent. As we expected, “Completeness” is the major issue for general LLM-agents to perform gene-editing experimental design tasks. They can usually provide general guides for designs while could not provide design details due to the lack of domain knowledge and external tools. On the opposite side, CRISPR-GPT showed much better “Completeness” performance scores in the design tasks, allowing the users to perform the gene-editing experiment solely based on the information provided by CRISPR-GPT. It's worth noting that both ChatGPT 3.5 and 4.0 outperform CRISPR-GPT in the “Completeness” performance score in “Q&A” mode. Such a result is due to the intentional tradeoff between “Completeness” and “Conciseness”. Answers directly generated by general LLM-agents usually include much irrelevant information in order to provide users with a more complete response. This usually confuses the users and makes it hard to catch the key information. In this case, we intentionally designed the CRISPR-GPT to provide concise accurate answers to the users across all different modes and accordingly CRISPR-GPT showed consistent better “Conciseness” performance scores.

Overall, through experts’ evaluation, we found CRISPR-GPT showed significantly improved performance over general LLM-agents for gene-editing experimental design tasks in all different aspects. Notwithstanding, CRISPR-GPT experienced difficulties in more complex gene editing scenarios and rare biological cases. It can be further extended and improved in the future with more up-to-date domain knowledge and better external tool sets.

3.3 CRISPR-GPT demonstrates its efficacy through real-world application.

To demonstrate CRISPR-GPT in assisting researchers with designing gene-editing experiments, we performed a gene knockout experiment in the human A375 cell line through continuous interaction with CRISPR-GPT (**Figure 7**).

In this experiment, we aimed to knock out a panel of 4 genes (TGFBR1, SNAI1, BAX, BCL2L1) individually in the human A375 cell line. To start, we selected the “Meta mode” to design the gene knockout experiment from scratch. Following the instruction of selecting the CRISPR system in the CRISPR-GPT, we selected AsCas12a as we hope to perform multiple-site edits and lower the potential off-target editing rate. For the delivery of the CRISPR system in A375 cells, we followed CRISPR-GPT's recommendation to use lentiviral transduction, ensuring stable expression of both the Cas enzyme and sgRNAs. Then, based on this information, we were able to obtain the Cas12a plasmid (previously owned). When it came to designing the sgRNAs, we

specifically targeted the human TGFBR1/SNAI1/BAX/BCL2L1 genes, fully aware of the ethical ramifications of human gene editing raised by CRISPR-GPT. CRISPR-GPT provided 4 sgRNA sequences for each gene (Exampled in **Figure 7**) from a published library so we were able to order the sequences for synthesis.

Later, CRISPR-GPT offered protocols for gRNA cloning. Detailed instructions were then provided for producing lentivirus through calcium phosphate transfection in HEK293T cells, using the necessary plasmids and viral packaging components. Following this, we exactly followed the protocol generated by CRISPR-GPT, through the transduction process, which involved cell culture procedures, the addition of lentivirus, and the use of polybrene to facilitate efficient transduction. To perform validation, we chose next-generation sequencing (NGS) for mutation detection and validation of the knockout in CRISPR-GPT, guided by the protocol provided by CRISPR-GPT agent. To prepare for NGS, we extracted genomic DNA from cells using the DNeasy Blood & Tissue Kit based on the protocols. For the crucial step of PCR primer design, we provided detailed sequence information to CRISPR-GPT, which automatically returned a set of primers designed with Primer3 to specifically amplify the target site. In the concluding stages of our experiment, CRISPR-GPT advised us to attach Illumina adaptors to the PCR products for library construction and emphasized the necessity of checking primer specificity with NCBI BLAST. This final validation step was crucial to prevent mis-priming and to ensure that the sequencing results would accurately reflect the intended genomic edits.

Finally, we analyzed the data from the NGS, and observed a consistent high rate of expected editing outcomes across all 4 targeted genes. Through this process, where CRISPR-GPT provided: (1) CRISPR system selection (2) guideRNA design (3) delivery system recommendation (4) plasmid and viral vector selection with cloning protocol (5) tissue culture, cell transduction procedures (6) cell harvesting and gene-editing efficiency quantification methods (7) sequencing primer design and readout validation protocol. Thus, the dynamic interaction between our expertise and CRISPR-GPT's computational guidance was instrumental in executing a precise and ethically considerate gene-editing experiment.

4. Safety and Ethical Concerns

Safety and ethical concerns arise when using AI tools to guide genome editing, ranging from the risk of illegally altering human genomes to privacy issues when user genome information is involved.

4.1 Mitigation of the risk of human heritable editing

Technologies such as CRISPR-Cas9, have made it possible to alter human genomes, which pose a number of ethical and safety risks. In particular, germline cell and embryo genome editing bring up a number of ethical challenges, including whether it would be permissible to use this technology to enhance normal human traits (such as height or intelligence). Based on concerns about ethics and safety, germline cell and embryo genome editing are currently illegal in the United States and many other countries. To ensure CRISPR-GPT follows the guidelines given in a moratorium[46] on heritable genome editing. CRISPR-GPT employs a mechanism to make sure in all tasks

users cannot bypass the existing step asking which organism they are editing. The agent would check if the editing target belonged to human tissues or organs. If it is found that the editing target is a human organ, it will trigger the following solution: Warning note when users proceed with designing human gene-editing exp. Link to this international moratorium with note. Ask users to confirm they understand the risk and have read this international guideline before proceeding.

4.2 Protection of user genome data privacy

Other concerns are related to user data privacy issues, especially when human genome sequence information might be exchanged by using AI tools. We follow the data privacy and HIPAA privacy rule in healthcare[47]. Although genome-scale sequences are fundamentally linked to identities, DNA segments of up to 20 bp length are considered safe and not able to identify human identity (REF). CPISPR-GPT is equipped with the following functionalities, to avoid supplying any identifiable private human/patient sequence to a public LLM model. Specifically, our solution is:

- CRISPR-GPT would never store any identifiable long genome sequence in the server that would potentially reveal patient private information.
- CRISPR-GPT implements a filter to detect if there is any >= 20bp of A/T/G/C/U sequence contained in the prompts before sending them to external LLMs. After detecting the existence of such a sequence, the agent would raise an error with a warning note, asking the user to manually delete such sequences in the input. In this way, it avoids leaking such sensitive information to the public LLM model.

5. Discussion

The CRISPR-GPT agent showcases the remarkable potential of LLMs in automating and enhancing the design process of complex biological experiments. By seamlessly integrating LLMs with domain knowledge, external tools, and a modular task execution system, CRISPR-GPT empowers researchers to navigate the intricate landscape of CRISPR gene editing experiments with unprecedented ease and efficiency. The multi-modal capabilities of CRISPR-GPT encompass meta-task pipelines, interactive prompts, and on-demand Q&A support. Researchers can leverage the agent's expertise to plan and execute gene editing experiments, from CRISPR system selection and guide RNA design to automated drafting of detailed protocol and validation strategies. This streamlined workflow not only accelerates the design process but also mitigates the risk of errors and oversights, thereby enhancing the quality and reproducibility of research outcomes.

While there exist LLM agent in other scientific domains such as chemistry, the complexities of biological experiments involving living materials demand a distinct set of considerations. Unlike chemical reactions, which often follow well-defined protocols, biological experiments require intricate procedures that account for the dynamic nature of living systems. CRISPR-GPT addresses this challenge by providing detailed, step-by-step guidance tailored to the specific experimental context, ensuring that researchers can navigate the nuances of working with living cells and organisms effectively.

Moreover, the free-style prompting and ad hoc Q&A capabilities of CRISPR-GPT set it apart from many existing agents. Researchers can pose unstructured queries and receive contextualized responses, facilitating a more natural and intuitive interaction with the agent. This feature is valuable in the face of unexpected challenges or unforeseen circumstances that may arise during the course of an experiment, enabling researchers to seek timely guidance and adapt their approach as needed.

Despite its impressive capabilities, CRISPR-GPT is not without limitations. While the agent can design individual components, such as guide RNAs and primers, it currently lacks the ability to generate complete constructs or vectors from natural language input. This limitation highlights an area for future development. For example, recent advancements in the field of modular design of gene-editing, such as FragMID[48] could be integrated with CRISPR-GPT to realize the potential for LLMs to empower researchers to explore and optimize CRISPR design and customized strategies, leading to more efficient gene-editing.

Looking ahead, the integration of CRISPR-GPT with automated laboratory platforms and robotics holds immense promise. By bridging computational design and physical execution, researchers could leverage the agent's expertise to orchestrate end-to-end automated experiments, minimizing manual intervention and accelerating the pace of discovery.

Reference

1. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**, 819–823 (2013).
2. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
3. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
4. Jinek, M. *et al.* RNA-programmed genome editing in human cells. *eLife* **2**, e00471 (2013).
5. Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096 (2014).

6.Mali, P., Esvelt, K. M. & Church, G. M. Cas9 as a versatile tool for engineering biology. *Nat. Methods* **10**, 957–963 (2013).

7. Hsu, P. D., Lander, E. S. & Zhang, F. Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* **157**, 1262–1278 (2014).

8. Sander, J. D. & Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* **32**, 347–355 (2014).

9. Shalem, O., Sanjana, N. E. & Zhang, F. High-throughput functional genomics using CRISPR–Cas9. *Nat. Rev. Genet.* **16**, 299–311 (2015).

10. Anzalone, A. V., Koblan, L. W. & Liu, D. R. Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* **38**, 824–844 (2020).

11. Pickar-Oliver, A. & Gersbach, C. A. The next generation of CRISPR–Cas technologies and applications. *Nat. Rev. Mol. Cell Biol.* **20**, 490–507 (2019).

12. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152**, 1173–1183 (2013).

13. Gilbert, L. A. *et al.* CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* **154**, 442–451 (2013).

14. Konermann, S. *et al.* Optical Control of Mammalian Endogenous Transcription and Epigenetic States. *Nature* **500**, 472–476 (2013).

15. Maeder, M. L. *et al.* CRISPR RNA-guided activation of endogenous human genes. *Nat. Methods* **10**, 977–979 (2013).

16. Perez-Pinera, P. *et al.* RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* **10**, 973–976 (2013).

17. Anzalone, A. V. *et al.* Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019).

18. Gaudelli, N. M. *et al.* Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**, 464–471 (2017).

19. Chowdhery, A. *et al.* PaLM: Scaling Language Modeling with Pathways. *J. Mach. Learn. Res.* **24**, 1–113 (2023).

20. Hoffmann, J. *et al.* Training Compute-Optimal Large Language Models. Preprint at https://doi.org/10.48550/arXiv.2203.15556 (2022).

21. OpenAI *et al.* GPT-4 Technical Report. Preprint at https://doi.org/10.48550/arXiv.2303.08774 (2024).

22. Anthropic. The Claude 3 Model Family: Opus, Sonnet, Haiku. (2024).

23. Anil, R. *et al.* PaLM 2 Technical Report. Preprint at https://doi.org/10.48550/arXiv.2305.10403 (2023).

24. Yao, S. *et al.* ReAct: Synergizing Reasoning and Acting in Language Models. in (2022).

25. Liu, R. *et al.* Mind's Eye: Grounded Language Model Reasoning through Simulation. in (2022).

26. Schick, T. *et al.* Toolformer: Language Models Can Teach Themselves to Use Tools. in (2023).

27. Cai, T., Wang, X., Ma, T., Chen, X. & Zhou, D. Large Language Models as Tool Makers. Preprint at https://doi.org/10.48550/arXiv.2305.17126 (2024).

28. Yang, C. *et al.* Large Language Models as Optimizers. Preprint at https://doi.org/10.48550/arXiv.2309.03409 (2023).

29. Li, T. *et al.* CancerGPT for few shot drug pair synergy prediction using large pretrained language models. *Npj Digit. Med.* **7**, 1–10 (2024).

30. Wu, S. *et al.* BloombergGPT: A Large Language Model for Finance. Preprint at https://doi.org/10.48550/arXiv.2303.17564 (2023).

31. Bran, A. M. *et al.* ChemCrow: Augmenting large-language models with chemistry tools. Preprint at https://doi.org/10.48550/arXiv.2304.05376 (2023).

32. Boiko, D. A., MacKnight, R., Kline, B. & Gomes, G. Autonomous chemical research with large language models. *Nature* **624**, 570–578 (2023).

33. Kim, H. K. *et al.* Deep learning improves prediction of CRISPR–Cpf1 guide RNA activity. *Nat. Biotechnol.* **36**, 239–241 (2018).

34. DeWeirdt, P. C. *et al.* Optimization of AsCas12a for combinatorial genetic screens in human cells. *Nat. Biotechnol.* **39**, 94–104 (2021).

35. Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).

36. Sanson, K. R. *et al.* Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nat. Commun.* **9**, 5416 (2018).

37. ChatGPT plugins. https://openai.com/blog/chatgpt-plugins.

38. Shen, Y. *et al.* HuggingGPT: Solving AI Tasks with ChatGPT and its Friends in Hugging Face. Preprint at https://doi.org/10.48550/arXiv.2303.17580 (2023).

39. Wang, G. *et al.* Voyager: An Open-Ended Embodied Agent with Large Language Models. Preprint at https://doi.org/10.48550/arXiv.2305.16291 (2023).

40. Hao, S., Liu, T., Wang, Z. & Hu, Z. ToolkenGPT: Augmenting Frozen Language Models with Massive Tools via Tool Embeddings. Preprint at https://doi.org/10.48550/arXiv.2305.11554 (2024).

41. Patil, S. G., Zhang, T., Wang, X. & Gonzalez, J. E. Gorilla: Large Language Model Connected with Massive APIs. Preprint at https://doi.org/10.48550/arXiv.2305.15334 (2023).

42. Untergasser, A. *et al.* Primer3—new capabilities and interfaces. *Nucleic Acids Res.* **40**, e115 (2012).

43. Yao, S. *et al.* ReAct: Synergizing Reasoning and Acting in Language Models. Preprint at https://doi.org/10.48550/arXiv.2210.03629 (2023).

44. Wei, J. *et al.* Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. Preprint at https://doi.org/10.48550/arXiv.2201.11903 (2023).

45. Zhou, D. *et al.* Least-to-Most Prompting Enables Complex Reasoning in Large Language Models. Preprint at https://doi.org/10.48550/arXiv.2205.10625 (2023).

46. Lander, E. S. *et al.* Adopt a moratorium on heritable genome editing. *Nature* **567**, 165–168 (2019).

47. Rights (OCR), O. for C. Summary of the HIPAA Privacy Rule. https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html (2008).

48. McGee, A. V. *et al.* Modular vector assembly enables rapid assessment of emerging CRISPR technologies. *BioRxiv Prepr. Serv. Biol.* 2023.10.25.564061 (2023) doi:10.1101/2023.10.25.564061.

49. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).

50. Cancellieri, S., Canver, M. C., Bombieri, N., Giugno, R. & Pinello, L. CRISPRitz: rapid, high-throughput and variant-aware in silico off-target site identification for CRISPR genome editing. *Bioinformatics* **36**, 2001–2008 (2020).

51. Giuliano, C. J., Lin, A., Girish, V. & Sheltzer, J. M. Generating Single Cell–Derived Knockout Clones in Mammalian Cells with CRISPR/Cas9. *Curr. Protoc. Mol. Biol.* **128**, e100 (2019).

52. Du, D. & Qi, L. S. CRISPR Technology for Genome Activation and Repression in Mammalian Cells. *Cold Spring Harb. Protoc.* **2016**, pdb.prot090175 (2016).

53. Huang, T. P., Newby, G. A. & Liu, D. R. Precision genome editing using cytosine and adenine base editors in mammalian cells. *Nat. Protoc.* **16**, 1089–1128 (2021).

54. Hanna, R. E. *et al.* Massively parallel assessment of human variants with base editor screens. *Cell* **184**, 1064-1080.e20 (2021).

55. Doman, J. L., Sousa, A. A., Randolph, P. B., Chen, P. J. & Liu, D. R. Designing and executing prime editing experiments in mammalian cells. *Nat. Protoc.* **17**, 2431–2468 (2022).

56. Hwang, G.-H. *et al.* PE-Designer and PE-Analyzer: web-based design and analysis tools for CRISPR prime editing. *Nucleic Acids Res.* **49**, W499–W504 (2021).

57. Chow, R. D., Chen, J. S., Shen, J. & Chen, S. A web tool for the design of prime-editing guide RNAs. *Nat. Biomed. Eng.* **5**, 190–194 (2021).

58. Morris, J. A., Rahman, J. A., Guo, X. & Sanjana, N. E. Automated design of CRISPR prime editors for 56,000 human pathogenic variants. *iScience* **24**, 103380 (2021).

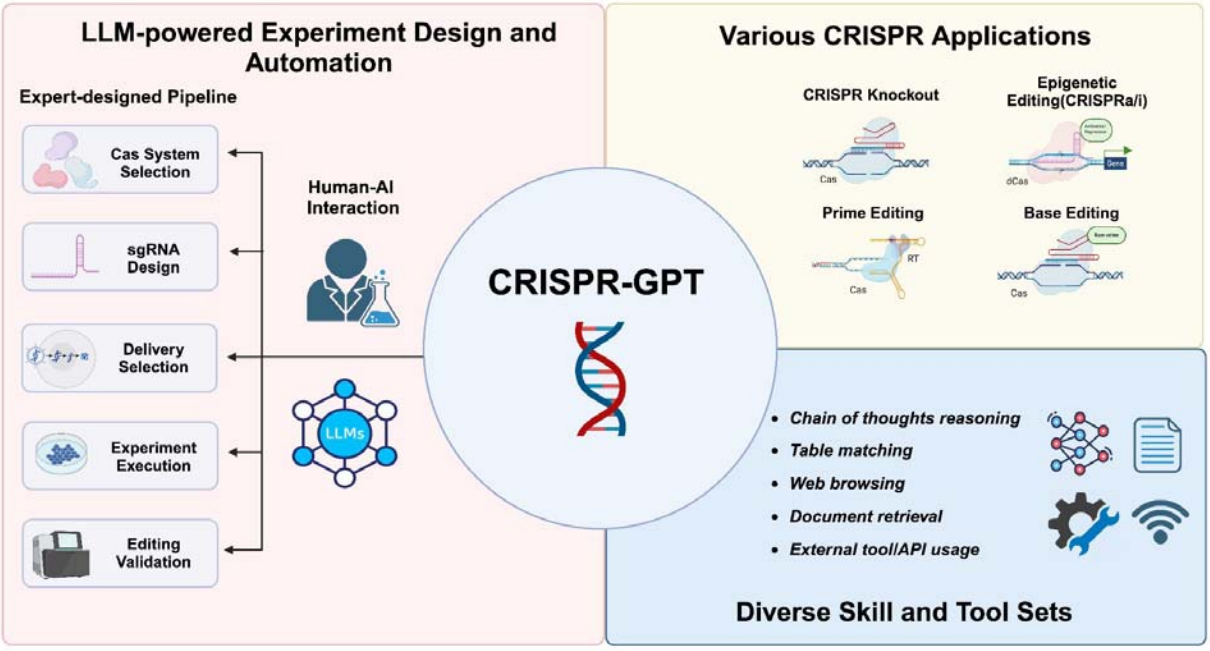

Figure 1: Overview of CRISPR-GPT Agent. The CRISPR-GPT is built upon an LLM-powered design and planning engine (left), which helps to complete 4 core meta-tasks (top right), as well as other auxiliary functions (freestyle Q&A, off-target prediction). CRISPR-GPT has integrated a set of useful skills and toolkits (bottom right) that the LLM agent would call when needed to facilitate human users across the different tasks and subtasks. Figure created with BioRender.com.

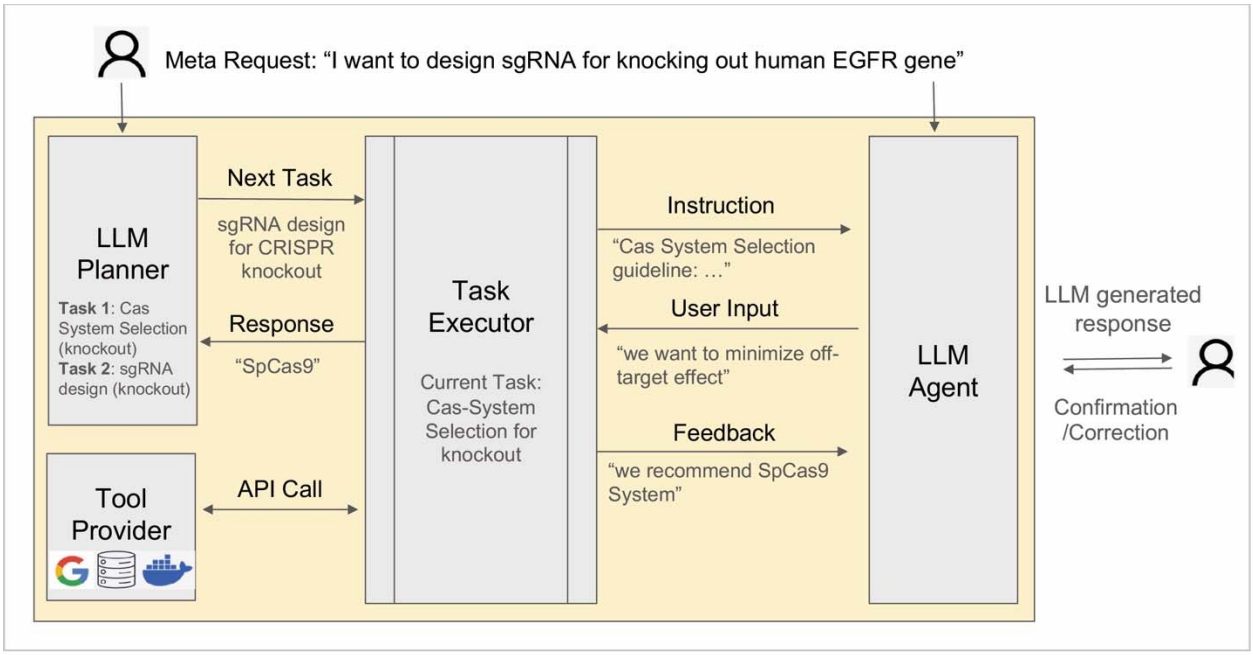

Figure 2. Components of CRISPR-GPT enable human-AI collaboration to automate gene-editing experimental designs across complex tasks. LLM Planner is responsible for configuring tasks based on the user's needs (4 predefined meta-tasks or LLM planned chain of tasks). **Tool Provider** connects the system to external APIs, tools, libraries, and documents. **Task Executor** is implemented as a state machine, responsible for providing instructions and feedback, receiving input from LLM Agent, and calling APIs via Tool Provider. **LLM Agent** is responsible for interacting with the task executor on behalf of the user, where the user can monitor the process and provide correction to the LLM agent if the generated content.

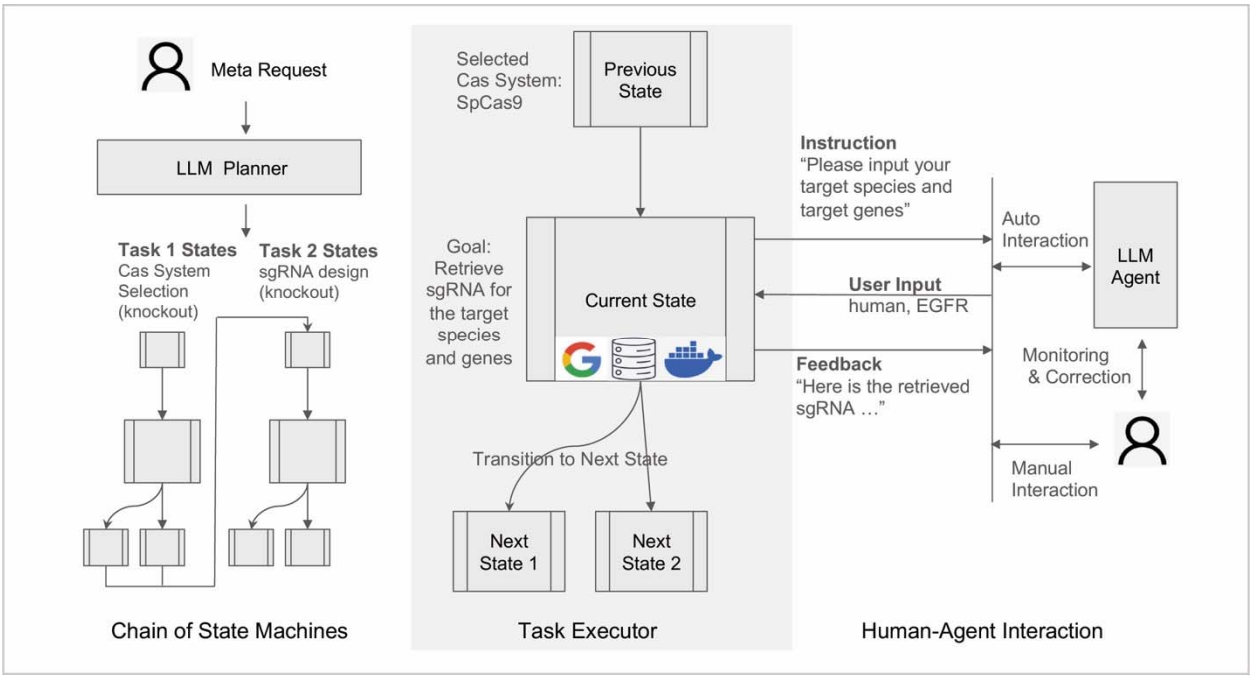

Figure 3. Task decomposition process and state machine implementation algorithm. (Left) Task decomposition; The LLMs can automatically perform task decomposition based on the user's request, the descriptions of the current supported tasks and the dependencies, and the internal knowledge inside the LLMs. The state machines of the selected tasks are chained together to fulfill the user's request. **(Right) State machines & LLM Agent; State machines** are the core of the Task Executor, where each state is responsible for one round of interaction with the user. The instruction is provided to the user first with sufficient information for the current decision-making step and the required inputs. After receiving the response from the user, it provides output and feedback, where APIs (e.g. program execution/web search/database retrieval) are potentially called during the execution of the state. Afterward, the state machine transits to the next state. **LLM Agent** generates responses to every step of the state machine on behalf of the user. The user monitors the whole process and provides corrections if the generated content is wrong or overrides the LLM Agent and manually interacts with the Task Executor.

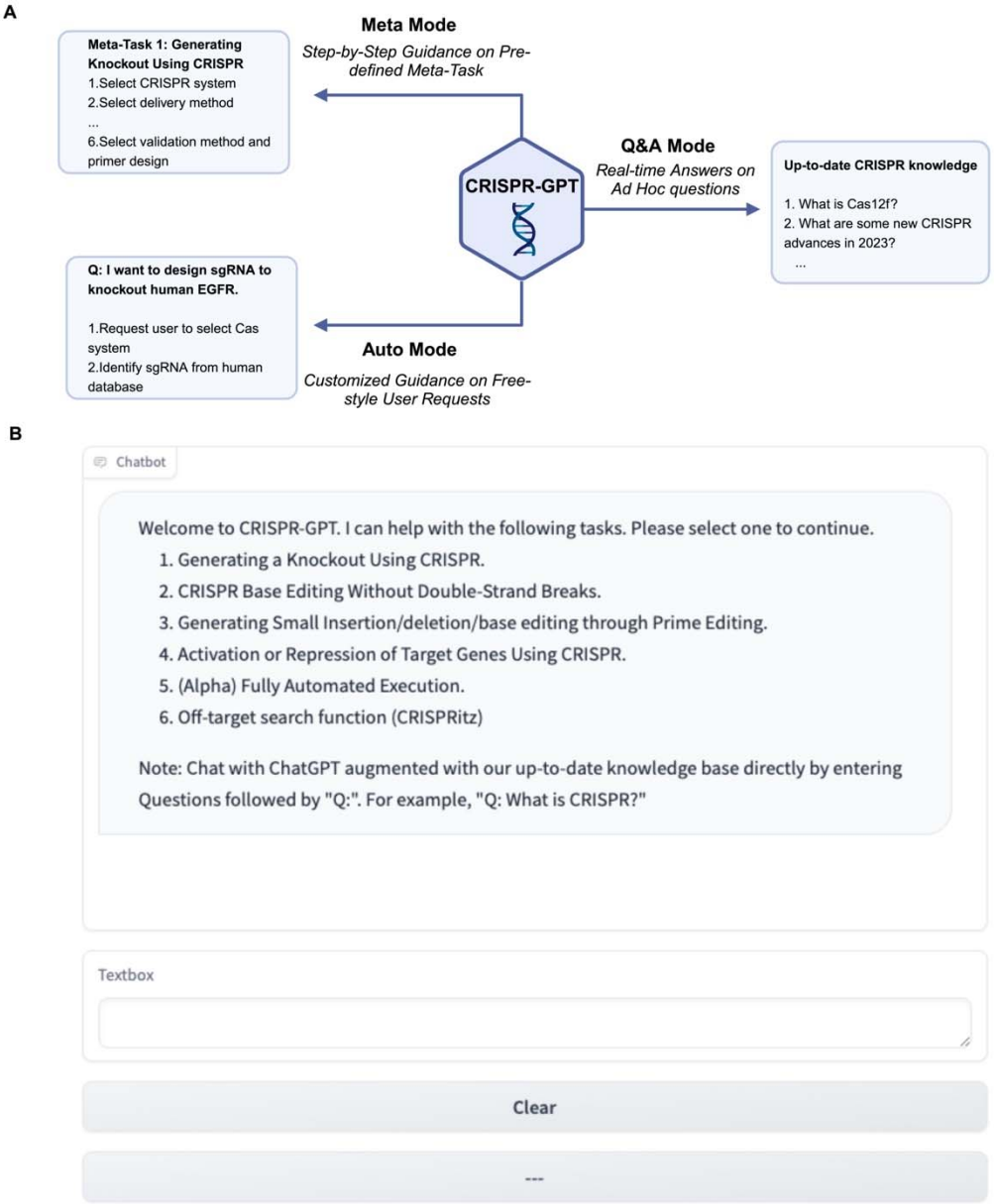

Figure 4. Overview of CRISPR-GPT's interactive modules for gene-editing experimental design. (**A**) schematics illustrating the functionalities of the three modules within CRISPR-GPT, accompanied by examples of their applications. (**B**) Web-interface of CRISPR-GPT, note No.1-4 is the "Meta Mode", No.5 is the "Auto Mode", No.6 is Off-target-prediction function and "Q: prompt would trigger the Q&A mode".

Meta-Tasks	Gene editing scenarios	Individual Design Tasks
CRISPR Knockout	Single/multiple genes knockout, deletion of gene fragments	CRISPR/Cas system selection[49]
		Delivery method selection
		sgRNA design for knockout[33–36]
		off-target evaluation[50]
		experimental protocol recommendation[51]
		validation protocol recommendation and primer design for sequencing[42,51]
CRISPR activation/interference	Gene activation and repression	CRISPR/Cas Activation/Interference system selection[52]
		Delivery method selection
		sgRNA design for activation/interference[33–36]
		off-target evaluation[50]
		experimental protocol recommendation[51]
		validation protocol recommendation and primer design for qPCR[42,51]
CRISPR Base Editing	Single base replacement from CG to AT or AT to CG and broad mutagenesis	Base editing system selection[53]
		Delivery method selection
		sgRNA design for base editing[54]
		off-target evaluation[50]
		experimental protocol recommendation[51]
		validation protocol recommendation and primer design for sequencing[42,51]
CRISPR Prime Editing	Small fragment insertion, replacement, and deletion	Prime editing system selection[55]
		Delivery method selection
		pegRNA design for prime editing[32,56–58]
		off-target evaluation[50]
		experimental protocol recommendation[51]
		validation protocol recommendation and primer design for sequencing[42,51]

Table 1. List of meta-mode tasks (4 major meta-task and 22 specific tasks)

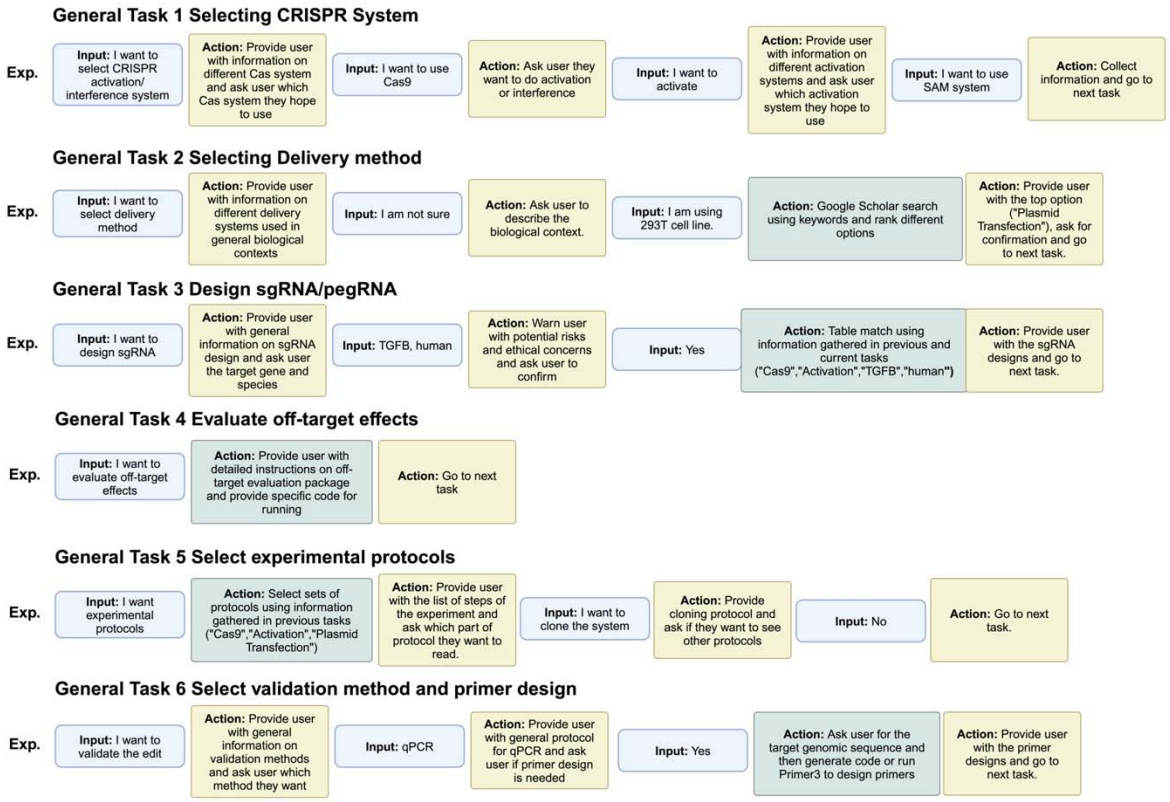

Figure 5. Example workflows outlining the general tasks involved in gene-editing experimental designs as facilitated by CRISPR-GPT.

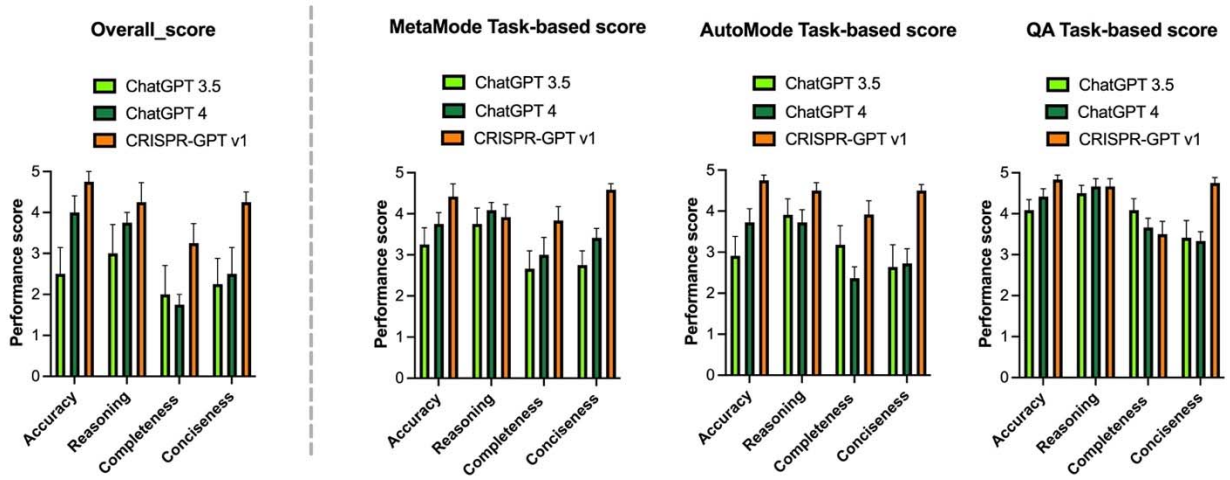

Figure 6. Evaluation results showing comparative performance of CRISPR-GPT and ChatGPT 3.5/4.0 in a range of gene-editing experiment design tasks across three different modes: MetaMode, AutoMode and QAMode. All tasks are scored in four different aspects: accuracy, reasoning, completeness, and conciseness. A score of 1 represents poor performance and 5 represents excellent performance. Detailed rubrics are listed in the **Supplementary Note 3**.

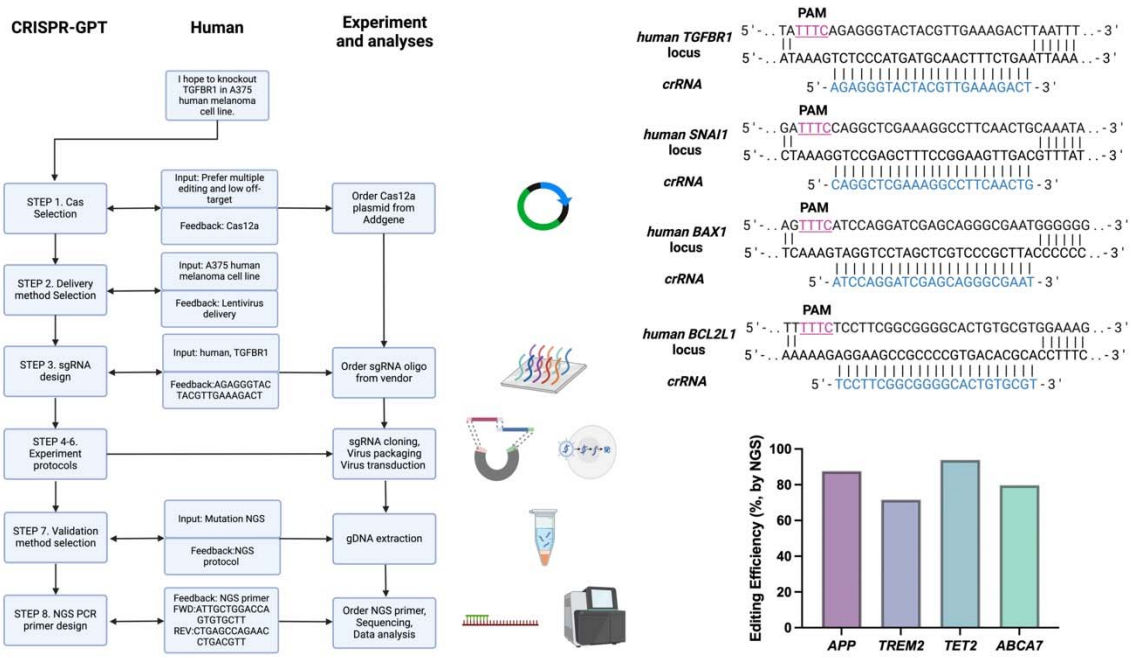

Figure 7. Wet-lab demonstration of human-AI collaboration in performing a gene-knockout experiment. (**Left**) Schematic showing the workflow of human-AI interaction. (**Top right**) Sample sgRNAs designed through CRISPR-GPT. (**Bottom right**) Editing outcome from next generation sequencing.